January 2015

A Thorough and Accessible Proof of the Erdo˝s-Kac Theorem Following Granville and Soundararajan

Thomas John Scheithauer
Eastern Kentucky University

A Thorough and Accessible Proof of the Erdős-Kac Theorem Following Granville and Soundararajan

By

T.J. Scheithauer

Thesis Approved:

Dr. Jason Gibson
Chair, Advisory Committee

Dr. Lisa W. Kay
Member, Advisory Committee

Dr. Rachel Bishop-Ross
Member, Advisory Committee

Dr. Jerry J. Pogatshnik
Dean, Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for an M.A. degree at Eastern Kentucky University, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made. Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or, in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature T. J. Schetthauer

Date 8-17-2015

A Thorough and Accessible Proof of the Erdős-Kac Theorem Following Granville and Soundararajan

By

Thomas J. Scheithauer

Bachelor Of Science, Math and Physics
University of the Cumberlands
Williamsburg, Kentucky
2005

Master of Arts in Teaching, Secondary Mathematics
University of the Cumberlands
Williamsburg, Kentucky
2007

Submitted to the Faculty of the Graduate School of
Eastern Kentucky University
in partial fulfillment of the requirements
for the degree of
MASTER OF ARTS
December, 2015

DEDICATION

I dedicate this thesis to my family for their unending love, patience, and support. Without them, I would never have completed this thesis.

I also dedicate this to all of the devoted and selfless teachers and professors who use their seemingly limitless patience, talents, and compassion to improve and inspire others.

ACKNOWLEDGMENTS

I am grateful and immensely indebted to several people who helped me complete this thesis.

Firstly, I would like to thank my advisor Dr. Jason Gibson. He selflessly and generously gave of his time and energy throughout this entire process. I've never met a more patient, kind, or knowledgeable educator and mentor. I'm grateful to have worked with him on this thesis. His abilities to teach and inspire were absolutely vital in my finishing this work.

I'm grateful to Dr. Lisa Kay and Dr. Rachel Bishop-Ross for agreeing to be on my review committee. I would like to further thank Dr. Lisa Kay for her constant encouragement and confidence in my abilities.

I would like to thank Dr. Jeffrey Neugebauer for selflessly helping me understand the underlying principles involved with some of the technical lemmata and for boosting my confidence in my abilities.

Finally, I would like to thank the Department of Mathematics and Statistics and the University of Eastern Kentucky for creating a program that allowed me to advance my education while maintaining full-time work as a math and physics teacher.

ABSTRACT

The Erdő-Kac Theorem states that, as n tends to infinity, the distribution of $\omega(n)$, the number of distinct prime divisors of n, becomes normally distributed with mean and variance $\log \log n$. Granville and Soundararajan gave a proof of the Erdős-Kac Theorem that avoided many of the specialized techniques present in several earlier approaches. This thesis includes considerable detail, prerequisite theorems, and instructive background in order to provide a self-contained exposition of the proof of Granville and Soundararajan.

Contents

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 An Introduction To The Erdős-Kac Theorem

The Erdős-Kac Theorem is a prime example of the surprising beauty in our world that only mathematics can discover. The theorem is simple enough that anyone with a basic understanding of the normal distribution can understand what the theorem states. However, the Erdős-Kac Theorem is also completely unintuitive. Its implications are profound and even bizarre in nature. The main purpose of this thesis is to present a highly accessible and reasonably self-contained proof of the Erdős-Kac Theorem. By including prerequisite theorems and proofs and significant detail in the proof of the Erdős-Kac Theorem, this thesis is designed to be instructive in nature.

Let $\omega(n)$ represent the number of unique prime divisors of the natural number n. For example,

$$\omega(150) = \omega(2 \cdot 3 \cdot 5^2) = 3, \text{ and}$$
$$\omega(5148) = \omega(2^2 \cdot 3^2 \cdot 11 \cdot 13) = 4.$$

The Erdős-Kac Theorem says that, as n becomes sufficiently large, the distribution of $\omega(n)$ is closely approximated by the normal distribution. Loosely speaking, the expected value of $\omega(n)$ is given by $E[\omega(n)] = \log\log n$, and the variance of $\omega(n)$ is given by $V[\omega(n)] = \log\log n$. Theorem 1.1 below is the Erdős-Kac Theorem [5].

Theorem 1.1. ***(The Erdős-Kac Theorem)***

Let $\tau, x \in \mathbb{R}$. Let $n \in \mathbb{N}$. Let $\omega(n)$ be the number of unique prime divisors of n. Then,

$$\lim_{x\to\infty} \frac{1}{x} \cdot \#\left\{ n : 3 \le n \le x, \text{ and } \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \le \tau \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\tau} \exp^{-t^2/2}\, dt. \tag{1.1}$$

1.2 The Erdős-Kac Theorem In Action

To best appreciate the implications of the Erdős-Kac Theorem, we consider some concrete examples. The data for these examples were generated by SAGE [13], and the graphs and tables were created using Microsoft Excel 2013. The variances in this section were calculated using the "var.p" command in Microsoft Excel 2013.

Example 1.1. This first example analyzes the distribution of $\omega(n)$ for n satisfying $1000 \le n \le 1500$. The values for $\omega(n)$ were generated by entering the code

```
n = 10^3
while n <= 10^3 + 500:
len(prime_divisors(n))
n += 1
```

on `cloud.sagemath.org` [13]. Table 1.1 and Figure 1.1 below summarize the results. Table 1.2 compares the mean and variance to the values predicted by the Erdős-Kac Theorem (EKT).

Table 1.1: The distribution of $\omega(n)$ for $1000 \leq n \leq 1500$

$\omega(n)$	Frequency	Relative Frequency
1	74	0.148
2	224	0.448
3	176	0.352
4	26	0.052

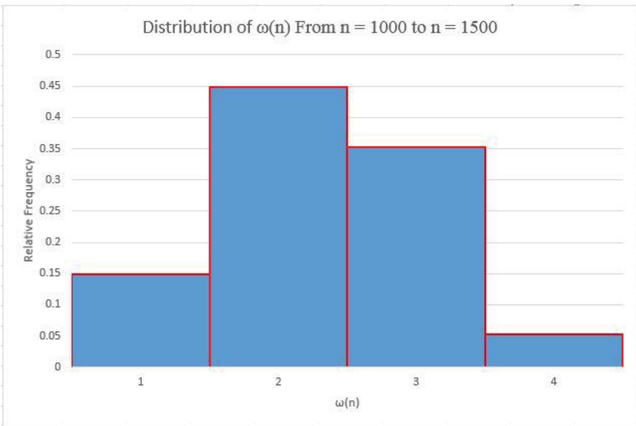

Figure 1.1: The distribution of $\omega(n)$ for $1000 \leq n \leq 1500$

Table 1.2: Comparing the mean and variance of $\omega(n)$ to the EKT predictions for $1000 \leq n \leq 1500$

Mean	2.308
Variance	0.613
$\log\log 1500$	1.990

As seen in Figure 1.1, the distribution doesn't look normal, but the relative frequency is highest for the two central values of $\omega(n)$. The EKT claims the normal distribution more accurately models the distribution of $\omega(n)$ as n increases. For this example, small n were used so the large relative discrepancies between the mean and variance to $\log\log 1500$ are not surprising.

Example 1.2. This example analyzes the distribution of $\omega(n)$ for n satisfying $10^6 \leq n \leq (10^6 + 10^3)$. The values for $\omega(n)$ were generated by entering the code

```
n = 10^6
while n <= 10^6 + 10^3:
len(prime_divisors(n))
n += 1
```

on `cloud.sagemath.org` [13]. Table 1.3 and Figure 1.2 below summarize the results. Table 1.4 compares the mean and variance to the values predicted by the EKT.

Table 1.3: The distribution of $\omega(n)$ for $10^6 \leq n \leq (10^6 + 10^3)$

$\omega(n)$	Frequency	Relative Frequency
1	75	0.075
2	268	0.268
3	371	0.371
4	227	0.227
5	58	0.058
6	2	0.002

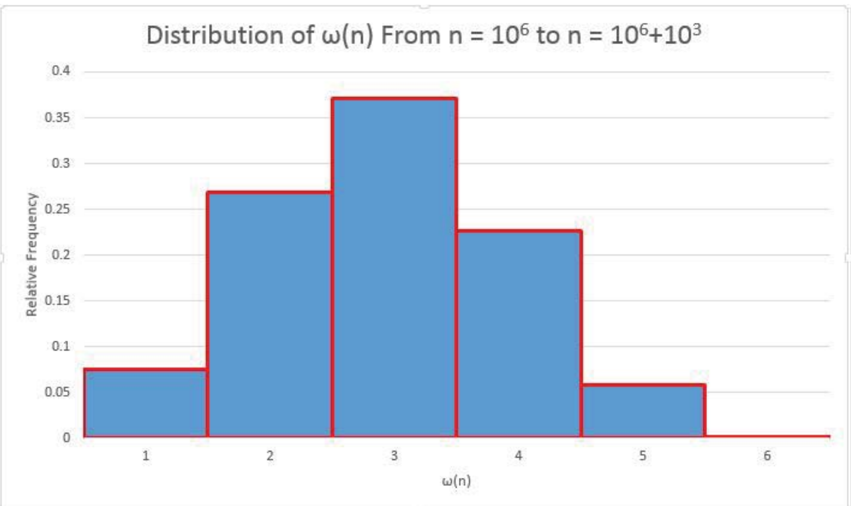

Figure 1.2: The distribution of $\omega(n)$ for $10^6 \leq n \leq (10^6 + 10^3)$

Table 1.4: Comparing the mean and variance of $\omega(n)$ to the EKT predictions for $10^6 \leq n \leq (10^6 + 10^3)$

Mean	2.931
Variance	1.039
$\log\log(10^6 + 10^3)$	2.626

Figure 1.2 is more representative of the signature "bell curve" of the normal distribution. Notice that the discrepancy between the mean and $\log\log(10^6 + 10^3)$ is smaller than in the previous example. However, the variance discrepancy is still somewhat large.

The following examples demonstrate how the distribution of $\omega(n)$ becomes "more normal" as n increases. Due to computational difficulties, the sample sizes for the two largest n are somewhat small. Subsequent commentary is not included in these examples.

Example 1.3. This example analyzes the distribution of $\omega(n)$ for n satisfying $1.5 \times 10^{12} \leq n \leq 1.5 \times 10^{12} + 10^4$. The values for $\omega(n)$ were generated by entering the code

```
n = 1.5*10^12
while n <= 1.5*10^12+ 10^4:
len(prime_divisors(n))
n += 1
```

on `cloud.sagemath.org` [13]. Table 1.5 and Figure 1.3 below summarize the results. Table 1.6 compares the mean and variance to the values predicted by the EKT.

Table 1.5: The distribution of $\omega(n)$ for $1.5 \times 10^{12} \leq n \leq 1.5 \times 10^{12} + 10^4$

$\omega(n)$	Frequency	Relative Frequency
1	373	0.0373
2	1648	0.1648
3	3034	0.3034
4	2817	0.2817
5	1578	0.1578
6	458	0.0458
7	83	0.0083
8	9	0.0009
9	1	0.0001

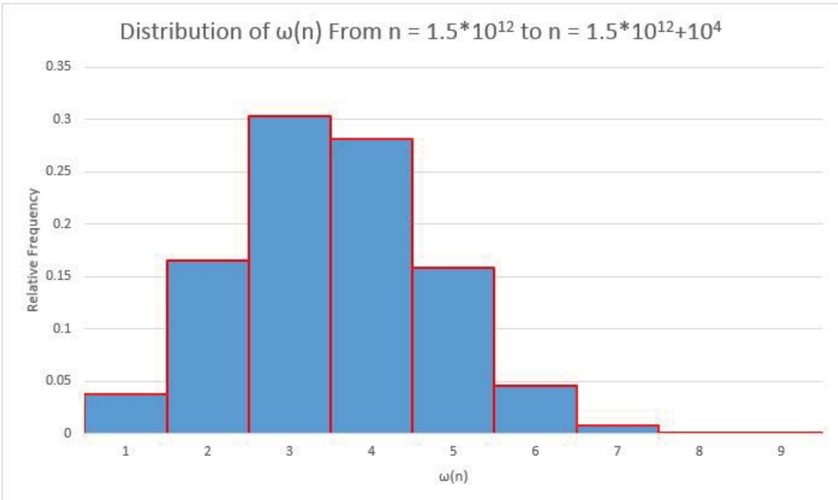

Figure 1.3: The distribution of $\omega(n)$ for $1.5 \times 10^{12} \leq n \leq 1.5 \times 10^{12} + 10^4$

Table 1.6: Comparing the mean and variance of $\omega(n)$ to the EKT predictions for $1.5 \times 10^{12} \leq n \leq 1.5 \times 10^{12} + 10^4$

Mean	3.534
Variance	1.513
$\log\log(10^{12} + 10^4)$	3.333

Example 1.4. This example analyzes the distribution of $\omega(n)$ for n satisfying $2 \times 10^{18} \leq n \leq 2 \times 10^{18} + 5 \times 10^4$. The values for $\omega(n)$ were generated by entering the code

```
n = 2*10^18
while n <= 2*10^18 + 5*10^4:
len(prime_divisors(n))
n += 1
```

on `cloud.sagemath.org` [13]. Table 1.7 and Figure 1.4 below summarize the results. Table 1.8 compares the mean and variance to the values predicted by the EKT.

Table 1.7: The distribution of $\omega(n)$ for $2 \times 10^{18} \leq n \leq (2 \times 10^{18} + 5 \times 10^4)$

$\omega(n)$	Frequency	Relative Frequency
1	1191	0.02382
2	5756	0.11512
3	11594	0.23188
4	13880	0.27760
5	10331	0.20662
6	5098	0.10196
7	1668	0.03376
8	394	0.00788
9	63	0.00126
10	6	0.00012

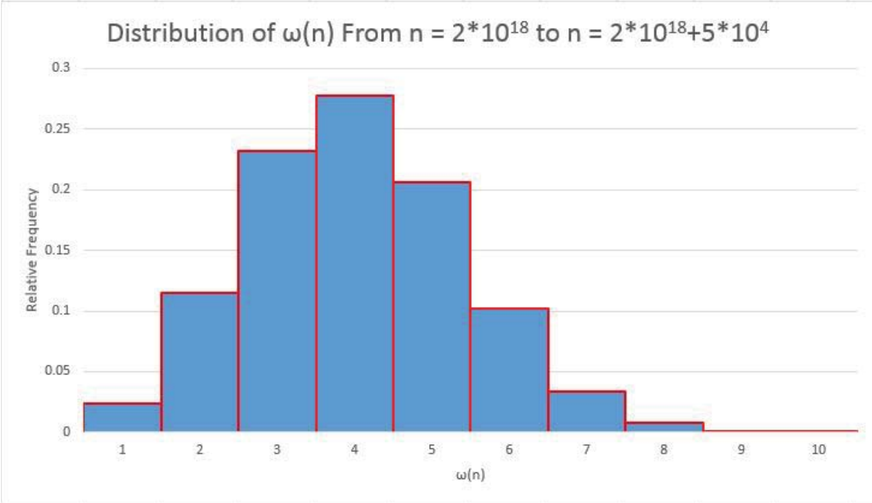

Figure 1.4: The distribution of $\omega(n)$ for $2 \times 10^{18} \leq n \leq (2 \times 10^{18} + 5 \times 10^4)$

Table 1.8: Comparing the mean and variance of $\omega(n)$ to the EKT predictions for $2 \times 10^{18} \leq n \leq (2 \times 10^{18} + 5 \times 10^4)$

Mean	4.017
Variance	1.987
$\log\log(2 \times 10^{18} + 5 \times 10^4)$	3.741

Example 1.5. This example analyzes the distribution of $\omega(n)$ for n satisfying $10^{50} \leq n \leq (10^{50} + 5 \times 10^3)$. The values for $\omega(n)$ were generated by entering the code

```
n = 10^50
while n <= 10^50 + 5*10^3:
len(prime_divisors(n))
n += 1
```

on `cloud.sagemath.org` [13]. Due to the computations timing out, the data for this example were generated 250 points at a time. Table 1.9 and Figure 1.5 below summarize the results. Table 1.10 compares the mean and variance to the values predicted by the EKT.

Table 1.9: The distribution of $\omega(n)$ for $10^{50} \leq n \leq (10^{50} + 5 \times 10^3)$

$\omega(n)$	Frequency	Relative Frequency
1	50	0.0100
2	241	0.0482
3	646	0.1292
4	1061	0.2122
5	1177	0.2354
6	892	0.1784
7	538	0.1076
8	250	0.0500
9	107	0.0214
10	33	0.0066
11	6	0.0012

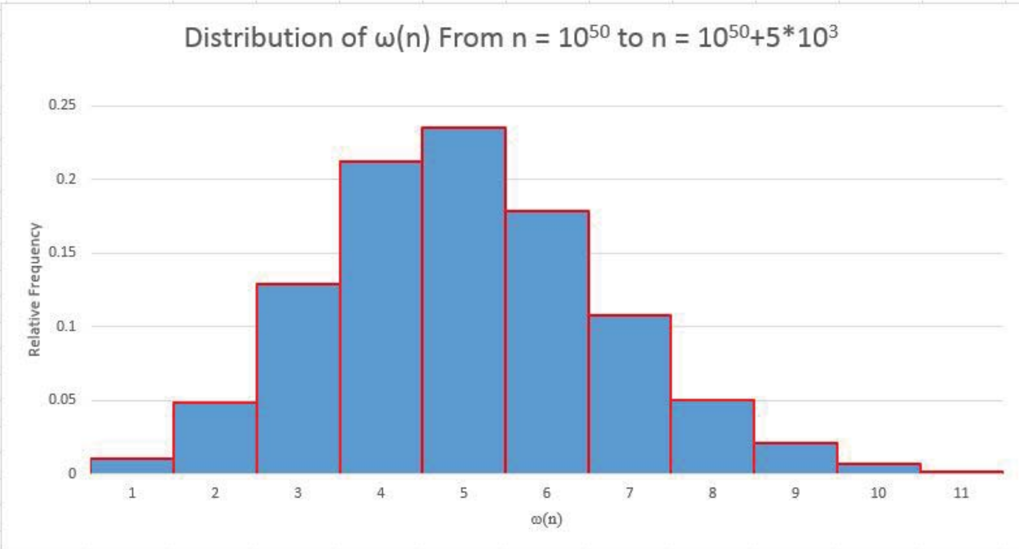

Figure 1.5: The distribution of $\omega(n)$ for $10^{50} \leq n \leq (10^{50} + 5 \times 10^3)$

Table 1.10: Comparing the mean and variance of $\omega(n)$ to the EKT predictions for for $10^{50} \leq n \leq (10^{50} + 5 \times 10^3)$

Mean	5.014
Variance	2.931
$\log\log(10^{50} + 5 \times 10^3)$	4.746

It is the hope of the author that the reader better appreciates what the EKT states. The fact that the distribution of $\omega(n)$ can be closely modeled by the normal distribution with a mean and variance that are calculated with relative ease is not intuitive. What is perhaps more bizarre is that, for large n, $\omega(n)$ is normally distributed regardless of n, and the methods used to calculated the mean and variance do not change.

1.3 The Structure Of The Remainder Of This Thesis

The main result of this thesis is a proof of the Erdős-Kac Theorem that follows Granville and Soundararajan [6]. Chapter 2 establishes theorems and lemmas that are needed in proving key results in the following chapters. The material in Chapter 2 is more broad in nature when compared to Chapters 3 and 4. Chapter 3 focuses on the normal distribution. Understanding the normal distribution and some of its key properties is necessary to understand the EKT. Chapter 4 is similar in nature to Chapter 2, except that its scope is more narrow in nature, focusing on lemmas and theorems typically attributed to number theory. The two main results in Chapter 4 used in the main proofs of this thesis are Mertens' Second Theorem and the gcd-restricted sum lemma. Finally, Chapter 5 provides a proof of the EKT that follows Granville and Soundararajan [6], which hinges on a proposition. A proof of this proposition, which also follows Granville and Soundararajan [6], is also included.

1.4 Final Introductory Remarks

In 1939, the original proof of the Erdős-Kac Theorem (EKT) utilized sieve methods, which are difficult and not accessible to a relatively broad audience (see [5]). For a similar result that also employs sieve methods, see Djanković [4]. In 1953, Delange [3], who was unaware of Erdős and Kac's discovery, proved the EKT using moments. In 1955, Halberstam [7], [8] also used moments to generalize Delange's work to include a larger set of functions. As mentioned in Granville and Soundararajan [6], Halberstam's and others' works involving moments utilized the binomial expansion of $(\omega(n) - \log\log x)^k$. Once expanded, several main terms had to be collected and carefully canceled out. This process was long and tedious, creating a drudgery of technical calculations.

What makes Granville and Soundararajan's proof [6] unique is the introduction of the function $f_r(n)$ (see equations (5.3) and (5.52)). By the way $f_r(n)$ is defined, the sum $\sum_{n \leq x} f_r(n)$ is small unless r is square-full. The arising novelty is that the main term is found without the tedious nature of previous work using moments.

In comparing Chapter 5 of this thesis to Granville and Soundararajan [6], it becomes readily apparent that Granville and Soundararajan omitted an immense amount of detail. The MathSciNet review of Granville and Soundararajan [6] highlights this lack of detail and their seemingly incorrect handling of the problem's history as concerns (for more information, see the review of [6] by Y.-F.S. Pétermann in Mathematical Reviews). On the other hand, not every paper can be a full treatise on a topic. In championing brevity, Granville and Soundararajan [6] possibly made some mistakes in the handling of error terms. This thesis provides the background necessary and the missing detail needed to follow Granville and Soundararajan's proof of the EKT. While the proof is considered easy in a relative sense when compared to sieve methods and previous works involving moments, it is still highly involved and technical in nature.

Chapter 2

Technical Lemmata

The purpose of this chapter is to state and prove various lemmas and theorems that are used in subsequent intermediate proofs and in the proof of the Erdős-Kac Theorem (EKT).

2.1 The Cauchy-Schwarz Inequality

The Cauchy-Schwarz inequality exists in various forms and is one of the most powerful and useful inequalities in mathematics. Lemma 2.1 is vital in the proof of Proposition 5.1. Since the EKT only involves real numbers, the proof of Lemma 2.1 only considers real numbers.

Lemma 2.1. *Let $\{a_1, a_2, \ldots, a_n\}$ and $\{b_1, b_2, \ldots, b_n\}$ each be finite sequences of real numbers. Then the Cauchy-Schwarz inequality states*

$$a_1b_1 + a_2b_2 + \ldots + a_nb_n \leq \sqrt{a_1^2 + a_2^2 + \ldots + a_n^2}\sqrt{b_1^2 + b_2^2 + \ldots + b_n^2}. \tag{2.1}$$

Proof. This proof is adapted from Chapter 1 of Steele [12]. The proof proceeds by induction.

Let $n = 1$. Then,

$$a_1 b_1 \leq a_1 b_1. \tag{2.2}$$

This result is trivial and does not satisfactorily establish the basis step.

Let $n = 2$. By the non-negativity of the square function, it holds that

$$0 \leq (a_1 b_2 - a_2 b_1)^2. \tag{2.3}$$

Expanding the right-hand side along with algebraic manipulation results in

$$0 \leq (a_1 b_2 - a_2 b_1)^2 \tag{2.4}$$

$$0 \leq a_1^2 b_2^2 - 2a_1 b_1 a_2 b_2 + a_2^2 + b_1^2 \tag{2.5}$$

$$2a_1 b_1 a_2 b_2 \leq a_1^2 b_2^2 + a_2^2 + b_1^2 \tag{2.6}$$

$$a_1^2 b_1^2 + 2a_1 b_1 a_2 b_2 + a_2^2 b_2^2 \leq a_1^2 b_1^2 + a_1^2 b_2^2 + a_2^2 b_1^2 + a_2^2 b_2^2 \tag{2.7}$$

$$(a_1 b_1 + a_2 b_2)^2 \leq \left(a_1^2 + a_2^2\right)\left(b_1^2 + b_2^2\right) \tag{2.8}$$

$$a_1 b_1 + a_2 b_2 \leq \sqrt{a_1^2 + a_2^2}\sqrt{b_1^2 + b_2^2}. \tag{2.9}$$

Thus the Cauchy-Schwarz inequality is true for $n = 2$, establishing the basis step.

For the induction step, suppose that

$$a_1 b_1 + a_2 b_2 + ... + a_m b_m \leq \sqrt{a_1^2 + a_2^2 + ... + a_m^2}\sqrt{b_1^2 + b_2^2 + ... + b_m^2}. \tag{2.10}$$

Consider $n = m + 1$. Grouping the first m terms in each sum and applying the induction hypothesis yields

$$(a_1 b_1 + a_2 b_2 + ... + a_m b_m) + a_{m+1} b_{m+1} \leq \sqrt{a_1^2 + a_2^2 + ... + a_m^2}\sqrt{b_1^2 + b_2^2 + ... + b_m^2} + (a_{m+1} b_{m+1}). \tag{2.11}$$

Define new variables α_n and β_n in order to apply the result from when $n = 2$ to inequality (2.11). Let α_m and β_m be defined by

$$\alpha_n = \sqrt{a_1^2 + a_2^2 + ... + a_m^2}$$
$$\beta_n = \sqrt{b_1^2 + b_2^2 + ... + b_m^2}.$$

Inequality (2.11) now takes the form

$$(a_1b_1 + a_2b_2 + ... + a_mb_m) + a_{m+1}b_{m+1} \leq \alpha_m\beta_m + a_{m+1}b_{m+1}. \tag{2.12}$$

The case when $n = 2$ established that

$$\alpha_n\beta_n + a_{m+1}b_{m+1} \leq \sqrt{\alpha_m^2 + a_{m+1}^2}\sqrt{\beta_m^2 + b_{m+1}^2}. \tag{2.13}$$

Combining inequalities (2.11) and (2.12) establishes that the Cauchy-Schwarz Inequality for real numbers holds when $n = m + 1$, concluding the proof. □

2.2 Abel Partial Summation

Abel partial summation transforms a finite sum of products of two terms by means of the partial sums of one of those terms. Abel summation is similar to integration by parts and can be used for expressing a sum that is difficult to calculate in terms of a more manageable sum. The proofs of the following lemma and theorem are adapted from Chapter I.0 of Tenenbaum [14]. Abel summation is employed in the proof of Mertens' theorems and in steps of various other proofs.

2.2.1 Abel Summation For Discrete Sums

Lemma 2.2. *Let $\{a_1, a_2, \ldots, a_N\}$ and $\{b_1, b_2, \ldots, b_N\}$ each be finite sequences of real numbers. For $n \geq 1$, let $A_n = \sum_{m=1}^{n} a_m$, where, by definition, $A_0 = a_0 = 0$.*

Then

$$\sum_{n=1}^{N} a_n b_n = \sum_{n=1}^{N-1} A_n \left(b_n - b_{n+1}\right) + A_N b_N. \tag{2.14}$$

Proof. To begin, consider the difference $A_n - A_{n-1}$. Apply the definition of A_n and expand the sums to obtain

$$A_n - A_{n-1} = \sum_{m=1}^{n} a_m - \sum_{m=1}^{n} a_{m-1} \tag{2.15}$$

$$= (a_1 - a_0) + (a_2 - a_1) + (a_3 - a_2) + (a_4 - a_3) + \cdots + (a_n - a_{n-1}) \tag{2.16}$$

$$= a_n - a_0 \tag{2.17}$$

$$= a_n. \tag{2.18}$$

Now consider the sum $\sum_{n=1}^{N} a_n b_n$. Substituting in (2.18) provides

$$\begin{aligned}\sum_{n=1}^{N} a_n b_n &= \sum_{n=1}^{N} \left(A_n - A_{n-1}\right) b_n \\ &= \sum_{n=1}^{N} A_n b_n - \sum_{n=1}^{N} A_{n-1} b_n \\ &= A_N b_N + \sum_{n=1}^{N-1} A_n b_n - \sum_{n=1}^{N-1} A_n b_{n+1} \\ &= A_N b_N + \sum_{n=1}^{N-1} A_n \left(b_n - b_{n+1}\right).\end{aligned}$$

□

2.2.2 Abel Summation Involving Continuous Functions

Abel Summation is not relegated strictly to discrete sums. It is applicable to continuously differentiable functions as well. Abel summation involving continuous functions is also employed in various proofs throughout the remainder of this thesis.

Theorem 2.1. *Let* $\{a_n\}_{n=1}^{\infty}$ *be a sequence of complex numbers. Let* $A(t) = \sum_{n\leq t} a_n$*, for* $t > 0$*. Let* $b(t)$ *be a continuously differentiable function on the interval* $[1, x]$*, where* $x > 1$ *is a real number. Then*

$$\sum_{1\leq n\leq x} a_n b(n) = A(x)b(x) - \int_1^x A(t)b'(t)dt.$$

Proof. Since $b(t)$ is continuously differentiable, the sum $\sum_{1\leq n\leq x} a_n b(n)$ can be expressed as a Riemann-Stieltjes integral. That is,

$$\sum_{1\leq n\leq x} a_n b(n) = \int_{1^-}^x b(t)dA(t). \tag{2.19}$$

Therefore, we seek to evaluate the sum by evaluating the integral. Proceed with integration by parts. Let $u = b(t)$. Then $du = b'(t)dt$. Let $dv = dA(t)$. Then $v = A(t)$. The integral in (2.19) can now be evaluated as

$$\int_{1^-}^x b(t)dA(t) = A(t)b(t)|_{1^-}^x - \int_1^x A(t)b'(t)dt \tag{2.20}$$

$$= \left(A(x)b(x) - \lim_{c\to 1^-} A(c)b(c)\right) - \int_1^x A(t)b'(t)dt \tag{2.21}$$

$$= A(x)b(x) - \int_1^x A(t)b'(t)dt. \tag{2.22}$$

Here, $\lim_{c\to 1^-} A(c)b(c) = 0$ because $\lim_{c\to 1^-} A(c) = \lim_{c\to 1^-} \sum_{n\leq c} a_n = 0$. This concludes the proof. □

2.3 The Binomial Theorem

The Binomial Theorem is used in a few subsequent prerequisite proofs and heavily in Chapter 5.

Theorem 2.2. *Let $x, y \in \mathbb{R}$. For all positive integers n,*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k,$$

where $\binom{n}{k}$ *denotes* $\frac{n!}{(n-k)!k!}$.

Proof. This proof was done independently by the author and verified by consulting Appendix A of Miller and Takloo-Bighash [9].

The proof proceeds by induction. For $n = 1$,

$$\begin{aligned}\sum_{k=0}^{1} \binom{n}{k} x^{n-k} y^k &= \binom{1}{0} x^{1-0} y^0 + \binom{1}{1} x^{1-1} y^1 \\ &= x + y = (x+y)^1.\end{aligned}$$

Thus the Binomial Theorem holds for $n = 1$.

Suppose that

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k.$$

It is useful to note that

$$(x+y)^{n+1} = (x+y)(x+y)^n. \tag{2.23}$$

Applying the induction hypothesis to (2.23) and then distributing yields

$$
\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n && (2.24)\\
&= (x+y)\left[\sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^k\right] && (2.25)\\
&= \sum_{k=0}^{n}\binom{n}{k}x^{n-k+1}y^k + \sum_{k=0}^{n}\binom{n}{k}x^{n-k}y^{k+1}. && (2.26)
\end{aligned}
$$

Expanding the sums in (2.26) and factoring by grouping provides

$$
\begin{aligned}
(x+y)^{n+1} \quad = \quad & \tbinom{n+1}{0}x^{n+1} + \left[\tbinom{n}{1}+\tbinom{n}{0}\right]x^n y + \left[\tbinom{n}{2}+\tbinom{n}{1}\right]x^{n-1}y^2\\
& + \left[\tbinom{n}{3}+\tbinom{n}{2}\right]x^{n-2}y^3 + \ldots + \left[\tbinom{n}{n}+\tbinom{n}{n-1}\right]xy^n + \tbinom{n+1}{n+1}y^{n+1}.
\end{aligned}
\qquad (2.27)
$$

With the exception of the x^{n+1} and y^{n+1} terms, notice that all of the coefficients are of the form $\binom{n}{k}+\binom{n}{k-1}$. Expanding the sum and using algebraic manipulation reveals that

$$
\begin{aligned}
\binom{n}{k}+\binom{n}{k-1} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!)(n-(k-1))!} && (2.28)\\
&= \frac{n!}{k(k-1)!(n-k)!} + \frac{n!}{(k-1)!(n+1-k)(n-k)!} && (2.29)\\
&= \frac{n!(n+1-k)+n!(k)}{k(k-1)!(n+1-k)(n-k)!} && (2.30)\\
&= \frac{(n+1)n!}{k!((n+1)-k)!} && (2.31)\\
&= \frac{(n+1)!}{k!((n+1)-k)!} && (2.32)\\
&= \binom{n+1}{k}. && (2.33)
\end{aligned}
$$

Apply the result in (2.33) to the coefficients in (2.27) to get

$$
\begin{aligned}
(x+y)^{n+1} &= \binom{n+1}{0}x^{n+1} + \binom{n+1}{1}x^{n}y + \binom{n+1}{2}x^{n-1}y^{2} + \dots \\
&\quad + \binom{n+1}{n-1}x^{2}y^{n-1} + \binom{n+1}{n}xy^{n} + \binom{n+1}{n+1}y^{n+1} \qquad (2.34)\\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}x^{n+1-k}y^{k}. \qquad (2.35)
\end{aligned}
$$

This completes the induction. □

2.4 The Gamma Function

Lemmas 2.3 and 2.4 are used in linking the constant C_k in Proposition 5.1 to the moments of the normal distribution in Theorem 3.5.

Definition 2.1. Let λ be a non-negative real number. The **Gamma function** $\Gamma(\lambda)$ is defined as

$$\Gamma(\lambda) = \int_0^\infty x^{\lambda-1}e^{-x}dx. \qquad (2.36)$$

Lemma 2.3. *For $n \in \mathbb{N}$,*

$$\Gamma(n) = (n-1)!. \qquad (2.37)$$

Proof. Both proofs in this section are adapted from Appendix G of Arfken and Hans [2].

If $n \in \mathbb{N}$, then

$$\Gamma(n) = \int_0^\infty x^{n-1}e^{-x}dx. \qquad (2.38)$$

Use integration by parts to express $\Gamma(n)$ as

$$
\begin{aligned}
&\Gamma(n) = \\
&\lim_{b\to\infty}(-1)^{n}e^{-x}\left(-x^{n-1}+(n-1)x^{n-2}-(n-1)(n-2)x^{n-3}+\dots+(n-1)!\right)\Big|_0^b.
\end{aligned}
\qquad (2.39)
$$

Consider that $e^0 = 1$ and that, for any $n \in \mathbb{N}$, $\lim_{x\to\infty} \left(x^n e^{-x}\right) = 0$ to obtain

$$\Gamma(n) = (n-1)!.$$

□

Lemma 2.4. *For* $\lambda \in \mathbb{R}^+$,

$$\Gamma(\lambda + 1) = \lambda \Gamma(\lambda). \tag{2.40}$$

Proof. For $\lambda \in \mathbb{R}^+$, using integration by parts results in

$$\begin{aligned}\Gamma(\lambda+1) &= \int_0^\infty x^\lambda e^{-x} dx \\ &= \underbrace{\lim_{b\to\infty} -e^{-x}x^\lambda\Big|_0^b}_{=0} + \lambda \underbrace{\int_0^\infty x^{\lambda-1}e^{-x}dx}_{=\Gamma(\lambda)} \\ &= \lambda\Gamma(\lambda).\end{aligned}$$

□

2.5 Big O Notation

A central theme of this thesis is the asymptotic equality of various functions. The use of big O notation is used frequently. To ensure clarity, this section defines what is meant by the use of big O notation.

The following definition of big O notation is quoted from the definition found in section 3.2 of Apostol [1].

Definition 2.2. If $g(x) > 0$ for all $x \geq a$, we write

$$f(x) = O(g(x)) \tag{2.41}$$

to mean that the quotient $\dfrac{f(x)}{g(x)}$ is bounded for $x \geq a$; that is, there exists a constant $M > 0$ such that

$$|f(x)| \leq M g(x) \quad \text{for all } x \geq a. \tag{2.42}$$

An equation of the form

$$f(x) = h(x) + O(g(x)) \tag{2.43}$$

means that $f(x) - h(x) = O(g(x))$. Note that $f(t) = O(g(t))$, for $t \geq a$, implies that

$$\int_a^x f(t)\, dt = O\left(\int_a^x g(t)\, dt\right) \quad \text{for } x \geq a. \tag{2.44}$$

Remark 2.1. For clarity of notation, please note that $f(x) \ll g(x)$ means $f(x) = O(g(x))$.

Chapter 3

The Normal Distribution

It is not possible to understand the Erdős-Kac Theorem (EKT) without understanding the normal distribution. This chapter establishes several key definitions, properties, and functions involving the normal distribution.

3.1 The Normal Distribution Density Function

Acknowledgment. The work in this section is adapted with permission from Siegrist [11] and expounded upon by the author.

Definition 3.1. A random variable Z is said to have a **normal distribution** if it has the probability density function $f(z)$ given by

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right], \tag{3.1}$$

where $z \in \mathbb{R}$, and σ and μ are both constant.

Remark 3.1. The expected value of Z, denoted as $E(Z)$, is equal to μ, and the variance of Z, denoted as $V(Z)$, is equal to σ^2. Proof that $E(Z) = \mu$ and $V(Z) = \sigma^2$ is included later in the next section.

Lemma 3.1. *The function $f(z)$, as defined in Definition 3.1, satisfies the properties of a probability density function: $f(z) \geq 0$ for all z such that $-\infty \leq z \leq \infty$, and $\int_{-\infty}^{\infty} f(z)dz = 1$.*

Proof. Since $\frac{1}{\sigma\sqrt{2\pi}} > 0$ and $\exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] > 0$ for all $z \in \mathbb{R}$, it follows that $f(z) \geq 0$ for all z such that $-\infty < z < \infty$.

It remains to prove that

$$\int_{-\infty}^{\infty} f(z)\, dz = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz = 1. \tag{3.2}$$

In order to make evaluating the integral in (3.2) less cumbersome, let $x = \frac{z-\mu}{\sigma}$. It follows that $dx = \frac{1}{\sigma}dz$. Note that $x \to \infty$ as $z \to \infty$, and $x \to -\infty$ as $z \to -\infty$. The integral can now be rewritten as

$$\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x)^2\right] \sigma dx.$$

Let

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x)^2\right] \sigma dx && (3.3)\\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx. && (3.4)\end{aligned}$$

Then

$$\begin{aligned} I^2 &= \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx\right)\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx\right) && (3.5)\\ &= \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx\right)\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy\right) && (3.6)\\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(x^2+y^2\right)} dxdy. && (3.7)\end{aligned}$$

Convert I^2 to polar coordinates such that $x = r\cos\theta$ and $y = r\sin\theta$, where

$r \in [0, \infty)$ and $\theta \in [0, 2\pi)$. Then $x^2 + y^2 = r^2$, and $dxdy = rdrd\theta$, and

$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} re^{-\frac{1}{2}r^2} drd\theta. \tag{3.8}$$

To finish evaluating I^2, let $u = \frac{r^2}{2}$. Then $du = rdr$. Implementing these substitutions yields

$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-u} dud\theta \tag{3.9}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} 1 d\theta \tag{3.10}$$

$$= \frac{1}{2\pi} 2\pi \tag{3.11}$$

$$= 1. \tag{3.12}$$

Since I cannot be negative, and $I^2 = 1$, it follows that $I = 1$ and the proof is finished. □

3.2 Properties Of The Normal Distribution Density Function

Acknowledgment. The work in this section is adapted with permission from Siegrist [11] and expounded upon by the author.

The normal density function, $f(z)$, has some useful identifying properties that are outlined in the following lemma.

Lemma 3.2. *The normal density function*

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right], \quad z \in \mathbb{R},$$

has the following four properties.

I. $f(z)$ is symmetric about $z = \mu$.

II. $f(z)$ increases and then decreases with its maximum value at $z = \mu$.

III. $f(z)$ has inflection points at $z = \mu \pm \sigma$.

IV. $\lim_{z \to \pm\infty} f(z) = 0$.

Proof.

I. To demonstrate symmetry about $z = \mu$, it must be shown that $f(\mu + z) = f(\mu - z)$, for all $z \in \mathbb{R}$. Evaluating $f(\mu + z)$ and $f(\mu - z)$ results in

$$f(\mu + z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\mu + z - \mu}{\sigma}\right)^2\right] = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2\right],$$

and

$$f(\mu - z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\mu - z - \mu}{\sigma}\right)^2\right] = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2\right].$$

II. $f'(z) = -\left(\frac{z-\mu}{\sigma^2}\right) f(z)$. Note that $f'(z) > 0$ only when $z < \mu$, $f'(z) < 0$ only when $z > \mu$, and $f'(z) = 0$ only when $z = \mu$. Therefore $f(z)$ increases over the interval $(-\infty, \mu)$, reaches its maximum value at $z = \mu$, and decreases over the interval (μ, ∞).

III. $f''(z) = f(z)\left(-\frac{1}{\sigma^2} + \frac{(z-\mu)^2}{\sigma^4}\right)$. Note that $f''(z) = 0$ only when $z = \mu \pm \sigma$. Note that $f''(z) > 0$ over $(-\infty, \mu - \sigma) \cup (\mu + \sigma, \infty)$, meaning that $f(z)$ is concave upward over this union of intervals. Similarly, note that $f''(z) < 0$

over the interval $(\mu - \sigma, \mu + \sigma)$ meaning that $f(z)$ is concave downward over this interval. Therefore, $f(z)$ changes concavity at the points $z = \mu \pm \sigma$.

IV. Note that $\lim_{z \to -\infty} -\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2 = -\infty$, and $\lim_{z \to \infty} -\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2 = -\infty$. This makes it easier to see that $\lim_{z \to -\infty} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] = 0$, and $\lim_{z \to \infty} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] = 0$. Thus, $\lim_{z \to -\infty} f(z) = 0$, and $\lim_{z \to \infty} f(z) = 0$.

□

3.3 The Normal Distribution Function

Acknowledgment. The work in this section is adapted with permission from Siegrist [11] with details expounded upon by the author.

As claimed in Section 3.1, the expected value of a continuous random variable Z that has a normal probability distribution is μ, and the variance of Z is σ^2. This section formally defines the expected value and variance of Z and proves the claims made in Section 3.1. Background definitions and theorems used in proving these claims are provided.

Definition 3.2. For the continuous random variable Z with the probability density function $f(z)$, the **probability distribution function** of Z, denoted as $F(z)$, is defined as

$$F(z) = \int_{-\infty}^{z} f(t)dt. \tag{3.13}$$

Definition 3.3. The **expected value** of a continuous random variable Z with probability density function $f(z)$ is

$$E(Z) = \int_{-\infty}^{\infty} z f(z)dz, \tag{3.14}$$

provided that the integral exists.

Definition 3.4. The variance of a continuous random variable Z, with probability density function $f(z)$ and expected value $E(Z) = \mu$, is

$$V(Z) = E\left[(Z-\mu)^2\right]. \tag{3.15}$$

The following theorem and subsequent lemma are needed in order to prove Theorem 3.3, which is the result-of-interest in this section.

Theorem 3.1. *Let $g_1(Z)$, $g_2(Z)$, $\ldots$, $g_k(Z)$ be functions of the continuous random variable Z. Then*

$$E(g_1(Z)+g_2(Z)+\cdots+g_k(Z)) = E(g_1(Z))+E(g_2(Z))+\cdots+E(g_k(Z)). \tag{3.16}$$

Proof. By Definition 3.3,

$$E(g_1(Z)+g_2(Z)+\cdots+g_k(Z)) = \int_{-\infty}^{\infty}(g_1(z)+g_2(z)+\cdots+g_k(z))f(z)\,dz, \tag{3.17}$$

where $f(z)$ is the probability density function of Z. Distributing the $f(z)$ term and using the fact that $\int(g_1(z)f(z)+g_2(z)f(z)+\cdots+g_k(z))\;dz = \int g_1(z)f(z)\,dz + \int g_2(z)f(z)\,dz+\cdots+\int g_k(z)f(z)\,dz$ results in

$$\begin{aligned}
E(g_1(Z)+g_2(Z)+\cdots+g_k(Z)) &= \int_{-\infty}^{\infty}(g_1(z)+g_2(z)+\cdots+g_k(z))f(z)\,dz \\
&= \int_{-\infty}^{\infty}(g_1(z)f(z)+g_2(z)f(z)\cdots+g_k(z)f(z))\,dz \\
&= \int_{-\infty}^{\infty}g_1(z)f(z)\,dz+\int_{-\infty}^{\infty}g_2(z)f(z)\,dz \\
&\quad \cdots+\int_{-\infty}^{\infty}g_k(z)f(z)\,dz \\
&= E(g_1(Z))+E(g_2(Z))+\cdots+E(g_k(Z)).
\end{aligned}$$

□

The following theorem is used briefly in the following corollary. Its brief use has resulted in the omission of the proof.

Theorem 3.2. *Let Z be a continuous random variable with probability density function $f(z)$. Let $g(Z)$ be a function of Z. Then the expected value of $g(Z)$ is*

$$E\left[g(Z)\right] = \int_{-\infty}^{\infty} g(z)f(z)\,dz, \tag{3.18}$$

provided that the integral exists.

Corollary 3.1. *A corollary to Theorems 3.1 and 3.2 is that, for continuous random variable Z, the variance can be expressed as*

$$V(Z) = E(Z^2) - \left(E(Z)\right)^2. \tag{3.19}$$

Proof. Begin with the definition of variance (see Definition 3.4) and apply Theorem 3.1 to obtain

$$\begin{aligned} V(Z) &= E\left[(Z-\mu)^2\right] && (3.20)\\ &= E\left[Z^2 - 2\mu Z + \mu^2\right] && (3.21)\\ &= E(Z^2) - E(2\mu Z) + \left[E(Z)\right]^2 && (3.22) \end{aligned}$$

Recall that $\mu = E(Z)$ in the definition of variance. By Theorem 3.2, the $E(2\mu Z)$ term equals $-2(E(Z))^2$. This can be shown explicitly as

$$\begin{aligned} -E(2\mu Z) &= -\int_{-\infty}^{\infty} 2\mu z f(z) && (3.23)\\ &= -2\mu \int_{-\infty}^{\infty} z f(z) && (3.24)\\ &= -2E(Z)E(Z) && (3.25)\\ &= -2(E(Z))^2. && (3.26) \end{aligned}$$

Substituting (3.26) into (3.22) results in

$$V(Z) = E(Z^2) - 2[E(Z)]^2 + [E(Z)]^2 = E(Z^2) - [E(Z)]^2, \tag{3.27}$$

which concludes the proof. □

The normal distribution function can be utilized to prove the following theorem pertaining to the expected value of Z and the variance of Z.

Theorem 3.3. *Let Z be a random variable with a normal probability distribution. Then the expected value of Z is*

$$E(Z) = \mu. \tag{3.28}$$

The variance of Z is

$$V(Z) = \sigma^2. \tag{3.29}$$

Proof. The expected value of Z is given by

$$E(Z) = \int_{-\infty}^{\infty} z f(z)\, dz = \int_{-\infty}^{\infty} z \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz. \tag{3.30}$$

Let $u = \frac{z-\mu}{\sigma}$. Then $du = \frac{1}{\sigma}dz$ and $z = u\sigma + \mu$. The integral can now be rewritten as

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (u\sigma + \mu) e^{-\frac{1}{2}u^2}\, du \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{\infty} u\sigma e^{-\frac{1}{2}u^2}\, du + \int_{-\infty}^{\infty} \mu e^{-\frac{1}{2}u^2}\, du \right) \qquad (3.31) \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_{-\infty}^{0} u\sigma e^{-\frac{1}{2}u^2}\, du + \int_{0}^{\infty} u\sigma e^{-\frac{1}{2}u^2}\, du + \int_{-\infty}^{\infty} \mu e^{-\frac{1}{2}u^2}\, du \right). \qquad (3.32) \end{aligned}$$

In seeking to evaluate the integral in equation (3.3), the result in equation (3.11) implies that $\int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2}\, du = \sqrt{2\pi}$. In order to evaluate the first two integrals, let $v = -\frac{1}{2}u^2$. Then $v \to -\infty$ as $u \to -\infty$, $v \to -\infty$ as $u \to \infty$, and $v \to 0$

as $u \to 0$. Further, $dv = -u\,du$. Make all of these substitutions to obtain

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{0} u\sigma e^{-\frac{1}{2}u^2}\,du + \int_{0}^{\infty} u\sigma e^{-\frac{1}{2}u^2}\,du + \int_{-\infty}^{\infty} \mu e^{-\frac{1}{2}u^2} du\right) \\
&= \frac{1}{\sqrt{2\pi}}\left(-\int_{-\infty}^{0} \sigma e^{-v}\,dv - \int_{0}^{-\infty} \sigma e^{-v}\,dv + \mu\sqrt{2\pi}\right) && (3.33) \\
&= \frac{1}{\sqrt{2\pi}}\left(-\int_{-\infty}^{0} \sigma e^{-v}\,dv + \int_{-\infty}^{0} \sigma e^{-v}\,dv + \mu\sqrt{2\pi}\right) && (3.34) \\
&= \frac{1}{\sqrt{2\pi}}\left(0 + \mu\sqrt{2\pi}\right) && (3.35) \\
&= \mu. && (3.36)
\end{aligned}$$

By Corollary 3.1 the variance $V(Z)$ is given by

$$V(Z) = E(Z^2) - [E(Z)]^2\,. \tag{3.37}$$

Theorem 3.2 can be applied to the $E(Z^2)$ term, and the preceding result in this proof, that $E(Z) = \mu$, can be applied to the $[E(Z)]^2$ term. Doing so provides

$$V(Z) = \int_{-\infty}^{\infty} z^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz - \mu^2. \tag{3.38}$$

The task is to show that the above integral is equal to $\sigma^2 + \mu^2$. As before, let $u = \frac{z-\mu}{\sigma}$. Then $du = \frac{1}{\sigma}dz$ and $z = u\sigma + \mu$. As $z \to -\infty$, $u \to -\infty$, and as $z \to \infty$, $u \to \infty$. The integral in (3.38) can now be rewritten as

$$\begin{aligned}
&\int_{-\infty}^{\infty} z^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (u\sigma + \mu)^2\, e^{-\frac{1}{2}u^2}\,du && (3.39) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(u^2\sigma^2 + 2\mu\sigma + \mu^2\right) e^{-\frac{1}{2}u^2}\,du && (3.40) \\
&= \frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty} u^2\sigma^2 e^{-\frac{1}{2}u^2}\,du + \int_{-\infty}^{\infty} 2\mu\sigma e^{-\frac{1}{2}u^2}\,du + \int_{-\infty}^{\infty} \mu^2 e^{-\frac{1}{2}u^2}\,du\right). && (3.41)
\end{aligned}$$

With techniques similar to those already used, it can be shown that

$$\int_{-\infty}^{\infty} u^2 e^{-\frac{1}{2}u^2}\,du = \sqrt{2\pi}, \tag{3.42}$$

$$\int_{-\infty}^{\infty} u e^{-\frac{1}{2}u^2}\,du = 0, \tag{3.43}$$

$$\text{and} \int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2}\,du = \sqrt{2\pi}. \tag{3.44}$$

Substitute these three values into (3.41) to obtain

$$\frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{\infty} u^2\sigma^2 e^{-\frac{1}{2}u^2}\,du + \int_{-\infty}^{\infty} 2\mu\sigma e^{-\frac{1}{2}u^2}\,du + \int_{-\infty}^{\infty} \mu^2 e^{-\frac{1}{2}u^2}\,du\right)$$

$$= \frac{1}{\sqrt{2\pi}}\left(\sigma^2\sqrt{2\pi} + 0 + \mu^2\sqrt{2\pi}\right) \tag{3.45}$$

$$= \sigma^2 + \mu^2 \tag{3.46}$$

Substitute this result into (3.38) to obtain $V(Z) = \sigma^2 + \mu^2 - \mu^2 = \sigma^2$. □

3.4 Moments Of The Normal Distribution Function

The proofs outlined in Chapter 5 depend on computing moments and comparing the results to the moments of the normal distribution function. Theorems 3.5 and 3.6 are the main interest of this section. Subsections 3.4.1 and 3.4.2 combine to provide the reader with an alternate means to establish Theorem 3.5 and are optional for the reader to consider.

Acknowledgment. The work in this section is adapted with permission from Siegrist [11] with details expounded on by the author.

Definition 3.5. The n^{th} central moment of continuous random variable Z is defined as

$$E\left[(Z-\mu)^n\right] = \int_{-\infty}^{\infty} (z-\mu)^n f(z)dz. \tag{3.47}$$

The purpose of this section is to establish two theorems involving the n^{th} central moment of continuous random variable Z.

Theorem 3.4. *For* $n \in \mathbb{N}$,

$$E\left[(Z-\mu)^{n+1}\right] = n\sigma^2 E\left[(Z-\mu)^{n-1}\right]. \tag{3.48}$$

Proof. Apply Definition 3.3 to $E\left[(Z-\mu)^{n+1}\right]$ and use the fact that $x^{n+1} = x^n \cdot x$ to obtain

$$E\left[(Z-\mu)^{n+1}\right] = \int_{-\infty}^{\infty} (z-\mu)^{n+1} f(z)dz \tag{3.49}$$

$$= \int_{-\infty}^{\infty} (z-\mu)^{n} (z-\mu) f(z)dz. \tag{3.50}$$

Recall that $f'(z) = -\left(\frac{z-\mu}{\sigma^2}\right) f(z)$, meaning $-\sigma^2 f'(z) = (z-\mu)f(z)$. Making this substitution provides

$$\int_{-\infty}^{\infty} (z-\mu)^{n} (z-\mu) f(z)dz = -\sigma^2 \int_{-\infty}^{\infty} (z-\mu)^{n} f'(z)dz. \tag{3.51}$$

Use integration by parts where $u = (z-\mu)^n$, $du = n(z-\mu)^{n-1}\, dz$, $dv = f'(z)\, dz$, and $v = f(z)$. The integral in (3.51) can be evaluated as

$$-\sigma^2 \int_{-\infty}^{\infty} (z-\mu)^{n} f'(z)\, dz = -\sigma^2 (z-\mu)^n f(z)|_{-\infty}^{\infty} + \sigma^2 \int_{-\infty}^{\infty} n(z-\mu)^{n-1} f(z)\, dz. \tag{3.52}$$

Since $f(z)$ can be differentiated an unlimited number of times, L'Hôpital's rule can be applied (with some algebraic manipulation) $n+1$ times to show that the first term equals zero. The remaining integral can be rewritten to obtain

$$\sigma^2 \int_{-\infty}^{\infty} n(z-\mu)^{n-1} f(z)dz = n\sigma^2 \int_{-\infty}^{\infty} (z-\mu)^{n-1} f(z)dz \tag{3.53}$$

$$= n\sigma^2 E\left[(Z-\mu)^{n-1}\right]. \tag{3.54}$$

$\square$

Theorem 3.5. *Let Z be a continuous random variable that has a normal probability distribution. Let $E(Z-\mu)^n$ represent the n^{th} central moment of Z. Then, for $n \in \mathbb{N}$,*

$$E\left[(Z-\mu)^{2n+1}\right] = 0, \tag{3.55}$$

and

$$E\left[(Z-\mu)^{2n}\right] = (1 \cdot 3 \cdots (2n-1))\,\sigma^{2n} = \frac{(2n)!\sigma^{2n}}{n!2^n}. \tag{3.56}$$

Remark 3.2. Theorem 3.5 says that all odd-valued central moments of Z are zero and all even-valued central moments of Z are equal to the right-hand side of (3.56). This theorem is vital in seeing that the proof of Theorem 5.1 is indeed a proof of the Erdős-Kac Theorem.

Proof. To prove (3.55) begin with the definition of the central moment of Z. Replacing n with $2n+1$ results in

$$E\left[(Z-\mu)^{2n+1}\right] = \int_{-\infty}^{\infty} (z-\mu)^{2n+1} f(z)\,dz. \tag{3.57}$$

Rewrite (3.57) as

$$E\left[(Z-\mu)^{2n+1}\right] = \int_{-\infty}^{\mu} (z-\mu)^{2n+1} f(z)\,dz + \int_{\mu}^{\infty} (z-\mu)^{2n+1} f(z)\,dz. \tag{3.58}$$

By property I of Lemma 3.2, $f(z)$ is symmetric about μ. Therefore,

$$(-(z-\mu))^{2n+1} = -(z-\mu)^{2n+1}, \tag{3.59}$$

and

$$\int_{-\infty}^{\mu} (z-\mu)^{2n+1} f(z)\,dz = -\int_{\mu}^{\infty} (z-\mu)^{2n+1} f(z)\,dz. \tag{3.60}$$

Thus, $E\left[(Z-\mu)^{2n+1}\right] = 0$, for $n \in \mathbb{N}$.

To prove (3.56), proceed using induction. Let $n = 1$. Then (3.55) yields

$$E\left[(Z-\mu)^2\right] = \int_{-\infty}^{\infty} (z-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz. \tag{3.61}$$

Let $u = \frac{z-\mu}{\sigma}$. Then $(z-\mu) = \sigma u$,and $du = \frac{1}{\sigma}\,dz$. As $z \to -\infty$, $u \to -\infty$, and as $z \to \infty$, $u \to \infty$. Make these substitutions into (3.61) to obtain

$$E\left[(Z-\mu)^{2n}\right] = \int_{-\infty}^{\infty} (z-\mu)^2 \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz \tag{3.62}$$

$$= \int_{-\infty}^{\infty} u^2\sigma^3 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\,du \tag{3.63}$$

$$= \sigma^2 \int_{-\infty}^{\infty} u^2 \left(\sigma \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}u^2}\right) du \tag{3.64}$$

$$= \sigma^2 \int_{-\infty}^{\infty} u^2 \Phi(u)\,du. \tag{3.65}$$

Here $\Phi(u)$ is the standard normal distribution function for the variable u. Thus, $\displaystyle\int_{-\infty}^{\infty} u^2\Phi(u)du = E(U^2) = V(U) - (E(U))^2 = 1 - 0^2 = 1$, leaving

$$E\left[(Z-\mu)^2\right] = \sigma^2. \tag{3.66}$$

Noting that $1 = \dfrac{(2(1))!}{(1!)(2!)}$ allows $E\left[(Z-\mu)^{2n}\right]$ to be expressed as

$$E\left[(Z-\mu)^{2n}\right] = \sigma^2 \tag{3.67}$$

$$= \frac{(2\,(1))!\sigma^{2(1)}}{(1!)(2!)} \tag{3.68}$$

$$= \frac{(2n)!\sigma^{2n}}{(n!)(2^n)}. \tag{3.69}$$

Thus, (3.56) holds true for $n = 1$, which establishes the basis step.

Now suppose that $E\left[(Z-\mu)^{2n}\right] = \dfrac{(2n)!\sigma^{2n}}{(n!)2^n}$. Let $n = k+1$, where $k \in \mathbb{N}$. Applying (3.48) from Theorem 3.4 provides

$$
\begin{aligned}
E\left[(Z-\mu)^{2(k+1)}\right] &= (2k+1)\sigma^2 E\left[(Z-\mu)^{2k}\right] && (3.70)\\
&= (2k+1)\sigma^2 \frac{(2k)!\sigma^{2k}}{(k!)2^k} && (3.71)\\
&= \frac{(2k+1)!\sigma^{2(k+1)}}{(k!)2^k} \cdot \frac{2(k+1)}{2(k+1)} && (3.72)\\
&= \frac{(2(k+1))!\sigma^{2(k+1)}}{(k+1)!2^{k+1}}, && (3.73)
\end{aligned}
$$

which completes the induction. □

The following theorem highlights an extremely important property of all moment-generating functions and is found in section 6.5 of Wackerly et. al [15].

Theorem 3.6. *Let $m_X(t)$ and $m_Y(t)$ denote the respective moment-generating functions of continuous random variables X and Y. If $m_X(t)$ and $m_Y(t)$ both exist and $m_X(t) = m_Y(t)$ for all values of t, then X and Y have the same probability distribution.*

The key idea in Theorem 3.6 is that each probability distribution has one and only one unique moment-generating function. The proof of Theorem 3.6 is extensive and beyond the scope of this thesis; therefore, it is omitted.

3.4.1 Moment-Generating Functions

The following definition and proceeding explanation originate from section 3.9 of Wackerly et. al [15] and are expounded upon by the author.

Definition 3.6. Let Z be a continuous random variable with probability density function $f(z)$. The **moment-generating function** $m(t)$ of Z is

$$m(t) = E(e^{tZ}). \tag{3.74}$$

Applying Theorem 3.2 to this definition results in

$$m(t) = E\left(e^{tZ}\right) = \int_{-\infty}^{\infty} e^{tz} f(z)\, dz. \tag{3.75}$$

To gain further insight into the moment-generating function, write the series expansion of e^{tz}. That is,

$$e^{tz} = 1 + tz + \frac{(tz)^2}{2!} + \frac{(tz)^2}{3!} + \cdots + \frac{(tz)^n}{n!} + \ldots \tag{3.76}$$

Substitute (3.76) into the integrand in (3.75) to get

$$\begin{aligned}
m(t) &= \int_{-\infty}^{\infty} \left(1 + tz + \frac{(tz)^2}{2!} + \frac{(tz)^3}{3!} + \cdots + \frac{(tz)^n}{n!} + \ldots\right) f(z)\, dz && (3.77)\\
&= \int_{-\infty}^{\infty} f(z)\, dz + \int_{-\infty}^{\infty} tz f(z)\, dz + \int_{-\infty}^{\infty} \frac{(tz)^2}{2!} f(z)\, dz && (3.78)\\
&\quad + \int_{-\infty}^{\infty} \frac{(tz)^3}{3!} f(z)\, dz + \cdots + \int_{-\infty}^{\infty} \frac{(tz)^n}{n!} f(z)\, dz + \ldots\\
&= \int_{-\infty}^{\infty} f(z)\, dz + t\int_{-\infty}^{\infty} z f(z)\, dz + t^2 \int_{-\infty}^{\infty} \frac{z^2}{2!} f(z)\, dz && (3.79)\\
&\quad + t^3 \int_{-\infty}^{\infty} \frac{z^3}{3!} f(z)\, dz + \cdots + t^n \int_{-\infty}^{\infty} \frac{z^n}{n!} f(z)\, dz + \ldots\\
&= 1 + tE(Z) + t^2 E(Z^2) + t^3 E(Z^3) + \cdots + t^n E(Z^n) + \ldots && (3.80)
\end{aligned}$$

Each $E(Z^n)$ factor is the n^{th} central moment of Z about the origin. Therefore, in the series expansion of $m(t)$, the n^{th} term has the n^{th} central moment of Z in its coefficient (starting at $n = 0$). In other words, $m(t)$ is a function that "contains" all of the central moments of Z about the origin. The reader may notice that the n^{th} central moment is equal to the n^{th} derivative of $m(t)$ with respect to t at $t = 0$.

3.4.2 The Moment-Generating Function Of The Normal Distribution

Theorem 3.7. *Let Z be a continuous random variable with a normal probability distribution, and with $E(Z) = \mu$ and $V(Z) = \sigma^2$. Then the moment-generating* *function $m(t)$ of Z is*

$$m(t) = \exp\left[\frac{2}{\sigma^2}t^2\right]. \tag{3.81}$$

Proof. The following proof is adapted from section 4.9 of Wackerly et. al [15].

Since Z is a normally distributed continuous random variable, its corresponding probability density function is (see Definition 3.1)

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right], \tag{3.82}$$

where $z \in \mathbb{R}$, and μ and σ are both constant.

Substituting (3.82) into the definition of the moment-generating function (3.74) yields

$$\begin{aligned}
m(t) &= E\left(e^{tZ}\right) && (3.83)\\
&= \int_{-\infty}^{\infty} \frac{e^{tz}}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz. && (3.84)\\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[tz - \frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2\right] dz && (3.85)\\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\frac{2\sigma^2 tz}{2\sigma^2} - \frac{1}{2\sigma^2}(z-\mu)^2\right] dz && (3.86)\\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}\left((z-\mu)^2 - 2\sigma^2 tz\right)\right] dz. && (3.87)
\end{aligned}$$

Let $u = z - \mu$. Then $du = dz$, and

$$\begin{aligned}
m(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}\left(u^2 - 2\sigma^2 t(u+\mu)\right)\right] du && (3.88)\\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}\left(u^2 - 2\sigma^2 tu - 2\sigma^2\mu\right)\right] du. && (3.89)
\end{aligned}$$

Next, complete the square of $(u^2 - 2\sigma^2 tu - 2\sigma^2\mu)$ to get

$$\begin{aligned}
u^2 - 2\sigma^2 tu - 2\sigma^2\mu &= u^2 - 2\sigma^2 tu - 2\sigma^2\mu + \sigma^4 t^2 - \sigma^4 t^2 && (3.90)\\
&= \left(u - \sigma^2 t\right)^2 - \sigma^4 t^2. && (3.91)
\end{aligned}$$

Now, (3.89) can be rewritten as

$$m(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2\sigma^2}\left(\left(u - \sigma^2 t\right)^2 - \sigma^4 t^2\right)\right] du. \tag{3.92}$$

Use of some algebraic manipulation and the fact that $e^{-x+y} = e^{-x}e^{y}$ provides

$$m(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{\left(u - \sigma^2 t\right)^2}{2\sigma^2} + \frac{\sigma^4 t^2}{2\sigma^2}\right] du \tag{3.93}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{\left(u - \sigma^2 t\right)^2}{2\sigma^2}\right] \exp\left[\frac{\sigma^4 t^2}{2\sigma^2}\right] du \tag{3.94}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{\left(u - \sigma^2 t\right)^2}{2\sigma^2}\right] \exp\left[\frac{\sigma^2 t^2}{2}\right] du \tag{3.95}$$

$$= \exp\left[\frac{\sigma^2 t^2}{2}\right] \left(\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{\left(u - \sigma^2 t\right)^2}{2\sigma^2}\right] du\right). \tag{3.96}$$

The portion of (3.96) enclosed in parentheses is the integral of the normal density function for random variable U (with $E(U) = \sigma^2 t$ and $V(U) = \sigma^2$) over all possible u, and it is thus equal to 1. Therefore,

$$m(t) = \exp\left[\frac{\sigma^2 t^2}{2}\right]. \tag{3.97}$$

□

Chapter 4

Number-theoretic Tools

The Erdős-Kac Theorem (EKT) is truly a blend of statistics and number theory. The two main points of interest in this chapter are Theorem 4.5 and Lemma 4.1. Theorem 4.5 and Lemma 4.1 are so vital in proving the EKT that all of the number-theoretic background necessary in their proofs is established first. The results in the initial sections are needed to prove Mertens' theorems. The results immediately following Mertens' theorems are used to prove the gcd-restricted sum lemma.

4.1 The p-adic Valuation Of $n!$

Definition 4.1. For each prime number p, the p-adic valuation v_p is defined as the arithmetic function that associates to each integer n the exponent of p in the canonical factorization of n.

The following theorem and its corollary are taken from Chapter I.0 of Tenenbaum [14].

Theorem 4.1. *Let $n \geq 1$. For each prime p,*

$$v_p(n!) = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor. \tag{4.1}$$

Proof. This proof was done independently by the author and verified by consulting Tenenbaum [14].

For $n \in \mathbb{N}$, write $n!$ as

$$n! = n(n-1)(n-2)(n-3)(n-4)\cdots(3)(2)(1). \tag{4.2}$$

For prime p such that $p \leq n$, then at least one term is divisible by p in the expansion of $n!$ because at least one term must equal p. If the expansion of $n!$ is written in sequential order as in (4.2) above, then p will divide every p^{th} factor counting up from 1. That is, every successive p^{th} factor in the expansion of $n!$ will contain one higher power of p in its respective canonical factorization. The expression $\left\lfloor \frac{n}{p} \right\rfloor$ thus provides the number of terms in the expansion of $n!$ that are divisible by p.

Similar reasoning can be extended to p^k to observe that the expression $\left\lfloor \frac{n}{p^k} \right\rfloor$ provides the number of terms in the sequential expansion of $n!$ that are divisible by p^k. As k increases, eventually p^k will be greater than n and $\left\lfloor \frac{n}{p^k} \right\rfloor = 0$.

In the case where p is greater than n, $\sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor = 0$, because n/p^k will be less than one, making $\left\lfloor \frac{n}{p^k} \right\rfloor = 0$ for all $k \in \mathbb{N}$.

The sum $\sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor$ thus counts the largest value of m ($m \in \mathbb{N}$) such that p^m divides $n!$. In other words, $\sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor$ is the p-adic valuation of $n!$. □

Example 4.1. A concrete example may prove instructive in this instance. Suppose that we want to evaluate $v_5(100!)$. In other words, we are interested in finding the largest possible α such that 5^α divides $100!$. We first proceed with a bit of a brute force method. We know that 5 divides 20 terms in the expansion of $100!$, namely $5, 10, 15, \ldots, 100$. We also know that 5^2 divides 4 terms in the expansion of 100!, namely $25, 50, 75, 100$. We know that 5^k, $k \geq 3$, divides 0 terms in the expansion of 100! because $5^k > 100$. Thus $v_5(100!) = 20 + 4 = 24$. Employing the result in Theorem 4.1 gives us

$$\begin{aligned} v_5(100!) &= \sum_{k=1}^{\infty} \left\lfloor \frac{100}{5^k} \right\rfloor \\ &= \left\lfloor \frac{100}{5} \right\rfloor + \left\lfloor \frac{100}{5^2} \right\rfloor + \left\lfloor \frac{100}{5^3} \right\rfloor + \left\lfloor \frac{100}{5^4} \right\rfloor + \cdots \\ &= 20 + 4 + 0 + 0 + \cdots \\ &= 24. \end{aligned}$$

Corollary 4.1. *Let $n \geq 1$. For each prime p,*

$$\frac{n}{p} - 1 < v_p(n!) \leq \frac{n}{p} + \frac{n}{p(p-1)}.$$

Proof. Let $n \in \mathbb{N}$. By Theorem 4.1, for any prime p, the smallest possible value that $v_p(n!)$ can take on is $\left\lfloor \frac{n}{p} \right\rfloor$. Using this fact along with $\lfloor x \rfloor > x - 1$ results in

$$\left\lfloor \frac{n}{p} \right\rfloor > \frac{n}{p} - 1. \tag{4.3}$$

By Theorem 4.1, $v_p(n!)$ can be expressed as

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \left\lfloor \frac{n}{p^3} \right\rfloor + \left\lfloor \frac{n}{p^4} \right\rfloor + \left\lfloor \frac{n}{p^5} \right\rfloor + \cdots \leq \frac{n}{p} + \frac{n}{p^2} + \frac{n}{p^3} + \frac{n}{p^4} + \frac{n}{p^5} + \cdots \tag{4.4}$$

. Here the the inequality holds because, $\lfloor x \rfloor \leq x$ for any positive real number x.

Note that

$$\frac{n}{p} + \frac{n}{p^2} + \frac{n}{p^3} + \frac{n}{p^4} + \frac{n}{p^5} + \cdots = \frac{n}{p} + \frac{n}{p} \underbrace{\left(\frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \frac{1}{p^4} + \ldots \right)}_{\text{Geometric Series}} \tag{4.5}$$

$$= \frac{n}{p} + \frac{n}{p} \left(\frac{1}{(p-1)} \right). \tag{4.6}$$

Combining (4.4) with (4.6) results in

$$v_p(n!) \leq \frac{n}{p} + \frac{n}{p\,(p-1)}. \tag{4.7}$$

Finally (4.3) and (4.7) can be combined to get that

$$\frac{n}{p} - 1 < v_p(n!) \leq \frac{n}{p} + \frac{n}{p\,(p-1)}. \tag{4.8}$$

□

4.2 A Comparison Of A Sum To An Integral

Before proving Mertens' First Theorem, Theorem 4.2 below is used to derive a weak Stirling approximation for $\log n!$. This approximation of $\log n!$ plays a pivotal role in proving Mertens' First Theorem.

Theorem 4.2. *Let f be a real monotonic function on the interval $[a, b]$, with $a, b \in \mathbb{Z}$. Then there exists some real number $\theta = \theta(a, b)$, $0 \leq \theta \leq 1$, such that*

$$\sum_{a<n\leq b} f(n) = \int_a^b f(t)dt + \theta(f(b) - f(a)). \tag{4.9}$$

Proof. This proof is adapted from page 4 of Tenenbaum [14], with more details explicitly provided by the author.

Begin by noting that

$$\sum_{a<n\leq b} f(n) = \int_a^b f(t)d\lfloor t\rfloor. \tag{4.10}$$

This equation holds true because the sum only adds values of $f = f(t)$ (over the interval from a to b) when t is an integer. Likewise, the integral term is the Riemann-Stieltjes integral equivalent of this sum. The integral is weighted against the floor of t, meaning it will only pick up the integrand when $\lfloor t\rfloor$ changes value, which occurs at integer values of t.

Now apply (4.10) to the difference between the discrete sum and the integral, obtaining

$$\sum_{a<n\leq b} f(n) - \int_a^b f(t)dt = \int_a^b f(t)d\lfloor t\rfloor - \int_a^b f(t)dt. \tag{4.11}$$

Since the integrands are identical, the right-hand side of (4.11) it can be written

$$\int_a^b f(t)d\lfloor t\rfloor - \int_a^b f(t)dt = \int_a^b f(t)d\left(\lfloor t\rfloor - t\right). \tag{4.12}$$

Note that $\lfloor t\rfloor - t = -\{t\}$, where $\{t\}$ denotes the fractional part of t. Then,

$$\int_a^b f(t)d\left(\lfloor t\rfloor - t\right) = -\int_a^b f(t)d\{t\}. \tag{4.13}$$

Use integration by parts with $u = f(t)$, $du = df(t)$, $dv = d\{t\}$, and $v = \{t\}$ to obtain

$$-\int_a^b f(t)d\{t\} = -f(b)\{b\} + f(a)\{a\} + \int_a^b \{t\}df(t) \tag{4.14}$$

$$= -f(b)(0) + f(a)(0) + \int_a^b \{t\}df(t) \tag{4.15}$$

$$= \int_a^b \{t\}df(t). \tag{4.16}$$

Recall that $\{a\} = 0 = \{b\}$ because a and b are integers.

Combining (4.11) and (4.16), we have

$$\sum_{a<n\leq b} f(n) = \int_a^b f(t)dt + \int_a^b \{t\}df(t). \tag{4.17}$$

In Theorem 4.2 f is required to be monotonic. For simplicity of reasoning, suppose that f is monotonic increasing. Then $df(t)$ is always positive. Consider the sum that the Riemann-Stieltjes integral in (4.16) represents,

$$\int_a^b \{t\}df(t) = \lim_{n\to\infty} \sum_{i=0}^{n-1} \{x_i\} \left(f(x_{i+1}) - f(x_i)\right) \tag{4.18}$$

$$\leq \lim_{n\to\infty} \sum_{i=0}^{n-1} 1 \left(f(x_{i+1}) - f(x_i)\right) \tag{4.19}$$

$$= f(b) - f(a). \tag{4.20}$$

Remember that $0 \leq \{t\} < 1$. Combining this result with (4.17) yields

$$\sum_{a<n\leq b} f(n) = \int_a^b f(t)dt + \theta \left(f(b) - f(a)\right), \tag{4.21}$$

where $0 \leq \theta \leq 1$. The result in (4.21) states that the sum and integral will not differ by a value greater in magnitude than the difference between $f(b)$ and $f(a)$. Therefore, the value of θ will depend on a and b. That is, $\theta = \theta(a, b)$. □

Corollary 4.2. *For integer* $n \geq 1$*,*

$$\log n! = n \log n - n + 1 + \theta \log n, \tag{4.22}$$

with $\theta = \theta(n) \in [0, 1]$*.*

Proof. First, expand $n!$ and use the fact that $\log(ab) = \log a + \log b$ to get

$$\begin{aligned} \log n! &= \log(n(n-1)(n-2)\cdots 1) && (4.23)\\ &= \log n + \log(n-1) + \log(n-2) + \cdots + \log 1 && (4.24)\\ &= \sum_{1<k\leq n} \log k. && (4.25) \end{aligned}$$

Now apply Theorem 4.2 to (4.25) to obtain

$$\begin{aligned} \log n! &= \sum_{1<k\leq n} \log k && (4.26)\\ &= \int_1^n \log t dt + \theta(\log n - \log 1) && (4.27)\\ &= (t\log t - t)|_1^n + \theta \log n && (4.28)\\ &= n\log n - n - (1\log 1 - 1) + \theta\log n && (4.29)\\ &= n\log n - n + 1 + \theta \log n, && (4.30) \end{aligned}$$

concluding the proof. □

Example 4.2. To illustrate Corollary 4.2, consider a concrete example. Use of the software package SAGE [13] confirms that

$$\log(3000!) \approx 21024.024853.$$

Further,

$$3000\log 3000 - 3000 + 1 \approx 21020.102703,$$

and

$$\begin{aligned} \log(3000!) - (3000\log 3000 - 3000 + 1) &\approx 3.92215\\ &\approx 0.4898788 \log 3000. \end{aligned}$$

4.3 An Upper Bound On The Product Of Primes

The upper bound established in Theorem 4.3 is employed in the proof of Mertens' First Theorem.

Theorem 4.3. *Let* $\prod_{p \le n} p$ *represent the product of all primes p that are less than or equal to n. Then, for* $n \ge 1$,

$$\prod_{p \le n} p \le 4^n. \tag{4.31}$$

Proof. This proof is adapted from Chapter I.0 of Tenenbaum [14] with details added by the author.

The proof proceeds by induction on n. However, the proof treats the case when n is even separately from when n is odd. Therefore, the basis step is established for each case. For $n = 2$,

$$\prod_{p \le 2} p = 2. \tag{4.32}$$

Since $2 \le 4^2 = 16$, the basis step is established for when n is even.

Because the smallest prime is 2, the product is empty and understood to equal 1 for $n = 1$. This trivial case does not establish the basis step. Accordingly, let $n = 3$. Then,

$$\prod_{p \le 3} p = 3 \cdot 2 = 6. \tag{4.33}$$

Since $6 \le 4^3 = 64$, the basis step is established for when n is odd.

Suppose that $\prod_{p \le n} p \le 4^n$.

Let n be odd, and write $n = 2m + 1$, where $m \in \mathbb{N}$. By the Binomial Theorem (Theorem 2.2),

$$2^{2m+1} = (1+1)^{2m+1} = \sum_{k=0}^{2m+1} \binom{2m+1}{k} 1^{n-k} 1^k = \sum_{k=0}^{2m+1} \binom{2m+1}{k}. \tag{4.34}$$

Notice that

$$\binom{2m+1}{m} = \frac{(2m+1)!}{m!(m+1)!} = \frac{(2m+1)!}{(m+1)!m!} = \binom{2m+1}{m+1}. \tag{4.35}$$

In other words, the coefficient $\binom{2m+1}{m}$ appears twice in the binomial expansion of 2^{2m+1}. This means that

$$\binom{2m+1}{m} \leq \frac{1}{2}\left(2^{2m+1}\right). \tag{4.36}$$

Using the fact that $(m+1)! = (m+1)m!$ allows $\binom{2m+1}{m}$ to be expressed as

$$\binom{2m+1}{m} = \frac{(2m+1)!}{(m+1)(m!)^2}. \tag{4.37}$$

All of the coefficients of the binomial expansion of 2^{2m+1} are integers, meaning $(m+1)$ divides $(2m+1)!$. The smallest value that m can take on is 1. When $m = 1$,

$$\frac{(2m+1)!}{(m+1)} = 3 > m+1. \tag{4.38}$$

This quotient only increases as m increases. In general, for $m \geq 1$,

$$\frac{(2m+1)!}{m+1} > m+1. \tag{4.39}$$

Similarly, $\binom{2m+1}{m} = 3$ when $m = 1$, and it also strictly increases as m increases. All integers greater than or equal to three are divisible by at least one prime number. The results in (4.38) and (4.39) show that $\binom{2m+1}{m}$ is divisible by a prime number that is greater than $m+1$. The largest possible prime divisor of

$$(2m+1)! = (2m+1)(2m)(2m-1)\dots(2)(1) \tag{4.40}$$

is $2m+1$. Therefore, $\prod_{m+1<p\leq 2m+1} p$ divides $\binom{2m+1}{m}$. The smallest possible prime

that could divide $\binom{2m+1}{m}$ is 2. This fact and inequality (4.36) combine to give

$$\left(\prod_{m+1<p\leq 2m+1} p\right) \Bigg| \binom{2m+1}{m} \leq \frac{1}{2}\left(2^{2m+1}\right). \tag{4.41}$$

Notice that $\frac{2^{2m+1}}{2} = 2^{2m} = 4^m$. Therefore,

$$\left(\prod_{m+1<p\leq 2m+1} p\right) \Bigg| \binom{2m+1}{m} \leq 4^m. \tag{4.42}$$

Since $\displaystyle\prod_{m+1\leq p\leq 2m+1} p$ divides $\binom{2m+1}{m}$, then

$$\prod_{m+1<p\leq 2m+1} p \leq 4^m \tag{4.43}$$

holds as well.

Apply the induction hypothesis to the case $n > m+1$ to obtain

$$\prod_{p\leq m+1} p \leq 4^{m+1}. \tag{4.44}$$

Multiplying (4.43) by (4.44) yields

$$\prod_{p\leq m+1} p \prod_{m+1<p\leq 2m+1} p \leq 4^{m+1}4^m. \tag{4.45}$$

Observe that

$$\prod_{p\leq m+1} p \prod_{m+1<p\leq 2m+1} p = \prod_{p\leq n} p, \tag{4.46}$$

and

$$4^{m+1}4^m = 4^{2m+1} = 4^n. \tag{4.47}$$

Therefore, when n is odd,

$$\prod_{p\leq n} p \leq 4^n. \tag{4.48}$$

This concludes the proof for the case when n is odd.

Let n be even. Let $m = n + 2$. Then m is even. If $n + 2$ is even, then $n + 2$ is not prime and

$$\prod_{p \leq m} p = \prod_{p \leq n+2} p = \prod_{p \leq n+1} p. \tag{4.49}$$

Since $n + 2$ is even, $n + 1$ is odd. The theorem for odd numbers has already been proven. Therefore,

$$\prod_{p \leq n+2} p = \prod_{p \leq n+1} p \leq 4^{n+1}. \tag{4.50}$$

For $n \geq 1$, $4^{n+1} < 4^{n+2}$. Combining this observation along with (4.50) results in

$$\prod_{p \leq n+2} p \leq 4^{n+2}. \tag{4.51}$$

This completes the induction for when n is even and completes the proof. □

4.4 Proof Of Mertens' First Theorem

The prime number theorem asserts that the prime-counting function $\pi(x)$, which counts the number of primes less than or equal to x, has an asymptotic behavior (see Tenenbaum [14], Chapter 1). However, $\pi(x)$ is not easily analyzed. Mertens allowed for this asymptotic behavior to be observed through functions more accessible (easier to analyze efficiently) than $\pi(x)$, as shown in his first and second theorems below.

Mertens' First Theorem establishes a weighted measure for counting prime numbers. Similar sums, such as that in Mertens' Second Theorem, are easier to prove once Mertens' First Theorem is established.

Theorem 4.4. ***(Mertens' First Theorem)*** *For $x \geq 2$ and prime p,*

$$\sum_{p \leq x} \frac{\log p}{p} = \log x + O(1), \tag{4.52}$$

where $O(1)$ lies in the open interval $(-1 - \log 4, \log 4)$.

Proof. This proof is adapted from Chapter I.0 of Tenebaum [14] with additional details added by the author.

Let $n = \lfloor x \rfloor$. As shown in Corollary 4.2,

$$\log n! = n \log n - n + 1 - \theta \log n, \tag{4.53}$$

where $\theta = \theta_n \in [0, 1]$. Rewrite $n!$ as its corresponding prime factorization. That is, rewrite $n!$ as $n! = \prod_{i=1}^{m} p_i^{\alpha_1}$, where each unique prime p_i divides $n!$, to obtain

$$\begin{aligned} \log n! &= \log(p_1^{\alpha_1} p_2^{\alpha_2} p_3^{\alpha_3} \ldots p_m^{\alpha_m}) && (4.54)\\ &= \log(p_1^{\alpha_1}) + \log(p_2^{\alpha_2}) + \log(p_3^{\alpha_3}) + \cdots + \log(p_m^{\alpha_m}) && (4.55)\\ &= \alpha_1 \log p_1 + \alpha_2 \log p_2 + \alpha_3 \log p_3 + \cdots + \alpha_m \log p_m. && (4.56) \end{aligned}$$

Here, each α_i is the p_i-adic valuation of $n!$, denoted by $v_p(n!)$. Thus $\log n!$ can be expressed as,

$$\log n! = \sum_{p \le n} v_p(n!) \log p. \tag{4.57}$$

By Corollary 4.1, we have

$$\log n! < \sum_{p \le n} \frac{n}{p} \log p + \sum_{p \le n} \frac{n}{p} \frac{\log p}{(p-1)} = n \sum_{p \le n} \frac{\log p}{p} + n \sum_{p \le n} \frac{\log p}{p(p-1)}, \tag{4.58}$$

and

$$\log n! > \sum_{p \le n} \frac{n}{p} \log p - \sum_{p \le n} 1 \log p = n \sum_{p \le n} \frac{\log p}{p} - \sum_{p \le n} \log p. \tag{4.59}$$

First analyze (4.59). By Theorem 4.3, $\prod_{p \le n} p \le 4^n$. Take the logarithm of both sides, and use the two properties of logs that
$\log(p_1 p_2 \ldots p_n) = \log(p_1) + \log(p_2) + \ldots \log(p_n)$ and $\log x^n = n \log x$ to get

$$\begin{aligned} \log\left(\prod_{p \le n} p\right) &\le \log 4^n && (4.60)\\ \sum_{p \le n} \log p &\le n \log 4. && (4.61) \end{aligned}$$

Therefore,

$$n \sum_{p \le n} \frac{\log p}{p} - n \log 4 < \log(n!) \tag{4.62}$$

$$n \sum_{p \le n} \frac{\log p}{p} - n \log 4 < n \log n - n + 1 + \log n \tag{4.63}$$

$$n \sum_{p \le n} \frac{\log p}{p} - n \log 4 < n \log n, \tag{4.64}$$

where (4.64) holds true because $n > 1 + \log n$ for all $n > 1$. Dividing both sides of (4.64) by n and then adding $\log 4$ provides

$$\sum_{p \le x} \frac{\log p}{p} = \sum_{p \le n} \frac{\log p}{p} \tag{4.65}$$

$$\sum_{p \le n} \frac{\log p}{p} < \log n + \log 4 \tag{4.66}$$

$$\sum_{p \le n} \frac{\log p}{p} \le \log x + \log 4. \tag{4.67}$$

The result in (4.65) holds because $n = \lfloor x \rfloor$. Similarly, the fact that $n = \lfloor x \rfloor$ maintains the inequality in going from (4.66) to (4.67).

Since, for each prime p,

$$\log p > \frac{\log p}{p(p-1)} \text{ and } \sum_{p \le n} \log p \le n \log 4, \tag{4.68}$$

it follows that

$$n \sum_{p \le n} \frac{\log p}{p(p-1)} \le \log 4. \tag{4.69}$$

Then, from (4.53) and (4.58), it stands that

$$n \sum_{p \le n} \frac{\log p}{p} + \log 4 \ge n \sum_{p \le n} \frac{\log p}{p} + n \sum_{p \le n} \frac{\log p}{p(p-1)} > \log n! \tag{4.70}$$

$$n \sum_{p \le n} \frac{\log p}{p} + \log 4 > n \log n - n + 1 \tag{4.71}$$

$$\sum_{p \le n} \frac{\log p}{p} + \frac{\log 4}{n} > \log n - 1 + \frac{1}{n}. \tag{4.72}$$

Since $n = \lfloor x \rfloor$, the sum $\sum_{p \le x} \frac{\log p}{p}$ is equal to the sum $\sum_{p \le n} \frac{\log p}{p}$. Apply this observation and further algebraic manipulation to obtain

$$\sum_{p \le x} \frac{\log p}{p} > \log n + \frac{1}{n} - 1 - \frac{\log 4}{n} \tag{4.73}$$

$$\sum_{p \le x} \frac{\log p}{p} > \log n + \frac{1}{n} - (1 + \log 4) \tag{4.74}$$

$$\sum_{p \le x} \frac{\log p}{p} \ge \log x - (1 + \log 4). \tag{4.75}$$

Since both (4.67) and (4.75) hold true, the theorem follows, and the proof is finished. □

4.5 Proof Of Mertens' Second Theorem

Mertens' Second Theorem provides a means of counting primes by weighting them against their own reciprocals. It is similar in nature to Mertens' First Theorem but much more useful in the proofs found in Chapter 5.

Theorem 4.5. ***(Mertens' Second Theorem)*** *There exists a constant c_1 such that, for $x \ge 2$,*

$$\sum_{p \le x} \frac{1}{p} = \log \log x + c_1 + O\left(\frac{1}{\log x}\right). \tag{4.76}$$

Proof. This proof follows Chapter I.0 of Tenenbaum [14] with the considerable detail deduced and added by the author.

First note that

$$\sum_{p\le x}\frac{1}{p}=\sum_{p\le x}\left(\frac{1}{p}\cdot\frac{\log p}{\log p}\right)=\sum_{p\le x}\left(\frac{1}{\log p}\cdot\frac{\log p}{p}\right). \tag{4.77}$$

Use (4.77) and Riemann-Stieltjes integration to obtain

$$\sum_{p\le x}\frac{1}{p}=\int_{2^-}^{x}\frac{1}{\log t}d\left(\sum_{p\le t}\frac{\log p}{p}\right). \tag{4.78}$$

Before proceeding, it is useful to define $R(t)$ by

$$R(t)\equiv\sum_{p\le t}\frac{\log p}{p}-\log t. \tag{4.79}$$

This definition of $R(t)$ is convenient because Theorem 4.4 (Mertens' First Theorem) provides a good estimate of $R(t)$. That is, according to Theorem 4.4, $R(t)=O(1)$. In other words, the discrete sum $\sum_{p\le x}\frac{\log p}{p}$ is "reasonably close" to the smooth curve $\log t$.

Employ the definition of $R(t)$ along with the fact that

$$\sum_{p\le t}\frac{\log p}{p}=\sum_{p\le t}\frac{\log p}{p}-\log t+\log t$$

to evaluate (4.78). Doing so results in

$$\begin{aligned}\sum_{p\le x}\frac{1}{p}&=\int_{2^-}^{x}\frac{1}{\log t}d\left(\sum_{p\le t}\frac{\log p}{p}\right) && (4.80)\\ &=\int_{2^-}^{x}\frac{1}{\log t}d\left(\sum_{p\le t}\frac{\log p}{p}-\log t+\log t\right) && (4.81)\\ &=\int_{2^-}^{x}\frac{1}{\log t}d(\log t)+\int_{2^-}^{x}\frac{1}{\log t}d\left(\sum_{p\le t}\frac{\log p}{p}-\log t\right) && (4.82)\\ &=\int_{2}^{x}\frac{1}{t\log t}dt+\int_{2^-}^{x}\frac{1}{\log t}d(R(t)). && (4.83)\end{aligned}$$

Going from (4.81) to (4.82) is allowable since the integrands are the same. In going from (4.82) to (4.83), the definition of $R(t)$ and the fact that $d(\log t) = \frac{1}{t}dt$ were used.

First, focus on $\displaystyle\int_2^x \frac{1}{t\log t}dt$. Using a standard u-substitution with $u = \log t$ yields

$$\int_2^x \frac{1}{t\log t}dt = \log\log x - \log\log 2. \tag{4.84}$$

Now, focus on $\displaystyle\int_{2^-}^x \frac{1}{\log t}d(R(t))$. Using integration by parts with $u = \frac{1}{\log t}$, $du = -\frac{1}{t\log^2(t)}$, $dv = d(R(t))$, and $v = R(t)$ yields

$$\int_{2^-}^x \frac{1}{\log t}d(R(t)) = \frac{R(x)}{\log x} - \frac{R(2^-)}{\log 2} + \int_2^x \frac{R(t)}{t\log^2(t)}dt \tag{4.85}$$

$$= \frac{R(x)}{\log x} - \frac{R(2^-)}{\log 2} + \int_2^\infty \frac{R(t)}{t\log^2(t)}dt - \int_x^\infty \frac{R(t)}{t\log^2(t)}dt. \tag{4.86}$$

Here, $\lim\limits_{t\to 2^-} \log t = \log 2$ and $\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$, $(a < b < c)$, were used.

In order to make the remainder of the proof easier to follow, proceed with term-by-term analysis of (4.86). First, note that

$$\frac{R(2^-)}{\log 2} = \frac{1}{\log 2}\left(\lim_{t\to 2^-}\left[\sum_{p\le t}\frac{1}{p} - \log t\right]\right) \tag{4.87}$$

$$= \frac{1}{\log 2}(0 - \log 2) \tag{4.88}$$

$$= -1. \tag{4.89}$$

Since there are no primes less than 2, $\displaystyle\lim_{t\to 2^-}\sum_{p\le t}\frac{1}{p} = 0$.

Next, consider $\int_2^\infty \frac{R(t)}{t\log^2 t}dt$. Since $R(t) = O(1)$ is bounded, it's true for a positive constant K that $|R(t)| \le K$. Using the comparison test for convergence results in

$$\int_2^\infty \frac{|R(t)|}{t\log^2 t}dt \le K\int_2^\infty \frac{1}{t\log^2 t}dt = \frac{K}{\log 2} < \infty. \tag{4.90}$$

The integral involving K was evaluated using a u-substitution with $u = \log t$. This result shows that $\int_2^\infty \frac{R(t)}{t\log^2 t}dt$ is absolutely convergent and thus finite in value, allowing it to be "absorbed" into c_1 (as seen below).

Finally, consider the difference $\frac{R(x)}{\log x} - \int_x^\infty \frac{R(t)}{t\log^2 t}dt$. It is useful to define R by $R = \sup_{t\ge 2^-} |R(t)|$. Then

$$\left|\frac{R(x)}{\log x} - \int_x^\infty \frac{R(t)}{t\log^2 t}dt\right| \le \frac{R}{\log x} + R\int_x^\infty \frac{1}{t\log^2 t}dt \tag{4.91}$$

$$\le \frac{R}{\log x} + \frac{R}{\log x} \tag{4.92}$$

$$\le \frac{2R}{\log x}. \tag{4.93}$$

The integral term was evaluated as before using a u-substitution. From the definition of R and the bounds on $O(1)$ in (4.4), it follows that

$$\frac{2R}{\log x} < \frac{2(1+\log 4)}{\log x}. \tag{4.94}$$

This pins down the $O\left(\frac{1}{\log x}\right)$ term.

Combining the results in (4.84) through (4.94) with (4.83) provides

$$\int_2^x \frac{1}{t\log t}dt + \int_{2^-}^x \frac{1}{\log t}d(R(t)) = \log\log x - \log\log 2 - (-1) + \int_2^\infty \frac{R(t)}{t\log^2 t}dt + O\left(\frac{1}{\log x}\right). \tag{4.95}$$

To finish the proof, let

$$c_1 = 1 - \log\log 2 + \int_2^\infty \frac{R(t)}{t\log^2 t}dt. \tag{4.96}$$

□

4.6 Definitions And Prerequisite Theorems For A gcd-restricted Sum

Before introducing and proving the lemma of interest (see Lemma 4.1), the reader is reminded of a few key definitions and necessary theorems.

Definition 4.2. For $n \in \mathbb{N}$, The **Mőbius function** $\mu(n)$ is defined by $\mu(1) = 1$ and

$$\mu(n) = \begin{cases} (-1)^k, & n = p_1p_2\dots p_k,\ p_i \text{ distinct primes}, \\ 0 & \text{otherwise.} \end{cases}$$

Remark 4.1. Notice by this definition that:

$\mu(n) = 0,$ if and only if n has one or more square factors,

$\mu(n) = -1,$ if n is a product of an odd number of distinct primes to the first power,

and

$\mu(n) = 1,$ if n is a product of an even number of distinct primes to the first power.

The following theorem involving $\mu(n)$ is needed in proving Theorem 4.7, which is subsequently needed to prove Lemma 4.1. The proof of this theorem is adapted from section 2.3 of Apostol [1].

Theorem 4.6. *If* $n \geq 1$, *then*

$$\sum_{d|n} \mu(d) = \left\lfloor \frac{1}{n} \right\rfloor = \begin{cases} 1 & n = 1, \\ 0 & n > 1. \end{cases} \tag{4.97}$$

Proof. If $n = 1$, then $d = 1$, and $\mu(d) = 1 = \lfloor \frac{1}{n} \rfloor$. If $n > 1$, then write $n = \prod_{i=1}^{k} p_i^{\alpha_1} = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$, each p_i being a distinct prime factor of n. The only d's that contribute a nonzero value to $\sum_{d|n} \mu(d)$ are $d = 1$ and any product of distinct primes present in the canonical factorization of n. The sum can thus be expressed as

$$\begin{aligned} \textstyle\sum_{d|n} \mu(d) &= \mu(1) + \mu(p_1) + \mu(p_2) + \cdots + \mu(p_k) \\ &\quad + \mu(p_1 p_2) + \mu(p_2 p_3) + \cdots + \mu(p_{k-1} p_k) \\ &\quad + \cdots + \mu(p_1 p_2 \dots p_k). \end{aligned} \tag{4.98}$$

Recall that $\mu(1) = 1$ by definition. Further, $\mu(n) = -1$ when n is prime. Thus, there are $\binom{k}{1}$ terms equal to -1 in the sum. Similarly $\mu(n) = (-1)^2$ when n is the product of two primes. There are $\binom{k}{2}$ terms in the sum of this sort. Extending this reasoning over the entire sum and applying the binomial theorem (see Theorem 2.2) results in

$$\begin{aligned} \sum_{d|n} \mu(d) &= 1 + \binom{k}{1}(-1) + \binom{k}{2}(-1)^2 + \dots \binom{k}{k-1}(-1)^{k-1} + \binom{k}{k}(-1)^k \\ &= (1-1)^k \\ &= 0, \end{aligned}$$

which concludes the proof. □

Definition 4.3. The **totient function** $\phi(n)$ counts the number of natural numbers $a \in \{1, \dots, n\}$ such that $gcd(a, n) = (a, n) = 1$. In other words, the totient function counts the number of natural numbers (up to n) that are relatively prime to n.

4.6.1 Relating $\mu(n)$ And $\phi(n)$

In proving Lemma 4.1, the following theorem is used. The theorem and subsequent proof in this section are taken from section 2.4 of Apostol [1] with the details expounded upon by the author.

Theorem 4.7. *If $n \geq 1$, then*

$$\phi(n) = \sum_{d|n} \mu(d)\frac{n}{d}. \tag{4.99}$$

Proof. By the definition of $\phi(n)$, it follows that

$$\phi(n) = \sum_{k=1}^{n} \left\lfloor \frac{1}{(n,k)} \right\rfloor. \tag{4.100}$$

Here, $(n,k) = \gcd(n,k)$. If $(n,k) = 1$, then the summation "counts" k as being relatively prime to n by adding one to the total sum. If $(n,k) \geq 1$, then the sum does not count k as being relatively prime to n and zero is added to the total sum (because $\left\lfloor \frac{1}{(n,k)} \right\rfloor = 0$ in this case). Applying Theorem 4.6 with (n,k) in place of n yields

$$\phi(n) = \sum_{k=1}^{n} \left(\sum_{d|(n,k)} \mu(d) \right) = \sum_{k=1}^{n} \sum_{\substack{d|n \\ d|k}} \mu(d). \tag{4.101}$$

The fact that, provided $d \mid (n,k)$, then $d \mid n$ and $d \mid k$, was applied.

The stipulation on the second sum in (4.101) that $d \mid n$ indicates that the sum is taken over each fixed d that divides n. This sum includes all k in the range $1 \leq k \leq n$ that are multiples of d. Thus, write $k = qd$, where q is an integer. Then $1 \leq k \leq n$ if and only if $1 \leq q \leq \dfrac{n}{d}$. The result in (4.101) can thus be expressed as

$$\sum_{k=1}^{n} \sum_{\substack{d|n \\ d|k}} \mu(d) = \sum_{d|n} \sum_{q=1}^{n/d} \mu(d) = \sum_{d|n} \mu(d) \sum_{q=1}^{n/d} 1 = \sum_{d|n} \mu(d)\frac{n}{d}, \tag{4.102}$$

concluding the proof. □

4.7 A gcd-restricted Sum

Lemma 4.1 is used in the proof of Proposition 5.1. The result isn't intuitive, and, as shown below, the proof isn't very straightforward. Granville and Soundararajan [6] use this lemma without any commentary or mention of it.

Lemma 4.1. *Let* $r = \prod_{i=1}^{s} p_i^{\alpha_i}$ *be the prime factorization of* r*, and define* R *as* $R \equiv \prod_{i=1}^{s} p_i$*. That is,* R *is the product of all of the prime factors of* r *to the first power. Let* $\gcd(n, R) = (n, R) = d$*. Then*

$$\sum_{\substack{n \le x \\ (n,R)=d}} 1 = \left(\frac{x}{R}\right) \phi\left(\frac{R}{d}\right) + O\left(\tau\left(\frac{R}{d}\right)\right). \tag{4.103}$$

Proof. Since $d = \gcd(n, R) = (n, R)$, there exists an integer m such that $n = md$. Similarly, $d \mid R$, meaning that $R = \left(\frac{R}{d}\right) d$. Thus, the indices on the sum on the left-hand side of (4.103) can be manipulated as

$$\sum_{\substack{n \le x \\ (n,R)=d}} 1 = \sum_{\substack{md \le x \\ (md,R)=d}} 1 \tag{4.104}$$

$$= \sum_{\substack{md \le x \\ (md,(R/d)d)=d}} 1 \tag{4.105}$$

$$= \sum_{\substack{m \le x/d \\ d(m,R/d)=d}} 1 \tag{4.106}$$

$$= \sum_{\substack{m \le x/d \\ (m,R/d)=1}} 1 \tag{4.107}$$

By similar reasoning used in developing (4.101), (4.107) can be written as

$$\sum_{\substack{m \le x/d \\ (m,R/d)=1}} 1 = \sum_{m \le x/d} \sum_{k \mid (m,R/d)} \mu(k). \tag{4.108}$$

By the way R is defined, R is square-free, meaning that d must be as well. If $k \mid (m, R/d)$, then k must be square-free as well. Similarly, $k \mid m$, meaning m $\equiv 0 \mod k$. Thus, the sum in (4.108) be written as

$$\sum_{m \le x/d} \sum_{k \mid \gcd(m,R/d)} \mu(k) = \sum_{k \mid (R/d)} \mu(k) \sum_{\substack{m \le x/d \\ m \equiv 0 \bmod k}} 1. \tag{4.109}$$

The sum $\sum_{\substack{m \le x/d \\ m \equiv 0 \bmod k}} 1$ counts the natural numbers up to $\frac{x}{d}$ that are divisible by k. The exact count of this inner sum is given by $\left\lfloor \frac{x}{dk} \right\rfloor$. Writing $\frac{x}{dk}$ instead of $\left\lfloor \frac{x}{dk} \right\rfloor$ introduces an error no larger than $\left\{ \frac{x}{dk} \right\}$. The inner sum can thus be replaced by $(\frac{x}{dk} + O(1))$. There are m/k natural numbers divisible by k, going up to n. Using the fact that $m \le \frac{x}{d}$ and the observations about the inner sum results in

$$\sum_{k|(R/d)} \mu(k) \sum_{\substack{m \le x/d \\ m \equiv 0 \bmod k}} 1 = \sum_{k|(R/d)} \mu(k) \left(\frac{x}{dk} + O(1) \right) \tag{4.110}$$

$$= \frac{x}{d} \sum_{K|(R/d)} \frac{\mu(k)}{k} + \sum_{k|(R/d)} \mu(k) O(1) \tag{4.111}$$

$$= \frac{x}{R} \phi \left(\frac{R}{d} \right) + O \left(\sum_{k|(R/d)} 1 \right) \tag{4.112}$$

$$= \frac{x}{R} \phi \left(\frac{R}{d} \right) + O \left(\tau \left(\frac{R}{d} \right) \right). \tag{4.113}$$

In going from (4.111) to (4.112), Theorem 4.7 was used. The error term in (4.111) can be reduced to the error term in (4.112) because $|\mu(k)| \le 1$. The sum in the error term in (4.112) represents the number of divisors of R/d, which is $\tau \left(\frac{R}{d} \right)$. □

Chapter 5

Proving The Erdős-Kac Theorem

This chapter begins with Theorem 5.1, which is the main result of this thesis. Proposition 5.1 is necessary in proving Theorem 5.1. Theorem 5.1 is first proved while assuming that Proposition 5.1 is true. A proof of Proposition 5.1 follows. The statements of Theorem 5.1 and Proposition 5.1 are from Granville and Soundararajan [6].

Theorem 5.1. *For any natural number k let $C_k = \dfrac{\Gamma(k+1)}{2^{k/2}\Gamma\left(\frac{k}{2}+1\right)}$. Then uniformly for even natural numbers $k \leq (\log\log x)^{1/3}$*

$$\sum_{n \leq x} (\omega(n) - \log\log x)^k = C_k x (\log\log x)^{k/2} \left(1 + O\left(\frac{k^3}{\log\log x}\right)\right), \tag{5.1}$$

and uniformly for odd natural numbers k $(\log\log x)^{1/3}$

$$\sum_{n \leq x} (\omega(n) - \log\log x)^k \ll C_k x (\log\log x)^{k/2} \frac{k^{3/2}}{\sqrt{\log\log x}}. \tag{5.2}$$

Remark 5.1. In analyzing Propositions 2 and 3 in Granville and Soundararajan [6] and similar results in Rhoades [10] and Djanković [4], it is believed that the exponents on the O-term in (5.1) should be adjusted to the values expressed above. The details of proof of Proposition 5.1 agree with the adjusted exponents as well.

Remark 5.2. The sum in Theorem 5.1 represents the k^{th} central moment of

$\omega(n)$ (see Definition 3.5). The right-hand sides of (5.1) and (5.2) correspond to k^{th} central moments of the normal distribution (see Theorem 3.5) plus some reasonably bounded error terms. By Theorem 3.6, probability distributions all have unique moments. Therefore, (5.1) and (5.2) combine to give the same result as the Erdős-Kac Theorem.

In proving Theorem 5.1, the following proposition is used. Granville and Soundararajan [6] claim the novelty of their proof of Theorem 5.1 rests on the defined function $f_p(n)$. This proposition will first be assumed true and used to deduce Theorem 5.1. The proof of the proposition will then follow.

Proposition 5.1. *Let p be prime. Let $n \in \mathbb{N}$. Define $f_p(n)$ by*

$$f_p(n) = \begin{cases} 1 - \frac{1}{p} & \text{if } p \mid n \\ -\frac{1}{p} & \text{if } p \nmid n. \end{cases} \tag{5.3}$$

Let $z \geq 10^6$ be a real number. Then uniformly for even natural numbers $k \leq (\log\log z)^{1/3}$

$$\sum_{n \leq x} \left(\sum_{p \leq z} f_p(n) \right)^k = C_k x \left(\log\log z\right)^{k/2} \left(1 + O\left(\frac{k^3}{\log\log z}\right)\right) + O\left(3^k \pi(z)^k\right), \tag{5.4}$$

and uniformly for odd natural numbers $k \leq (\log\log z)^{1/3}$

$$\sum_{n \leq x} \left(\sum_{p \leq z} f_p(n) \right)^k \ll C_k x \left(\log\log z\right)^{k/2} \frac{k^{3/2}}{\sqrt{\log\log z}} + 3^k \pi(z)^k. \tag{5.5}$$

Remark 5.3. In deducing and explicitly writing the details of the proof of Proposition 5.1, the author found that the terms $O(2^k \pi(z)^k)$ and $2^k \pi(z)^k$ (as found in Granville and Soundararajan [6]) should be $O(3^k \pi(z)^k)$ and $3^k \pi(z)^k$ respectively.

5.1 Proof Of The Erdős-Kac Theorem

Proof. The proof is given in Granville and Soundararajan [6] with several details omitted. The proof is reproduced in this section with many of the details deduced and added by the author. The objective is to evaluate $\sum_{n \le x} (\omega(n) - \log\log x)^k$ for natural numbers $k \le (\log\log)^{1/3}$.

First analyze C_k. Using Lemmas 2.3 and 2.4, observe that

$$C_k = \frac{\Gamma(k+1)}{2^{k/2}\Gamma\left(\frac{k}{2}+1\right)} = \frac{k\Gamma(k)}{2^{k/2}\frac{k}{2}\Gamma\left(\frac{k}{2}\right)} = \frac{k(k-1)!}{2^{k/2}\frac{k}{2}\left(\frac{k}{2}-1\right)!} = \frac{k!}{2^{k/2}\left(\frac{k}{2}\right)!}. \tag{5.6}$$

The result in (5.6) matches the coefficient of σ^2 in Theorem 3.5 for even moments of the normal distribution by replacing $2n$ with k. In Theorem 5.1, equation (5.2) indicates that the odd moments are zero within some reasonably bounded error term. This matches the result of Theorem 3.5 for odd moments of the normal distribution. Similarly, comparing (5.1) to Theorem 3.5 (replacing $2n$ with k) shows that

$$E(\omega(n)) = \log\log x, \tag{5.7}$$

and

$$V(\omega(n)) = \log\log x. \tag{5.8}$$

Set $z = x^{1/k}$, and begin by analyzing the difference $\omega(n) - \log\log x$. By the definition of $f_p(n)$, for $p \le z$,

$$\sum_{p \le z} \left(f_p(n) + \frac{1}{p}\right) = \begin{cases} 1 & \text{if } p \mid n \\ 0 & \text{if } p \nmid n. \end{cases} \tag{5.9}$$

In other words, for $p \le z$,

$$\sum_{p \le z} \left(f_p(n) + \frac{1}{p}\right) = \omega(n). \tag{5.10}$$

However, it is possible that there are prime divisors of n that are larger than z. In this case, for $p > z$,

$$\omega(n) = \sum_{\substack{p|n \\ p>z}} 1. \tag{5.11}$$

Combining (5.10) and (5.11) gives

$$\omega(n) = \sum_{p \le z} \left(f_p(n) + \frac{1}{p} \right) + \sum_{\substack{p|n \\ p>z}} 1. \tag{5.12}$$

Substituting the right-hand side of (5.12) for $\omega(n)$ in the difference $\omega(n) - \log\log x$ and rearranging the terms results in

$$\omega(n) - \log\log x = \sum_{p \le z} \left(f_p(n) + \frac{1}{p} \right) + \sum_{\substack{p|n \\ p>z}} 1 - \log\log x \tag{5.13}$$

$$= \sum_{p \le z} f_p(n) + \sum_{\substack{p|n \\ p>z}} 1 + \left(\sum_{p \le z} \frac{1}{p} - \log\log x \right). \tag{5.14}$$

By Mertens' Second Theorem (see Theorem 4.5),

$$-\log\log x = -\sum_{p \le x} \frac{1}{p} + c_1 + O\left(\frac{1}{\log x} \right). \tag{5.15}$$

Substituting (5.15) into the right-hand side of (5.12) yields

$$\omega(n) - \log\log x = \sum_{p \le z} f_p(n) + \sum_{\substack{p|n \\ p>z}} 1 + \left(\sum_{p \le z} \frac{1}{p} - \sum_{p \le x} \frac{1}{p} + c_1 + O\left(\frac{1}{\log x} \right) \right). \tag{5.16}$$

The objective here is to consider how well $\sum_{p \le z} f_p(n)$ estimates $\omega(n) - \log\log x$. The terms contributing the most significant error are $\sum_{\substack{p|n \\ p>z}} 1$ and $\left(\sum_{p \le z} \frac{1}{p} - \sum_{p \le x} \frac{1}{p} \right)$. Since $p > z = x^{1/k}$ and $n \le x$, then there are at most k such primes p that divide

n. This means that $\sum_{\substack{p|n \\ p>z}} 1$ contributes error on the order of $O(k)$. Next observe that

$$\sum_{p \leq z} \frac{1}{p} - \sum_{p \leq x} \frac{1}{p} = - \sum_{z \leq p \leq x} \frac{1}{p}. \tag{5.17}$$

By Mertens' Second Theorem, the error introduced by (5.17) is at most on the order of $O(1)$. Since $O(1) \ll O(k)$, equation (5.15) can be expressed as

$$\omega(n) - \log \log x = \sum_{p \leq z} f_p(n) + O(k). \tag{5.18}$$

Using the result in (5.18), now analyze $(\omega(n) - \log \log x)^k$. Begin with

$$(\omega(n) - \log \log x)^k = \left(\sum_{p \leq z} f_p(n) + O(k) \right)^k. \tag{5.19}$$

Apply the Binomial Theorem (see Theorem 2.2) with $x = \omega(n)$ and $y = O(k)$ to obtain

$$\begin{aligned}
(\omega(n) - \log \log x)^k &= \left(\sum_{p \leq z} f_p(n) + O(k) \right)^k && (5.20)\\
&= \sum_{l=0}^{k} \binom{k}{l} \left(\sum_{p \leq z} f_p(n) \right)^l (O(k))^{k-l} && (5.21)\\
&= \sum_{p \leq z} (f_p(n))^k + \sum_{l=0}^{k-1} \binom{k}{l} \left(\sum_{p \leq z} f_p(n) \right)^l (O(k))^{k-l} && (5.22)\\
&= \sum_{p \leq z} (f_p(n))^k + O\left((ck)^{k-l} \binom{k}{l} \left| \sum_{p \leq z} f_p(n) \right|^l \right), && (5.23)
\end{aligned}$$

where c is some positive constant associated with the $O(k)$ term in (5.19). The use of absolute value bars on the sum of $f_p(n)$ is necessary because $f_p(n)$ can take on negative values.

Sum $(\omega(n) - \log\log x)^k$ over all integers $n \le x$ to get

$$\sum_{n\le x}\left((\omega(n)-\log\log x)^k\right) = \sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k + \sum_{n\le x}\left(O\left((ck)^{k-l}\binom{k}{l}\left|\sum_{p\le z} f_p(n)\right|^l\right)\right) \tag{5.24}$$

$$= \sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k + O\left(\sum_{n\le x}\left((ck)^{k-l}\binom{k}{l}\left|\sum_{p\le z} f_p(n)\right|^l\right)\right). \tag{5.25}$$

The first sum on the right-hand side of (5.25) is handled by Proposition 5.1.

The second sum on the right-hand side of (5.25) establishes the bound on Theorem 5.1. For $l \le k-1$, begin by estimating the expression

$$\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^l. \tag{5.26}$$

If l is even, then equation (5.26) is estimated by (5.4).

If l is odd, then apply the Cauchy-Schwarz Inequality (see Lemma 2.1) to get that

$$\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^l \le \left(\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^{l-1}\right)^{1/2}\left(\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^{l+1}\right)^{1/2}. \tag{5.27}$$

The exponents $l-1$ and $l+1$ are both even and (5.4) of Proposition 5.1 can be used once again. Using equation (5.4) to evaluate each sum results in

$$\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^l \le \left(C_{l-1}x(\log\log z)^{(l-1)/2}\left(1+O\left(\frac{(l-1)^3}{\log\log z}\right)\right)+O\left(3^{l-1}\pi(z)^{l-1}\right)\right)^{1/2} \left(C_{l+1}x(\log\log z)^{(l+1)/2}\left(1+O\left(\frac{(l+1)^3}{\log\log z}\right)\right)+O\left(3^{l+1}\pi(z)^{l+1}\right)\right)^{1/2}. \tag{5.28}$$

Notice that $C_{l-1}x\,(\log\log z)^{(l-1)/2}$ and $C_{l+1}x\,(\log\log z)^{(l+1)/2}$ are the dominant terms in (5.28). This allows (5.28) to be used to deduce that

$$\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^l \ll \left(C_{l-1}x\,(\log\log z)^{(l-1)/2}\right)^{1/2}\left(C_{l+1}x\,(\log\log z)^{(l+1)/2}\right)^{1/2} \tag{5.29}$$

$$\ll \left(C_{l-1}C_{l+1}x^2\,(\log\log x)^l\right)^{1/2} \tag{5.30}$$

$$\ll \sqrt{C_{l-1}C_{l+1}}x\,(\log\log z)^{l/2}. \tag{5.31}$$

Applying the result obtained in (5.31) to (5.25) provides

$$\sum_{n\le x}(\omega(n)-\log\log x)^k = \sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k + O\Bigg(\sum_{\substack{l=0\\ l\equiv 0 \mod 2}}^{k-1}(ck)^{k-l}\binom{k}{l} \left[C_l x\,(\log\log z)^{l/2}\left(1+O\left(\frac{l^3}{\log\log z}\right)\right)+O\left(3^l\pi(z)^l\right)\right]\Bigg)+ O\left(\sum_{\substack{l=0\\ l\equiv 1 \mod 2}}^{k-1}(ck)^{k-l}\binom{k}{l}\sqrt{C_{l-1}C_{l+1}}x\,(\log\log z)^{l/2}\right). \tag{5.32}$$

The task now is to eliminate negligible error terms. Observe that, for $l\le k\le(\log\log x)^{1/3}$,

$$3^l\pi(z)^l \ll 3^k\left(\frac{x^{1/k}}{\log z}\right)^k \tag{5.33}$$

$$\ll 3^k\left(\frac{x}{(\log z)^k}\right) \tag{5.34}$$

$$\ll x, \tag{5.35}$$

for z sufficiently large (as indicated in Proposition 5.1). The result in (5.35) indicates that $3^l\pi(z)^l$ is negligible in magnitude in comparison to the other terms in (5.32). Therefore, equation (5.32) can be reduced to

$$\sum_{n\leq x}(\omega(n)-\log\log x)^k = \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + O\left(\sum_{\substack{l=0\\ l\equiv 0 \mod 2}}^{k-1} (ck)^{k-l}\binom{k}{l}\left[C_l x(\log\log z)^{l/2}\left(1+O\left(\frac{l^3}{\log\log z}\right)\right)\right]\right) + O\left(\sum_{\substack{l=0\\ l\equiv 1 \mod 2}}^{k-1} (ck)^{k-l}\binom{k}{l}\sqrt{C_{l-1}C_{l+1}}x(\log\log z)^{l/2}\right). \quad (5.36)$$

Because $l \leq k$, it is true that

$$\frac{l^3}{\log\log z} \leq \frac{k^3}{\log\log z}. \quad (5.37)$$

Applying (5.37) to (5.36), distributing inside the brackets for the sum when l is even, and rearranging the terms yields

$$\sum_{n\leq}(\omega(n)-\log\log x)^k = \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + O\left(\sum_{\substack{l=0\\ l\equiv 0 \mod 2}}^{k-1} (ck)^{k-l}\binom{k}{l}\left[C_l x(\log\log z)^{l/2}\right]\right) + O\left(\sum_{\substack{l=0\\ l\equiv \mod 2}}^{k-1} (ck)^{k-l}\binom{k}{l}\sqrt{C_{l-1}C_{l+1}}x(\log\log z)^{l/2}\right) + O\left(\frac{k^3}{\log\log z}\sum_{\substack{l=0\\ l\equiv 0 \mod 2}}^{k-1} (ck)^{k-l}\binom{k}{l}\right). \quad (5.38)$$

Because of the constraint that $k \leq (\log\log z)^{1/3}$, the fourth term in (5.38) is negligible in magnitude in comparison to the other three terms. This reduces

(5.38) to

$$\sum_{n\leq} (\omega(n) - \log\log x)^k = \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + O\left(\sum_{\substack{l=0 \\ l\equiv 0 \mod 2}}^{k-1} (ck)^{k-l}\binom{k}{l}\left[C_l x\left(\log\log z\right)^{l/2}\right]\right) + O\left(\sum_{\substack{l=0 \\ l\equiv 1 \mod 2}}^{k-1} (ck)^{k-l}\binom{k}{l}\sqrt{C_{l-1}C_{l+1}}x\left(\log\log z\right)^{l/2}\right). \tag{5.39}$$

The error terms in (5.39) can now be expressed more concisely as

$$\sum_{n\leq x} (\omega(n) - \log\log x)^k = \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + O\left(\sum_{l=0}^{k-1} (ck)^{k-l}\binom{k}{l}\max(C_l, \sqrt{C_{l-1}C_{l+1}})x\frac{C_k}{C_k}\left(\log\log z\right)^{l/2}\right) \tag{5.40}$$

$$= \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + O\left(C_k x\sum_{l=0}^{k-1}\binom{k}{l}\frac{\max(C_l, \sqrt{C_{l-1}C_{l+1}})}{C_k}(ck)^{k-l}\left(\log\log z\right)^{l/2}\right) \tag{5.41}$$

$$= \sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k + O\left(C_k x\sum_{l=0}^{k-1}\binom{k}{l}\frac{\max(C_l, \sqrt{C_{l-1}C_{l+1}})}{C_k}\left(\frac{ck}{\left(\log\log z\right)^{1/2}}\right)^{k-l}\right). \tag{5.42}$$

To finish analyzing the error terms, use the fact that $k \leq (\log\log z)^{1/3}$ and apply the laws of exponents to the $(k-l)$ power obtaining

$$\sum_{n\leq x}(\omega(n)-\log\log x)^k=\sum_{n\leq x}\left(\sum_{p\leq z}f_p(n)\right)^k+O\left(C_k x\left(\log\log z\right)\right)^{k/2-1/6}. \quad (5.43)$$

Currently all of the results are expressed in terms of $\log\log z$, but Theorem 5.1 is in terms of $\log\log x$. To rectify this, observe that

$$z=x^{1/k} \quad (5.44)$$

$$\log z=\log x^{1/k} \quad (5.45)$$

$$\log z=\frac{\log x}{k}. \quad (5.46)$$

Make this substitution for each instance of $\log\log z$ to obtain

$$\log\log z=\log\left(\frac{\log x}{k}\right) \quad (5.47)$$

$$=\log\log x-\log k. \quad (5.48)$$

Recall that $k\leq(\log\log x)^{1/3}$ to get

$$\log\log z\leq\log\log x-\log\left(\log\log x\right)^{1/3} \quad (5.49)$$

$$\log\log z+\log\left(\log\log x\right)^{1/3}\leq\log\log x \quad (5.50)$$

$$\log\log z+O(\log k)=\log\log x. \quad (5.51)$$

Compared to the other error terms, $O(\log k)$ is negligible, allowing $\log\log x$ to be substituted for $\log\log z$.

Applying (5.51) and Proposition 5.1 to (5.43) results in Theorem 1.

□

5.2 Proof Of Proposition 5.1

The proof of Theorem 5.1 relies upon Proposition 5.1. The purpose of this section is to prove Proposition 5.1. The following proof follows Granville and Soundararajan [6] with the omitted details deduced and included by the author. The author disagrees with Granville and Soundararajan on the handling of one of the error terms (see Remark 5.3); however, this does not have a significant impact on the proof of Proposition 5.1 or Theorem 5.1.

Proof. Let $r = \prod_i p_i^{\alpha_i}$ be the prime factorization of r. Since $f_p(n)$ is multiplicative, then

$$f_r(n) = \prod_i f_{p_i}(n)^{\alpha_i}. \tag{5.52}$$

Let $n_x = \lfloor x \rfloor$. Let p_z be the largest prime such that $p_z \leq z$. Suppose $r = \prod_i^s q_i^{\alpha_i}$, where the q_i are distinct primes and $\alpha_i \geq 1$. Set $R = \prod_i^s q_i$. Then

$$\begin{aligned}
\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k &= \sum_{n\leq x}\left(f_{p_1}(n) + f_{p_2}(n) + \cdots + f_{p_z}(n)\right)^k &(5.53)\\
&= \left(f_{p_1}(n_1) + f_{p_2}(n_1) + \cdots + f_{p_z}(n_1)\right)^k + \\
&\quad \left(f_{p_1}(n_2) + f_{p_2}(n_2) + \cdots + f_{p_z}(n_2)\right)^k + \\
&\quad \cdots + \\
&\quad \left(f_{p_1}(n_x) + f_{p_2}(n_x) + \cdots + f_{p_z}(n_x)\right)^k . &(5.54)
\end{aligned}$$

Each grouped term raised to the k^{th} power in (5.54) forms a multinomial that, when expanded, will produce terms of the form $f_{p_1}(n)f_{p_2}(n)\ldots f_{p_k}(n)$. Applying this observation, the multinomial theorem, and (5.52) to (5.54) results in

$$\sum_{n\leq x}\left(\sum_{p\leq z} f_p(n)\right)^k = \sum_{p_1,p_2,\ldots,p_k\leq z}\ \sum_{n\leq x} f_{p_1}\ldots f_{p_k}(n). \tag{5.55}$$

For clarity of notation of the outer sum on the right-hand side of (5.55), note the following example. Suppose we want to evaluate $\sum_{p_1,p_2,\ldots,p_k \le z} p$. Then,

$$\sum_{p_1,p_2,\ldots,p_k \le z} p = \left(\sum_{p_1 \le z} p\right)\left(\sum_{p_2 \le z} p\right)\cdots\left(\sum_{p_k \le z} p\right), \tag{5.56}$$

where $p_1, p_2, \ldots, p_k$ are not ordered and are not necessarily unique.

Now, focus shifts to the sum $\sum_{n \le x} f_r(n)$. If $d = \gcd(n, R) = (n, R)$, then

$$f_r(n) = f_r(d). \tag{5.57}$$

This is true because prime factor q_i of r divides n if and only if q_i divides d. Applying (5.57) to the sum $\sum_{n \le x} f_r(n)$ yields

$$\sum_{n \le x} f_r(n) = \sum_{d|R} f_r(d) \sum_{\substack{n \le x \\ (n,R)=d}} 1. \tag{5.58}$$

The inner sum of (5.58) counts the number of times $f_r(d)$ is in the expansion of $f_r(n)$. Applying the gcd-lemma (see Lemma 4.1) to the inner sum gives

$$\sum_{n \le x} f_r(n) = \sum_{d|R} f_r(d) \left[\frac{x}{R}\phi\left(\frac{R}{d}\right) + O\left(\tau\left(\frac{R}{d}\right)\right)\right] \tag{5.59}$$

$$= \frac{x}{R}\sum_{d|R} f_r(d)\phi\left(\frac{R}{d}\right) + O\left(\sum_{d|R} f_r(d)\tau\left(\frac{R}{d}\right)\right). \tag{5.60}$$

In order to ease further analysis, define $G(r)$ by

$$G(r) \equiv \frac{1}{R}\sum_{d|R} f_r(d)\phi\left(\frac{R}{d}\right). \tag{5.61}$$

It follows that

$$\sum_{n \le x} f_r(n) = xG(r) + O\left(\sum_{d|R} f_r(d)\tau\left(\frac{R}{d}\right)\right). \tag{5.62}$$

The summand of $G(r)$ is a convolution of two multiplicative functions. Suppose that $r = q^\alpha$, where q is prime. Then $R = q$, and

$$G(r) = \frac{1}{R} \sum_{d|R} f_r(d) \phi\left(\frac{R}{d}\right) \tag{5.63}$$

$$= \frac{1}{q} \left(f_1(q)^\alpha(1)\phi(q) + f_q^\alpha(q)\phi(1)\right) \tag{5.64}$$

$$= \frac{1}{q} \left(\left[-\frac{1}{q}\right]^\alpha q \left[1 - \frac{1}{q}\right] + \left[1 - \frac{1}{q}\right]^\alpha \cdot 1\right) \tag{5.65}$$

$$= \left(\left[-\frac{1}{q}\right]^\alpha \left[1 - \frac{1}{q}\right] + \left[\frac{1}{q}\right] \left[1 - \frac{1}{q}\right]\right). \tag{5.66}$$

Therefore,

$$G(r) = \prod_{q^\alpha || r} \left(\frac{1}{q}\left(1 - \frac{1}{q}\right)^\alpha + \left(-\frac{1}{q}\right)^\alpha \left(1 - \frac{1}{q}\right)\right), \tag{5.67}$$

where $r = \prod_i^s q_i^{\alpha_i}$ as originally stated. For more details about establishing (5.67), the reader may wish to examine Theorems 2.13 and 2.14 in Apostol [1].

Next, pare down the error term in (5.62). Since $0 < |f_r(d)| < 1$,

$$O\left(\sum_{d|R} f_r(d)\tau\left(\frac{R}{d}\right)\right) \ll O\left(\sum_{d|R} \tau\left(\frac{R}{d}\right)\right). \tag{5.68}$$

Recall that R is defined to be square-free. Thus, there are 2^s square-free divisors of R, meaning $\tau(R) = 2^s$. Suppose a prime factor of R is fixed in d. Then there are $(s \cdot 2^{s-1})$ factors of R/d, or $\tau(R/d) = (s \cdot 2^{2-1})$. Suppose a product of two prime factors of R are fixed in d. There are $\binom{s}{2}$ ways to choose such factors, and $\tau(R/d) = \binom{s}{2}2^{s-2}$. Continuing this reasoning (keeping in mind there are at most

s prime factors that can be fixed in d) provides

$$\sum_{d|R} \tau\left(\frac{R}{d}\right) = \binom{s}{0}2^s + \binom{s}{1}2^{s-1} + \binom{s}{2}2^{s-2} + \dots \binom{s}{s}2^{s-s} \tag{5.69}$$

$$= \binom{s}{0}2^s1^0 + \binom{s}{1}2^{s-1}1^1 + \binom{s}{2}2^{s-2}1^2 + \dots \binom{s}{s}2^01^s \tag{5.70}$$

$$= \sum_{l=0}^{s} \binom{s}{l}2^l1^{s-l} \tag{5.71}$$

$$= (2+1)^s \tag{5.72}$$

$$= 3^s. \tag{5.73}$$

Here, the Binomial Theorem (see Theorem 2.2) was used in the last three lines. Equation (5.66) can now be written as

$$\sum_{n\le x} f_r(n) = xG(r) + O(3^s). \tag{5.74}$$

Observe that if $q^\alpha \mid\mid r$ and $\alpha = 1$, then

$$G(r) = -\frac{1}{q}\left(1-\frac{1}{q}\right) + \frac{1}{q}\left(1-\frac{1}{q}\right) = 0. \tag{5.75}$$

In other words, $G(r) = 0$ unless r is square-full.

The focus now returns to analyzing equation (5.55). Applying the result in (5.74) and the observation in (5.75) to the right-hand side of (5.55) yields

$$\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k = \sum_{\substack{p1,\dots,p_k\le z\\ p_1\dots p_k \text{ square-full}}} \left(x\, G(p_1\dots p_k) + O(3^k)\right) \tag{5.76}$$

$$= x \sum_{\substack{p1,\dots,p_k\le z\\ p_1\dots p_k \text{ square-full}}} G(p_1\dots p_k) + \sum_{\substack{p1,\dots,p_k\le z\\ p_1\dots p_k \text{ square-full}}} O(3^k) \tag{5.77}$$

$$= x \sum_{\substack{p1,\dots,p_k\le z\\ p_1\dots p_k \text{ square-full}}} G(p_1\dots p_k) + O\left(\sum_{p1,\dots,p_k\le z} 3^k\right) \tag{5.78}$$

$$= x \sum_{\substack{p1,\dots,p_k\le z\\ p_1\dots p_k \text{ square-full}}} G(p_1\dots p_k) + O\left(3^k\pi(z)^k\right). \tag{5.79}$$

The error term in (5.79) comes from the fact that there are $\pi(z)$ primes less than or equal to z and thus there are $\pi(z)^k$ summands in the O-term.

The primes $p_1 \dots p_k \leq z$ are not necessarily unique or ordered. Suppose $q_1 < q_2 < \cdots < q_s$ are the distinct primes in $p_1 \dots p_k \leq z$. Since $p_1 \dots p_k$ is square-full, then $s \leq k/2$. Thus the main term in (5.79) can be expressed as

$$\sum_{\substack{p1,\dots,p_k \leq z \\ p_1 \dots p_k \text{ square-full}}} G(p_1 \dots p_k) = \sum_{s \leq k/2} \sum_{q_1 < q_2 < \cdots < q_s \leq z} \sum_{\substack{\alpha_1,\dots\alpha_s \geq 2 \\ \sum_i \alpha_i = k}} \frac{k!}{\alpha_1! \dots \alpha_s!} G\left(q_1^{\alpha_1} \dots q_s^{\alpha_s}\right). \tag{5.80}$$

The expression $\dfrac{k!}{\alpha_1! \dots \alpha_s!}$ represents the multinomial coefficients on the expansion of $G\left(q_1^{\alpha_1} \dots g_s^{\alpha_s}\right)$. The multinomials must be summed over the distinct primes, of which there are $s \leq k/2$.

Suppose k is even. Then $s = k/2$, and there is a term in (5.80) with all $\alpha_i = 2$. This term contributes

$$\frac{k!}{2^{k/2}(k/2)!} \sum_{\substack{q_1,\dots,q_{k/2} \leq z \\ q_i \text{ distinct}}} \prod_{i=1}^{k/2} \frac{1}{q_i}\left(1 - \frac{1}{q_i}\right). \tag{5.81}$$

The fact that when $\alpha = 2$, $G(r) = \prod_{q^\alpha || r} \frac{1}{q}\left(1 - \frac{1}{q}\right)$, was used in place of $G\left(q_1^{\alpha_1} \dots q_s^{\alpha_s}\right)$. The $2^{k/2}$ arises from the fact that each $\alpha_i! = 2! = 2$ and that there are $k/2$ α_i-factors in the denominator. The distinct primes in the sum are not ordered. There are $s! = (k/2)!$ ways to order the $k/2$ distinct primes. Thus, sum is $(k/2)!$ factorial times the value that (5.80) indicates. This is accounted for with the $(k/2)!$ added to the denominator.

Upper and lower bounds on the sum in (5.81) are now sought. Each factor in the product $\prod_{i=1}^{k/2} \frac{1}{q_i}\left(1 - \frac{1}{q_i}\right)$ is less than 1. Therefore,

$$\prod_{i=1}^{k/2} \frac{1}{q_i}\left(1 - \frac{1}{q_i}\right) \leq \frac{1}{p}\left(1 - \frac{1}{p}\right), \tag{5.82}$$

where p could be any of the q_i's. To further maximize the sum, remove the condition that the q's are unique. The sum in (5.81) is thus bounded above by

$$\left(\sum_{p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right)\right)^{k/2}, \tag{5.83}$$

where p is prime.

In seeking a lower bound, consider the case when $q_1,\ldots,q_{k/2-1}$ are given. Then the sum over $q_{k/2}$ (which is the inner-most sum in (5.81)) is at least

$$\sum_{\pi_{k/2}\le p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right), \tag{5.84}$$

where $\pi_{k/2}$ represents the $(k/2)^{th}$ smallest prime number. This forms a lower bound because p is restricted in its minimum value, and $\frac{1}{p}\left(1-\frac{1}{p}\right)$ strictly decreases over the primes. Repeating this reasoning over the remaining $(k/2-1)$ sums provides a lower bound of

$$\left(\sum_{\pi_{k/2}\le p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right)\right)^{k/2}. \tag{5.85}$$

By Merten's Second Theorem (see Theorem 4.5),

$$\begin{aligned}\sum_{p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right) &= \sum_{p\le z}\frac{1}{p}-\sum_{p\le z}\frac{1}{p^2} && (5.86)\\ &= \log\log z + c_1 + O(1) - \sum_{p\le z}\frac{1}{p^2} && (5.87)\\ &= \log\log z + O(1). && (5.88)\end{aligned}$$

Similarly,

$$\begin{aligned}\sum_{\pi_{k/2}\le p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right) &= \sum_{p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right)-\sum_{p\le \pi_{k/2}}\frac{1}{p}\left(1-\frac{1}{p}\right) && (5.89)\\ &= \log\log z + O(1) - \big(\log\log \pi_{k/2} + O(1)\big) && (5.90)\\ &= \log\log z + O(1) - \left(\log\log k + O(1)\right). && (5.91)\end{aligned}$$

The upper and lower bounds in (5.82) and (5.84) establish that

$$\left(\sum_{\pi_{k/2}\le p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right)\right)^{k/2} \le \sum_{\substack{q_1,\ldots,q_{k/2}\le z\\ q_i \text{ distinct}}}\prod_{i=1}^{k/2}\frac{1}{q_i}\left(1-\frac{1}{q_i}\right) \le \left(\sum_{p\le z}\frac{1}{p}\left(1-\frac{1}{p}\right)\right)^{k/2}. \tag{5.92}$$

Apply the implications of Mertens' Second Theorem in (5.88) and (5.91) to (5.92) to get

$$\sum_{\substack{q_1,\ldots,q_{k/2}\le z\\ q_i \text{ distinct}}}\prod_{i=1}^{k/2}\frac{1}{q_i}\left(1-\frac{1}{q_i}\right) = (\log\log z + O(1+\log\log k))^{k/2}. \tag{5.93}$$

To estimate the terms $s < k/2$, first observe that

$$0 \le G\left(q_1^{\alpha_1}\ldots q_s^{\alpha_s}\right) \le \frac{1}{q_1\ldots q_s}. \tag{5.94}$$

In order to understand the validity of (5.94), recall that $G(r) = 0$ if r is not square-full. If r is square-full, consider the expression

$$\frac{1}{q}\left(1-\frac{1}{q}\right)^{\alpha} + \left(-\frac{1}{q}\right)^{\alpha}\left(1-\frac{1}{q}\right). \tag{5.95}$$

In the current context, each $\alpha \ge 2$. Then, for a given prime, the expression in (5.95) decreases as α increases. Therefore, for a given prime, (5.95) is a maximum value when $\alpha = 2$. Let $\alpha = 2$ and observe that

$$\frac{1}{q}\left(1-\frac{1}{q}\right)^{2} + \left(-\frac{1}{q}\right)^{2}\left(1-\frac{1}{q}\right) = \frac{1}{q}-\frac{1}{q^2} \le \frac{1}{q}. \tag{5.96}$$

For legibility, let A represent the right-hand side of (5.80). That is, let

$$A = \sum_{s\le k/2}\ \sum_{q_1<q_2<\cdots<q_s\le z}\ \sum_{\substack{\alpha_1,\ldots\alpha_s\ge 2\\ \sum_i \alpha_i = k}} \frac{k!}{\alpha_1!\ldots\alpha_s!}G\left(q_1^{\alpha_1}\ldots q_s^{\alpha_s}\right).$$

Apply (5.94) to the right-hand side of (5.80) to get

$$A \leq \sum_{s<k/2} \sum_{q_1<q_2<\cdots<q_s\leq z} \sum_{\substack{\alpha_1,\ldots\alpha_s\geq 2\\ \sum_i \alpha_i=k}} \frac{k!}{\alpha_1!\ldots\alpha_s!}\frac{1}{q_1\ldots q_s} \tag{5.97}$$

$$A \leq \sum_{s<k/2} \frac{k!}{s!}\left(\sum_{q\leq z}\frac{1}{q}\right)^s \sum_{\substack{\alpha_1,\ldots\alpha_s\geq 2\\ \sum_i \alpha_i=k}} \frac{1}{\alpha_1!\ldots\alpha_s!}. \tag{5.98}$$

The introduction of $s!$ in (5.99) follows for the same reason $(k/2)!$ was introduced in (5.81).

Use Mertens' Second Theorem (see Theorem 4.5) to obtain

$$\left(\sum_{q\leq z}\frac{1}{q}\right)^s = (\log\log z + O(1))^s. \tag{5.99}$$

To evaluate the inner-most sum, the number of ways that k can be composed as, $k = \alpha_1 + \cdots + \alpha_s$ with each $\alpha_i \geq 2$, must be found. Subtract one from each α_i, of which there are s, to write

$$k - s = (\alpha_1 - 1) + (\alpha_2 - 1) + \cdots + (\alpha_s - 1) \tag{5.100}$$

$$k - s = \beta_1 + \beta_2 + \cdots + \beta_s, \tag{5.101}$$

with each $\beta_i \geq 1$. There are $\binom{k-s-1}{s-1}$ ways to compose $k - s$ as $\sum_i \beta_i$. Adding s to each side of (5.101) only adds a set constant to both sides. This means that there are also $\binom{k-s-1}{s-1}$ ways to compose k as $\sum_i \alpha_1$. To ensure the inequality in (5.95) holds, replace each α_i with the minimum possible value of 2. There are s α_i-terms, introducing $(2!)^s = 2^s$ in the denominator of the inner-most sum in (5.95). Apply the observations about the number of compositions of k and using 2^s along with (5.95) to (5.99) to get

$$A \leq \sum_{s<k/2} \frac{k!}{s!2^s}\binom{k-s-1}{s-1}(\log\log z + O(1))^s. \tag{5.102}$$

All that remains is to show that (5.79), (5.93), and (5.102) combine to give Proposition 5.1. Begin with the sum in (5.79). For k even, the main term of (5.79) is given by (5.93), which is

$$\frac{k!}{(k/2)!2^{k/2}}x(\log\log z + O(1+\log\log k))^{k/2} = C_k x(\log\log z + O(1+\log\log k))^{k/2}, \tag{5.103}$$

where $C_k = \frac{k!}{(k/2)!2^{k/2}}$. Factor the $\log\log z$ term to obtain

$$\begin{aligned} &C_k x(\log\log z + O(1+\log\log k))^{k/2} \\ &= C_k x\,(\log\log z)^{k/2}\left[1 + O\left(\frac{1+\log\log k}{\log\log z}\right)\right]^{k/2} \qquad (5.104)\\ &= C_k x\,(\log\log z)^{k/2}\left[1 + O\left(\sum_{s=1}^{k/2}\binom{k/2}{s}\left(\frac{\log\log k}{\log\log z}\right)^{s}\right)\right]. \qquad (5.105)\end{aligned}$$

The Binomial Theorem (see Theorem 2.2) was used in going from (5.104) to (5.105). Because $k \le (\log\log z)^{1/3}$, the first term in the sum in (5.105) is the largest. The sum is thus composed of $k/2$ terms, each of which is no larger than $\frac{k\log\log k}{\log\log z}$. Use the bound that $\log\log k \ll k$ to get

$$\begin{aligned}\sum_{s=1}^{k/2}\binom{k/2}{s}\left(\frac{\log\log k}{\log\log z}\right)^{s} &\le \frac{k}{2}\frac{k\log\log k}{\log\log z} \qquad (5.106)\\ &\le \frac{k}{2}\frac{k\cdot k}{\log\log z} \qquad (5.107)\\ &\le \frac{k^3}{\log\log z}. \qquad (5.108)\end{aligned}$$

Therefore, for k even, the main term in (5.79) is

$$(\log\log z)^{k/2}\left[1 + O\left(\frac{k^3}{\log\log z}\right)\right]. \tag{5.109}$$

For the remaining terms, those with $s < k/2$, equation (5.102) indicates terms bounded by

$$\sum_{s<k/2} \frac{k!}{s!2^s} \binom{k-s}{s} (\log\log z + O(1))^s \tag{5.110}$$

must be accounted for. Let T denote the largest integer less than $k/2$. Then, multiply by $\frac{C_k}{C_k}$ and extract the dominant power of $\log\log z$ by writing

$$\begin{aligned}
&\sum_{s<k/2} \frac{k!}{s!2^s} \binom{k-s}{s} (\log\log z + O(1))^s \\
&\ll C_k (\log\log z)^T \sum_{s\le T} \frac{k!}{C_k s!2^s} \binom{k-s}{s} (\log\log z)^{s-T} \qquad (5.111)\\
&\ll C_k (\log\log z)^T \sum_{s\le T} \frac{k!}{C_k s!2^s} \binom{k-s}{s} \left(\frac{1}{\log\log z}\right)^{T-s}. \qquad (5.112)
\end{aligned}$$

Continue by using the facts that $C_k = \dfrac{\Gamma(k+1)}{2^{k/2}\Gamma(k/2+1)}$ and $k^3 \le \log\log z$ to write

$$\begin{aligned}
&\sum_{s<k/2} \frac{k!}{s!2^s} \binom{k-s}{s} (\log\log z + O(1))^s \\
&\ll C_k (\log\log z)^T \sum_{s\le T} \frac{k!}{C_k s!2^s} \binom{k-s}{s} \left(\frac{1}{k^3}\right)^{T-s} \qquad (5.113)\\
&\ll C_k (\log\log z)^T \sum_{s\le T} \frac{k!}{s!2^s} \frac{2^{k/2}\Gamma(k/2+1)}{\Gamma(k+1)} \binom{k-s}{s} \left(\frac{1}{k^3}\right)^{T-s} \qquad (5.114)\\
&\ll C_k (\log\log z)^T \sum_{s\le T} \frac{2^{k/2}}{2^s} \frac{\Gamma(k/2+1)}{s!} \binom{k-s}{s} \left(\frac{1}{k^3}\right)^{T-s}. \qquad (5.115)
\end{aligned}$$

If k is even, then $T = \frac{k}{2} - 1$, and (5.115) becomes

$$\begin{aligned}
&C_k (\log\log z)^T \sum_{s\le T} \frac{2^{k/2}}{2^s} \frac{\Gamma(k/2+1)}{s!} \binom{k-s}{s} \left(\frac{1}{k^3}\right)^{T-s} \\
&\ll C_k \frac{(\log\log z)^{k/2}}{\log\log z} \sum_{s\le T} \frac{2^{k/2}}{2^s} \frac{\Gamma(\frac{k}{2}+1)}{s!} \binom{k-s}{s} \left(\frac{1}{k^3}\right)^{\frac{k}{2}-1-s} \qquad (5.116)\\
&\ll C_k \frac{(\log\log z)^{k/2}}{\log\log z} k^3 \sum_{s\le T} \frac{2^{k/2}}{2^s} \frac{\Gamma(\frac{k}{2}+1)}{s!} \binom{k-s}{s} \left(\frac{1}{k^3}\right)^{k/2-s}. \qquad (5.117)
\end{aligned}$$

If k is odd, then $T = \frac{k-1}{2}$, and (5.115) becomes

$$C_k(\log\log z)^T \sum_{s\le T} \frac{2^{k/2}}{2^s}\frac{\Gamma(\frac{k}{2}+1)}{s!}\binom{k-s}{s}\left(\frac{1}{k^3}\right)^{T-s}$$

$$\ll C_k \frac{(\log\log z)^{k/2}}{\sqrt{\log\log z}} \sum_{s\le T} \frac{2^{k/2}}{2^s}\frac{\Gamma(\frac{k}{2}+1)}{s!}\binom{k-s}{s}\left(\frac{1}{k^3}\right)^{\frac{(k-1)}{2}-s} \tag{5.118}$$

$$\ll C_k \frac{(\log\log z)^{k/2}}{\sqrt{\log\log z}} k^{3/2} \sum_{s\le T} \frac{2^{k/2}}{2^s}\frac{\Gamma(\frac{k}{2}+1)}{s!}\binom{k-s}{s}\left(\frac{1}{k^3}\right)^{k/2-s}. \tag{5.119}$$

The sum

$$\sum_{s\le T} \frac{2^{k/2}}{2^s}\frac{\Gamma(\frac{k}{2}+1)}{s!}\binom{k-s}{s}\left(\frac{1}{k^3}\right)^{k/2-s} \tag{5.120}$$

is common to both (5.117) and (5.119). To obtain the desired result, observe that

$$\sum_{s\le T} \frac{2^{k/2}}{2^s}\frac{\Gamma(\frac{k}{2}+1)}{s!}\binom{k-s}{s}\left(\frac{1}{k^3}\right)^{k/2-s} \ll 1. \tag{5.121}$$

Proposition 5.1 now immediately follows from (5.109), (5.119), and (5.121).

□

Bibliography

[1] Tom M. Apostol, *Introduction to analytic number theory*, Springer-Verlag, New York-Heidelberg, 1976, Undergraduate Texts in Mathematics. MR 0434929 (55 #7892)

[2] George B. Arfken and Hans J. Weber, *Mathematical methods for physicists*, fourth ed., Academic Press, Inc., San Diego, CA, 1995. MR 1423357 (98a:00001)

[3] Hubert Delange, *Sur le nombre des diviseurs premiers de n*, C. R. Acad. Sci. Paris **237** (1953), 542–544. MR 0057282 (15,201f)

[4] Goran Djanković, *The Erdős-Kac theorem for curvatures in integral Apollonian circle packings*, Publications De L'Institut Mathématique **89** (2001), 11–17.

[5] P. Erdös and M. Kac, *The Gaussian law of errors in the theory of additive number theoretic functions*, Amer. J. Math. **62** (1940), 738–742. MR 0002374 (2,42c)

[6] Andrew Granville and K. Soundararajan, *Sieving and the Erdős-Kac theorem*, Equidistribution in number theory, an introduction, NATO Sci. Ser. II Math. Phys. Chem., vol. 237, Springer, Dordrecht, 2007, pp. 15–27. MR 2290492 (2008b:11103)

[7] H. Halberstam, *On the distribution of additive number-theoretic functions*, J. London Math. Soc. **30** (1955), 43–53. MR 0066406 (16,569g)

[8] ———, *On the distribution of additive number-theoretic functions. II*, J. London Math. Soc. **31** (1956), 1–14. MR 0073626 (17,461d)

[9] Steven J. Miller and Ramin Takloo-Bighash, *An invitation to modern number theory*, Princeton University Press, Princeton, NJ, 2006, With a foreword by Peter Sarnak. MR 2208019 (2006k:11002)

[10] Robert C. Rhoades, *Statistics of prime divisors in function fields*, Int. J. Number Theory **5** (2009), no. 1, 141–152. MR 2499026 (2010c:11124)

[11] Kyle Siegrist, *Virtual laboratories in probability and statistics*, `http://www.math.uah.edu/stat/index.html`, 1997-2014, Online; Last accessed 01192015.

[12] J. Michael Steele, *The Cauchy-Schwarz master class*, MAA Problem Books Series, Mathematical Association of America, Washington, DC; Cambridge University Press, Cambridge, 2004, An introduction to the art of mathematical inequalities. MR 2062704 (2005a:26035)

[13] W. A. Stein et al., *Sage Mathematics Software (Version 6.8)*, 2015, `http://cloud.sagemath.org`.

[14] Gérald Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge Studies in Advanced Mathematics, vol. 46, Cambridge University Press, Cambridge, 1995, Translated from the second French edition (1995) by C. B. Thomas. MR 1342300 (97e:11005b)

[15] Mendenhall William Wackerly, Dennis and Richard L. Scheaffer, *Mathematical statistics with applications*, seventh ed., Duxbury Press, North Scituate, Mass., 2008.