1-1-2022

# ON THE DISTRIBUTION OF THE NUMBER OF PRIME FACTORS OF AN INTEGER

Jacob Mitchell Derrick

# ON THE DISTRIBUTION OF THE NUMBER OF PRIME FACTORS OF AN INTEGER

THESIS

A Thesis
presented in partial fulfillment of requirements
for the degree of Master of Science
in the Department of Mathematics
The University of Mississippi

by

JACOB DERRICK

May 2022

## ABSTRACT

The distribution of the prime numbers has intrigued number theorists for centuries. As our understanding of this distribution has evolved, so too have our methods of analyzing the related arithmetic functions. If we let $\omega(n)$ denote the number of distinct prime divisors of a natural number $n$, then the celebrated Erdős–Kac Theorem states that the values of $\omega(n)$ are normally distributed (satisfying a central limit theorem as $n$ varies). This result is considered the beginning of Probabilistic Number Theory. We present a modern proof of the Erdős–Kac Theorem using a moment based argument due to Granville and Soundararajan, which we explain in full detail. We also use similar techniques to study the second moment of $\omega(n)$, refining a classical result of Turán.

ACKNOWLEDGEMENTS

I would first like to thank Dr. Khan and Dr. Milinovich. Not only did they share their passion for number theory with me, but their investment in me as a student has made me a better mathematician and has given me the tools to continue that growth.

Additionally, I would like to thank Dr. Clark. I truly appreciate his time and the help he has given while on my committee as well as in his courses, all of which prepared me in writing this thesis.

I am also extremely grateful for my family and friends who have supported and encouraged me throughout my time in the University of Mississippi. Specifically, I would like to thank Phil Kains. His friendship and advice were instrumental throughout this entire process.

Lastly, I must thank my wonderful wife, Alison Derrick. Her love, patience, and support throughout my time at the University of Mississippi allowed me to constantly keep pushing forward.

TABLE OF CONTENTS

# 1 INTRODUCTION

A classical problem in number theory is to understand the distribution of the number of prime divisors of an integer. For $n \in \mathbb{N}$, define

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

where the sum counts the distinct prime divisors $p$ of $n$. For example, $\omega(6) = 2$, $\omega(9) = 1$, and $\omega(1000) = \omega(2^3 5^3) = 2$. Note that $\omega(p) = 1$ for any prime $p$. It is also well-known that

$$\omega(n) = O\left(\frac{\log n}{\log\log n}\right)$$

for any $n \in \mathbb{N}$; see [7, Thm. 2.10].

In this thesis, we study the distribution of $\omega(n)$ when picked uniformly at random from the set $\{1, \dots, x\}$ for large $x$. In Chapter 4, we show that the mean of $\omega(n)$ satisfies

$$\frac{1}{x}\sum_{n \le x} \omega(n) = \log\log x + O(1)$$

for large $x$. Thus, we can expect the typical integer $n \le x$ to have about $\log\log x$ distinct prime divisors with an error only up to some constant. It would be impossible to proceed any further into this field without mention of a foundational theorem of probabilistic number theory by Godfrey Hardy and Srinivasa Ramanujan, which they proved in [6].

**Theorem 1.0.1** (Hardy-Ramanujan Theorem)**.** *For almost all integers $n \leq x$, the function $\omega(n)$ has normal order* $\log\log n$. *That is to say*

$$|\omega(n) - \log\log n| < \epsilon \log\log n$$

*for any $\epsilon > 0$ and all but $o_\epsilon(x)$ integers up to $x$.*

Roughly speaking, they showed that there tends to be little difference between $\omega(n)$ and $\log\log n$, that difference typically only being about $\sqrt{\log\log n}$. This proves to be important as we can now begin consider what the distribution of $\omega(n)$ might be. Before we move to this question we must mention Paul Turán's contribution. Turán gave an incredibly elementary proof of the Hardy-Ramanujan Theorem [8] by adding a probabilistic twist and studying the second moment of $\omega(n)$. His proof relies on the estimate

$$\sum_{n \leq x} \omega(n)^2 = x(\log\log x)^2 + O(x \log\log x).$$

In Chapter 3, we refine this calculation by finding explicit constants $A$ and $B$ such that

$$\frac{1}{x}\sum_{n \leq x} \omega(n)^2 = (\log\log x)^2 + A\log\log x + B + O\left(\frac{\log\log x}{\log x}\right).$$

See Theorem 3.3.1 for a precise statement of this result. Although this may be known to experts, we were unable to find this result in the literature.

Because the curiosity of mathematicians knows no bounds these results, which give us both a mean and second moment respectively, only spark more intrigue into the mysteries of $\omega(n)$. Thus, we now find ourselves staring straight at the aforementioned question regarding the distribution of $\omega(n)$ armed with intuition borrowed from the field of probability. Knowing that $\omega(n)$ and $\log\log n$ are only ever about $\sqrt{\log\log n}$ apart, would this have an identifiable distribution? Paul Erdős and Mark Kac were the first to answer this question in their influential paper [3], where they proved the following result.

**Theorem 1.0.2** (Erdős-Kac Theorem)**.** *For each* $\alpha \in \mathbb{R}$ *we have,*

$$\frac{1}{x} \sum_{\substack{n \leq x \\ \omega(n) - \log\log x \leq \alpha\sqrt{\log\log x}}} 1 \longrightarrow \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\alpha} e^{\frac{-t^2}{2}} dt$$

*as* $x \to \infty$.

This incredible finding did many things. Namely it proved Kac's suspicions that $\omega(n)$ was distributed similarly to that of a normal distribution, implying that

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

is approximately normally distributed with mean 0 and variance 1 when $n$ is chosen uniformly at random, and it birthed the field of probabilistic number theory as we know it.

In Chapter 4, we present a modern proof of the Erdős-Kac Theorem due to a paper by Andrew Granville and Kannan Soundararajan [5], which illustrates a way to compute all moments of $\omega(n)$. As well as giving an exposition of their work, we will provide the full details, which were not given in [5]. We hope that this chapter will be a useful addition to the literature for mathematicians interested in and attempting to understand this approach to the famous Erdős-Kac Theorem.

# 2 PRELIMINARY RESULTS

Though by no means intuitive, all of our findings can remarkably be derived from just a few fundamental number theory-based estimates and some observations based in analysis. Below we provide the needed results.

## 2.1 Number theory lemmas

Starting with the prerequisite number theory knowledge needed, we introduce Chebyshev's upper bound to the prime counting function $\pi(x)$ defined as

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

for a prime $p$.

**Lemma 2.1.1** (Chebyshev's upper bound for $\pi(x)$)**.** *For $x \geq 2$, we have*

$$\pi(x) := \sum_{p \leq x} 1 = O\left(\frac{x}{\log x}\right).$$

*Proof.* This is Corollary 2.6 [7]. □

Another lemma that will prove to be crucial to us is Mertens' estimates.

**Lemma 2.1.2** (Mertens' estimates)**.** *For $x \geq 2$, we have*

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

*and*

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

*where $b$ can be expressed using Euler's constant $\gamma_0$ as below*

$$b = \gamma_0 - \sum_{p}\sum_{k=2}^{\infty}\frac{1}{kp^k}.$$

*Proof.* See [7]. □

Lastly, we will need an observation regarding the Möbius function, which is defined as

$$\mu(n) := \begin{cases} 0 & \textit{if } n \textit{ has one or more repeated prime factors} \\ 1 & \textit{if } n = 1 \\ (-1)^k & \textit{if } n \textit{ is a product of } k \textit{ distinct primes.} \end{cases}$$

**Lemma 2.1.3.** *For the Möbius function $\mu(n)$,*

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

.

*Proof.* For $n = 1$,

$$\sum_{d|1}\mu(d) = \mu(1) = 1.$$

Now assume $n > 1$. We may rewrite $n = \prod_{i=1}^{r} p_i^{\alpha_i}$. It follows $d \mid n$ if and only if $d = \prod_{i=1}^{r} p_i^{\beta_i}$ for $\beta_i \leq \alpha_i$ and $i \in \{1, \ldots, r\}$. However, $\mu(d) = 0$ if $\beta_r > 1$. Therefore,

$$\begin{aligned}\sum_{d|n} \mu(d) &= \sum_{\substack{\beta_1,\ldots,\beta_r \\ \beta_i \leq 1}} \mu\left(\prod_{i=1}^{r} p_i^{\beta_i}\right) \\ &= (-1)^0 \binom{r}{0} + (-1)^1 \binom{r}{1} + (-1)^2 \binom{r}{2} + \cdots + (-1)^r \binom{r}{r}.\end{aligned}$$

Here we use the binomial expansion theorem, and our right-hand side is simply,

$$(1-1)^r = 0.$$

So our lemma holds for either case. □

## 2.2 Analysis lemmas

In this section we state and prove analytical results that will be utilized later in this thesis.

**Lemma 2.2.1.** *For $u \in \mathbb{C}$ with $|u| \leq \frac{1}{2}$, we have*

$$\left|\log\left(\tfrac{1}{1-u}\right)\right| \leq \frac{3}{2}|u|.$$

*Proof.* We begin with the well-known Taylor series,

$$\log\left(\tfrac{1}{1-u}\right) = u + \frac{u^2}{2} + \frac{u^3}{3} + \cdots ,$$

which is valid for $|u| < 1$. Therefore, for $|u| \leq \frac{1}{2}$, we have

$$\begin{aligned}\left|\log\left(\tfrac{1}{1-u}\right)\right| &\leq |u| + \frac{|u|^2}{2} + \frac{|u|^3}{3} + \cdots \\ &\leq |u| + \frac{1}{2}(|u|^2 + |u|^3 + \cdots) \\ &= |u| + \frac{|u|^2}{2}(1 + |u| + |u|^2 + \cdots) \\ &= |u| + \frac{|u|^2}{2}\left(\frac{1}{1-|u|}\right).\end{aligned}$$

The last step comes from rewriting the geometric series. Thus, by strategically using substitutions for $u \leq \frac{1}{2}$ we have

$$\begin{aligned}\left|\log\left(\tfrac{1}{1-u}\right)\right| &\leq |u| + \frac{|u|}{4}\left(\frac{1}{1-\frac{1}{2}}\right) \\ &= \frac{3|u|}{2}\end{aligned}$$

as claimed. □

Now we move to an intricate integral that will prove invaluable to our refined estimate of the second moment of $\omega(n)$.

**Lemma 2.2.2.** *For $x \geq 4$, we have*

$$\int_2^{\sqrt{x}} \frac{\log\log u}{u \log \frac{x}{u}} du = (\log 2)\log\log x - \frac{1}{2}(\log 2)^2 - \frac{\pi^2}{12} + O\left(\frac{1}{\log x}\right).$$

*Proof.* Since

$$\left|\int_2^{e} \frac{\log\log u \, du}{u \log \frac{x}{u}}\right| \ll \frac{1}{\log x}\int_2^{e} \frac{|\log\log u| \, du}{u} \ll \frac{1}{\log x},$$

we first rewrite the integral as

$$\int_{2}^{\sqrt{x}} \frac{\log\log u\,du}{u\log\frac{x}{u}} = \int_{e}^{\sqrt{x}} \frac{\log\log u\,du}{u\log\frac{x}{u}} + \int_{2}^{e} \frac{\log\log u\,du}{u\log\frac{x}{u}}$$
$$= \int_{e}^{\sqrt{x}} \frac{\log\log u\,du}{u\log\frac{x}{u}} + O\left(\frac{1}{\log x}\right).$$

We note that the integrand in this new integral is nonnegative over the range of integration from $e$ to $\sqrt{x}$. This proves helpful since $\log\log e = 0$. Next, we estimate this new integral by expanding the term $\frac{1}{\log\frac{x}{u}}$ as a geometric series, interchanging the resulting sum and integral, and then integrating term-by-term. Since $0 < \frac{\log u}{\log x} \leq \frac{1}{2}$, we have

$$\int_{e}^{\sqrt{x}} \frac{\log\log u}{u\log\frac{x}{u}}du = \int_{e}^{\sqrt{x}} \frac{\log\log u}{u\left(\log x - \log u\right)}du$$
$$= \frac{1}{\log x}\int_{e}^{\sqrt{x}} \frac{\log\log u}{u\left(1-\frac{\log u}{\log x}\right)}du$$
$$= \frac{1}{\log x}\int_{e}^{\sqrt{x}} \frac{\log\log u}{u}\sum_{k=0}^{\infty}\left(\frac{\log u}{\log x}\right)^{k}du$$
$$= \sum_{k=0}^{\infty}\frac{1}{(\log x)^{k+1}}\int_{e}^{\sqrt{x}} \frac{\log\log u}{u}(\log u)^{k}du.$$

The integrands and summands are nonnegative so we can use Tonelli's Theorem to justify the interchange of summation and integration. A standard calculus exercise shows that

$$\frac{d}{du}\left\{\frac{(\log u)^{k+1}}{k+1}\log\log u - \frac{(\log u)^{k+1}}{(k+1)^2}\right\} = \frac{\log\log u}{u}(\log u)^k,$$

so the Fundamental Theorem of Calculus gives

$$\begin{aligned}
&\int_e^{\sqrt{x}} \frac{\log\log u}{u}(\log u)^k du \\
&\quad = \frac{\big(\log(\sqrt{x})\big)^{k+1}\log\log(\sqrt{x})}{k+1} - \frac{\big(\log(\sqrt{x})\big)^{k+1}}{(k+1)^2} + \frac{1}{(k+1)^2} \\
&\quad = \frac{\big(\log\log x - \log 2\big)\,(\log x)^{k+1}}{2^{k+1}(k+1)} - \frac{(\log x)^{k+1}}{2^{k+1}(k+1)^2} + \frac{1}{(k+1)^2}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
\int_e^{\sqrt{x}} \frac{\log\log u\, du}{u\log\frac{x}{u}} &= \sum_{k=0}^{\infty}\frac{1}{(\log x)^{k+1}}\int_e^{\sqrt{x}}\frac{\log\log u}{u}(\log u)^k du \\
&= \big(\log\log x - \log 2\big)\sum_{k=0}^{\infty}\frac{1}{2^{k+1}(k+1)} - \sum_{k=0}^{\infty}\frac{1}{2^{k+1}(k+1)^2} \\
&\qquad + \sum_{k=0}^{\infty}\frac{1}{(k+1)^2(\log x)^{k+1}} \\
&= (\log\log x - \log 2)\sum_{n=1}^{\infty}\frac{1}{n\,2^n} - \sum_{n=1}^{\infty}\frac{1}{n^2 2^n} + O\left(\frac{1}{\log x}\right).
\end{aligned}$$

From Equations 1 and 2 of §0.241 of Gradshteyn and Ryzhik [4], we know that

$$\sum_{n=1}^{\infty}\frac{1}{n\,2^n} = \log 2 \quad \text{and} \quad \sum_{n=1}^{\infty}\frac{1}{n^2 2^n} = \frac{\pi^2}{12} - \frac{1}{2}(\log 2)^2.$$

Therefore,

$$\int_{e}^{\sqrt{x}} \frac{\log\log u\,du}{u\log\frac{x}{u}} = (\log 2)\log\log x - \frac{1}{2}(\log 2)^2 - \frac{\pi^2}{12} + O\Big(\frac{1}{\log x}\Big).$$

Combining our estimates, the lemma follows. □

# 3 THE VARIANCE OF $\omega(n)$

In this chapter we will discuss and prove increasingly more refined approximations of the second moment of $\omega(n)$, concluding with our own new result, which is the most precise. First it easily checked that

$$\sum_{n \le x} \omega(n) = \sum_{p \le x} \left\lfloor \frac{x}{p} \right\rfloor.$$

With this in mind we begin with a simple lemma that will prove useful throughout.

**Lemma 3.0.1.** *Let $p, q$ denote primes. For $x \ge 2$, we have*

$$\sum_{n \le x} \omega(n)^2 = \sum_{\substack{pq \le x \\ p \ne q}} \left\lfloor \frac{x}{pq} \right\rfloor + \sum_{p \le x} \left\lfloor \frac{x}{p} \right\rfloor.$$

*Proof.* By expressing $\omega(n)$ as a sum and then interchanging the order of summation, we have

$$\sum_{n \le x} \omega(n)^2 = \sum_{n \le x} \left( \sum_{p|n} 1 \right) \left( \sum_{q|n} 1 \right) = \sum_{p,q \le x} \sum_{\substack{n \le x \\ p|n \\ q|n}} 1.$$

There are two cases: either $p$ and $q$ are distinct primes or $p = q$. Therefore,

$$\sum_{n \le x} \omega(n)^2 = \sum_{\substack{pq \le x \\ p \ne q}} \sum_{\substack{n \le x \\ pq|n}} 1 + \sum_{p \le x} \sum_{\substack{n \le x \\ p|n}} 1 = \sum_{\substack{pq \le x \\ p \ne q}} \left\lfloor \frac{x}{pq} \right\rfloor + \sum_{p \le x} \left\lfloor \frac{x}{p} \right\rfloor,$$

as claimed. $\square$

## 3.1 Turán's proof

Though Turán's estimate for the second moment of $\omega(n)$ only required elementary knowledge, the importance of his result, which gave a simple proof of the Hardy and Ramanujan theorem, cannot be overstated.

**Theorem 3.1.1** (Turán's Theorem)**.** *For $x \geq 2$, we have*

$$\frac{1}{x}\sum_{n\leq x}\omega(n)^2 = (\log\log x)^2 + O(\log\log x).$$

*Proof.* Let $p, q$ denote primes and $x \geq 2$. Starting with Lemma 3.0.1, we have that

$$\sum_{n\leq x}\omega(n)^2 = \sum_{\substack{pq\leq x\\ p\neq q}}\left\lfloor\frac{x}{pq}\right\rfloor + \sum_{p\leq x}\left\lfloor\frac{x}{p}\right\rfloor.$$

The second sum on the right is simply the first moment of $\omega(n)$, which gives

$$\begin{aligned}\sum_{p\leq x}\left\lfloor\frac{x}{p}\right\rfloor &= \sum_{p\leq x}\left(\frac{x}{p} + O(1)\right)\\ &= x\sum_{p\leq x}\frac{1}{p} + O\left(\sum_{p\leq x}1\right).\end{aligned}$$

Here we apply Lemma 2.1.2 to get

$$\sum_{p\leq x}\left\lfloor\frac{x}{p}\right\rfloor = x\log\log x + O(x).$$

Now looking at the first sum on the right-hand side, we have the clever relation,

$$\left(\sum_{p\leq\sqrt{x}}\frac{1}{p}\right)^2 < \sum_{\substack{pq<x\\ p\neq q}}\frac{1}{pq} < \left(\sum_{p\leq x}\frac{1}{p}\right)^2.$$

Then by Lemma 2.1.2, we conclude that

$$\big(\log\log x + O(1)\big)^2 < \sum_{\substack{pq\le x\\ p\ne q}} \frac{1}{pq} < \big(\log\log x + O(1)\big)^2.$$

Hence,

$$\begin{aligned}\sum_{\substack{pq\le x\\ p\ne q}} \left\lfloor \frac{x}{pq} \right\rfloor &= x\sum_{\substack{pq\le x\\ p\ne q}} \frac{1}{pq} + O\left(\sum_{\substack{pq\le x\\ p\ne q}} 1\right)\\ &= x(\log\log x)^2 + O(x\log\log x),\end{aligned}$$

where

$$\sum_{\substack{pq\le x\\ p\ne q}} 1 \ll \sum_{p\le x} 1 \ll x.$$

Putting the sums together we have

$$\sum_{n\le x} \omega(n)^2 = x(\log\log x)^2 + O(x\log\log x),$$

the desired result. □

## 3.2 Montgomery and Vaughan's proof

In pursuit of an even tighter bound of the second moment of the function $\omega(n)$, we will now state a theorem inspired by Montgomery and Vaughan [7]. This approach was even suggested by Turán at the end of his paper [8].

**Theorem 3.2.1.** *For $x \ge 2$, we have*

$$\frac{1}{x}\sum_{n\le x} \omega(n)^2 = (\log\log x)^2 + (2b+1)\log\log x + O(1).$$

Before we prove this theorem, we will first state and prove a simple lemma.

**Lemma 3.2.2.** *For $x \geq 2$, we have*

$$\frac{1}{x}\sum_{p\leq x}\left\lfloor \frac{x}{p}\right\rfloor = \log\log x + b + O\left(\frac{1}{\log x}\right),$$

*where $b$ is the constant in Lemma 2.1.2.*

*Proof.* Since $0 \leq y - \lfloor y \rfloor < 1$, then by using Lemma 2.1.1 and Lemma 2.1.2, we see that

$$\begin{aligned}\sum_{p\leq x}\left\lfloor \frac{x}{p}\right\rfloor &= \sum_{p\leq x}\left(\frac{x}{p} + O(1)\right)\\ &= x\sum_{p\leq x}\frac{1}{p} + O\big(\pi(x)\big)\\ &= x\log\log x + b\,x + O\left(\frac{x}{\log x}\right)\end{aligned}$$

as claimed. □

*Proof of Theorem 3.2.1.* Once again, we start from Lemma 3.0.1:

$$\sum_{n\leq x}\omega(n)^2 = \sum_{\substack{pq\leq x\\ p\neq q}}\left\lfloor \frac{x}{pq}\right\rfloor + \sum_{p\leq x}\left\lfloor \frac{x}{p}\right\rfloor.$$

By adding and subtracting the terms where $p = q$ to the first sum on the right-hand side, we see that

$$\sum_{\substack{pq\leq x\\ p\neq q}}\left\lfloor \frac{x}{pq}\right\rfloor = \sum_{pq\leq x}\left\lfloor \frac{x}{pq}\right\rfloor - \sum_{p^2\leq x}\left\lfloor \frac{x}{p^2}\right\rfloor = \sum_{pq\leq x}\left\lfloor \frac{x}{pq}\right\rfloor - \sum_{p\leq \sqrt{x}}\left\lfloor \frac{x}{p^2}\right\rfloor.$$

Therefore,

$$\sum_{n\leq x}\omega(n)^2 = \sum_{p\leq x}\left\lfloor \frac{x}{p}\right\rfloor + \sum_{pq\leq x}\left\lfloor \frac{x}{pq}\right\rfloor - \sum_{p\leq \sqrt{x}}\left\lfloor \frac{x}{p^2}\right\rfloor. \tag{3.1}$$

From this and Lemma 3.2.2, we see that

$$\sum_{n\le x}\omega(n)^2 = x\log\log x + b\,x + \sum_{pq\le x}\left\lfloor\frac{x}{pq}\right\rfloor - \sum_{p\le\sqrt{x}}\left\lfloor\frac{x}{p^2}\right\rfloor + O\left(\frac{x}{\log x}\right).$$

To estimate the second sum on right hand-side, we find that

$$\begin{aligned}\sum_{p\le\sqrt{x}}\left\lfloor\frac{x}{p^2}\right\rfloor &= \sum_{p\le\sqrt{x}}\left(\frac{x}{p^2}+O(1)\right)\\ &= x\sum_{p\le\sqrt{x}}\frac{1}{p^2}+O\left(\sqrt{x}\right),\end{aligned}$$

since $\pi(\sqrt{x}) \le \sqrt{x}$. Moreover,

$$x\sum_{p\le\sqrt{x}}\frac{1}{p^2} \le x\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}\,x = O(x).$$

Therefore,

$$\sum_{n\le x}\omega(n)^2 = x\log\log x + \sum_{pq\le x}\left\lfloor\frac{x}{pq}\right\rfloor + O(x).$$

This leaves us with only one more sum to analyze. We have

$$\begin{aligned}\sum_{pq\le x}\left\lfloor\frac{x}{pq}\right\rfloor &= \sum_{pq\le x}\left(\frac{x}{pq}+O(1)\right)\\ &= x\sum_{pq\le x}\frac{1}{pq}+O(x),\end{aligned}$$

because the set of products of two primes that are less than or equal to $x$ is a subset of all positive integers that are less than or equal to $x$. Noticing that if $pq \le x$ then it follows that at least one of $p$ or $q$ must be less than or equal to $\sqrt{x}$. Hence, by the so-called hyperbola method (summing above the hyperbola), we see further that

$$
\begin{aligned}
\sum_{pq\le x}\frac{1}{pq} &= \left(\sum_{p\le x}\frac{1}{p}\right)^2 - \sum_{p\le x}\sum_{\substack{q\le x\\ pq>x}}\frac{1}{pq} \\
&= \left(\sum_{p\le x}\frac{1}{p}\right)^2 - \sum_{p\le\sqrt{x}}\frac{1}{p}\sum_{\frac{x}{p}<q\le x}\frac{1}{q} + \left(\sum_{\sqrt{x}<p<x}\frac{1}{p}\right)^2 - \sum_{q\le\sqrt{x}}\frac{1}{q}\sum_{\frac{x}{q}<p\le x}\frac{1}{p} \\
&= \left(\sum_{p\le x}\frac{1}{p}\right)^2 - 2\sum_{p\le\sqrt{x}}\frac{1}{p}\sum_{\frac{x}{p}<q\le x}\frac{1}{q} + \left(\sum_{\sqrt{x}<p<x}\frac{1}{p}\right)^2. \qquad (3.2)
\end{aligned}
$$

By Lemma 2.1.2, for $2 \le y \le \sqrt{x}$, we have

$$
\begin{aligned}
\sum_{\frac{x}{y}<p\le x}\frac{1}{p} &= \log\log x - \log\log\tfrac{x}{y} + O\left(\frac{1}{\log x}\right) \\
&= \log\left(\frac{\log x}{\log x - \log y}\right) + O\left(\frac{1}{\log x}\right) \\
&= \log\left(\frac{1}{1-\frac{\log y}{\log x}}\right) + O\left(\frac{1}{\log x}\right) \\
&\ll \frac{\log y}{\log x}.
\end{aligned}
$$

Here we used Lemma 2.2.1 to deduce the bound in the final step. Going back to (3.2) and using this estimate, the second sum is

$$
\sum_{p\le\sqrt{x}}\frac{1}{p}\sum_{\frac{x}{p}<q\le x}\frac{1}{q} \ll \sum_{p\le\sqrt{x}}\frac{1}{p}\left(\frac{\log p}{\log x}\right) = \frac{1}{\log x}\sum_{p\le\sqrt{x}}\frac{\log p}{p} \ll \frac{1}{\log x}(\log x) \ll 1,
$$

where we have used Lemma 2.1.2 to estimate the sum over primes. The third sum in (3.2) is

$$\left(\sum_{\sqrt{x}<p<x}\frac{1}{p}\right)^2 = \left(\log\log x - \log\log\sqrt{x} + O\left(\frac{1}{\log x}\right)\right)^2$$
$$= \left(\log 2 + O\left(\frac{1}{\log x}\right)\right)^2$$
$$\ll 1.$$

Therefore,

$$\sum_{pq\le x}\frac{1}{pq} = \left(\sum_{p\le x}\frac{1}{p}\right)^2 + O(1)$$
$$= \left(\log\log x + b + O\left(\frac{1}{\log x}\right)\right)^2 + O(1)$$
$$= (\log\log x)^2 + 2b\log\log x + O(1).$$

Combining our estimates, we see that

$$\sum_{n\le x}\omega(n)^2 = x\log\log x + x\,(\log\log x)^2 + 2bx\log\log x + O(x)$$
$$= x\,(\log\log x)^2 + (2b+1)x\log\log x + O(x).$$

This completes the proof. □

### 3.3 A refined estimate for the second moment of $\omega(n)$

Lastly, we look at our own original calculation of the second moment of $\omega(n)$ and the tightest bound of the error yet. Though all that is needed to arrive at our error is trivial, the result requires the careful manipulation of several moving pieces. We have written and proved each of these manipulations as their own lemmas to make all steps as clear to the reader as possible. We will conclude by combining all of our results together to achieve our goal of finding as minimal of an error as possible.

**Theorem 3.3.1.** *For $x \geq 4$, we have*

$$\sum_{n \leq x} \omega(n)^2 = x\,(\log\log x)^2 + (2b+1)\,x \log\log x + Cx + O\left(\frac{x \log\log x}{\log x}\right),$$

*where $b$ is the constant in Lemma 2.1.2, and*

$$C = b^2 + b - \frac{\pi^2}{6} - \sum_{p} \frac{1}{p^2}.$$

As we did with Theorem 3.2.1, we will first state and prove simple lemmas. Our first result is an adaptation of Lemma 3.0.1, incorporating both (3.1) and the hyperbola method in a different manner than used in proving the previous theorem. The main idea of this new method is to estimate the sum

$$\sum_{pq \leq x} \left\lfloor \frac{x}{pq} \right\rfloor$$

in a more precise way. Similar to what we did in the proof of Lemma 3.2.1, we note that since $pq \leq x$, at least one of $p$ or $q$ has to be less than or equal to $\sqrt{x}$. So, without loss of generality we can assume that one of the primes $p$ or $q$ is less than or equal to $\sqrt{x}$. This simple observation allows us to refine our previous calculation in numerous places.

**Lemma 3.3.2.** *Let $p, q$ denote primes. For $x \geq 4$, we have*

$$\sum_{n \leq x} \omega(n)^2 = 2 \sum_{p \leq \sqrt{x}} \sum_{q \leq \frac{x}{p}} \left\lfloor \frac{x}{pq} \right\rfloor - \sum_{p \leq \sqrt{x}} \sum_{q \leq \sqrt{x}} \left\lfloor \frac{x}{pq} \right\rfloor + \sum_{p \leq x} \left\lfloor \frac{x}{p} \right\rfloor - \sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p^2} \right\rfloor. \tag{3.3}$$

*Proof.* From (3.1), we have

$$\sum_{n \leq x} \omega(n)^2 = \sum_{pq \leq x} \left\lfloor \frac{x}{pq} \right\rfloor + \sum_{p \leq x} \left\lfloor \frac{x}{p} \right\rfloor - \sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p^2} \right\rfloor,$$

for $x \geq 4$. Applying the hyperbola method (or the inclusion-exclusion principle) to the first sum on the right-hand side, we see that

$$\sum_{pq \leq x} \left\lfloor \frac{x}{pq} \right\rfloor = \sum_{p \leq \sqrt{x}} \sum_{q \leq \frac{x}{p}} \left\lfloor \frac{x}{pq} \right\rfloor + \sum_{q \leq \sqrt{x}} \sum_{p \leq \frac{x}{q}} \left\lfloor \frac{x}{pq} \right\rfloor - \sum_{p \leq \sqrt{x}} \sum_{q \leq \sqrt{x}} \left\lfloor \frac{x}{pq} \right\rfloor$$

$$= 2 \sum_{p \leq \sqrt{x}} \sum_{q \leq \frac{x}{p}} \left\lfloor \frac{x}{pq} \right\rfloor - \sum_{p \leq \sqrt{x}} \sum_{q \leq \sqrt{x}} \left\lfloor \frac{x}{pq} \right\rfloor.$$

After combining these estimates, the lemma follows. □

We now estimate each of the terms on the right-hand side of (3.3). In Lemma 3.2.2, for $x \geq 2$, we estimated the third sum on the right-hand side of (3.3) and proved that

$$\sum_{p \leq x} \left\lfloor \frac{x}{p} \right\rfloor = x \log \log x + b\, x + O\left( \frac{x}{\log x} \right).$$

In the next lemma, we estimate the fourth sum on the right-hand side of (3.3).

**Lemma 3.3.3.** *For $x \geq 4$, we have*

$$\frac{1}{x} \sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p^2} \right\rfloor = \sum_{p} \frac{1}{p^2} + O\left( \frac{1}{\sqrt{x}} \right).$$

*Proof.* We see that

$$\sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p^2} \right\rfloor = \sum_{p \leq \sqrt{x}} \left( \frac{x}{p^2} + O(1) \right)$$

$$= x \sum_{p \leq \sqrt{x}} \frac{1}{p^2} + O\big(\pi(\sqrt{x})\big)$$

$$= x \sum_{p} \frac{1}{p^2} - x \sum_{p > \sqrt{x}} \frac{1}{p^2} + O\left( \frac{\sqrt{x}}{\log x} \right),$$

where we have added and subtracted the terms with $p \geq \sqrt{x}$ and applied Lemma 2.1.1 to the error term. Note that

$$x \sum_{p > \sqrt{x}} \frac{1}{p^2} \leq x \sum_{n > \sqrt{x}} \frac{1}{n^2} \ll x \int_{\sqrt{x}}^{\infty} \frac{du}{u^2} = \sqrt{x},$$

so we have

$$\sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p^2} \right\rfloor = x \sum_{p} \frac{1}{p^2} + O(\sqrt{x}).$$

Thus, proving the lemma. □

Next, we estimate the second sum on the right-hand side of (3.3).

**Lemma 3.3.4.** *For $x \geq 4$, we have*

$$\frac{1}{x} \sum_{p \leq \sqrt{x}} \sum_{q \leq \sqrt{x}} \left\lfloor \frac{x}{pq} \right\rfloor = (\log \log x)^2 + 2\,(b - \log 2) \log \log x$$
$$+ (b - \log 2)^2 + O\left(\frac{\log \log x}{\log x}\right).$$

*Proof.* Notice that by Lemma 2.1.2, we have

$$\begin{aligned} \sum_{p \leq \sqrt{x}} \sum_{q \leq \sqrt{x}} \left\lfloor \frac{x}{pq} \right\rfloor &= \sum_{p \leq \sqrt{x}} \sum_{q \leq \sqrt{x}} \left( \frac{x}{pq} + O(1) \right) \\ &= x \sum_{p \leq \sqrt{x}} \frac{1}{p} \sum_{q \leq \sqrt{x}} \frac{1}{q} + O\left( \sum_{p \leq \sqrt{x}} \sum_{q \leq \sqrt{x}} 1 \right) \\ &= x \left( \sum_{p \leq \sqrt{x}} \frac{1}{p} \right)^2 + O\left( \left( \sum_{p \leq \sqrt{x}} 1 \right)^2 \right). \end{aligned}$$

Here we are able to apply both Lemma 2.1.2 and 2.1.1 respectively to the above terms to get

$$\begin{aligned}
&x\left(\log\log\sqrt{x}+b+O\left(\frac{1}{\log\sqrt{x}}\right)\right)^2+O\left(\left(\frac{\sqrt{x}}{\log\sqrt{x}}\right)^2\right)\\
&=x\left(\log\log x-\log 2+b+O\left(\frac{1}{\log x}\right)\right)^2+O\left(\frac{x}{\log^2 x}\right)\\
&=x\left(\log\log x\right)^2+2\left(b-\log 2\right)x\log\log x\\
&\qquad+\left(b-\log 2\right)^2 x+O\left(\frac{x\log\log x}{\log x}\right),
\end{aligned}$$

as claimed. □

Before we can estimate the first term on the right-hand side of (3.3), we first prove a preliminary estimate. This result is where the proof of our theorem differs the most from previous investigations as we will now estimate the first term on the right-hand side of (3.3) using Lemma 2.2.2.

**Lemma 3.3.5.** *For $x \geq 4$, we have*

$$\begin{aligned}
\frac{1}{x}\sum_{p\leq\sqrt{x}}\sum_{q\leq\frac{x}{q}}\left\lfloor\frac{x}{pq}\right\rfloor &= \left(\log\log x\right)^2+\left(2b-\log 2\right)\log\log x\\
&\qquad+b\left(b-\log 2\right)+\frac{1}{2}\left(\log 2\right)^2-\frac{\pi^2}{12}+O\left(\frac{\log\log x}{\log x}\right).
\end{aligned}$$

*Proof.* We have

$$\begin{aligned}
\sum_{p\leq\sqrt{x}}\sum_{q\leq\frac{x}{q}}\left\lfloor\frac{x}{pq}\right\rfloor &= \sum_{p\leq\sqrt{x}}\sum_{q\leq\frac{x}{q}}\left(\frac{x}{pq}+O(1)\right)\\
&= x\sum_{p\leq\sqrt{x}}\frac{1}{p}\sum_{q\leq\frac{x}{p}}\frac{1}{q}+O\left(\sum_{p\leq\sqrt{x}}\sum_{q\leq\frac{x}{p}}1\right).
\end{aligned}$$

Using Lemmas 2.1.1 and 2.1.2, we estimate the error term by

$$\sum_{p\le\sqrt{x}}\ \sum_{q\le\frac{x}{p}} 1 \ll \sum_{p\le\sqrt{x}} \pi\left(\tfrac{x}{p}\right) \ll \sum_{p\le\sqrt{x}} \frac{x}{p\log\frac{x}{p}} \ll \frac{x}{\log x}\sum_{p\le\sqrt{x}}\frac{1}{p} \ll \frac{x\log\log x}{\log x}.$$

Using Lemma 2.1.2 again, we find that

$$\begin{aligned}\sum_{p\le\sqrt{x}}\frac{1}{p}\ \sum_{q\le\frac{x}{p}}\frac{1}{q} &= \sum_{p\le\sqrt{x}}\frac{1}{p}\left(\log\log\tfrac{x}{p} + b + O\left(\frac{1}{\log\frac{x}{p}}\right)\right)\\ &= \sum_{p\le\sqrt{x}}\frac{1}{p}\log\log\tfrac{x}{p} + b\sum_{p\le\sqrt{x}}\frac{1}{p} + O\left(x\sum_{p\le\sqrt{x}}\frac{1}{p\log\frac{x}{p}}\right).\end{aligned}$$

The error term above is handled easily as:

$$\sum_{p\le\sqrt{x}}\frac{1}{p\log\frac{x}{p}} \ll \frac{1}{\log x}\sum_{p\le\sqrt{x}}\frac{1}{p} \ll \frac{\log\log x}{\log x}.$$

Therefore,

$$\sum_{p\le\sqrt{x}}\ \sum_{q\le\frac{x}{q}}\left\lfloor\frac{x}{pq}\right\rfloor = x\sum_{p\le\sqrt{x}}\frac{1}{p}\log\log\tfrac{x}{p} + x\,b\sum_{p\le\sqrt{x}}\frac{1}{p} + O\left(\frac{x\log\log x}{\log x}\right). \tag{3.4}$$

To estimate the second term on the right-hand side, we see that

$$\begin{aligned}x\,b\sum_{p\le\sqrt{x}}\frac{1}{p} &= x\,b\left(\log\log\sqrt{x} + b + O\left(\frac{1}{\log\sqrt{x}}\right)\right)\\ &= x\,b\left(\log\log x - \log 2 + b + O\left(\frac{1}{\log x}\right)\right) \qquad (3.5)\\ &= b\,x\log\log x + b\,(b-\log 2)\,x + O\left(\frac{1}{\log x}\right).\end{aligned}$$

To estimate the first term on the right-hand side, we use the method of partial summation. Since

$$\frac{d}{du}\Big\{\log\log\big(\tfrac{x}{u}\big)\Big\} = \frac{-1}{u\log\frac{x}{u}}, \tag{3.6}$$

we see that

$$\begin{aligned}\sum_{p\le\sqrt{x}}\frac{1}{p}\log\log\tfrac{x}{p} &= \int_{2^-}^{\sqrt{x}}\log\log\big(\tfrac{x}{u}\big)\,d\Bigg(\sum_{p\le u}\frac{1}{p}\Bigg)\\ &= \Bigg(\sum_{p\le u}\frac{1}{p}\Bigg)\log\log\tfrac{x}{u}\Bigg|_{u=2^-}^{u=\sqrt{x}} + \int_{2}^{\sqrt{x}}\Bigg(\sum_{p\le u}\frac{1}{p}\Bigg)\frac{du}{u\log\frac{x}{u}}.\end{aligned}$$

Now, the first term above is

$$\begin{aligned}&\Bigg(\sum_{p\le u}\frac{1}{p}\Bigg)\log\log\tfrac{x}{u}\Bigg|_{u=2^-}^{u=\sqrt{x}}\\ &\quad= \Bigg(\sum_{p\le\sqrt{x}}\frac{1}{p}\Bigg)\log\log\sqrt{x}\\ &\quad= \Bigg(\log\log\sqrt{x}+b+O\bigg(\frac{1}{\log\sqrt{x}}\bigg)\Bigg)(\log\log x-\log 2) \qquad (3.7)\\ &\quad= \Bigg(\log\log x-\log 2+b+O\bigg(\frac{1}{\log x}\bigg)\Bigg)(\log\log x-\log 2)\\ &\quad= (\log\log x)^2+(b-2\log 2)\log\log x\\ &\qquad\qquad -\log 2(b-\log 2)+O\bigg(\frac{\log\log x}{\log x}\bigg),\end{aligned}$$

while the integral is

$$\begin{aligned}\int_2^{\sqrt{x}} \left(\sum_{p\le u}\frac{1}{p}\right)\frac{du}{u\log\frac{x}{u}} &= \int_2^{\sqrt{x}} \left(\log\log u + b + O\left(\frac{1}{\log u}\right)\right)\frac{du}{u\log\frac{x}{u}} \\ &= \int_2^{\sqrt{x}} \frac{\log\log u}{u\log\frac{x}{u}}du + b\int_2^{\sqrt{x}}\frac{du}{u\log\frac{x}{u}} \\ &\quad + O\left(\int_2^{\sqrt{x}}\frac{du}{u\log u\log\frac{x}{u}}\right).\end{aligned}$$

The error term here is

$$\ll \frac{1}{\log x}\int_2^{\sqrt{x}}\frac{du}{u\log u} \ll \frac{\log\log x}{\log x}. \tag{3.8}$$

In Lemma 2.2.2, we showed that

$$\int_2^{\sqrt{x}} \frac{\log\log u}{u\log\frac{x}{u}}du = (\log 2)\log\log x - \frac{1}{2}(\log 2)^2 - \frac{\pi^2}{12} + O\left(\frac{1}{\log x}\right), \tag{3.9}$$

and by (3.6) we have

$$\begin{aligned} b\int_2^{\sqrt{x}}\frac{du}{u\log\frac{x}{u}} &= b\int_1^{\sqrt{x}}\frac{du}{u\log\frac{x}{u}} + O\left(\frac{1}{\log x}\right) \\ &= -b\,\log\log\frac{x}{u}\Big|_{u=1}^{u=\sqrt{x}} + O\left(\frac{1}{\log x}\right) \\ &= b\left(\log\log x - \log\log\sqrt{x}\right) + O\left(\frac{1}{\log x}\right) \\ &= b\log 2 + O\left(\frac{1}{\log x}\right). \end{aligned} \tag{3.10}$$

Therefore, combining (3.7), (3.8), (3.9), and (3.10), we deduce that

$$\sum_{p\le\sqrt{x}} \frac{1}{p}\log\log\tfrac{x}{p} = (\log\log x)^2 + (b-\log 2)\log\log x$$
$$+\frac{1}{2}(\log 2)^2 - \frac{\pi^2}{12} + O\Big(\frac{\log\log x}{\log x}\Big).$$

In light of the expression in (3.4), adding this result to the estimate in (3.5) and then simplifying, the lemma now follows. □

We are now in position to prove Theorem 3.3.1.

*Proof of Theorem 3.3.1.* By Lemmas 3.2.2, 3.3.2, 3.3.3, 3.3.4, and 3.3.5, we have

$$\begin{aligned}\sum_{n\le x}\omega(n)^2 &= 2\sum_{p\le\sqrt{x}}\sum_{q\le\frac{x}{q}}\Big\lfloor\frac{x}{pq}\Big\rfloor - \sum_{p\le\sqrt{x}}\sum_{q\le\sqrt{x}}\Big\lfloor\frac{x}{pq}\Big\rfloor + \sum_{p\le x}\Big\lfloor\frac{x}{p}\Big\rfloor - \sum_{p\le\sqrt{x}}\Big\lfloor\frac{x}{p^2}\Big\rfloor\\
&= 2\Big\{x\,(\log\log x)^2 + (2b-\log 2)\,x\,\log\log x\\
&\qquad + x\Big(b\,(b-\log 2) + \frac{1}{2}(\log 2)^2 - \frac{\pi^2}{12}\Big) + O\Big(\frac{x\log\log x}{\log x}\Big)\Big\}\\
&\quad - \Big\{x\,(\log\log x)^2 + 2\,(b-\log 2)\,x\log\log x\\
&\qquad + (b-\log 2)^2\,x + O\Big(\frac{x\log\log x}{\log x}\Big)\Big\}\\
&\quad + \Big\{x\log\log x + b\,x + O\Big(\frac{x}{\log x}\Big)\Big\}\\
&\quad - \Big\{x\sum_p\frac{1}{p^2} + O(\sqrt{x}\,)\Big\}\\
&= x\,(\log\log x)^2 + (2b+1)\,x\log\log x + Cx + O\Big(\frac{x\log\log x}{\log x}\Big),\end{aligned}$$

where the constant

$$C = b^2 + b - \frac{\pi^2}{6} - \sum_p\frac{1}{p^2}.$$

This completes the proof of the theorem. □

## 3.4 A refined estimate for the variance of $\omega(n)$

With this in mind, we have now calculated estimates of everything needed to find the variance of $\omega(n)$. Recalling the formula for the variance of a random variable is simply

$$\mathrm{Var}\,(X) = \mathrm{E}\,[(X-\mu)^2],$$

we now have all of the needed estimates.

**Theorem 3.4.1.** *For $x \geq 4$,*

$$\mathrm{Var}\,\big(\omega(n)\big) = \log\log x + b - \sum_{n} \frac{1}{n^2} - \sum_{p} \frac{1}{p^2} + O\left(\frac{\log\log x}{\log x}\right).$$

*Proof.* First let us find a stronger estimate for the average of $\omega(n)$ then we have previously stated.

$$\begin{aligned}\sum_{n\leq x} \omega(n) &= \sum_{p\leq x} \left\lfloor \frac{x}{p} \right\rfloor \\ &= \sum_{p\leq x} \left(\frac{x}{p} + O(1)\right) \\ &= x\sum_{p\leq x} \frac{1}{p} + O\left(\sum_{p\leq x} 1\right).\end{aligned}$$

Now using Lemmas 2.1.2 and 2.1.1 respectively on the terms above we have

$$\sum_{n\leq x} \omega(n) = x\log\log x + xb + O\left(\frac{x}{\log x}\right).$$

The importance of this step is to ensure that we have an error term for our average of $\omega(n)$ that will not interfere with the error we found for the second moment. Thus,

$$\begin{aligned}
\mathrm{Var}\,(\omega(n)) &= \frac{1}{x}\sum_{n\le x}\left(\omega(n) - \frac{1}{x}\sum_{n\le x}\omega(n)\right)^2 \\
&= \frac{1}{x}\sum_{n\le x}\left(\omega(n) - \log\log x - b + O\left(\frac{1}{\log x}\right)\right)^2 \\
&= \frac{1}{x}\sum_{n\le x}\omega(n)^2 - \frac{1}{x}\left(2\log\log x + 2b + O\left(\frac{1}{\log x}\right)\right)\sum_{n\le x}\omega(n) \\
&\qquad + \left(\log\log x + b + O\left(\frac{1}{\log x}\right)\right)^2 \\
&= \frac{1}{x}\sum_{n\le x}\omega(n)^2 - \left(\log\log x + b + O\left(\frac{1}{\log x}\right)\right)^2 \\
&= \frac{1}{x}\sum_{n\le x}\omega(n)^2 - (\log\log x)^2 - 2b\log\log x - b^2 + O\left(\frac{\log\log x}{\log x}\right).
\end{aligned}$$

Replacing $\frac{1}{x}\sum_{n\le x}\omega(n)^2$ with our result from Theorem 3.3.1 and combining terms we see that

$$\mathrm{Var}\,(\omega(n)) = \log\log x + b - \frac{\pi^2}{6} - \sum_p \frac{1}{p^2} + O\left(\frac{\log\log x}{\log x}\right)$$

where $\frac{\pi^2}{6} = \sum_n \frac{1}{n^2}$, proving the theorem. □

This final approximation of the variance of $\omega(n)$ aligns with what we would have expected from the variance based on Turán's Theorem. Ours is just a more precise statement of his result.

# 4 GRANVILLE AND SOUNDARARAJAN'S PROOF OF THE ERDŐS-KAC THEOREM

In this section we give a proof of the classical Erdős-Kac Theorem based on a modern treatment by Granville and Soundararajan [5]. The proof of Granville and Soundararajan proceeds by computing the moments of

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

for $n \leq x$. They showed that these moments match the moments of a random variable with standard normal distribution. From this, the Erdős-Kac theorem follows since it is known that the normal distribution is completely characterized by its moments [1, Thm. 30.1]. The moments approach to proving the Erdős-Kac theorem was actually first accomplished by Delange [2], but the proof of Granville and Soundararajan is much simpler. It is also more powerful as it yields the $k$-th moment uniformly for $k \leq (\log\log x)^{\frac{1}{3}}$. However in this thesis, we are content to establish the $k$-th moment for any *fixed* natural number $k$, which is strong enough to conlcude the Erdős-Kac Theorem still. Although the paper by Granville and Soundararajan is beautiful, it is rather terse. Our exposition provides more details to the arguments given in that paper. We hope that this will be useful for those trying to learn the field. Below is the main theorem of [5].

**Theorem 4.0.1** (Granville and Soundararajan)**.** *Let* $x \geq 3$. *For any fixed natural number* $k$, *let* $C_k = \dfrac{\Gamma(k+1)}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}+1\right)}$. *For even* $k$ *we have*

$$\text{(a)} \sum_{n \leq x} (\omega(n) - \log\log x)^k = C_k x (\log\log x)^{\frac{k}{2}} \left(1 + O\left(\frac{1}{\sqrt{\log\log x}}\right)\right),$$

*and for odd* $k$ *we have*

$$\text{(b)} \sum_{n \leq x} (\omega(n) - \log\log x)^k \ll x (\log\log x)^{\frac{k}{2}} \frac{1}{\sqrt{\log\log x}}.$$

Recall that for $k$ even, the $k$-th moment of a standard normal variable is precisely $C_k$. For $k$ odd, the $k$-th moment of a standard normal variable is 0. Asymptotically speaking, our results agree with this since (b) implies that the $k$-th moment of $\frac{\omega(n)-\log\log x}{\sqrt{\log\log x}}$ decays similar to $(\log\log x)^{-\frac{1}{2}}$, as $x \to \infty$, when $k$ is odd.

Define

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

where the reasoning for this definition will be explained in Section 4.1. Now we extend this definition totally multiplicatively in the subscript. If $r \geq 1$ has prime factorization $r = \prod_i (q_i)^{\alpha_i}$ for distinct primes $q_i$ and $\alpha_i \geq 1$, then define

$$\begin{aligned} f_r(n) &= f_{q_1^{\alpha_1} q_2^{\alpha_2} q_3^{\alpha_3} \ldots}(n) \\ &= \prod_i \left(f_{q_i}(n)\right)^{\alpha_i}. \end{aligned} \tag{4.1}$$

Theorem 4.0.1 will arise from the following result. It will become apparent in the next section why the moments given in Proposition 4.0.2 can serve as a substitute for the moments of Theorem 4.0.1.

**Proposition 4.0.2.** *Let $x, z \geq 3$. For fixed even natural numbers $k$,*

$$\text{(a)} \sum_{n \leq x} \left( \sum_{p \leq z} f_p(n) \right)^k = C_k x (\log\log z)^{\frac{k}{2}} \left( 1 + O\left( \frac{1}{\log\log z} \right) \right) + O(\pi(z)^k)$$

*while, for fixed odd numbers $k$, we have*

$$\text{(b)} \sum_{n \leq x} \left( \sum_{p \leq z} f_p(n) \right)^k \ll x (\log\log z)^{\frac{k}{2}} \frac{1}{\sqrt{\log\log z}} + \pi(z)^k.$$

*Here $\pi(z)$ denotes the number of primes less than or equal to $z$.*

### 4.1 Main idea of Granville and Soundararajan's result

A natural question to ask is, why do we want to work with the moments of Proposition 4.0.2 in the first place? Let us address this first before moving on. If we think of a prime $p$ dividing $n$ with probability $\frac{1}{p}$ independently of other primes, then we have $\mathbb{E}(f_p) = 0$. Thus $\sum_{p \leq z} f_p(n)$ is a sum of independent random variables of mean 0. By the central limit theorem we would then expect this sum to tend towards a normal distribution, which is what we want. Of course, this is only a model, for $f_p(n)$ is not really a random variable. But the independence idea described above can actually be realized on average over $n$, up to an "error term". If one has faith in this model and goes ahead with all the number theoretic calculations, then one would hope that everything will work out "in the wash" and the normal distribution will arise asymptotically. This is exactly what happens.

### 4.2 Deducing Theorem 4.0.1 from Proposition 4.0.2

In this section, we assume Proposition 4.0.2, and deduce Theorem 4.0.1 from it. We now begin building up to this. The following lemma motivates why the moments given in Proposition 4.0.2 can serve as a substitute for the moments of Theorem 4.0.1.

**Lemma 4.2.1.** *For $z = x^{\frac{1}{k}}$ and $n \leq x$, we have*

$$\omega(n) - \log\log x = \sum_{p \le z} f_p(n) + O(1).$$

*Proof.* We have

$$\begin{aligned}\omega(n) - \log\log x &= \sum_{p|n} 1 - \log\log x \\ &= \sum_{\substack{p|n \\ p>z}} 1 + \sum_{\substack{p|n \\ p\le z}} 1 - \log\log x\end{aligned}$$

by separating the primes, $p$, $p > z$ where and $p \le z$. Furthermore, we can rewrite this as

$$\begin{aligned}\sum_{\substack{p|n \\ p>z}} 1 + \sum_{\substack{p|n \\ p\le z}} 1 - \log\log x &= \sum_{\substack{p|n \\ p>z}} 1 + \sum_{\substack{p|n \\ p\le z}} 1 + \left(\sum_{\substack{p|n \\ p\le z}} \frac{1}{p} - \sum_{\substack{p|n \\ p\le z}} \frac{1}{p}\right) - \log\log x \\ &= \sum_{\substack{p|n \\ p>z}} 1 + \sum_{\substack{p|n \\ p\le z}} \left(1 - \frac{1}{p}\right) + \sum_{\substack{p|n \\ p\le z}} \frac{1}{p} - \log\log x.\end{aligned}$$

Now the sum $\sum_{\substack{p|n \\ p\le z}} \frac{1}{p}$ can be extended to all $p \le z$ as long as we subtract away the sum $\sum_{\substack{p\nmid n \\ p\le z}} \frac{1}{p}$. Thus, the above expression is as follows

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

Looking at our second and third summation above, we see this is exactly $\sum_{p\le z} f_p(n)$. This gives us

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

Hence, by Lemma 2.1.2 we have on the right-hand side

$$\sum_{p \le z} f_p(n) + \sum_{\substack{p \mid n \\ p > z}} 1 + O(1).$$

The second sum is bounded (in terms of $k$, which is fixed) since an integer $n \le x$ cannot have more than $k$ prime divisors larger than $z = x^{\frac{1}{k}}$. Thus

$$\sum_{\substack{p \mid n \\ p > z}} 1 \ll 1.$$

So we have

$$\omega(n) - \log\log x = \sum_{p \le z} f_p(n) + O(1),$$

proving the lemma. □

We now state and prove an extension of the prior lemma.

**Lemma 4.2.2.** *We have*

$$(\omega(n) - \log\log x)^k = \left( \sum_{p \le z} f_p(n) \right)^k + O\left( \left| \max_{0 \le \ell \le k-1} \sum_{p \le z} f_p(n) \right|^\ell \right).$$

*Proof.* From Lemma 4.2.1 we have

$$(\omega(n) - \log\log x)^k = \left( \sum_{p \le z} f_p(n) + O(1) \right)^k.$$

Here we can apply the Binomial Expansion Theorem to get

$$\begin{aligned} &\sum_{\ell=0}^{k} \left[ \binom{k}{\ell} \left( \sum_{p \le z} f_p(n) \right)^\ell O(1)^{k-\ell} \right] \\ &= \binom{k}{k} \left( \sum_{p \le z} f_p(n) \right)^k + \sum_{\ell=0}^{k-1} \left[ \binom{k}{\ell} \left( \sum_{p \le z} f_p(n) \right)^\ell O(1)^{k-\ell} \right], \end{aligned}$$

where the last line simply comes from writing out the $\ell = k$ term of the summation. Thus, we have

$$(\omega(n) - \log\log x)^k = \left(\sum_{p\le z} f_p(n)\right)^k + O\left(\left|\max_{0\le\ell\le k-1} \sum_{p\le z} f_p(n)\right|^\ell\right)$$

as stated. □

The previous lemma suggests where the main term and error term of our moments of $\sum_{n\le x}(\omega(n) - \log\log x)$ will arise from. So let us first treat the error term we found in Lemma 4.2.2.

**Lemma 4.2.3.** *We have*

$$O\left(\left|\max_{0\le\ell\le k-1} \sum_{p\le z} f_p(n)\right|^\ell\right) = O\left(x(\log\log z)^{\frac{k-1}{2}}\right).$$

*Proof.* Let $\ell \le k-1$.

**Case 1** (Assume $\ell$ is even):

Clearly

$$\left|\sum_{p\le z} f_p(n)\right|^\ell = \left(\sum_{p\le z} f_p(n)\right)^\ell.$$

Therefore, it follows directly that

$$\begin{aligned} O\left(\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^\ell\right) &= O\left(\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^\ell\right) \\ &= O\left(x(\log\log z)^{\frac{\ell}{2}} + \pi(z)^\ell\right) \end{aligned}$$

by applying part (a) of Proposition 4.0.2. We now make a quick note that

$$\pi(z)^\ell = \left(\sum_{p\le z} 1\right)^\ell \le \left(\sum_{n\le z} 1\right)^k = \left(\sum_{n\le x^{\frac{1}{k}}} 1\right)^k \le \left(x^{\frac{1}{k}}\right)^k = x.$$

Hence,

$$O\left(\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^{\ell}\right) = O\Big(x(\log\log z)^{\frac{\ell}{2}}\Big) = O\Big(x(\log\log z)^{\frac{k-1}{2}}\Big)$$

since $\ell \le k-1$.

**Case 2** (Assume $\ell$ is odd):

Now when $\ell$ is odd,

$$\begin{aligned}\left|\sum_{p\le z} f_p(n)\right|^{\ell} &= \left(\left|\sum_{p\le z} f_p(n)\right|^{\ell-1}\right)^{\frac{1}{2}}\left(\left|\sum_{p\le z} f_p(n)\right|^{\ell+1}\right)^{\frac{1}{2}}\\ &= \left(\left(\sum_{p\le z} f_p(n)\right)^{\ell-1}\right)^{\frac{1}{2}}\left(\left(\sum_{p\le z} f_p(n)\right)^{\ell+1}\right)^{\frac{1}{2}}\end{aligned}$$

since $\ell-1$, $\ell+1$ are even. For the sake of being concise, let's define

$$\alpha := \left(\left(\sum_{p\le z} f_p(n)\right)^{\ell-1}\right)^{\frac{1}{2}}$$

and

$$\beta := \left(\left(\sum_{p\le z} f_p(n)\right)^{\ell+1}\right)^{\frac{1}{2}}.$$

Therefore, for an odd natural number $\ell$, we've shown

$$\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^{\ell} = \sum_{n\le x}\alpha\beta.$$

This allows us to use the Cauchy-Schwarz lemma. We have

$$\begin{aligned}\left|\sum_{n\le x}\alpha\beta\right| &\le \sqrt{\sum_{n\le x}\alpha^2}\sqrt{\sum_{n\le x}\beta^2}\\ &= \left(\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^{\ell-1}\right)^{\frac{1}{2}}\left(\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^{\ell+1}\right)^{\frac{1}{2}}\end{aligned}$$

where the last line simply comes from using our definitions of $\alpha$ and $\beta$. This rewrite of our original sum $\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^{\ell}$ proves crucial as both $\ell-1,\ \ell+1$ are clearly even, allowing us to apply part (a) of Proposition 4.0.2 to both of these sums. Hence, it is now easy to see that

$$\begin{aligned}
&O\left(\sum_{n\le x}\left|\sum_{p\le z} f_p(n)\right|^{\ell}\right)\\
&=O\left(\left(\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^{\ell-1}\right)^{\frac{1}{2}}\left(\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^{\ell+1}\right)^{\frac{1}{2}}\right)\\
&=O\left(\left(\sqrt{x(\log\log z)^{\frac{\ell-1}{2}}+\pi(z)^{\ell-1}}\right)\left(\sqrt{x(\log\log z)^{\frac{\ell+1}{2}}+\pi(z)^{\ell+1}}\right)\right),
\end{aligned}$$

using part (a) of Proposition 4.0.2. Simplifying the last expression, we are left with

$$O\left(x(\log\log z)^{\frac{\ell}{2}}\right)=O\left(x(\log\log z)^{\frac{k-1}{2}}\right),$$

since $\ell \le k-1$. □

We now have all the tools necessary to prove Theorem 4.0.1 using Proposition 4.0.2.

*Proof of Theorem 4.0.1, assuming Proposition 4.0.2.* Similar to the prior lemma, we will prove this theorem in two cases.

**Case 1** (Assume $k$ is even): By lemmas 4.2.2 and 4.2.3

$$(\omega(n)-\log\log x)^k=\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k+O\left(x(\log\log z)^{\frac{k-1}{2}}\right).$$

Since we are assuming that $k$ is even, then we can apply part (a) of Proposition 4.0.2 to the right-hand side giving us

$$C_k x(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\log\log z}\right)\right)+O(\pi(z)^k)+O\left(x(\log\log z)^{\frac{k-1}{2}}\right)$$
$$=C_k x(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\sqrt{\log\log z}}\right)\right).$$

Recall that $z=x^{\frac{1}{k}}$. After making this substitution for $z$ we quickly see that

$$(\omega(n)-\log\log x)^k = C_k\left(\log\log x^{\frac{1}{k}}\right)^{\frac{k}{2}}\left[1+O\left(\frac{1}{\sqrt{\log\log x^{\frac{1}{k}}}}\right)\right]$$
$$=C_k(\log\log x-\log k)^{\frac{k}{2}}\left[1+O\left(\frac{1}{\sqrt{\log\log x-\log k}}\right)\right].$$

Keeping in mind that $k$ is fixed, we have

$$C_k(\log\log x)^{\frac{k}{2}}\left[1+O\left(\frac{1}{\log\log x}\right)\right]^{\frac{k}{2}}\left[1+O\left(\frac{1}{\sqrt{\log\log x}}\right)\right]$$
$$=C_k(\log\log x)^{\frac{k}{2}}\left[1+O\left(\frac{1}{\log\log x}\right)\right]\left[1+O\left(\frac{1}{\sqrt{\log\log x}}\right)\right]$$
$$=C_k(\log\log x)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\sqrt{\log\log x}}\right)\right).$$

Hence, the theorem holds in Case 1.

**Case 2** (Assume $k$ is odd): Once again from lemmas 4.2.2 and 4.2.3 we start with

$$(\omega(n)-\log\log x)^k=\sum_{n\le x}\left(\sum_{p\le z}f_p(n)\right)^k+O\left(x(\log\log z)^{\frac{k-1}{2}}\right).$$

Because we are assuming that $k$ is odd, this time we apply part (b) of Proposition 4.0.2 to the right-hand side to get

$$\sum_{n\le x}\Bigg(\sum_{p\le z} f_p(n)\Bigg)^k + O\Big(x(\log\log z)^{\frac{k-1}{2}}\Big)$$
$$\ll x(\log\log z)^{\frac{k}{2}}\frac{1}{\sqrt{\log\log z}} + \pi(z)^k + x(\log\log z)^{\frac{k-1}{2}}.$$

We again make the same substitution for $z$ as we did in the previous case. Therefore, the expression above is

$$\begin{aligned}
&\ll x(\log\log x-\log k)^{\frac{k}{2}}\frac{1}{\sqrt{\log\log x-\log k}}\\
&\ll x(\log\log x)^{\frac{k}{2}}\left(1+\frac{1}{\log\log x}\right)^{\frac{k}{2}}\left(\frac{1}{\sqrt{\log\log x}}\right)\\
&\ll x(\log\log x)^{\frac{k}{2}}\left(1+\frac{1}{\log\log x}\right)\left(\frac{1}{\sqrt{\log\log x}}\right)\\
&\ll x(\log\log x)^{\frac{k}{2}}\left(\frac{1}{\sqrt{\log\log x}}\right).
\end{aligned}$$

Hence, we have proved that Proposition 4.0.2 implies Theorem 4.0.1. □

## 4.3 Proof of Proposition 4.0.2

By definition after expanding out the left-hand side,

$$\sum_{n\le x}\Bigg(\sum_{p\le z} f_p(n)\Bigg)^k = \sum_{p_1,\dots,p_k\le z}\Bigg(\sum_{n\le x} f_{p_1\cdots p_k}(n)\Bigg). \tag{4.2}$$

For $r\ge 1$ with a prime factorization of $r := \prod_{i=1}^s q_i^{\alpha_i}$ for distinct primes $q_i$ and $\alpha_i\ge 1$, let us denote the square-free part of $r$ by $R$. Thus $R := \prod_{i=1}^s q_i$. This leads us to our next lemma.

**Lemma 4.3.1.** *We have*

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

*Proof.* Since $R$ is a square-free product of unique primes then if $d = (n, R)$, then $d$ is also product of square-free, unique primes. These primes are shared by $R$ and $n$; hence, they are shared by $r$ and $n$ as well. Furthermore, for $n = \prod_j p_j^{\beta_j}$ and $r = \prod_{i=1}^{s} q_i^{\alpha_i}$, we have

$$\begin{aligned} f_r(n) &= \prod_i f_{q_i}(n)^{\alpha_i} \\ &= \prod_i f_{q_i}\left(\prod_j p_j^{\beta_j}\right)^{\alpha_i}. \end{aligned}$$

Note $q_i \mid p^\beta$ if and only if $q_i \mid p$. So our right-hand side can be rewritten as

$$\prod_i f_{q_i}\left(\prod_j p_j\right)^{\alpha_i}.$$

Furthermore, $q_i \mid \prod_j p_j$ if and only if $q_i = p_j$ for some $j$, which leads us to

$$\begin{aligned} f_r(n) &= \prod_i f_{q_i}\left(\prod_j p_j\right)^{\alpha_i} \\ &= \prod_i f_{q_i}\Big((R, n)\Big)^{\alpha_i} \\ &= f_r(d). \end{aligned}$$

Thus,

$$\begin{aligned}\sum_{n\le x} f_r(n) &= \sum_{\substack{n\le x\\ d=(n,R)}} \sum_{d|R} f_r(n)\\ &= \sum_{\substack{n\le x\\ d=(n,R)}} \sum_{d|R} f_r(d)\\ &= \sum_{d|R} f_r(d) \sum_{\substack{n\le x\\ d=(n,R)}} 1\end{aligned}$$

as stated. □

Here Lemma 2.1.3 proves to be important as it can be used to build a "delta" function to pick out the integer $n = 1$. Let $\varphi(n)$ denote the Euler totient function (which counts the number of integers in the closed interval $[1, n]$ which are coprime to $n$) and let $\tau(n)$ denote the divisor function (which counts the number of positive divisors of $n$). We want to work towards

**Lemma 4.3.2.**

$$\sum_{n\le x} f_r(n) = xG(r) + O(1)$$

*for*

$$G(r) := \frac{1}{R}\sum_{d|R} f_r(d)\varphi\left(\tfrac{R}{d}\right).$$

First let us start with a simpler version of Lemma 4.3.2.

**Lemma 4.3.3.**

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

*Proof.* From Lemma 4.3.1 we have

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

We begin by analyzing the inner sum on the right-hand side. For $d|R$, we have

$$\sum_{\substack{n\le x\\ d=(n,R)}} 1 = \sum_{\substack{n\le x\\ d|n\\ \left(\frac{n}{d},\frac{R}{d}\right)=1}} 1.$$

Now we use the Möbius function to express the condition $\left(\frac{n}{d},\frac{R}{d}\right)=1$. We have that the sum above equals

$$\sum_{\substack{n\le x\\ d|n}} \sum_{e|\left(\frac{n}{d},\frac{R}{d}\right)} \mu(e),$$

which after exchanging the order of summation is

$$\begin{aligned}\sum_{e|\frac{R}{d}}\mu(e)\sum_{\substack{n\le x\\ ed|n}}1 &= \sum_{e|\frac{R}{d}}\mu(e)\left[\frac{x}{ed}+O(1)\right]\\ &= \sum_{e|\frac{R}{d}}\frac{\mu(e)R}{ed}\frac{x}{R}+O\left(\sum_{e|\frac{R}{d}}1\right).\end{aligned}$$

Now recall a basic identity for the Euler totient function (which follows immediately by multiplicativity and the value of $\varphi$ at the prime powers):

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

Using this identity, we finally get

$$\sum_{\substack{n\le x\\ d=(n,R)}} 1 = \sum_{e|\frac{R}{d}}\frac{\mu(e)R}{ed}\frac{x}{R}+O\left(\sum_{e|\frac{R}{d}}1\right)$$

where our error term comes from using Lemma 2.1.3, and our right-hand side is

$$= \varphi\big(\tfrac{R}{d}\big)\frac{x}{R} + O\Big(\tau\big(\tfrac{R}{d}\big)\Big).$$

Inserting back this expression for the inner sum, we get

$$\begin{aligned}
\sum_{n\le x} f_r(n) &= \sum_{d|R} f_r(d) \sum_{\substack{n\le x\\ d=(n,R)}} 1\\
&= \frac{x}{R}\sum_{d|R} f_r(d)\varphi\big(\tfrac{R}{d}\big) + O\left(\sum_{d|R} f_r(d)\tau\big(\tfrac{R}{d}\big)\right)\\
&= xG(r) + O\left(\sum_{d|R} f_r(d)\tau\big(\tfrac{R}{d}\big)\right)
\end{aligned}$$

as claimed. □

To get Lemma 4.3.2 we look at the error from Lemma 4.3.3.

**Lemma 4.3.4.** *We have*

$$O\left(\sum_{d|R} f_r(d)\tau\big(\tfrac{R}{d}\big)\right) = O(1).$$

*Proof.* First we will show that

$$\sum_{d|R} f_r(d)\tau\big(\tfrac{R}{d}\big)$$

is multiplicative in $R$. Precisely, we will show that

$$\sum_{d|R} f_r(d)\tau\big(\tfrac{R}{d}\big) = \prod_{i=1}^{s}\left[\sum_{d_i|q_i} f_{{q_i}^{\alpha_i}}(d_i)\tau\left(\tfrac{q_i}{d_i}\right)\right]$$

by working in reverse and using that $\tau(n)$ is multiplicative in $n$ and $f_r(n)$ is multiplicative as well by definition. Multiplying together the values of $\tau$ using multiplicativity, we get that the right-hand side above equals

$$\prod_{i=1}^{s}\left[\sum_{d_i|q_i} f_{q_i{}^{\alpha_i}}(d_i)\tau\left(\tfrac{q_i}{d_i}\right)\right] = \tau\left(\tfrac{R}{d}\right)\left[\left(\sum_{d_1|q_1} f_{q_1{}^{\alpha_1}}(d_1)\right)\cdots\left(\sum_{d_s|q_s} f_{q_s{}^{\alpha_s}}(d_s)\right)\right]$$

$$= \left[\tau\left(\tfrac{R}{d}\right)\right]\sum_{\substack{d_1|q_1\\ \vdots\\ d_s|q_s}}\left(\prod_{i=1}^{s} f_{q_i{}^{\alpha_i}}(d)\right)$$

$$= \sum_{d|R} f_{\prod q_i{}^{\alpha_i}}(d)\tau\left(\tfrac{R}{d}\right),$$

proving that we have multiplicativity. Furthermore, for $q$ prime, we have $d \mid q$ if and only if $d = 1$ or $d = q$, and $\tau(q) = 2$. Thus, we see that

$$\left|\sum_{d|q_i} f_{q_i^{\alpha_i}}(d)\tau\left(\tfrac{q_i}{d}\right)\right| = \left|f_{q_i^{\alpha_i}}(1)\tau(q_i) + f_{q_i^{\alpha_i}}(q_i)\tau(1)\right|$$

$$= \left|\left(-\frac{1}{q_i}\right)^{\alpha_i}\tau(q_i) + \left(1-\frac{1}{q_i}\right)^{\alpha_i}\right|$$

$$\leq 1.$$

Using this we have,

$$\left|\sum_{d|R} f_{\prod_{i=1}^{s} q_i^{\alpha_i}}(d)\tau\left(\tfrac{R}{d}\right)\right| = \left|\prod_{i=1}^{s}\sum_{d|q_i} f_{q_i^{\alpha_i}}(d)\tau\left(\tfrac{q_i}{d}\right)\right|$$

$$\leq \prod_{i=1}^{s} 1 = 1,$$

which proves our lemma. □

We are now able to prove Lemma 4.3.2.

*Proof of Lemma 4.3.2.* By lemmas 4.3.1, 4.3.3, and 4.3.4, we have that

$$\begin{aligned}\sum_{n\leq x} f_r(n) &= \sum_{d|R} f_r(d) \sum_{\substack{n\leq x \\ d=(n,R)}} 1 \\ &= xG(r) + O\left(\sum_{d|R} f_r(d)\tau\left(\tfrac{R}{d}\right)\right) \\ &= xG(r) + O(1)\end{aligned}$$

as stated. □

Let's now make an observation about $G(r)$.

**Lemma 4.3.5.** *We have*

$$G(r) = \prod_{q^\alpha \| r}\left[\frac{1}{q}\left(1-\frac{1}{q}\right)^\alpha + \left(1-\frac{1}{q}\right)\left(-\frac{1}{q}\right)^\alpha\right].$$

*Proof.* We begin this proof similarly to Lemma 4.3.4 by first showing that $G(r)$ is multiplicative. Define $d := d_1 d_2 \cdots d_s$. Then

$$\begin{aligned}\prod_{i=1}^{s}\left(G(q_i^{\alpha_i})\right) &= G(q_1^{\alpha_1})G(q_2^{\alpha_2})\cdots G(q_s^{\alpha_s}) \\ &= \left[\frac{1}{q_1}\sum_{d_1|q_1} f_{q_1^{\alpha_1}}(d_1)\varphi\left(\tfrac{q_1}{d_1}\right)\right]\left[\frac{1}{q_2}\sum_{d_2|q_2} f_{q_2^{\alpha_2}}(d_2)\varphi\left(\tfrac{q_2}{d_2}\right)\right]\cdots \\ &\quad\cdots\left[\frac{1}{q_s}\sum_{d_s|q_s} f_{q_s^{\alpha_s}}(d_s)\varphi\left(\tfrac{q_s}{d_s}\right)\right].\end{aligned}$$

Noting that $\varphi(n)$ is multiplicative then we have

$$\begin{aligned}&\frac{1}{R}\left[\left(\sum_{d_1|q_1} f_{q_1^{\alpha_1}}(d_1)\right)\left(\sum_{d_2|q_2} f_{q_2^{\alpha_2}}(d_2)\right)\cdots\left(\sum_{d_s|q_s} f_{q_s^{\alpha_s}}(d_s)\right)\right]\varphi\left(\tfrac{R}{d}\right)\\ &=\frac{1}{R}\sum_{\substack{d_1|q_1\\ \vdots\\ d_s|q_s}}\left[\prod_{i=1}^{s}\left(f_{q_i}(d_i)\right)^{\alpha_i}\right]\varphi\left(\tfrac{R}{d}\right).\end{aligned}$$

Using the multiplicativity of $f_r(n)$, we get that the above expression is

$$\begin{aligned}&\frac{1}{R}\sum_{d|R} f_{\prod q_i^{\alpha_i}}(d)\varphi\left(\tfrac{R}{d}\right)\\ &= G\left(\prod_{i=1}^{s} q_i^{\alpha_i}\right).\end{aligned}$$

Thus $G(r)$ is multiplicative. Furthermore, at prime powers we have

$$\begin{aligned}G(q_i^{\alpha_i}) &= \frac{1}{q_i}\sum_{d|q_i} f_{q_i^{\alpha_i}}(d)\varphi\left(\tfrac{q_i}{d}\right)\\ &= \frac{1}{q_i}\left(1-\frac{1}{q_i}\right)^{\alpha_i}\varphi\left(\frac{q_i}{q_i}\right)+\frac{1}{q_i}\left(-\frac{1}{q_i}\right)^{\alpha_i}\varphi\left(\tfrac{q_i}{1}\right).\end{aligned}$$

Now evaluating $\varphi(q_i) = q_i - 1$, the above expression is

$$\begin{aligned}&\frac{1}{q_i}\left(1-\frac{1}{q_i}\right)^{\alpha_i}+\frac{1}{q_i}\left(-\frac{1}{q_i}\right)^{\alpha_i}(q_i-1)\\ &=\frac{1}{q_i}\left(1-\frac{1}{q_i}\right)^{\alpha_i}+\left(1-\frac{1}{q_i}\right)\left(-\frac{1}{q_i}\right)^{\alpha_i}.\end{aligned}$$

With this we see that

$$
\begin{aligned}
G(r) = G\left(\prod_{i=1}^{s} q_i^{\alpha_i}\right) &= \prod_{i=1}^{s} G(q_i^{\alpha_i}) \\
&= \prod_{i=1}^{s}\left[\frac{1}{q_i}\left(1-\frac{1}{q_i}\right)^{\alpha_i} + \left(1-\frac{1}{q_i}\right)\left(-\frac{1}{q_i}\right)^{\alpha_i}\right] \\
&= \prod_{q^\alpha \| r}\left[\frac{1}{q}\left(1-\frac{1}{q}\right)^{\alpha} + \left(1-\frac{1}{q}\right)\left(-\frac{1}{q}\right)^{\alpha}\right],
\end{aligned}
$$

giving us the wanted result. □

Keeping this lemma in mind, we see a crucial property of $G(r)$ given in the next lemma. We say that $r$ is square-full if every prime divisor of $r$ occurs with exponent at least two. That is, $q|r \implies q^2|r$ for all primes $q$.

**Lemma 4.3.6.** *If $r$ is not square-full, then $G(r) = 0$.*

*Proof.* We have seen that

$$G(r) = \prod_{q^\alpha \| r}\left[\frac{1}{q}\left(1-\frac{1}{q}\right)^{\alpha} + \left(1-\frac{1}{q}\right)\left(-\frac{1}{q}\right)^{\alpha}\right].$$

If $r$ is not square-full, then it has a prime factor $q$ with corresponding exponent $\alpha = 1$. Now it remains to observe that

$$\left(-\frac{1}{q}\right)\left(1-\frac{1}{q}\right)^{1} + \left(1-\frac{1}{q}\right)\left(\frac{1}{q}\right)^{1} = 0.$$

□

Now we are ready to use our tools. We return to (4.2), where we saw

$$\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k = \sum_{p_1,\ldots,p_k\le z}\left(\sum_{n\le x} f_{p_1\cdots p_k}(n)\right).$$

To this we can apply Lemma 4.3.2 and Lemma 4.3.6 to get that the above $k$-th moment equals

$$\begin{aligned}
&\sum_{p_1,\ldots,p_k\leq z}\big(xG(p_1\cdots p_k)+O(1)\big)\\
&=\sum_{\substack{p_1,\ldots,p_k\leq z\\ p_1\cdots p_k \text{ square-full}}} xG(p_1\cdots p_k)+O\Bigg(\sum_{p_1,\ldots,p_k\leq z}1\Bigg)\\
&=\sum_{\substack{p_1,\ldots,p_k\leq z\\ p_1\cdots p_k \text{ square-full}}} xG(p_1\cdots p_k)+O\big(\pi(z)^k\big). \qquad (4.3)
\end{aligned}$$

From here forward, we will begin referencing the first term above as the “main term.” Let $q_1<\cdots<q_s$ be the distinct primes of the square-full number $p_1\cdots p_k$ in the argument of the main term. Note that the exponents of each prime $q_i$ must satisfy $\alpha_i\geq 2$, by the square-full assumption. This implies that $s\leq\frac{k}{2}$. So we may rewrite our main term as follows

$$\sum_{s\leq\frac{k}{2}}\sum_{q_1<\cdots<q_s\leq z}\sum_{\substack{\alpha_1\cdots\alpha_s\geq 2\\ \sum_i\alpha_i=k}}\frac{k!}{\alpha_1!\cdots\alpha_s!}G(q_1^{\alpha_1}\cdots q_s^{\alpha_s}) \qquad (4.4)$$

where $\frac{k!}{\alpha_1!\cdots\alpha_s!}$ comes from the number of ways to divide $k$ different objects into $s$ groups of sizes $\alpha_1,\ldots,\alpha_s$.

We now want to find those values in our summation that make the largest contribution to the overall value. Our strategy for this is to first calculate the value of the term when $s=\frac{k}{2}$ (which occurs only when $k$ is even) and then approximate the value for those terms where $s<\frac{k}{2}$. We denote the former by $M_{\frac{k}{2}}$ and the total of the rest of the terms by $M_{<\frac{k}{2}}$. Thus (4.4) equals $M_{\frac{k}{2}}+M_{<\frac{k}{2}}$

**Lemma 4.3.7.** *When $k$ is even, we have*

$$M_{\frac{k}{2}}=\frac{k!}{2^{\frac{k}{2}}\big(\frac{k}{2}\big)!}\big(\log\log z+O(1)\big)^{\frac{k}{2}}.$$

*Proof.* Recall that $M_{\frac{k}{2}}$ is the term of (4.4) with $s = \frac{k}{2}$. In this term we have $\alpha_i = 2$ for all $i$. Thus

$$M_{\frac{k}{2}} = \frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!} \sum_{\substack{q_1,\ldots,q_{k/2}\leq z \\ q_i \text{ distinct}}} \left[ \prod_{i=1}^{\frac{k}{2}} \frac{1}{q_i}\left(1 - \frac{1}{q_i}\right) \right]$$

since there are $\frac{k}{2}$ many $\alpha_i' s$, which can be ordered $\frac{k}{2}!$ times. Let us first work towards finding an upper bound. By ignoring the distinctness condition of the $q_i' s$ we see that our sum is bounded from above by

$$\frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!} \left( \sum_{p\leq z} \frac{1}{p}\left(1 - \frac{1}{p}\right) \right)^{\frac{k}{2}} = \frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!} \left( \sum_{p\leq z} \frac{1}{p} - \sum_{p\leq z} \frac{1}{p^2} \right)^{\frac{k}{2}}.$$

The first sum on the right-hand side is $\log\log z + O(1)$ by Lemma 2.1.2. Also

$$\sum_{p\leq z} \left(\frac{1}{p}\right)^2 \ll \sum_{n=1}^{\infty} \frac{1}{n^2} \ll 1.$$

Hence, we have an upper bound for $M_{\frac{k}{2}}$ of

$$\frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!} \big(\log\log z + O(1)\big)^{\frac{k}{2}}.$$

Now onto finding a lower bound. Clearly,

$$\sum_{\substack{q_1,\ldots,q_{k/2}\leq z \\ q_i \text{ distinct}}} \left[ \prod_{i=1}^{\frac{k}{2}} \frac{1}{q_i}\left(1 - \frac{1}{q_i}\right) \right] = \sum_{q_1\leq z} \frac{1}{q_1}\left(1 - \frac{1}{q_1}\right) \cdots \sum_{q_{\frac{k}{2}}\leq z} \frac{1}{q_{\frac{k}{2}}}\left(1 - \frac{1}{q_{\frac{k}{2}}}\right) \tag{4.5}$$

for $q_{\frac{k}{2}}$ distinct from $q_{\frac{k}{2}-1}, \ldots, q_1$. From here we want to find a lower bound for the last sum of (4.5) and then recursively apply that bound for each sum in the product. The term

$$\frac{1}{q_{\frac{k}{2}}}\left(1 - \frac{1}{q_{\frac{k}{2}}}\right)$$

is the smallest it can possibly be when $q_{\frac{k}{2}}$ is the largest prime possible. Now $q_{\frac{k}{2}}$ can take on any prime value *except* the ones already taken up by $q_1, \ldots, q_{\frac{k}{2}-1}$ since the sum over $q_{\frac{k}{2}}$ is innermost and there is a distinctness requirement. Thus the smallest values of $\frac{1}{q_{\frac{k}{2}}}\left(1 - \frac{1}{q_{\frac{k}{2}}}\right)$ will arise when $q_1, \ldots, q_{\frac{k}{2}-1}$ have taken up the smallest prime values available, leaving only larger prime values for $q_{\frac{k}{2}}$ to take on. Let $\pi_n$ be the $n$th smallest prime. By this logic, we have

$$\sum_{q_{\frac{k}{2}} \leq z} \frac{1}{q_{\frac{k}{2}}}\left(1 - \frac{1}{q_{\frac{k}{2}}}\right) \geq \sum_{\pi_{\frac{k}{2}} \leq q_{\frac{k}{2}} \leq z} \frac{1}{q_{\frac{k}{2}}}\left(1 - \frac{1}{q_{\frac{k}{2}}}\right)$$

$$= \sum_{\pi_{\frac{k}{2}} \leq q_{\frac{k}{2}} \leq z} \frac{1}{q_{\frac{k}{2}}} - \sum_{\pi_{\frac{k}{2}} \leq q_{\frac{k}{2}} \leq z} \left(\frac{1}{q_{\frac{k}{2}}}\right)^2.$$

Applying Lemma 2.1.2 to the first term gives

$$\log\log z + O(1) - \big(\log\log \pi_{\frac{k}{2}} + O(1)\big) = \log\log z + O(1),$$

since $\pi_{\frac{k}{2}} \leq k$ and $k$ is fixed. Furthermore,

$$\sum_{\pi_{\frac{k}{2}} \leq q_{\frac{k}{2}} \leq z} \left(\frac{1}{q_{\frac{k}{2}}}\right)^2 \ll \sum_{n=1}^{\infty} \frac{1}{n^2} \ll 1,$$

as already observed. Thus, when combined we see that

$$\sum_{\pi_{\frac{k}{2}} \leq q_{\frac{k}{2}} \leq z} \frac{1}{q_{\frac{k}{2}}}\left(1 - \frac{1}{q_{\frac{k}{2}}}\right) \geq \log\log z + O(1).$$

Repeating this process for all of the sums given in the product (4.5) gives the lower bound

$$\frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}\big(\log\log z + O(1)\big)^{\frac{k}{2}}.$$

Thus, our lower and upper bounds are equal (asymptotically) so, when $k$ is even, $M_{\frac{k}{2}}$ is

$$\frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}\sum_{\substack{q_1,\dots,q_{k/2}\\ q_i \text{ distinct}}}\prod_{i=1}^{\frac{k}{2}}\frac{1}{q_i}\left(1-\frac{1}{q_i}\right)$$
$$=\frac{k!}{2^{\frac{k}{2}}\left(\frac{k}{2}\right)!}\big(\log\log z+O(1)\big)^{\frac{k}{2}}$$

as stated. ☐

Now we move to the last step needed to prove Granville and Soundararajan's beautiful theorem: that is to approximate$M_{<\frac{k}{2}}$. For this last part of the proof we will be utilizing some elementary combinatorics.

**Lemma 4.3.8.** *We have*

$$M_{<\frac{k}{2}} \ll \max_{s<\frac{k}{2}}\bigg(\log\log z\bigg)^{s}.$$

*Proof.* Let us begin with the simple observation that

$$0 \le G(q_1^{\alpha_1}\cdots q_s^{\alpha_s}) \le \frac{1}{q_1\cdots q_s}.$$

Therefore, from (4.4) we have that $M_{<\frac{k}{2}}$ is bounded above by

$$\sum_{s<\frac{k}{2}}\ \sum_{q_1<\cdots<q_s\le z}\ \sum_{\substack{\alpha_1\cdots\alpha_s\ge 2\\ \sum_i \alpha_i=k}}\frac{k!}{\alpha_1!\cdots\alpha_s!}G(q_1^{\alpha_1}\cdots q_s^{\alpha_s})$$
$$=\sum_{s<\frac{k}{2}}k!\sum_{q_1<\cdots<q_s\le z}G(q_1^{\alpha_1}\cdots q_s^{\alpha_s})\sum_{\substack{\alpha_1\cdots\alpha_s\ge 2\\ \sum_i \alpha_i=k}}\frac{1}{\alpha_1!\cdots\alpha_s!}$$
$$=\sum_{s<\frac{k}{2}}\frac{k!}{s!}\sum_{q\le z}G\Bigg(\prod_{q^\alpha\|r}q^\alpha\Bigg)\sum_{\substack{\alpha_1\cdots\alpha_s\ge 2\\ \sum_i \alpha_i=k}}\frac{1}{\alpha_1!\cdots\alpha_s!}.$$

A word of explanation for the last line: since we are no longer forcing the primes $q_i$ in ascending order we have to divide by $s!$. Now by noting our prior observation and the fact that we have a product of $s$ terms we see that

$$\begin{aligned}
&\sum_{s<\frac{k}{2}} \frac{k!}{s!} \sum_{q\leq z} G\left(\prod_{q^\alpha \| r} q^\alpha\right) \sum_{\substack{\alpha_1 \cdots \alpha_s \geq 2 \\ \sum_i \alpha_i = k}} \frac{1}{\alpha_1! \cdots \alpha_s!} \\
&\leq \sum_{s<\frac{k}{2}} \frac{k!}{s!} \left(\sum_{q\leq z} \frac{1}{q}\right)^s \sum_{\substack{\alpha_1,\dots,\alpha_s \geq 2 \\ \sum_i \alpha_i = k}} \frac{1}{\alpha_1! \dots \alpha_s!}. \qquad (4.6)
\end{aligned}$$

The number of ways that $k$ can be written as $\alpha_1 + \cdots + \alpha_s$ for $\alpha_i \geq 2$ is the same as the number of partitions of $k - s$ into $s$ positive natural numbers. This is because

$$k = \alpha_1 + \cdots + \alpha_s \iff k - s = (\alpha_1 - 1) + \cdots + (\alpha_s - 1).$$

Now by the famous combinatorics "stars and bars" problem, which tells us that the number of such partitions of $k - s$ into $s$ natural numbers is $\binom{k-s-1}{s-1}$. It is also simple to see that

$$\frac{1}{\alpha_1 \cdots \alpha_s} \leq \frac{1}{2^s},$$

as $\alpha_i \geq 2$ by the square-full condition. Hence, by using Lemma 2.1.2 we have that (4.6) is bounded above by

$$\leq \sum_{s<\frac{k}{2}} \frac{k!}{s!2^s} \binom{k-s-1}{s-1} \left(\log\log z + O(1)\right)^s,$$

which gives the bound we wanted, since $k$ is fixed. □

We are now able to prove Proposition 4.0.2.

*Proof of Proposition 4.0.2.* Recall that from (4.3) and (4.4) we have

$$\sum_{n\leq x}\Bigg(\sum_{p\leq z} f_p(n)\Bigg)^k = \sum_{\substack{p_1,\ldots,p_k\leq z\\ p_1\cdots p_k\text{square-full}}} xG(p_1\cdots p_k) + O\big(\pi(z)^k\big)$$

$$=\sum_{s\leq\frac{k}{2}}\ \sum_{q_1<\cdots<q_s\leq z}\ \sum_{\substack{\alpha_1\cdots\alpha_s\geq 2\\ \sum_i\alpha_i=k}} \frac{k!}{\alpha_1!\cdots\alpha_s!}G(q_1^{\alpha_1}\cdots q_s^{\alpha_s}) + O\big(\pi(z)^k\big)$$

where we have been referring to the summation on the right-hand side as the main term. From Lemma 4.3.7 we see that when $k$ is even the largest part of the main term is when $s=\frac{k}{2}$, which contributes a value of

$$x\frac{k!}{2^{\frac{k}{2}}(\frac{k}{2})!}\big(\log\log z + O(1)\big)^{\frac{k}{2}}$$

$$= xC_k(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\log\log z}\right)\right)^{\frac{k}{2}}$$

$$= xC_k(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\log\log z}\right)\right).$$

Now let us define

$$\ell := \begin{cases} \frac{k-2}{2} & \textit{if } k \textit{ is even}\\ \frac{k-1}{2} & \textit{if } k \textit{ is odd}\end{cases}$$

where $\ell$ is the greatest integer strictly less than $\frac{k}{2}$. Then by Lemma 4.3.8 the remaining parts of the main term, that is $M_{<\frac{k}{2}}$, only contribute a value

$$O\big(x(\log\log z)^{\ell}\big).$$

It is crucial for us to note that

$$xC_k(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\log\log z}\right)\right)$$
$$\gg x(\log\log z)^{\ell}.$$

Thus, all of the terms where $s < \frac{k}{2}$ are "inconsequential" with respect to magnitude when compared to the largest term, $M_{\frac{k}{2}}$ arising from $s = \frac{k}{2}$, which once again only happens when $k$ is even.

Putting everything together, we get for $k$ is even:

$$\begin{aligned}\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k &= xC_k(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\log\log z}\right)\right)\\ &\quad + O\big(x(\log\log z)^{\ell}\big) + O\big(\pi(z)^k\big)\\ &= xC_k(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\log\log z}\right)\right)\\ &\quad + O\Big(x(\log\log z)^{\frac{k-2}{2}}\Big) + O\big(\pi(z)^k\big)\\ &= xC_k(\log\log z)^{\frac{k}{2}}\left(1+O\left(\frac{1}{\log\log z}\right)\right) + O\big(\pi(z)^k\big),\end{aligned}$$

and when $k$ is odd

$$\begin{aligned}\sum_{n\le x}\left(\sum_{p\le z} f_p(n)\right)^k &\ll (\log\log z)^{\frac{k-1}{2}} + \pi(z)^k\\ &\ll x(\log\log z)^{\frac{k}{2}}\frac{1}{\sqrt{\log\log z}} + \pi(z)^k,\end{aligned}$$

proving Proposition 4.0.2.

□

BIBLIOGRAPHY

[1] Patrick Billingsley, *Probability and measure*, 2nd ed., Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986. MR830424

[2] H. Delange, *Sur le nombre des diviseurs premiers de $n$*, Acta Arith. **7** (1961/62), 191–215, DOI 10.4064/aa-7-2-191-215 (French). MR136588

[3] 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, DOI 10.2307/2371483. MR2374

[4] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed., Elsevier/Academic Press, Amsterdam, 2007. Translated from the Russian; Translation edited and with a preface by Alan Jeffrey and Daniel Zwillinger; With one CD-ROM (Windows, Macintosh and UNIX). MR2360010

[5] 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, DOI 10.1007/978-1-4020-5404-4-2. MR2290492

[6] G. H. Hardy and S. Ramanujan, *The normal number of prime factors of a number $n$ [Quart. J. Math.* **48** *(1917), 76–92]*, Collected papers of Srinivasa Ramanujan, AMS Chelsea Publ., Providence, RI, 2000, pp. 262–275. MR2280878

[7] Hugh L. Montgomery and Robert C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics, vol. 97, Cambridge University Press, Cambridge, 2007. MR2378655

[8] Paul Turán, *On a Theorem of Hardy and Ramanujan*, J. London Math. Soc. **9** (1934), no. 4, 274–276, DOI 10.1112/jlms/s1-9.4.274. MR1574877

# VITA

The author graduated from West Union Attendance Center in 2016. He then attended Blue Mountain College where he graduated with a B.S. in Mathematics in 2019. During this time he was awarded the Anne Sanford Mathematics Award in 2018 and 2019 from his department.