8-2018

Brun's 1920 Theorem on Goldbach's Conjecture

James A. Farrugia

Utah State University

BRUN'S 1920 THEOREM ON GOLDBACH'S CONJECTURE

by

James A. Farrugia

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Mathematics

Approved:

David E. Brown, Ph.D. Major Professor	James S. Cangelosi, Ph.D Committee Member
Dariusz M. Wilczynski, Ph.D. Committee Member	Mark R. McLellan, Ph.D. Vice President for Research Dean of the School of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2018

ABSTRACT

Brun's 1920 Theorem on Goldbach's Conjecture

by

James A. Farrugia, Master of Science

Utah State University, 2018

Major Professor: Dr. David E. Brown
Department: Mathematics and Statistics

One form of Goldbach's Conjecture asserts that every even integer greater than 4 is the sum of two odd primes. In 1920 Viggo Brun proved that every sufficiently large even number can be written as the sum of two numbers, each having at most nine prime factors. This thesis explains the overarching principles governing the intricate arguments Brun used to prove his result.

Though there do exist accounts of Brun's methods, those accounts seem to miss the forest for the trees. In contrast, this thesis explains the relatively simple structure underlying Brun's arguments, deliberately avoiding most of his elaborate machinery and idiosyncratic notation. For further details, the curious reader is referred to Brun's original paper (in French).

Brun constructs two main sieves. For each of these sieves, he establishes a lower bound for the number of elements N that fall through the sieve. Then, he uses additional results by Stirling and Mertens, along with innovative algebraic manipulations to construct improved lower bounds for N. Subsequently, Brun applies his earlier results, *mutatis mutandis*, to Merlin's double sieve and obtains the result that allows him to prove his theorem on Goldbach's Conjecture.

In distilled form, Brun's arguments run as follows: start with a lower bound of the form $N > M - R$, which this thesis calls "the Fundamental Inequality." Then, bound M below by A and bound R above by B, to show that $M > A - B$. Thus, $A - B$ is the improved lower bound for N.

(95 pages)

PUBLIC ABSTRACT

Brun's 1920 Theorem on Goldbach's Conjecture

James A. Farrugia

One form of Goldbach's Conjecture asserts that every even integer greater than 4 is the sum of two odd primes. In 1920 Viggo Brun proved that every sufficiently large even number can be written as the sum of two numbers, each having at most nine prime factors. This thesis explains the overarching principles governing the intricate arguments Brun used to prove his result.

Though there do exist accounts of Brun's methods, those accounts seem to miss the forest for the trees. In contrast, this thesis explains the relatively simple structure underlying Brun's arguments, deliberately avoiding most of his elaborate machinery and idiosyncratic notation. For further details, the curious reader is referred to Brun's original paper (in French).

ACKNOWLEDGMENTS

Thanks to everyone who helped.

James A. Farrugia

Contents

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

1.1 Overview

A conjecture named after Christian Goldbach (1690-1764) asserts that every even integer greater than four is the sum of two odd primes [53]. This conjecture has not yet been proved. However, many results related to Goldbach's Conjecture have been proved, one of which was published by Viggo Brun in 1920 [3]: *every sufficiently large even number can be written as the sum of two numbers, each of which has at most nine prime factors*. This thesis gives a simplified explanation of how Brun achieved this result, which we call *Brun's theorem*.

Although accounts of Brun's theorem do exist in the literature, those accounts seem intended for advanced undergraduates or graduate students [7, 17, 19, 21, 30, 39, 43]. The goal of this thesis is to explain in a simplified but faithful way the methods Brun used to obtain his result on Goldbach's Conjecture, and thus make Brun's original arguments accessible to a wider audience.

1.1.1 What Brun Did to Achieve his Result

At the risk of oversimplification, Brun

1. modified the sieve of Eratosthenes [40] (see Section 1.2) by
    - using a double sieve, based on work by Jean Merlin [22], to provide results about pairs of numbers whose sum is an even number (see Section 1.9), and
    - reducing the size of the largest prime used for sifting, with the result that certain numbers (those with at most nine primes factors) fall through his sieve (see Section 1.8)

2. made and exploited observations about the component calculations used in the sieve of Eratosthenes-Legendre to develop various lower-bound estimates for the number of numbers that survive his sifting process (see Sections 1.4, 3.2, 3.3, and 3.4), and

3. used certain asymptotic approximations (formulas by Mertens and Stirling [49]) to show that one of his improved lower-bound estimates allowed him to establish his result on Goldbach's Conjecture (see Sections 3.3 - 3.5).

1.1.2 Preview of the Fundamental Inequality: $N > M - R$

In §3 and §4 of [3], Brun develops two sieves and shows that the number N of elements that fall through each sieve is bounded below by the difference of a main term M and a remainder term R. Brun then uses inequalities of the form $N > M - R$ as his points of departure for calculations that yield an improved lower bound for the number of elements falling through his sieves.

In each of his two main sieves in §3 and §4 of [3], Brun's lengthy calculations can be put into a simple conceptual scheme: he bounds the main term M below by A and the remainder term R above by B, thus arriving at an improved lower bound $N > A - B$.

Because Brun starts with lower bounds expressed as inequalities of the form $N > M - R$, and because Brun's subsequent calculations can be put into the simple conceptual scheme described in the previous paragraph, we refer to the inequality $N > M - R$ as the *Fundamental Inequality* (see also Section 1.6). This conceptual scheme is the guiding thread through Brun's calculations that can help readers see the forest for the trees.

In §3 and §4 of [3] Brun shows that the two sieves he uses do actually yield lower bounds for N, the number of elements that fall through the sieve. In Section 3.3.2 we explain, using an argument different from the one Brun uses, why his first sieve gives a lower bound for N. In Section 3.4.2 we explain, using more complicated arguments that follow those used by Brun, why his second sieve gives a lower bound for N.

1.1.3 Plan

The remainder of this chapter deals with the following:

- The sieve of Eratosthenes

- The sieve of Eratosthenes-Legendre
- The Legendre formula
- The "best" lower bound for N
- The Fundamental Inequality $N > M - R$
- Sieve integrity when truncating the Legendre formula
- Why some composites fall through Brun's sieves
- Brun's use of Merlin's double sieve
- A contextual overview of Brun's approach

Notation is introduced along with examples, and the relevant sieve machinery is built up slowly. Chapter 2 discusses some history and significance of Goldbach's Conjecture, Brun's work, and related results. Chapter 3 provides a discussion of the methods Brun used to achieve his 1920 result on Goldbach's Conjecture. The discussion is intended to provide a perspective from which readers can see Brun's work as "nothing more" than an elaborate modification of the fundamental inequality $N > M - R$. Many details, most of which tend to obscure the main threads of Brun's arguments, are left for the reader to pursue, either in Brun's original paper in French [3], or in the English translation by Rui [44] (*caveat lector*, since much is lost in translation).

1.2 The Sieve of Eratosthenes

In everyday terminology a *sieve* is a device that separates coarser materials from finer ones by trapping the coarser material in the sieve and allowing the finer material to pass through the sieve. Depending on one's intention, one's interest may be in the material that gets trapped (think of panning for gold) or in the material that passes through (think of sifting flour). In number theory, a sieve is an operation on a sequence of numbers (often a finite sequence of positive integers) that yields a subset, or an estimate of the size of a subset, of the original sequence. Interest in the results of a number-theoretic sieve usually centers on the finer elements (primes or composites containing only a few prime factors) that fall through the sieve, or the number of such elements.

The *sieve of Eratosthenes* is a millennia-old technique for finding primes numbers, given that we already know certain smaller primes. The sieve of Eratosthenes is named after Eratosthenes, a Greek mathematician from the 3rd century BC, who may be best known for his estimate of the circumference of the earth (and less well known for his role as librarian of the ancient library at Alexandria). The earliest known description of the sieve of Eratosthenes is by Nicomachus (ca. 60 AD - ca. 120 AD) in his *Introduction to Arithmetic* [40], though the sieve described by Nicomachus appears to be not quite the same as modern versions of the sieve of Eratosthenes. (In particular, it sifts the odd positive integers greater than 1 (see [40], p. 204), whereas modern versions typically sift consecutive positive integers less than or equal to some number x).

The numbers that fall through the sieve of Eratosthenes are primes and the number 1 (see Example 1.2.0.1). Our chief interest is in the *number* of elements that fall through the sieve. However, in certain modifications of the sieve of Eratosthenes, such as that done by Brun, we are also interested in the *nature* of the numbers that pass through the sieve, since in Brun's modified sieve, not all the numbers that fall through the sieve are necessarily prime. In Brun's theorem, these numbers may have up to nine prime factors, and his "coarser sieve" can be considered to be a byproduct or cost of his technique.

The sieve of Eratosthenes plays an important role in our investigation because it forms the basis of methods subsequently developed by Legendre (see Sections 1.3 and 1.4), which in turn were extensively manipulated by Brun (though Brun does not mention Legendre's work in his 1920 paper).

Example 1.2.0.1. *Consider the finite sequence of positive integers* $1, 2, 3, \ldots, 37$. *Suppose we want to "sift out" the composite numbers of this sequence to allow only the primes and the number* 1 *to fall through the sieve. How might we do that?*

The sieving operation we perform on the original sequence can be considered in two steps. First, we identify the composite numbers in the sequence. Second, we remove those composite numbers from the original sequence - i.e., we trap them in the sieve – so that the numbers that fall through the sieve are just the prime numbers from 1 through 37 and the number 1.

Elaborating on the first step, we might ask, "How can we identify the composite numbers between 1 and 37 (inclusive)?" Since, by the fundamental theorem of arithmetic, any composite number can be expressed uniquely (up to rearrangement of factors) as a product of primes, we can identify composite numbers between 1 and 37 by first considering which prime numbers can be factors of numbers in that range. Certainly we need not consider any primes greater than 37, since no such prime can be a factor of 37. It turns out that we also need not consider any primes greater than $\sqrt{37}$. We prove this claim below.

So, to identify the composite numbers between 1 and 37, it suffices to form multiples of the primes less than or equal to $\sqrt{37}$, i.e., multiples of 2, 3, and 5. We can see in Figure 1.2.0.1 the result of using the sieve of Eratosthenes on the sequence $1, 2, 3, \ldots, 37$ from Example 1.2.0.1.

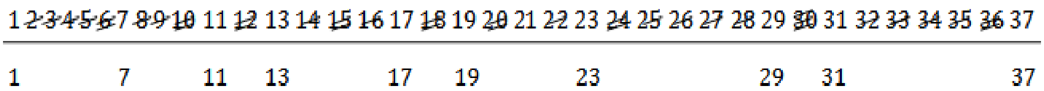


FIGURE 1.2.0.1: Primes and 1 fall through the sieve of Eratosthenes

Lines through numbers in the top row indicate that those numbers, as multiples of 2, 3 or 5, are trapped in the sieve. More than one line through a number indicates that the number is a multiple of more than one of 2, 3, and 5. As we shall see in Section 1.3, the number of times a given composite number is crossed out is significant. By keeping track of how many times a given composite is crossed out, we can develop a formula to compute, in simple cases, the number of elements that pass through a sieve.

Next we prove the general case of the claim mentioned above, namely, that when sifting the sequence $1, 2, 3, \ldots, 37$, we need only consider multiples of primes less than or equal to $\sqrt{37}$.

Setup: Let x be an arbitrary but fixed positive integer, and consider the sequence $1, 2, 3, \ldots, x$ of consecutive positive integers less than or equal to x. Let $p_1, p_2, \ldots p_r$ be consecutive primes with $p_r = p(\sqrt{x})$ the largest prime less than or equal to $\sqrt{x}$. Define the output of the sifting function $S(1, x, p_1, p_2, \ldots, p_r)$ to be the elements of the sequence $1, 2, 3, \ldots, x$ that pass through a sieve that traps multiples of the primes $p_1, p_2, \ldots, p_r$

that lie between 1 and x inclusive, while allowing the remaining numbers of the sequence $1, 2, 3, \ldots, x$ to pass through.

Claim 1.2.1. When sifting the sequence of consecutive integers from 1 through x by the consecutive primes $p1, p2, \ldots, p_r$, where $p_r \leq \sqrt{x}$, we need only consider multiples of primes less than or equal to $\sqrt{x}$.

Proof. Suppose, to the contrary, that to trap all the composite numbers from 1 through x in the sieve, we need to sift by at least one prime greater than $\sqrt{x}$. That is, suppose there exists a composite number $c \leq x$ that is not trapped in the sieve when sifting by the primes $p_1, p_2, \ldots, p_r$. That means c is not a multiple of *any* of the primes $p_1, p_2, \ldots, p_r$, for if it were, it would be trapped in the sieve. Thus, each of c's prime factors must be greater than p_r and therefore greater than $\sqrt{x}$. But that would mean that c itself is greater than x, contradicting the constraint that $c \leq x$. Therefore, the original supposition is false, and the claim is true. □

The sifting operation can also be further broken down into different steps that correspond to trapping multiples of individual primes. Viewing the details of these steps will allow us to see certain key internal workings of the sieve machinery. An amplification of the component machinery of the sieve of Eratosthenes, called the sieve of Eratosthenes-Legendre (see Section 1.3), will show us that various components of this machinery can be manipulated to achieve certain desired results. In fact, Brun manipulated these internal components by actually removing some of them, with the result that his sieve machinery still works, but gives a different kind of result from that obtained by the operation of the traditional sieve of Eratosthenes.

1.3 The Sieve of Eratosthenes-Legendre

The sieve of Eratosthenes-Legendre refers to the sieve of Eratosthenes from the perspective of a formula credited to Legendre [1] for calculating the number of elements, $N(1, x, p_1, p_2, \ldots, p_r)$, that pass through a sieve, where $N(1, x, p_1, p_2, \ldots, p_r)$ equals the number of elements that fall through a sieve of positive integers (not necessarily consecutive) from 1 through x that are sifted by multiples of the primes $p_1, \ldots, p_r$. The primes $p_1, \ldots, p_r$ are called the *sifting primes*.

[1] The credit accorded to Legendre for this formula comes from his work in the second edition of his *Essai sur la théorie des nombres* in 1808 [34] (pp. 399-401), where he gave a method of counting elements in a sequence by considering how many times certain prime factors divided those elements.

One way to consider the functioning of the formula for $N(1, x, p_1, p_2, \ldots, p_r)$ is as a stepwise procedure that computes a running count of the number of elements trapped in the sieve and then adjusts that count at each step. See Example 1.4.0.1.

Consider the number of elements that fall through the traditional sieve of Eratosthenes. We'll calculate this number via a sequence of steps that track a running count of the elements so far trapped in the sieve. We start with a full count of all the integers in the given sequence, and we adjust this count by considering how many times an element is trapped in the sieve at each step of the process. In Example 1.2.0.1, for instance, 6 is trapped twice - once as a multiple of 2 and once again as a multiple of 3. The formula we'll use adjusts its running count by alternately including or excluding the number of certain composites that are trapped at a given stage of the sieve process - hence its usual name, the 'Principle of Inclusion-Exclusion' or 'PIE' for short [48]. The PIE can be seen as a formulaic distillation of the argument Legendre used. Thus, when we use the PIE to find $N(1, x, p_1, p_2, \ldots, p(\sqrt{x}))$, the number of elements that pass through the sieve is equal to the number of elements passing through the traditional sieve of Eratosthenes. Example 1.4.0.1 discusses an implementation of the PIE for $N(1, 37, p_1, p_2, \ldots, p_r)$, where the r-th sifting prime is p_r, which equals in this case $p(\sqrt{37}))$, the largest prime less than or equal to $\sqrt{37}$, i.e., 5.

Suppose for a moment that we have in hand a formula for the PIE and that we've used it to calculate $N(1, x, p_1, p_2, \ldots, p_r)$ for specific values of x and $p_1, \ldots, p_r$. That is all well and good. But, *just because we have a closed formula for calculating $N(1, x, p_1, p_2, \ldots, p_r)$ in particular cases does not mean we know or can easily determine the size of $N(1, x, p_1, p_2, \ldots, p_r)$ in all cases that may be of interest.* In fact, according to Tenenbaum and Mèndes France,

> Brun's motivation was to make Legendre's formula for the sieve of Eratosthenes usable, for in its basic form it contains too many terms to permit manageable calculations ([52], p. 25).

The reason we want to know the size of $N(1, x, p_1, p_2, \ldots, p_r)$ is to make sure that for all cases under consideration (i.e., for sufficiently large even numbers) the size of $N(1, x, p_1, p_2, \ldots, p_r)$ is at least large enough to allow some of the "right kind of numbers" to fall through the sieve, where by the "right kind of numbers" we mean pairs of numbers that have at most nine prime factors and whose sum is an even number x.

There are at least two relevant ways of estimating the size of $N(1, x, p_1, p_2, \ldots, p_r)$: through asymptotics (which deal with the eventual size of $N(1, x, p_1, p_2, \ldots, p_r)$ as x becomes large) and through bounds that guarantee that enough of the right kind of numbers pass through his sieve. Brun obtains his results by a skillful use of both these ways of estimating $N(1, x, p_1, p_2, \ldots, p_r)$.

1.4 The Legendre Formula

The *Legendre formula* (1.4.0.2) is a method that Legendre used in [34] to calculate the number of elements passing through a sieve, though he did not use sieve terminology per se. Legendre's method was later recognized to be essentially the Principle of Inclusion-Exclusion (PIE), and so the Legendre formula as we know it today *is* just an implementation of the PIE that allows us to determine the number of elements that pass through a sieve. Next, we show how to use the Legendre formula on the sequence $1, 2, 3, \ldots, 37$.

Example 1.4.0.1. *Use the Legendre formula to find* $N(1, 37, p_1, p_2, \ldots, p(\sqrt{37}))$*, where* $p(\sqrt{37}) = 5$*, the largest prime less than or equal to* $\sqrt{37}$*.*

This example is the same as Example 1.2.0.1, but now we focus on counting the *number* of elements that fall through the sieve, and we pay attention to the *number of times* certain composite numbers are "crossed out" in the sifting process.

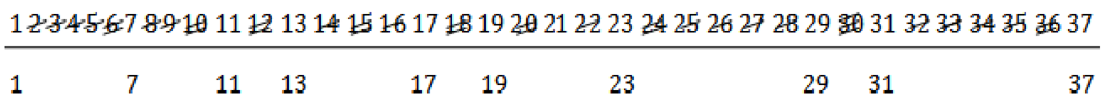


FIGURE 1.4.0.1: Finding $N(1, 37, 2, 3, 5)$

Each strike-through in the top row of numbers in Figure 1.4.0.1 tracks one step of our running count. To count the number of multiples of a given number d that are in the sequence $1, 2, 3, 4, \ldots, 37$, we divide 37 by d and then take the greatest integer part of the result, denoted $\left\lfloor \frac{37}{d} \right\rfloor$. We start our running count at 37, because at this step we trap no numbers, so they all fall through. Next we subtract from 37 the number of all the multiples of 2, 3, and 5 that are in the top row of Figure 1.4.0.1. Since this last steps overcounts the number of times we trap multiples of 2, 3, and 5, we then adjust the running count by adding back the appropriate multiples of pairs of primes less than $\sqrt{37}$. Finally, since at this point we have added to our count the product $2 \cdot 3 \cdot 5$ one time too many, we subtract it back out.

$$
\begin{aligned}
N(1, 37, p_1, p_2, \ldots, p(\sqrt{37})) = \lfloor 37 \rfloor & \\
& - \left\lfloor \frac{37}{2} \right\rfloor - \left\lfloor \frac{37}{3} \right\rfloor - \left\lfloor \frac{37}{5} \right\rfloor \\
& + \left\lfloor \frac{37}{2 \cdot 3} \right\rfloor + \left\lfloor \frac{37}{2 \cdot 5} \right\rfloor + \left\lfloor \frac{37}{3 \cdot 5} \right\rfloor \\
& - \left\lfloor \frac{37}{2 \cdot 3 \cdot 5} \right\rfloor \\
= 10 &
\end{aligned} \tag{1.4.0.1}
$$

In general, the pattern of the Legendre formula for $N(1, x, p_1, p_2, \ldots, p_r)$, where $p_r = p(\sqrt{x})$ is the largest prime less than or equal to $\sqrt{x}$, is:

$$
\begin{aligned}
N(1, x, p_1, p_2, \ldots, p_r) = \lfloor x \rfloor & \\
& - \sum_{p \leq \sqrt{x}} \left\lfloor \frac{x}{p} \right\rfloor \\
& + \sum_{p_1 < p_2 \leq \sqrt{x}} \sum \left\lfloor \frac{x}{p_1 p_2} \right\rfloor \\
& - \sum_{p_1 < p_2 < p_3 \leq \sqrt{x}} \sum \sum \left\lfloor \frac{x}{p_1 p_2 p_3} \right\rfloor \\
& + \ldots + (-1)^r \sum_{p_1 < \ldots < p_r \leq \sqrt{x}} \sum \cdots \sum \left\lfloor \frac{x}{p_1 p_2 \ldots p_r} \right\rfloor
\end{aligned} \tag{1.4.0.2}
$$

With two additional notations, we can represent the Legendre formula differently, in a way that will provide certain insights into the value of $N(1, x, p_1, p_2, \ldots, p_r)$ (see Section 1.4.1)[2].

First, we combine the different summations in 1.4.0.2 into just one summation. We do this by first setting $P = \prod_{p \leq \sqrt{x}} p$, the product of all the sifting primes. Then, a number d divides P iff $d = 1$, or d is prime, or d is the product of unique primes in P. With this

[2] I follow fairly closely the presentation of [43], though I elaborate on some details omitted there.

new notation 1.4.0.2 becomes

$$N(1, x, p_1, p_2, \ldots, p_r) = \sum_{d|P} (-1)^k \left\lfloor \frac{x}{d} \right\rfloor, \tag{1.4.0.3}$$

where r equals the number of primes less than or equal to $\sqrt{x}$, and k equals the number of distinct prime factors in d.

Second, we consolidate the alternating $+$ and $-$ signs by defining a function that keeps track of these signs. As it happens, there is a well-known function in number theory that does precisely what we want: the Möbius μ function, defined in 1.4.0.4 below. This function has deep connections with other areas of number theory (see, for example [31]), but we consider it here simply as a way to express the Legendre formula in a form that allows for more detailed analysis:

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{if } a^2|n \text{ for some } a > 1 \\ (-1)^k & \text{if } n \text{ has } k \text{ distinct prime factors.} \end{cases} \tag{1.4.0.4}$$

With the definitions of $P = \prod_{p \le \sqrt{x}} p$ and the Möbius function in hand, we can express the Legendre formula succinctly as

$$N(1, x, p_1, p_2, \ldots, p_r) = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor. \tag{1.4.0.5}$$

Next, we re-write 1.4.0.5 in order to show a major limitation of the sieve of Eratosthenes-Legendre as it is implemented via the Legendre formula.

1.4.1 Chief Limitation of the Legendre Formula

The chief limitation of the Legendre formula for Brun's purposes is that without further modification it yields an "error term" that is too large, which in turn makes difficult the estimation of a lower bound for $N(1, x, p_1, p_2, \ldots, p_r)$, the number of elements passing through a sieve. (Recall that Brun wants a lower bound estimate for $N(1, x, p_1, p_2, \ldots, p_r)$ so that he can be sure enough of the "right kind" of numbers pass through his sieve, where the "right kind" of numbers are such that they contain at most nine prime factors and there are pairs of them whose sum is an even number.)

What do we mean by the "error term" in the Legendre formula, why is it too large, and how does a large error term cause problems in finding adequate lower-bound estimates for $N(1, x, p_1, p_2, \ldots, p_r)$? We answer these questions in turn.

1.4.1.1 What is the Error Term in the Legendre Formula?

The error term in the Legendre formula is the second term on the right-hand side of 1.4.1.2 below, namely, $\sum_{d|P} \mu(d) \left(\left\lfloor \frac{x}{d} \right\rfloor - \frac{x}{d} \right)$.

Recall the succinct form of the Legendre function given in 1.4.0.5:

$$N(1, x, p_1, p_2, \ldots, p_r) = \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor. \tag{1.4.1.1}$$

We can re-write the integer part of $\frac{x}{d}$ as the quotient minus the remainder term

$$\left\lfloor \frac{x}{d} \right\rfloor = \frac{x}{d} - \left(\frac{x}{d} - \left\lfloor \frac{x}{d} \right\rfloor \right),$$

which on substituting into 1.4.1.1 and rearranging terms, yields

$$N(1, x, p_1, p_2, \ldots, p_r) = x \sum_{d|P} \frac{\mu(d)}{d} + \sum_{d|P} \mu(d) \left(\left\lfloor \frac{x}{d} \right\rfloor - \frac{x}{d} \right) \tag{1.4.1.2}$$

since

$$\begin{aligned} N(1, x, p_1, p_2, \ldots, p_r) &= \sum_{d|P} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \\ &= \sum_{d|P} \mu(d) \left(\frac{x}{d} - \left(\frac{x}{d} - \left\lfloor \frac{x}{d} \right\rfloor \right) \right) \\ &= \sum_{d|P} \mu(d) \frac{x}{d} - \sum_{d|P} \mu(d) \left(\frac{x}{d} - \left\lfloor \frac{x}{d} \right\rfloor \right) \\ &= x \sum_{d|P} \frac{\mu(d)}{d} + \sum_{d|P} \mu(d) \left(\left\lfloor \frac{x}{d} \right\rfloor - \frac{x}{d} \right). \end{aligned} \tag{1.4.1.3}$$

There are at least two ways of understanding why the error term has the name that it does. First, it describes the amount left over after dividing P by the numbers d. Second, the term $x \sum_{d|P} \dfrac{\mu(d)}{d}$, often called the *main term*, contains the bulk of the expression in the last line of 1.4.1.3, because of the presence of x (which is assumed to be large) as a factor. One often considers using this main term to estimate the size of $N(1, x, p_1, p_2, \ldots, p_r)$, provided that what is left over (i.e., the error term) is not too large. As we shall see, the problem of a potentially large error term is a major limitation of the Legendre formula.

Brun manipulated the Legendre formula by removing terms from the main term and the error term. Thus, he was able to find a better (i.e., larger) lower bound for the value of $N(1, x, p_1, p_2, \ldots, p_r)$.

Because Brun modified the Legendre formula by limiting the primes that go into the Legendre formula, and because the Legendre formula can be expressed via the Möbius μ function, it can be and is often said that Brun achieved his results by "truncating the Möbius function."

1.4.1.2 What is Meant by Saying "the Error Term is Too Large"?

Saying "the error term is too large" means that as x grows large the error term could swamp the value of the main term in the Legendre expansion.

Explaining exactly why the error term could swamp the value of the main term requires a bit of development. We begin by developing the main term for Example 1.4.0.1, where we used the Legendre formula to find $N(1, 37, p_1, p_2, \ldots, p(\sqrt{37}))$, where $p(\sqrt{37}) = 5$.

Example 1.4.1.1. *Let $A = 1, 2, 3, \ldots 37$. Use the Legendre formula in the form of 1.4.1.2 to calculate $N(1, x, p_1, p_2, \ldots, p_r)$.*

We already know that $N(1, x, p_1, p_2, \ldots, p_r) = \#\{1, 7, 11, 13, 17, 19, 23, 29, 31, 37\} = 10$. Let's see how Equation 1.4.1.2 allows us to generate that result. The main term works out to be $9\frac{13}{15}$:

$$\begin{aligned}
x \sum_{d|P} \frac{\mu(d)}{d} &= 37 \sum_{d|2\cdot 3\cdot 5} \frac{\mu(d)}{d} \\
&= 37 \left(\frac{\mu(1)}{1} - \frac{\mu(2)}{2} - \frac{\mu(3)}{3} - \frac{\mu(5)}{5} \right. \\
&\qquad + \frac{\mu(2\cdot 3)}{2\cdot 3} + \frac{\mu(2\cdot 5)}{2\cdot 5} + \frac{\mu(3\cdot 5)}{3\cdot 5} \\
&\qquad \left. + \frac{\mu(2\cdot 3\cdot 5)}{2\cdot 3\cdot 5} \right) \\
&= 37 (1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} + \frac{1}{2\cdot 3} + \frac{1}{2\cdot 5} + \frac{1}{3\cdot 5} - \frac{1}{2\cdot 3\cdot 5}) \\
&= 9\frac{13}{15}.
\end{aligned} \tag{1.4.1.4}$$

It is important to point out that the sum $\sum_{d|2\cdot3\cdot5}\dfrac{\mu(d)}{d}$ in 1.4.1.4 can be written in two different but equivalent ways (1.4.1.5 and 1.4.1.6 below), both of which Brun regularly uses in his paper (see Figure 3.2.1.1 for one of these). The equivalence of these expressions is significant, because Brun uses one or the other to suit the purpose at hand, and it is through these expressions that we can see how Brun manipulates the terms in the Legendre formula to achieve his result on Goldbach's Conjecture. We have the following equivalent expressions for $\sum_{d|2\cdot3\cdot5}\dfrac{\mu(d)}{d}$:

$$\begin{aligned}\sum_{d|2\cdot3\cdot5}\frac{\mu(d)}{d} &= 1-\frac{1}{2}-\frac{1}{3}-\frac{1}{5}\\ &\quad+\frac{1}{3\cdot2}\\ &\quad+\frac{1}{5\cdot2}+\frac{1}{5\cdot3}(1-\frac{1}{2})\end{aligned} \tag{1.4.1.5}$$

and

$$\sum_{d|2\cdot3\cdot5}\frac{\mu(d)}{d} = (1-\frac{1}{2})(1-\frac{1}{3})(1-\frac{1}{5}). \tag{1.4.1.6}$$

The reader can easily verify the equivalence of these two expansions.

The error term works out to be, as it must, $\frac{2}{15}$:

$$\begin{aligned}\sum_{d|P} \mu(d)\left(\left\lfloor\frac{x}{d}\right\rfloor - \frac{x}{d}\right) &= \mu(1)\left(\left\lfloor\frac{37}{1}\right\rfloor - \frac{37}{1}\right)\\
&\quad + \mu(2)\left(\left\lfloor\frac{37}{2}\right\rfloor - \frac{37}{2}\right)\\
&\quad + \mu(3)\left(\left\lfloor\frac{37}{3}\right\rfloor - \frac{37}{3}\right)\\
&\quad + \mu(5)\left(\left\lfloor\frac{37}{5}\right\rfloor - \frac{37}{5}\right)\\
&\quad + \mu(2\cdot 3)\left(\left\lfloor\frac{37}{2\cdot 3}\right\rfloor - \frac{37}{2\cdot 3}\right)\\
&\quad + \mu(2\cdot 5)\left(\left\lfloor\frac{37}{2\cdot 5}\right\rfloor - \frac{37}{2\cdot 5}\right)\\
&\quad + \mu(3\cdot 5)\left(\left\lfloor\frac{37}{3\cdot 5}\right\rfloor - \frac{37}{3\cdot 5}\right)\\
&\quad + \mu(2\cdot 3\cdot 5)\left(\left\lfloor\frac{37}{2\cdot 3\cdot 5}\right\rfloor - \frac{37}{2\cdot 3\cdot 5}\right)\\
&= 0 - 1(-\frac{1}{2}) - 1(-\frac{1}{3}) - 1(-\frac{2}{5})\\
&\quad + 1(-\frac{1}{6}) + 1(-\frac{7}{10}) + 1(-\frac{7}{15})\\
&\quad - 1(-\frac{7}{30})\\
&= \frac{1}{2} + \frac{1}{3} + \frac{2}{5} - \frac{1}{6} - \frac{7}{10} + \frac{1}{30}\\
&= \frac{2}{15}.
\end{aligned} \tag{1.4.1.7}$$

Although in this example we are able to calculate both the main term and error term exactly, the issue is that as x becomes large, the magnitude of the error term could dwarf the magnitude of the main term. Why? Because the error term will include 2^r terms (where r is the number of sifting primes), each of which is bounded in absolute value by 1. In other words, the bound on the error is exponential in the number of sifting primes used in the sieve[3].

Thus, even though we can't say much at this point that is useful about a lower bound for N, we do know this much:

$$N(1, x, p_1, p_2, \ldots, p_r) > x \sum_{d|p_1 p_2 \ldots p_r} \frac{\mu(d)}{d} - 2^r. \tag{1.4.1.8}$$

[3]This fact suggests that one way to reduce the error term is to use fewer sifting primes than are used in the traditional sieve of Eratosthenes. Indeed, Brun does reduce the number of sifting primes (by choosing a smaller maximum sifting prime), though at the cost of allowing some composite numbers through the sieve (see Section 1.8).

We'll call the right side of this inequality the *default, or "worst" lower bound* for N based on the Legendre formula.

In the above example, notice that there are eight terms in the sum involved in the error term. Each of these eight terms has a maximum value bounded in absolute value by 1, because each of these terms is a remainder of the form $(\lfloor \frac{x}{d} \rfloor - \frac{x}{d})$. For smaller values of x, such as $x = 37$ in example 1.4.1.1, none of these terms is very near its theoretical maximum value, and so the error term itself $(\frac{2}{15})$ is smaller than the theoretical maximum bound of 2^3. However, certain theoretical results discussed in Section 1.4.1.3 show that as x becomes large, the error term is of the same order of magnitude as the main term, which makes estimating the value of $N(1, x, p_1, p_2, \ldots, p_r)$ problematic.

1.4.1.3 Why is a Large Error Term Problematic?

A large error term is problematic because its value hinders our ability to estimate the value of $N(1, x, p_1, p_2, \ldots, p_r)$.

To see why a large error term causes problems in estimating lower bounds for $N(1, x, p_1, p_2, \ldots, p_r)$, consider the alternate formulation of the Legendre formula, alluded to above, which is given by substituting the equivalent value $\prod_{p|P}(1 - 1/p)$ for the main term $\sum_{d|P} \dfrac{\mu(d)}{d}$. [4] Thus we obtain from 1.4.1.2

$$N(1, x, p_1, p_2, \ldots, p_r) = x \prod_{p|P}(1 - 1/p) + \sum_{d|P} \mu(d) \left(\left\lfloor \frac{x}{d} \right\rfloor - \frac{x}{d} \right) \tag{1.4.1.9}$$

This form of the Legendre formula is seen often in the modern literature on sieves, and the particular form of the main term is also one that Brun uses in §3 and §4 of his paper, which are discussed in more detail in Chapter 3.

As we saw earlier, the value of $N(1, x, p_1, p_2, \ldots, p(\sqrt{x}))$ equals the number of primes between $\sqrt{x}$ and x, plus one. The standard notation for the number of primes

[4]To see these two expressions are equivalent, consider the terms in the expansion of $\prod_{p|P}(1 - 1/p)$. Specifically, consider which terms are in the expansion of $\prod_{p|P}(1 - 1/p)$, how many terms of each kind there are, and the sign of each term. In the expansion of the product there is only one 1 and it has a positive sign, corresponding to $\mu(1)$. There are $\binom{P}{1} = P$ terms of the form $\frac{1}{p}$, each with a negative sign, yielding $(-1)^1 \sum_{p|P} \frac{1}{p}$, corresponding to $\sum_p \frac{\mu(p)}{p}$ for individual primes p. Similarly, there are $\binom{P}{r}$ terms of the form $\frac{1}{p_1 p_2 \cdots p_r}$, which are counted by the summation and given the appropriate sign (negative or positive according to whether r is odd or even) by the μ function.

less than or equal to a positive number x is $\pi(x)$. So, $N(1, x, p_1, p_2, \ldots, p_r) = \pi(x) - \pi(\sqrt{x}) + 1$. Thus, substituting $\pi(x) - \pi(\sqrt{x}) + 1$ into the left-hand side of Equation 1.4.1.9, we obtain

$$\pi(x) - \pi(\sqrt{x}) + 1 = x \prod_{p|P} (1 - 1/p) + \sum_{d|P} \mu(d) \left(\left\lfloor \frac{x}{d} \right\rfloor - \frac{x}{d} \right). \tag{1.4.1.10}$$

As mentioned above, each term in the sum that forms the error term is less than 1. Further, by the nature of the alternating signs of the terms in that sum, we can expect some reduction from the value of 2^r. However, as x increases the error term will still be too large to make the Legendre formula useful for computing $\pi(x) - \pi(\sqrt{x}) + 1$.

There are well-known asymptotic estimates for both $\pi(x) - \pi(\sqrt{x}) + 1$ and $\prod_{p|P}(1 - 1/p)$, which when taken together show that the error term $\sum_{d|P} \mu(d)(\lfloor \frac{x}{d} \rfloor - \frac{x}{d})$ has an asymptotic value proportional to the main term $\prod_{p|P}(1 - 1/p)$. The first estimate comes from the Prime Number Theorem [37], which shows that $\lim_{x\to\infty} \frac{\pi(x)}{\frac{x}{\log x}} = 1$ and hence (with some additional work) that $\pi(x) - \pi(\sqrt{x}) + 1 \sim \frac{x}{\log x}$. Additionally, a formula from Mertens [52] shows that $x \prod_{p|P}(1 - 1/p) \sim C \frac{x}{\log x}$, for a constant C. These two results taken together show that what we have been calling the error term is also asymptotic to a constant times $\frac{x}{\log x}$ and thus (asymptotically) of the same order of magnitude as the main term. For additional details, see [43].

1.4.2 Of What Use, then, is the Legendre Formula?

The above considerations show us that as x gets large, the error term is of an order of magnitude that should not be ignored or discounted when estimating the main term.

The previous discussion shows that the sieve of Eratosthenes, by itself, sheds no more light on the value of $N(1, x, p_1, p_2, \ldots, p_r)$ than does the Prime Number Theorem.

> At the moment we are in the rather depressing position of having a method which fails to give us good estimates for the number $\pi(x)$ of primes up to x, but worse yet, the only reason we even know that it must inevitably fail is because of other techniques, coming from analytic number theory, succeed (in proving the Prime Number Theorem), thereby telling us so. ([17], p. 3).

However, the Legendre formula nevertheless serves an essential purpose for Brun, because

- it is from the Legendre formula that the inequality for the default ("worst") lower bound is derived;
- the inequality for the "worst" lower bound has the form of the Fundamental Inequality $N > M - R$;
- Brun's two main sieves (in §3 and §4 of his paper) have initial lower bounds of the same form, $N > M - R$; and
- Brun's subsequent calculations for his two main sieves show that he bounds the main term M below and the remainder term R above, which results, taken together allow him to improve the lower bounds of each of his sieves (subject to the assumptions he makes and the additional results he uses to obtain his results).

Thus, the Legendre formula can be seen as the genesis of the default lower bound, which, having the same form ($N > M - R$) as Brun's initial lower bounds, sets the stage for Brun's subsequent calculations that improve the lower bounds for his two sieves in a way that is, at least conceptually, quite straightforward. So, far from being a roadblock, the Legendre formula opens up the possibility for Brun to create a path that, albeit full of switchbacks, eventually leads him to his result.

1.5 The "Best" Lower Bound for N

Recall from the discussion after 1.4.1.8 that the default or "worst" lower bound for N based on the Legendre formula is $N(1, x, p_1, p_2, \ldots, p_r) > x \sum_{d|p_1p_2\ldots p_r} \dfrac{\mu(d)}{d} - 2^r$.

It's natural to ask how or whether that lower bound can be improved. In particular, if we look at the terms in the expansion of the first term of that lower bound, a natural question is, "Can we manipulate terms in the expansion to find the *best* lower bound possible for $N(1, x, p_1, p_2, \ldots, p_r)$?"

To answer this question, consider the following example.

Example 1.5.0.1. *Find the default lower bound for* $N(1, 37, 2, 3, 5)$ *and then remove terms from the expansion formula to give an improved (increased) lower bound.*

The default lower bound is just the main term minus the worst-case bound on the error term (which equals 2^r, where r is the number of sifting primes used). This lower bound is shown using the first form of the equivalent main-term expansions in 1.4.1.5 and previous results.

$$\begin{aligned} N(1,37,2,3,5) > 37\Big(1 &- \frac{1}{2} - \frac{1}{3} - \frac{1}{5} \\ &+ \frac{1}{3\cdot 2} \\ &+ \frac{1}{5\cdot 2} + \frac{1}{5\cdot 3}(1-\frac{1}{2})\Big) - 2^3 \\ &= 9\frac{13}{15} - 8 \\ &= 1\frac{13}{15} \end{aligned} \tag{1.5.0.1}$$

To increase the difference $37\Big(1-\frac{1}{2}-\frac{1}{3}-\frac{1}{5}+\frac{1}{3\cdot 2}+\frac{1}{5\cdot 2}+\frac{1}{5\cdot 3}(1-\frac{1}{2})\Big)-8$ by removing positive terms from the outermost parentheses, we must decrease the first term, $37\Big(1-\frac{1}{2}-\frac{1}{3}-\frac{1}{5}+\frac{1}{3\cdot 2}+\frac{1}{5\cdot 2}+\frac{1}{5\cdot 3}(1-\frac{1}{2})\Big)$, by an amount less than the amount by which we decrease the second term, 8.

Thus we'll increase this difference if the terms we remove from the outermost parentheses are such that when multiplied by 37 they yield a value less than the number of terms removed (since that number is then subtracted from 8).

From the expansion in the above example, the only terms that can be removed to increase the lower bound for $N(1,37,2,3,5)$ are the two terms $\frac{1}{5\cdot 3}$ and $\frac{1}{5\cdot 3\cdot 2}$ in the tail end of the expansion of the main term, since only for those terms is 37 times the terms removed less than the number of terms removed. Specifically, $37(\frac{1}{5\cdot 3}(1-\frac{1}{2})) = \frac{37}{30} < 2$. Thus we reduce the main term by $1\frac{7}{30}$ and the error term by 2 (since there are now only 6 terms in the expansion), so the increased lower bound is now $N(1,37,2,3,5) > (9\frac{13}{15} - 1\frac{7}{30}) - (8-2) = 8\frac{19}{30} - 6 = 2\frac{19}{30}$.

The above example illustrates the method to find the best lower bound estimate for $N(1,x,p_1,\ldots,p_r)$ by removing terms from the expansion of $N(1,x,p_1,p_2,\ldots,p_r)$. Brun discusses similar examples in §2, where he removes terms from an expansion to increase the lower bound for $N(1,x,p_1,p_2,\ldots,p_r)$. However, his examples seem more complicated than they need to be to support the point he makes that in other examples he will compute the lower bound in a simpler way (p. 107, bottom).

Why wouldn't Brun continue to work with this method for obtaining "la meilleure limite inférieure" ([3], p. 9) for $N(1, x, p_1, p_2, \ldots, p_r)$, even though it may be somewhat complicated? One plausible answer is that, because as r (the number of sifting primes) grows large, each of 2^r terms would need to be checked with increasingly long calculations to see whether its removal would increase the lower bound.

1.6 Revisiting the Fundamental Inequality: $N > M - R$

Again, recall from 1.4.1.8 the default lower bound for N:

$$N(1, x, p_1, p_2, \ldots, p_r) > x \sum_{d|p_1 p_2 \ldots p_r} \frac{\mu(d)}{d} - 2^r. \tag{1.6.0.1}$$

Section 1.4.1.3 showed that this default lower bound is of limited direct use, since as x gets large, the error or remainder term in the Legendre expansion is of the same order of magnitude as the main term, thus rendering it difficult to estimate the value of N. Further, Section 1.5 showed that computing a "best" lower bound (by starting with the default lower bound and manipulating the main and remainder terms) is computationally prohibitive for large x.

At least two methods can be used to manipulate the right-hand side of the above inequality (or similar inequalities that Brun uses as points of departure for his calculations in §3 and §4 - see 3.3.2 and 3.4.2) to ensure that N is still greater than the manipulated right-hand side, i.e., that we still have a lower bound for N.

One method would subtract a positive amount from the right-hand side of 1.6.0.1. This could be done by dropping some positive terms from the expansion of the main term and then adjusting the remainder term so that the net result is either a decrease in the right-hand side of 1.6.0.1, or an increase that still yields a lower bound for N. But this approach is unsatisfactory for two reasons: 1) it would involve complex computations that would be prohibitive to carry out for a general case, and 2) if it *decreases* the lower bound for N, it moves the bound in the wrong direction: Brun wants to be sure that N is *large* enough, so that enough of the right kind of numbers fall through the sieve.

The second method is much simpler conceptually: just bound appropriately the main term and the error term of the right-hand side of 1.6.0.1. This can be done by rewriting 1.6.0.1 as $N > M - R$, and then bounding M below by some number, say, A and bounding R above by some number, say, B. The result will be $N > A - B$, showing that we still have a lower bound for N. Brun's complicated derivations in §3 and §4 can be cast in this simple conceptual scheme, as noted in 1.1.2.

Notice, however, that if we are given no further information, although we can be sure that the new inequality $N > A - B$ does still give a lower bound for N, we cannot know much about that new lower bound for N unless we incorporate other information. Brun does incorporate other information, particularly in §4 of his paper, to show that the lower bound thus achieved for N is large enough for his purposes.

1.7 Sieve Integrity when Truncating the Legendre Formula

An important unanswered question remains: How do we know that by dropping certain terms from the main term of the Legendre formula Brun is not in effect calculating results for a different sieve that would give results different from those of the original sieve that uses the non-truncated formula for $N(1, x, p_1, p_2, \ldots, p_r)$?

That is, why is it legitimate to truncate the Legendre formula (1.4.0.2) - which we know gives the true result for the number of elements falling through the sieve - and make claims based on the truncated Legendre formula about the results of the full formula (i.e., the results of the original sieve)? Aren't Brun's modifications to the Legendre formula effectively creating a different sieve that gives different results for $N(1, x, p_1, p_2, \ldots, p_r)$? We use a familiar example to answer these questions.

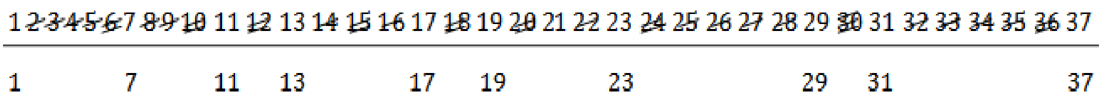


Figure 1.7.0.1: Finding $N(1, 37, 2, 3, 5)$

We can see in Figure 1.7.0.1 that the operation of the sieve (i.e., the trapping of multiples of 2, 3, and 5) is represented graphically by the number of strike-throughs of composite numbers. The sieve's operation can also be displayed arithmetically by the expansion of the Legendre formula. Indeed, the Legendre formula was introduced above

1.4.0.2 as a method of accounting to track the number of times composite numbers were trapped in the sieve.

$$\begin{aligned} N(1,37,2,3,5) >\Big(1 &- \frac{1}{2} - \frac{1}{3} - \frac{1}{5} \\ &+ \frac{1}{3\cdot 2} \\ &+ \frac{1}{5\cdot 2} + \frac{1}{5\cdot 3}(1-\frac{1}{2})\Big) - 2^3. \end{aligned} \tag{1.7.0.1}$$

Now consider the following observations. The number 30 is trapped in the sieve three times, once as a multiple of 2, once as a multiple of 3, and once as a multiple of 5; but *it needs to be trapped only once for the sieve to work correctly* (by not letting 30 fall through). The term $37(\frac{1}{5\cdot 3\cdot(-2)})$ is the term in the expansion that corresponds to trapping 30 three times. We can remove this term from from the expansion and still be faithful to the operation of the sieve, provided that we keep in the expansion some terms that correspond to the trapping of 2, 3, and 5.

Note that $37(\frac{1}{5\cdot 3}(1-\frac{1}{2}))$ deals not only with the trapping of 30 (by the component $-\frac{1}{5\cdot 3\cdot 2}$), but also with the trapping of 15 (by the the component $\frac{1}{5\cdot 3}$), and the trapping of 15 has already been reflected by including the terms $37(\frac{1}{5})$ and $37(\frac{1}{3})$. So, provided that those terms (or other terms that deal with the trapping of 3 and 5) remain in the expansion, we can remove the term $37\frac{1}{5\cdot 3}$ from the expansion and still be faithful to the operation of the sieve.

Not only can we remove $37(\frac{1}{5\cdot 3}(1-\frac{1}{2}))$ from the expansion of the main term, but by removing it we also achieve the “best” lower bound for N, as was shown in Example 1.5.0.1.

Further, because other terms that deal with the trapping of 30 and 15 remain in the formula, we are remaining faithful to the operation of the original sieve: those elements that were trapped in the original sieve are also trapped when we use the truncated formula; and, the truncated formula introduces no new trapped numbers, because it is a subformula of the original formula. Therefore, the improved lower bound achieved in the truncated formula is indeed a lower bound, which (in this case) happens to be the “best” lower bound (obtained by adjusting values in the main term and remainder terms of the default lower bound), for $N(1, x, p_1, p_2, \ldots, p_r)$, the number of elements that fall through the original sieve.

1.8 Why Some Composites Fall Through Brun's Sieves

Our work with $N(1, x, p_1, p_2, \ldots, p_r)$ so far has considered that our largest sifting prime is the largest prime less than or equal to $\sqrt{x}$. In that case, all the numbers except 1 that fall through the sieve are primes. However, it may be that with a smaller value of p_r some composite numbers also fall through the sieve.

Example 1.8.0.1. *Let the sequence to be sifted be the sequence of consecutive integers* $1, 2, 3, \ldots, 37$*. Sift this sequence by multiples of* 2 *and* 3*, but not also by multiples of* 5 *(unless they were already sifted out as multiples of* 2 *or* 3*). That is, compute* $N(1, 2, 3, \ldots, 37)$.

1 ~~2~~ ~~3~~ ~~4~~ ~~5~~ ~~6~~ 7 ~~8~~ ~~9~~ ~~10~~ 11 ~~12~~ 13 ~~14~~ ~~15~~ ~~16~~ 17 ~~18~~ 19 ~~20~~ 21 ~~22~~ 23 ~~24~~ 25 ~~26~~ ~~27~~ ~~28~~ 29 ~~30~~ 31 ~~32~~ ~~33~~ ~~34~~ 35 ~~36~~ 37

1 7 11 13 17 19 23 **25** 29 31 **35** 37

FIGURE 1.8.0.1: Finding $N(1, 37, 2, 3)$: the 2-primes 25 and 35 fall through

Observe that now, because 5 is not used as a sifting prime, any multiple of 5 in the given sequence that is not also a multiple of the sifting primes 2 or 3 will fall through the sieve. In this example, two such multiples of 5 are 25 and 35.

The properties of 25 and 35 that allow those to numbers to pass through the sieve can also be considered more generally. Consider a composite number c in the sequence of consecutive integers from 1 through x. Let that sequence be sifted by the sifting primes $p_1, p_2, \ldots, p_r$, where the largest sifting prime $p_r = p(x^{\frac{1}{k+1}})$, for some nonnegative integer k. If c has up to k (not necessarily distinct) prime factors, each of which is greater than $p(x^{\frac{1}{k+1}})$, then c will fall through the sieve. The reason is simply that the sifting primes $p_1, p_2, \ldots p_r$ "never reach" c, because none of c's prime factors is a multiple of a sifting prime.

A composite number with up to k (not necessarily distinct) prime factors is called a *k-prime.* Recall that Brun's theorem deals with 9-primes. The reason his sieve allows 9-primes to fall through is that his largest sifting prime is $p(x^{\frac{1}{10}})$.

1.9 Brun's Use of Merlin's Double Sieve

Another aspect of Brun's theorem is that the sum of certain pairs of numbers falling through his sieve is an even number. This section shows how the traditional sieve of Eratosthenes can be modified to create a sieve such that there are certain pairs of numbers falling through the sieve whose sum is an even number. First, we give a simple example of the way that Brun views the sifting process.

Example 1.9.0.1. *Let the sequence to be sifted be the sequence of consecutive integers* $1, 2, 3, \ldots, 37$*. Sift this sequence by multiples of* 2*,* 3*, and* 5 *to find* $N(1, 37, 2, 3, 5)$*.*

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
	2		4		6		8		10		12		14		16		18		20		22		24		26		28		30		32		34		36	
		3			6			9			12			15			18			21			24			27			30			33			36	
				5					10					15					20					25					30					35		
1						7				11		13				17		19				23						29		31						37

FIGURE 1.9.0.1: Brun's way of showing a sieve

Here Brun lists the sequence to be sifted on the top row. The bottom row shows the output of the sieve[5]. The rows in between show the composite numbers - multiples of 2, 3, and 5 - that are trapped in the sieve. With this setup, Brun asks, essentially, "How many numbers of the top row do not appear in any of the intermediate rows?" Specifically, he asks, "How many terms different from all the terms of the other lines does the first line contain?" (p. 101). That number is exactly the number of elements that fall through the sieve; those elements are shown in the bottom row.

Brun uses a similar way of picturing a sieve when he modifies his sieve so that some of the numbers that fall through his double sieve come in pairs that add up to an even number. Brun credits Merlin ([22]) with this idea, which he calls an "emploi double" or "double use" of the sieve of Eratosthenes. We'll refer to this kind of sieve simply as "Brun's double sieve." An example follows.

Example 1.9.0.2. *Let the sequence to be sifted be* $1, 2, 3, \ldots, 38$*. Use a double sieve, with sifting primes* 2*,* 3*, and* 5*, to sift this sequence.*

[5] Brun did not use a bottom row; it is used here to convey the sense of primes "falling through" the sieve.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38
0		2		4		6		8		10		12		14		16		18		20		22		24		26		28		30		32		34		36		38
0		2		4		6		8		10		12		14		16		18		20		22		24		26		28		30		32		34		36		38
0			3			6			9			12			15			18			21			24			27			30			33			36		
		2			5			8			11			14			17			20			23			26			29			32			35			
0					5					10					15					20					25					30					35			
			3					8					13					18					23					28					33					
	1						7												19												31						37	

FIGURE 1.9.0.2: Brun's way of showing his double sieve

The idea is that by sifting the sequence twice – once starting at 0 and moving from left to right, and once starting at 38 and moving from right to left – the sieve yields numbers such that certain pairs of them add up to 38. To see why this claim is true, we think of the first sieve as operating from left to right in the usual way, and the second operating from right to left by subtracting from 38 multiples of 2, 3, and 5.

Claim 1.9.1. In Example 1.9.0.2, if a number p passes through both sieves, then p is prime, $38 - p$ is prime, $\sqrt{38} \leq p < 38$, and $\sqrt{38} \leq 38 - p < 38$.

Proof. Suppose p passes through both sieves. Since p passes through the first sieve, p is either 1 or a prime in $[\sqrt{38}, 38)$, by the operation of the traditional sieve of Eratosthenes. Without loss of generality [6] assume that $p \neq 1$; hence, we suppose p is a prime that lies in $[\sqrt{38}, 38)$. It remains to show that $38 - p$ is prime and that it also lies in $[\sqrt{38}, 38)$.

Since p falls through both sieves, it falls through the second sieve. Therefore, p is *not* of the form $38 - 2k_1$, or $38 - 3k_2$, or $38 - 5k_3$, since those elements are exactly the ones trapped by the second sieve (assume sensible bounds on k_1, k_2, k_3). Therefore, $38 - p$ is not of the form $2k_1$, and $38 - p$ is not of the form $3k_2$, and $38 - p$ is not of the form $5k_3$ (again with sensible bounds on k_1, k_2, k_3). Thus, $38 - p$ is not a composite number that has 2, 3, or 5 as a factor. This means that $38 - p$ is a prime that lies in $[\sqrt{38}, 38)$. □

Next we consider a double sieve operating on a sequence of odd numbers (in effect, "pre-sifting" a sequence of consecutive integers by 2). Here is what that sieve looks like.

[6] The reason has to do with Brun's context of "sufficiently large even numbers." Although for a given even number, e.g., 36, the number 1 may fall through both sieves and be outside the range $[\sqrt{36}, 36)$, we can always find a larger even number (e.g., 40) for which the second-row sieve traps the number 1, thus preventing 1 from passing through both sieves.

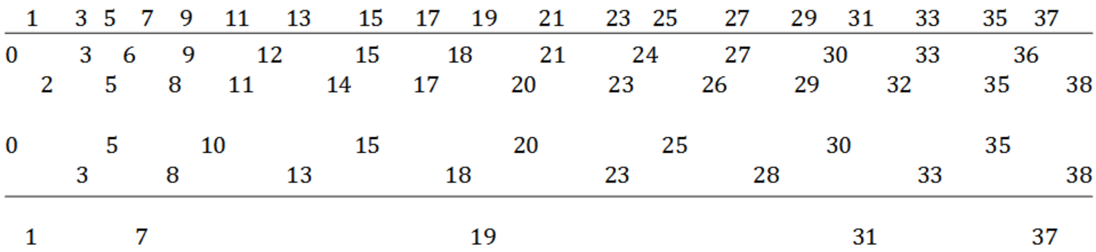


FIGURE 1.9.0.3: Double sieve starting with sequence of consecutive odd numbers

Naturally, the results are the same as in the previous example. But something more interesting happens when we sift the same odd sequence using only 3 as the sole sifting prime.

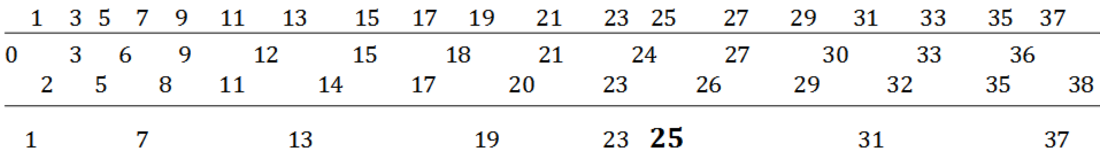


FIGURE 1.9.0.4: Double sieve allowing the 2-prime 25 to fall through

The results are still such that certain pairs of the numbers that fall through the sieve add up to 38, but now the 2-prime 25 is among them. This result follows by the same logic we discussed in the previous section, i.e., if the largest sifting prime is $p(x^{\frac{1}{k+1}})$, then some k-primes can pass through the sieve.

This last example, where the sifting sequence is a sequence of consecutive odd numbers and the largest sifting prime allows some k-primes to fall through the sieve is very much like an example that Brun uses on p. 127 in §6 of his paper, right before he begins to pull together his results to prove his famous theorem.

1.10 Contextual View of Brun's Approach

Based on the details of the discussion thus far, we are able to appreciate Brun's approach in a somewhat fuller context.

- We have seen how the traditional sieve of Eratosthenes can be used to find certain prime numbers if we know other smaller prime numbers.

- We have shown how the Legendre formula uses the Principle of Inclusion/Exclusion (PIE) to calculate $N(1, x, p_1, p_2, \ldots, p_r)$, where the sequence of consecutive integers $1, 2, 3, \ldots, x$ is sifted by the sifting primes $2, 3, 5, \ldots, p_r$, where $p_r = p(\sqrt{x})$ is the largest prime less than or equal to x.

- We re-formulated the Legendre formula as a main term plus an error term, showing that a certain "default" error term (2^r) was too large for the Legendre formula to be useful in estimating the size of $N(1, x, p_1, p_2, \ldots, p_r)$.

- We also argued that even if the default error term can be improved, it is still, as x grows large, of the same order of magnitude as the main term; hence, it will be of little help in estimating the size of $N(1, x, p_1, p_2, \ldots, p_r)$.

- We mentioned the importance to Brun of finding a suitable lower-bound estimate for the Legendre formula: he wants to make sure enough numbers "of the right kind" fall through his sieve, so that there are at least some pairs of numbers having at most nine prime factors each falling through his sieve whose sum is an even number.

- We gave a simple conceptual framework for Brun's work based on the Fundamental Inequality $N > M - R$: Brun essentially just bounds M below and R above to achieve his improved lower bounds for N.

- We mentioned that Brun's manipulations demonstrate that not only do his two sieves in §3 and §4 provide lower bounds for N, but they also give definite improvements that work under certain conditions.

- We saw that when the largest sifting prime is reduced from $p(x^{\frac{1}{2}})$, as in the traditional sieve of Eratosthenes, to $p(x^{\frac{1}{k+1}})$, where $k > 1$, certain composite numbers (that have at most k prime factors) can also fall through the sieve.

- We showed how Brun used ideas from Merlin ([22]) to create a double sieve with pairs of numbers falling through whose sum is an even number.

It turns out that Brun investigates three different methods for determining a lower bound for $N(1, x, p_1, p_2, \ldots, p_r)$, each of which is an improvement on the default ("worst") lower bound (which has an error or remainder term of 2^r). All three methods involve discarding particular terms from both the main term and the error term of the Legendre formula.

The first method, which gives “la meilleure limite inférieure” ([3], p. 9), discards terms using a particular process that calculates certain properties of the discarded terms. This lower bound was discussed in the example in Section 1.5. See also Section 3.2.1.1 below. Brun says that later in his paper he wants to select the discarded terms in simpler ways.

The second method, which discards terms “in small print” (this mysterious phrase is explained in Chapter 3) specifies a way of truncating the main term of the Legendre formula at a certain point of its expansion. Brun shows that by this method he can reduce the error term in the Legendre formula from 2^r to r^{m+1}, where m is an odd number less than r. But, he says without explanation, the growth of the error term is still “too great for our purpose” ([44], p. 114). Chapter 3 of this thesis gives one reason why this improved error term is still problematic.

The third method that Brun uses to discard terms is to remove certain terms “on the right of the vertical lines” ([44], p. 110); this cryptic phrase is also explained in Chapter 3. This method is similar to the second method, but in addition to truncating the main term of the Legendre formula at a certain point of its expansion, he also eliminates certain terms from the terms that remain after the initial truncation. In other words, his third method does what his second method does, and then some.

In his third method Brun also uses certain asymptotic results (theorems from Stirling and Mertens) to show how his estimates of a lower bound for his modified Legendre expansion behave for large enough x.

Brun develops these latter two methods for a *single* sieve, i.e., a sieve that is not a double sieve. Later, in §6 of his paper he applies his previous results, with the necessary modifications, to the case of a double sieve, finally achieving his theorem about Goldbach’s Conjecture.

CHAPTER 2

GOLDBACH'S CONJECTURE AND RELATED RESULTS

2.1 Goldbach's Conjecture

The first recorded mention of Goldbach's Conjecture is evidently the statement by Euler in his letter to Goldbach dated June 30, 1742. Euler says, possibly in reply to a remark by Goldbach, "That . . . every even number is the sum of two primes I feel is a quite certain theorem, although I cannot demonstrate it" ([18], my translation). For Euler and Goldbach this conjecture would have applied to the even number 2, since in those days the number 1 was commonly considered to be a prime. It would also have applied to the even number 4, though in many modern formulations of the conjecture, 4 is omitted for the sake of simplicity, with the result that the conjecture is often stated as "every even number greater than 4 is the sum of two odd primes."

It is difficult to say with any certainty when this conjecture began to be generally noticed by mathematicians. It was evidently first published by Waring - without attribution, comment, or proof - in his *Meditationes Algebraicae* in 1770 (see p. 217 of [54]). The relevant correspondence of Goldbach and Euler was published later, in the mid-nineteenth century [18]. Although Goldbach's Conjecture was known to mathematicians in the nineteenth century, the following questions by Poincaré indicate that as late as 1894 the source of the conjecture and the existing support for it were not common knowledge among mathematicians: "Where has Goldbach published his famous empirical theorem: every even number is the sum of two primes? ... What confirmations have been found for it?"(cited in [8], my translation).

Since the time early work began on Goldbach's Conjecture, simple calculations have been part of the historical accumulation of evidence in support of Goldbach's Conjecture [14]. But other kinds of arguments - based on theoretical results, related conjectures, or heuristic considerations - have been employed as well. Some of these other arguments are described in the next section. In his 1920 paper, Brun used arguments based both on counting techniques and on theoretical considerations to establish his result on Goldbach's Conjecture.

Several recent results are "close" to Goldbach's Conjecture, one of which has apparently proved what is known as "Golbach's Ternary Conjecture"': that every odd number greater than 5 is the sum of three primes [27][1]. Additionally, in [13] Goldbach's Conjecture is verified up to $4 \cdot 10^{18}$. A strong theoretical result remarkably "close" to Goldbach's Conjecture was published in 1973 by Chen and showed that every sufficiently large even number is the sum of a prime and a number with at most 2 prime factors [6].

2.2 Some Evidence that Goldbach's Conjecture is True

2.2.1 Empirical Evidence

One of the first things you may observe if you set out to test Goldbach's Conjecture is that even numbers greater than 4 do indeed seem to be expressible as the sum of two primes. What is more, the even numbers greater than 20 seem to be expressible as the sum of two primes *in more than one way*. For instance, $22 = 11 + 11 = 19 + 3$.

Illustrations of this observation are given in Figures 2.2.1.1 and 2.2.1.2, where n on the horizontal access denotes an even number and $r(n)$ on the vertical access denotes the number of distinct representations of n as the sum of two primes. (Two representations like $10 = 3 + 7$ and $10 = 7 + 3$ are not considered distinct and so are counted as a single representation.) The graphs in Figures 2.2.1.1, 2.2.1.2, 2.2.2.3, and 2.2.2.4 were created by the author using SageMath [11], after generating the needed primes with primegen [12].

The above observation is not new. It was made at least as far back as 1855, and in 1896 Sylvester [51] appears to credit Euler himself with the observation. In 1855, Desboves asserted that he verified this observation up to $10,000$ [10]. In one sense, this result seems to have come early in the set of results that support Goldbach's Conjecture, since most of the progress in addressing Goldbach's Conjecture has come in the twentieth century. In another sense, though, when one considers that the work of Desboves was published more than a century after Goldbach's Conjecture was formulated by Euler, it seems to arrive late.

[1] This has been published in the ArXiv (arXiv:1501.05438 [math.NT]).

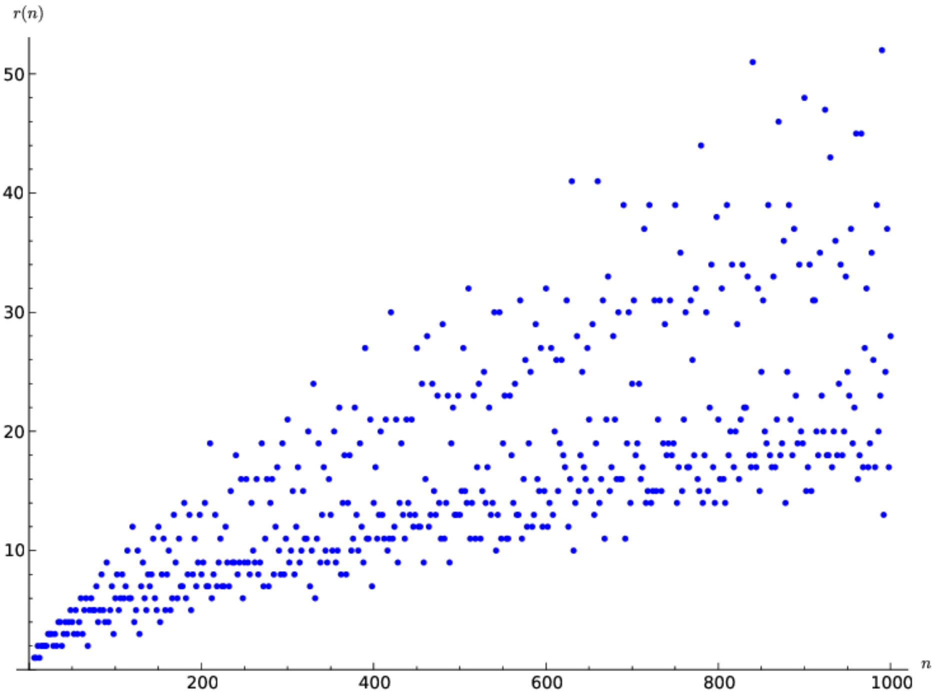


FIGURE 2.2.1.1: $r(n)$: number of distinct representations of n as sum of two odd primes

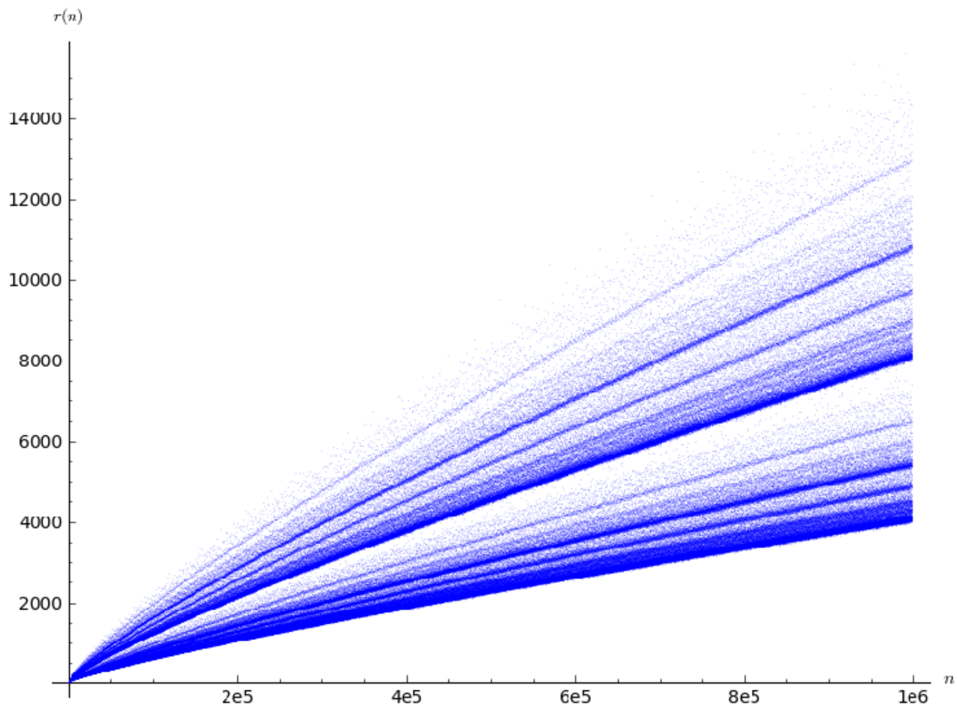


FIGURE 2.2.1.2: As n increases, $r(n)$ reveals its signature pattern

But considering that Goldbach's Conjecture evidently did not find its way into print until Waring published it in 1770 (without attribution, proof, or comment) in his *Meditationes Algebraicae* (see p. 217 of [54]), it is understandably difficult to assess how much empirical work was done on Goldbach's Conjecture in the hundred years or so after it first appeared in Euler's letter. Indeed, as mentioned above, as late as 1894 Poincaré was asking where Goldbach's Conjecture had been published [8].

The pattern shown in Figures 2.2.1.1 and 2.2.1.2, which is part of the pattern of an infinite sequence of numbers, has been called 'Goldbach's Comet' and is further described in entry A002375 from *The On-Line Encyclopedia of Integer Sequences* [47].

Two nineteenth-century works that provide empirical support of Goldbach's Conjecture are those by Desboves in 1855 [10] (already mentioned) and Cantor in 1894 [5], whose table gave the decompositions of even numbers up to 1000 as the sum of two primes. Other empirical work, along with some theoretical work, was done by Haufsner in 1892 ([25]) and Haussner in 1897 ([26]).

Recent empirical support for Goldbach's Conjecture comes from a number of computational studies that establish that all even numbers up to a certain point can be expressed as the sum of two primes. Most recently (as of 2014) the largest number for which Goldbach's Conjecture has been computationally verified is $4 \cdot 10^{18}$ [13].

Although the graphs in Figures 2.2.1.1 and 2.2.1.2 show a pattern that seems to hold for all large even n, and although Goldbach's Conjecture has been verified for even numbers greater than a billion billion, such evidence does not a proof make. Indeed, particularly suggestive patterns dealing with prime numbers have been known to break down as the numbers under consideration become very large (far larger than a mere billion billion). For discussion of a famous counterexample, see Derbyshire's account of the Skewes number in [9].

Because simple tests of Goldbach's Conjecture for small even numbers are easy to carry out by hand, it is quite plausible that the first investigations into Goldbach's Conjecture were done empirically, rather than theoretically. Indeed, Goldbach himself made simple empirical calculations to support a related conjecture he made to Euler, namely, that every integer that can be written as the sum of two primes can also be written as the sum of as many primes as one wishes, until all terms are units [18].

Arguments based on counting also played a fundamental role in asserting another famous conjecture about prime numbers, to which we now turn.

2.2.2 The Prime Number Theorem and Heuristics

In the 1790s, both Legendre and Gauss used counting arguments to support their versions of a conjecture, which was not proved until 1896 and is now called the Prime Number Theorem [37].

The Prime Number Theorem (PNT) asserts an asymptotic value for the number of primes less than or equal to a positive number x, typically denoted $\pi(x)$. In one formulation (there are several), the Prime Number Theorem says that $\lim_{x\to\infty} \dfrac{\pi(x)}{\frac{x}{\log x}} = 1$. In other words, the larger x gets, the closer $\pi(x)$ is to $\frac{x}{\log x}$.

Thus, the PNT gives an asymptotic approximation for $\pi(x)$, which for particular finite values of x can also be used as an estimate of $\pi(x)$, the number of primes less than or equal to x. But the PNT by itself does not give us specific details about the nature of the primes or how they're distributed. So, the PNT seemingly could offer little or no support for Goldbach's Conjecture.

For example, consider Figure 2.2.2.1, which depicts the primes from 3 to 50 and also shows how certain pairs of those primes align to give a sum of 50, indicated by the lines connecting the primes in the top and bottom rows. Shorter lines from the top or bottom row indicate primes in one row or the other, but not both.

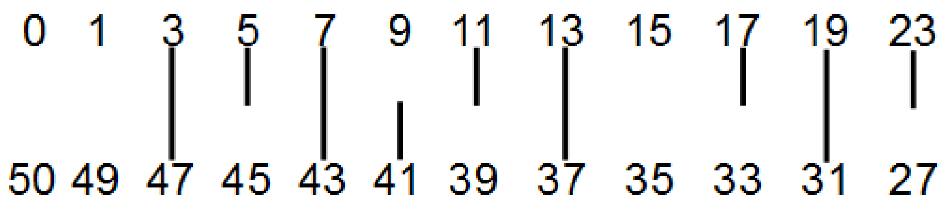


Figure 2.2.2.1: Consider $r(50)$: with "matching pairs of primes" connected

From Figure 2.2.2.1 we can see that in order for two primes to sum to an even number n there must be at least one prime between $n/2$ and n, because if there were no prime between those numbers (25 and 50 in Figure 2.2.2.1), there would be no way to find a pair of primes that add up to n. (Fortunately, a result known as Bertrand's Postulate, proved by Chebychev (and later by Ramanujan and Erdös among others), assures us that this is the case). But, in order for Goldbach's Conjecture to hold, not only does there need to be at least one prime p between $n/2$ and n, this prime also needs to be a 'matching' prime for some prime between between 3 and $n/2$. In terms of Figure 2.2.2.1, what this means is that not only do we need to have a prime in the bottom half of the diagram, but we also need a 'matching' prime in the top half of the diagram such that the sum of those two primes is 50.

Because the Prime Number Theorem doesn't address the distribution of primes with the kind of detail just discussed, it seems that it would have little application to Goldbach's Conjecture. Yet, it does.

In particular, in 1871 Sylvester used certain assertions related to the Prime Number Theorem to offer a theoretical argument, supplemented with heuristic notions from probability, to highlight the role of the ratio $\frac{x}{\log x}$ in the context of Goldbach's Conjecture [50]. Actually, the ratio mentioned in the Sylvester paper was $\frac{x}{(\log x)^2}$, which is just the product of $\frac{1}{\log x}$ and $\frac{x}{\log x}$, the latter ratio being, by the PNT, the approximate number of primes less than or equal to x.

Sylvester's work was mentioned by Hardy and Littlewood in a paper from 1923 [24], where they gave a conjecture (called 'Conjecture A') for the asymptotic value of the number of representations of an even number as the sum of two primes (what we call $r(n)$ in Figure 2.2.2.2).

The asymptotic estimate for $r(n)$ given by Hardy and Littlewood in 'Conjecture A' contains three factors: an initial constant, a term involving $\frac{n}{(\log n)^2}$, (with its echoes of the Prime Number Theorem), and a product term credited to Sylvester, which according to Hardy and Littlewood in some manner accounts for the irregularity in the distribution of primes.

To see the significance of the ratio $\frac{n}{(\log n)^2}$ in the calculation of $r(n)$, consider Figures 2.2.2.2 and 2.2.2.3, where $\frac{n}{(\log n)^2}$ is plotted in red against the dusty blue bands of $r(n)$.

Notice that the plot of $\frac{n}{(\log n)^2}$ tracks the arc of the data quite well, though we can see that it is too high to provide a lower bound for $r(n)$. (An easily conjectured lower bound for $r(n)$ is $\frac{\sqrt{2}}{2}\frac{\mathbf{n}}{(\mathbf{\log n})^2} - 4$, which Figure 2.2.2.4 illustrates for even numbers up to $1,000,000$).

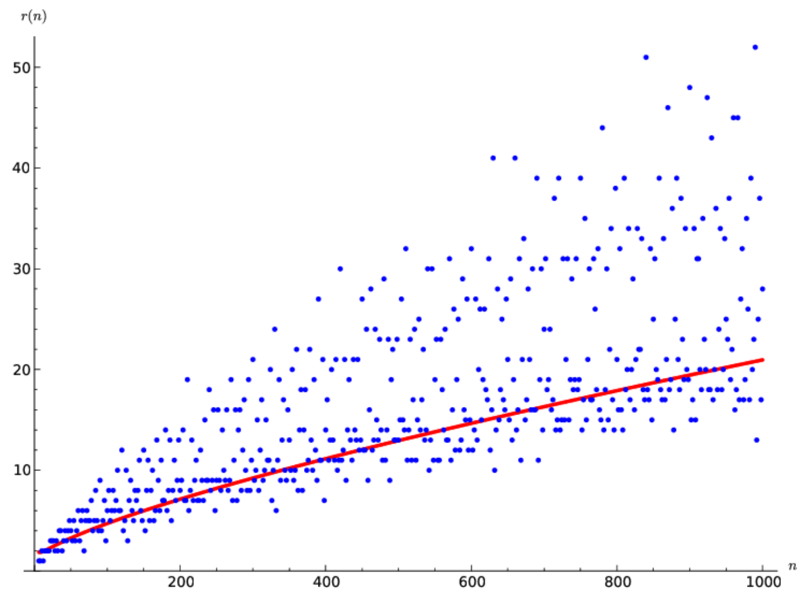


FIGURE 2.2.2.2: Red line $\frac{n}{(\log n)^2}$ tracks the data but is too high for a lower bound

The discussion in this section thus suggests that theoretical considerations based in part on the Prime Number Theorem can be combined with empirical data to create a compelling case for the truth of Goldbach's Conjecture.

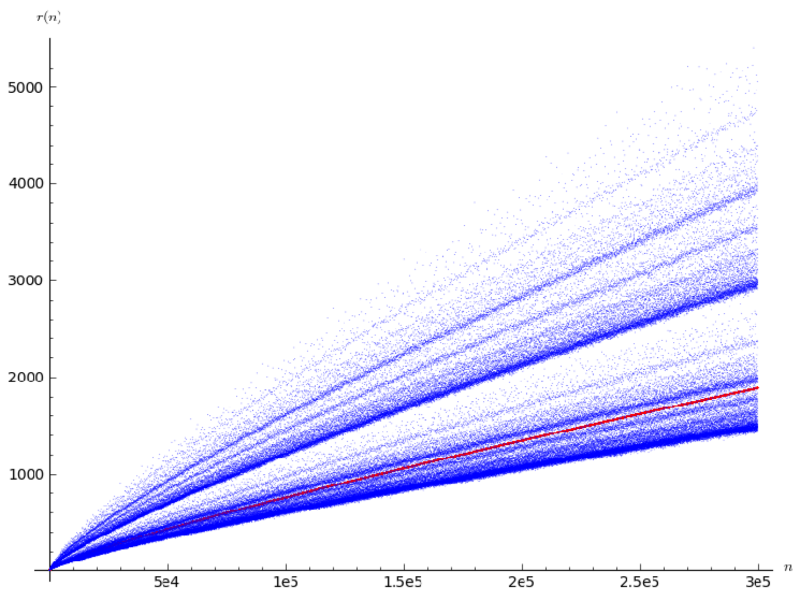


FIGURE 2.2.2.3: Red line $\frac{n}{(\log n)^2}$ tracks the data but is too high for a lower bound

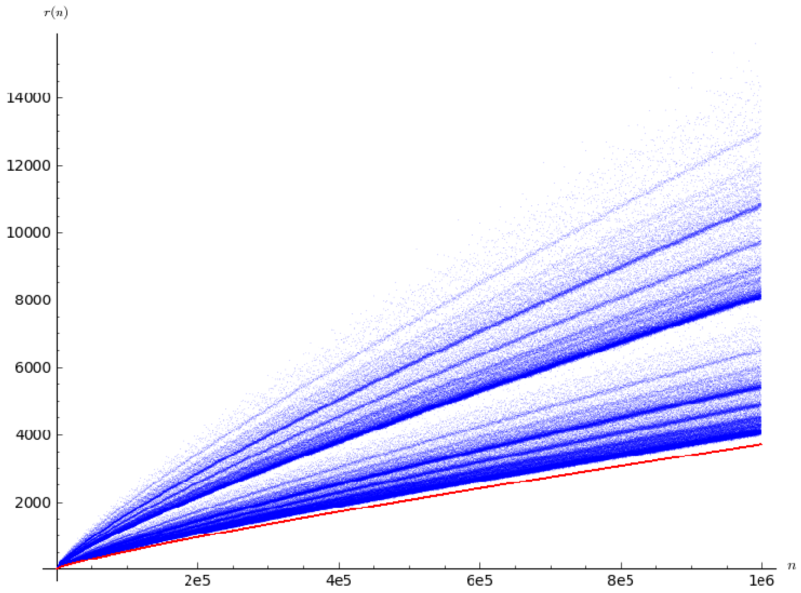


FIGURE 2.2.2.4: Red line $\frac{\sqrt{2}}{2}\frac{\mathbf{n}}{(\mathbf{\log n})^{\mathbf{2}}} - 4$, a plausible lower bound for r(n)

2.2.3 Sieve Methods

As mentioned in Chapter 1, sieve methods in number theory are operations on sequences of integers that yield a subsequence or an estimate of the size of a subsequence of the original sequence. Other writers have explained the purpose of sieves: "The basic purpose for which the sieve was invented was the successful estimation of the number of primes in interesting integer sequences" [16], and "The aim of sieve theory is to construct estimates for the number of integers remaining in a set after members of certain arithmetic progressions have been discarded" [36].

Sieves in mathematics also have applications to areas other than the theory of prime numbers "ranging from a topic as old as squarefree numbers to one as new as quantum ergodicity" ([17], p. 305). For examples of these other applications the interested reader is referred to [17] and [32].

Number-theoretic sieves were also used recently to establish three of the "closest" results to Goldbach's Conjecture mentioned earlier: Chen's result that every sufficiently large even number can be written as the sum of a prime and a number with at most two prime factors [6]; Helfgott's result on the Ternary Goldbach Conjecture [27], and Oliveira e Silva's empirical result that establishes the truth of Goldbach's Conjecture up through $4 \cdot 10^{18}$ [13].

Between 1920 and 1973 numerous other results related to Goldbach's Conjecture have been published. See [53] for some of the most significant of these.

2.2.4 The Circle Method

Hardy and Littlewood's 'Conjecture A' was mentioned in Section 1.2.2. To justify this conjecture they used a method called the 'circle method', which had its genesis in joint work by Hardy and Ramanujan. This method uses complex analysis to estimate the number of ways to decompose a positive integer into the sum of other positive integers. Discussion of this method is out of the scope of this thesis. The reader can find an account of this method in Nathanson [39]. In 2015, the circle method was used along with sieve methods and exponential sums to prove Goldbach's Ternary Conjecture: every number greater than 5 is the sum of three primes [27].

2.2.5 Evidence Based on the Density of Primes

A third important strand in the theoretical fabric supporting the truth of Goldbach's Conjecture has to do with so-called 'density arguments,' an early and significant example of which was given by Schnirelman in 1933 [45]: There is a constant C with the property that every integer greater than or equal to 4 is the sum of at most C primes [38]. Discussion of this area of research is also out of the scope of this thesis. The interested reader is referred to [38].

2.2.6 Lower Bounds, Upper Bounds, and Asymptotic Estimates

On the one hand, the above discussion shows that the empirical results concerning Goldbach's Conjecture are exact and suggestive, as far as they go. But they don't go far enough. The theoretical support for Goldbach's Conjecture, though exact in its foundations and its reach, seems frustratingly approximate in its ability to specify the kind of fine-grained understanding of the distribution of primes that would appear to be needed to prove Goldbach's Conjecture. The reader may well conclude that somewhere, somehow, certain fundamental facts about the prime numbers remain elusive.

2.3 Significance: Goldbach's Conjecture, Sieve Methods, Brun's Work

2.3.1 Significance of Goldbach's Conjecture

Goldbach's Conjecture is important less for its namesake, Christian Goldbach [41] - who is today chiefly remembered for this conjecture [20] - than for the mathematics and the mathematicians involved in addressing the conjecture. In particular, Goldbach's Conjecture is significant for at least four reasons: the nature of the problem itself, related problems, the mathematics that has been used to address it, and the fact that it has engaged the attention of some prominent mathematicians [14, 42, 53].

Goldbach's Conjecture is a simple statement about primes that remains unproved more than a quarter millennium after it was first written down in Euler's letter of 1742. Although it is a simple statement about primes - certainly simpler than the statement of the Riemann Hypothesis - the primes themselves are both frustratingly irregular and yet curiously regular in their distribution [55]. The conjecture would seem to depend on the distribution of prime numbers, because unless the primes are distributed in a certain way, the conjecture is false. The fact that it hasn't been proved yet, along with the fact that researchers believe the current techniques do not suffice to prove it, suggests that some new mathematics (or a new combination of existing mathematics) is needed to prove it, assuming that it can, indeed, be either proved or disproved. These facts alone can provide sufficient motivation for mathematicians to address the conjecture.

Goldbach's Conjecture is related to another well-known conjecture that has not yet been proved, the Twin Primes Conjecture, which contends that there are infinitely pairs of primes that differ by two. Both the Twin Primes Conjecture and Goldbach's Conjecture were mentioned by Landau in 1912 as examples of problems he considered "unassailable" [33]. One way to see that these two problems are related is to note that Brun uses (in his papers of 1915 and 1920) similar sieve machinery to discuss both problems [1, 3]. Additionally, although it is not wise to claim that either of these problems is related to any formulas or hypotheses of Riemann, when Hilbert gave his 1900 address to the International Congress of Mathematicians he said (in the problem where he discussed the Riemann Hypothesis), "After an exhaustive discussion of Riemann's prime number formula, perhaps we may sometime be in a position to attempt the rigorous solution of Goldbach's problem" [28]. The fact that Goldbach's Conjecture, along with the Twin Primes Conjecture, was mentioned by Hilbert in the context of the Riemann's 1859 paper is not insignificant.

Since the early twentieth century Goldbach's Conjecture has been investigated using new mathematics that includes sieve methods and the circle method. Subsequently, sieves have been used in areas as diverse as probability and discrete groups, where in the latter case "[t]he basic motivation is that any discrete set with interesting structure can be investigated by ideas that are related to sieve [sic]" [32].

Finally, Goldbach's Conjecture has received attention and serious study by well-known mathematicians. Among the names of the mathematicians who have considered various aspects of Goldbach's Conjecture are Euler, Cantor, Brun, Hardy and Littlewood, and Landau.

2.3.2 Significance of Sieve Methods

Brun's early work in sieves marked the beginning of much work on sieves in the twentieth century. Sieves have also been used to address other conjectures about primes (e.g., the Twin Primes Conjecture and the conjecture that there are infinitely many primes of the form $n^2 + 1$ [35]). They have also, as mentioned above, been applied in other areas and problems of mathematics [17, 32]. One reason for the versatility of sieve methods is that they can be given very general specifications [32] and can produce general results [23].

2.3.3 Significance of Brun's Sieve Methods

Brun's work on sieves is widely acknowledged and highly regarded by researchers in number theory (e.g., [17, 21, 23]). Three of his papers are of particular interest. In 1915 Brun introduced a sieve method that enabled him to prove that the infinite sum of the reciprocals of the twin primes is finite or converges [1]. In 1919 he *announced* the result concerning Goldbach's Conjecture ([2]) that he would later *prove* in his paper of [3], which is examined in some detail in Chapter 3.

One indication of high regard researchers have for Brun's work comes from an address on sieve methods to the International Congress of Mathematicians [2] in 1978 (the year Brun died), which was dedicated to the memory of Vigo Brun and which begins as follows: "In the early twenties of this century Viggo Brun . . . introduced a method which proved to be one of the most fruitful tools in elementary number theory" [29].

Some other recognition given to Brun's work includes:

[2] The International Congress of Mathematicians is the same conference where, in 1912, Landau gave an address in which he spoke of the Goldbach conjecture as "unangreifbar," i.e., "unassailable," [33]. Also, in a 1966 address to this congress, Vigo Brun himself delivered his "Reflections on the Sieve of Eratosthenes" [4].

- "The first to devise an effective sieve mechanism that goes substantially beyond the sieve of Eratosthenes was Viggo Brun . . . " [23].
- The study of Brun's later work on his sieve method "is essential as a viable tool in sieve theory that cannot be ignored" [7].
- "Sieve Methods in Number Theory have roots which can be traced back to antiquity, but the modern era may be said to have begun with the papers of Viggo Brun, in particular with his article 'Le Crible d'Eratosthène et le Théorème de Goldbach' in 1920" [21].
- "The sieve of Eratosthenes lay virtually ignored for some two thousand years. Despite the few touches applied by Legendre, the modern subject of sieve methods really begins with Viggo Brun" [17].
- In reference to two results from Brun's 1920 paper, "Brun had obtained these powerful results by means of skillful improvements of a sieve method that goes substantially beyond his historic source: the sieve of Eratosthenes (3rd century B.C.)" [46].
- Although improvements to Brun's methods were made by other authors, ". . . the authors had to employ other arithmetical devices; this shows how fruitful Brun's ideas have been and still are since they have found access into the repertoire of methods of number theory" [46].
- "Brun's [sieve] method is perhaps our most powerful elementary tool in number theory" [15].

CHAPTER 3
BRUN'S THEOREM ON GOLDBACH'S CONJECTURE

3.1 Overview of "Le Crible d'Eratosthène et le Théorème de Goldbach"

In "Le Crible d'Eratosthène et le Théorème de Goldbach" [3] Brun developed a new sieve method. His sieve employed Merlin's double application of the sieve of Eratosthenes, exploited certain asymptotic results, and manipulated the terms of the Legendre formula to establish a lower bound for the number of elements that fall through his sieve - all of which taken together enabled him to prove that every sufficiently large even number can be written as the sum of two 9-primes (positive integers with at most nine prime factors each).

"Le Crible d'Eratosthène et le Théorème de Goldbach" [3] contains seven sections, four of which (Sections 2, 3, 4, and 6) are discussed in some detail later in this chapter.

The rest of this section provides a brief overview of the seven sections in Brun's paper, with a focus on the results that are most germane to this thesis. (Brun has several notable results in his paper, in addition to his theorem on Goldbach's Conjecture, but those results are not treated in this thesis.) All page numbers given below refer to the English translation by Rui [44], except when the context makes clear that the original French version is being discussed, which is necessary to do from time to time to shed light on certain issues of language and typography.

In §1 (pp. 99-100) Brun states Goldbach's Conjecture, mentions the Twin Primes Conjecture[1], and says that one now (in 1920) has a starting point for addressing these problems. According to Brun, the first to draw attention to this fact was Jean Merlin (see [22], which was submitted by Hadamard on behalf of Merlin, who died in WWI). The double sieve was discussed in Section 1.9; it will be discussed again in Section 3.5 where we discuss §6 of Brun's paper.

[1] The Twin Prime Conjecture asserts that there are infinitely many pairs of primes with difference 2.

Section 2 of Brun's paper (pp. 100-110) introduces Brun's notational framework, as well as several examples of truncating the Legendre formula, en route to a method that will allow Brun to obtain a useful lower bound for the number of items that fall through his sieve. In this section Brun uses two different but equivalent notations for the expansion of the main term in his lower bound formula (see 1.4.1.5 and 1.4.1.6), which allow him some flexibility as he derives certain results in §3 and §4 of his paper.

In Section 3 (pp. 110-114) Brun develops in more detail the method of his second example from §2. By using Stirling's formula, he improves his initial lower-bound estimate for the number of elements that fall through his sieve, though he says at the end of this section that the remainder term in this newer estimate is "still too great for our purpose" (p. 114).

In Section 4 (pp. 114-123) Brun develops in more detail the method of his third example in §2. This method is perhaps best seen as an elaborate extension of the method he illustrated in §3: it provides Brun with a useful lower bound for the number of elements that fall through his sieve. In elaborating this method Brun exploits Stirling's formula again, as well as two theorems from Mertens. This section creates the machinery that is later applied to the double sieve in §6 to prove his theorem.

Section 5 (pp. 123-125) deals with a generalization of his developments from §1-4. This generalization appears to be not directly relevant to the establishment of Brun's theorem; therefore, it is not covered further in this chapter.

In Section 6 (pp. 125-131) Brun pulls together the results from earlier in his paper to prove his famous result on Goldbach's Conjecture:

> One can write [every] even number x, greater than x_0, as a sum of two numbers, whose numbers of prime factors do not exceed nine. x_0 denotes a determinable number and the prime factors can be different or not (p. 131).

Section 7 (pp. 131-136) contains additional results, which, since they occur after the proof of Brun's theorem in Section 6, are not discussed in this thesis.

3.2 Different Lower Bounds and Brun's Generalized Notation

In §2 of his paper, Brun develops his notation (see Section 3.2.2 below), discusses the best[2] lower bound for the Legendre formula $N(1, x, p_1, p_2, \ldots, p_r)$ (see 1.4.0.2), and gives three examples of how to calculate three other lower bounds for the Legendre formula. Next we discuss the different lower bounds. Then, we describe some of Brun's general notation.

3.2.1 Different Lower Bounds for the Legendre Formula

As Brun works to establish different lower bounds for the number of terms that fall through a sieve, he contrasts his results with what we have been calling the "worst" lower bound for the Legendre formula, given in 1.4.1.8 and repeated for convenience below.

$$N(1, x, p_1, p_2, \ldots, p_r) > x \sum_{d|p_1p_2\ldots p_r} \frac{\mu(d)}{d} - 2^r \tag{3.2.1.1}$$

Note that instead of using 2^r to refer to the number of terms in the expansion of the sum, Brun more often uses R to refer to this number (which can vary as the sum he is considering varies), as in

$$N(1, x, p_1, p_2, \ldots, p_r) > x \sum_{d|p_1p_2\ldots p_r} \frac{\mu(d)}{d} - R. \tag{3.2.1.2}$$

The different lower bounds he constructs are all improvements, in particular ways, on this default lower bound.

3.2.1.1 Searching for Improved Lower Bounds

The method Brun uses to calculate the "la meilleure limite inférieure" ([3], p. 9) for $N(1, x, p_1, p_2, \ldots, p_r)$ is discussed, along with several examples, on pp. 106 and 107. Since we have already discussed this method in Section 1.5, and since Brun does not deal with it again in sections 3, 4, or 6, we do not discuss it further.

Although we have also already discussed the method underlying the first of his three examples (Example 1 on p. 108), we will look at it again below, because considering it in the context of Brun's §2 helps us see the progression of Brun's examples, which develop lower bounds that are increasingly better, and such that the last of which

[2] That is, he describes the best lower bound obtainable via operations on the default lower bound formula that remove certain terms from its expansion.

is sufficient, along with the double sieve, to allow him to establish his theorem on Goldbach's Conjecture.

3.2.1.2 The Default ("Worst") Lower Bound, Revisited

In Chapter 1, Section 1.4.1.8, we discussed the worst-case lower bound for the number of items that fall through a sieve. That lower bound is obtained by dropping no terms from the main-term expansion and using 2^r, the number of terms in the expansion of the Legendre formula, as the bound on the maximum absolute value of the error term in the Legendre formula, where r is the number of sifting primes. The example Brun gives for this default lower bound on page 108, and is reproduced from the original in Figure 3.2.1.1.

Ex. 1)

$$
\begin{aligned}
N(1,x,2,3,5,7) > x\Bigg[1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} - \frac{1}{7}
&+ \frac{1}{3\cdot 2} \\
&+ \frac{1}{5\cdot 2} + \frac{1}{5\cdot 3}\left(1 - \frac{1}{2}\right) \\
&+ \frac{1}{5\cdot 2} + \frac{1}{7\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{7\cdot 5}\left(1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{3\cdot 2}\right)\Bigg] - 16 \\
&= x\left[\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{7}\right)\right] - 2^4
\end{aligned}
$$

FIGURE 3.2.1.1: Default lower bound for $N(1, x, 2, 3, 5, 7)$: dropping no terms from the expansion of the main term

Note the two equivalent forms that Brun uses for the expansion of the main term, which follow the pattern established in formulas 1.4.1.5 and 1.4.1.6, the equivalence of which the reader has no doubt previously verified. The reason that $N(1, x, 2, 3, 5, 7)$ is *greater than* rather than *equal to* the right-hand expression is that Brun is considering the worst-case scenario in making this particular lower-bound estimate: the lower bound for $N(1, x, 2, 3, 5, 7)$ is greater than the main term minus the worst-case value for remainder R, which is bounded in magnitude by 2^4. That is, the absolute value of each summand of the remainder term $\sum_{d|2\cdot4\cdot5\cdot7}\mu(d)\left(\left\lfloor\frac{x}{d}\right\rfloor-\frac{x}{d}\right)$ in the expansion of the Legendre formula is bounded by 1, and there are 16 terms in its expansion. (Brun's examples do not show any details of the error term; they simply give a value for R, which equals the number of terms in the main-term expansion of whichever example Brun happens to be dealing with.)

3.2.1.3 Discarding Terms in "Small Print"

In the second of the three examples on pp. 108 - 110 (see Figure 3.2.1.2) Brun shows that a lower bound for $N(1,x,p_1,p_2,\ldots,p_r)$ is obtained by dropping particular terms from the Legendre expansion, in particular, those terms "in small print."

Ex. 2) $N(1,x,2,3,5,7,11) > x\Bigg[1-\frac{1}{2}-\frac{1}{3}-\frac{1}{5}-\frac{1}{7}-\frac{1}{11}$

$+\frac{1}{3\cdot2}$

$+\frac{1}{5\cdot2}+\frac{1}{5\cdot3}\left(1-\frac{1}{2}\right)$

$+\frac{1}{7\cdot2}+\frac{1}{7\cdot3}\left(1-\frac{1}{2}\right)+\frac{1}{7\cdot5}\left[1-\frac{1}{2}-\frac{1}{3}+\frac{1}{3\cdot2}\right]$

$+\frac{1}{11\cdot2}+\frac{1}{11\cdot3}\left(1-\frac{1}{2}\right)+\frac{1}{11\cdot5}\left(1-\frac{1}{2}-\frac{1}{3}+\frac{1}{3\cdot2}\right)+\frac{1}{11\cdot7}\left(1-\frac{1}{2}-\frac{1}{3}-\frac{1}{5}+\frac{1}{3\cdot2}+\frac{1}{5\cdot2}+\frac{1}{5\cdot3}\left(1-\frac{1}{2}\right)\right)\Bigg]-26$

où les termes écartés sont ajoutés en petit. On peut aussi écrire:

$$N(1,x,2,3,5,7,11) > x\left[\left(1-\frac{1}{2}\right)\left(1-\frac{1}{3}\right)\left(1-\frac{1}{5}\right)\left(1-\frac{1}{7}\right)\left(1-\frac{1}{11}\right)-\right.$$

$$-\left(\frac{1}{7\cdot5\cdot3\cdot2}+\frac{1}{11\cdot5\cdot3\cdot2}+\frac{1}{11\cdot7\cdot3\cdot2}+\frac{1}{11\cdot7\cdot5\cdot2}+\frac{1}{11\cdot7\cdot5\cdot3}\right)+\left.\left(\frac{1}{11\cdot7\cdot5\cdot3\cdot2}\right)\right]-$$

$$-\left(1+5+\frac{5\cdot4}{1\cdot2}+\frac{5\cdot4\cdot3}{1\cdot2\cdot3}\right)=x[0,2078-0,0121+0,0004]-26=0,1961\,x-26$$

Ici nous avons écarté tous termes de la forme $\frac{1}{p_a\cdot p_b\cdot p_c\cdot p_d}$ et de la forme $\frac{1}{p_a\cdot p_b\cdot p_c\cdot p_d\cdot p_e}$.

FIGURE 3.2.1.2: Lower bound for $N(1,x,2,3,5,7,11)$ obtained by dropping terms "in small print"

The typography and language of the original French are key to understanding what Brun means here. Notice that in Figure 3.2.1.2 some terms in the right-hand side of the inequality are in smaller typeface than other terms are. This typography works hand in glove with Brun's language - "les termes écartés sont ajoutés en petit" - to show the reader which terms he is talking about: the terms that are in *small print*!

By contrast, Rui's translation and its associated typography obscure rather than illuminate which terms are being set aside (Figure 3.2.1.3).

Eg. 2) $$N(1, x, 2, 3, 5, 7, 11) > x\Big[1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} - \frac{1}{7} - \frac{1}{11}$$
$$+ \frac{1}{3\cdot 2} + \frac{1}{5\cdot 2} + \frac{1}{5\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{7\cdot 2} + \frac{1}{7\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{7\cdot 5}\left(1 - \frac{1}{2} - \frac{1}{3}\right)$$
$$+ \frac{1}{11\cdot 2} + \frac{1}{11\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{11\cdot 5}\left(1 - \frac{1}{2} - \frac{1}{3}\right) + \frac{1}{11\cdot 7}\left(1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5}\right)\Big]$$
$$- 26 \quad ,$$

where the terms set aside are added on a small scale. One can also write

$$N(1, x, 2, 3, 5, 7, 11) > x\Big[\left(1 - \frac{1}{2}\right)\left(1 - \frac{1}{3}\right)\left(1 - \frac{1}{5}\right)\left(1 - \frac{1}{7}\right)\left(1 - \frac{1}{11}\right)$$
$$- \left(\frac{1}{7\cdot 5\cdot 3\cdot 2} + \frac{1}{11\cdot 5\cdot 3\cdot 2} + \frac{1}{11\cdot 7\cdot 3\cdot 2} + \frac{1}{11\cdot 7\cdot 5\cdot 2} + \frac{1}{11\cdot 7\cdot 5\cdot 3}\right)$$
$$+ \left(\frac{1}{11\cdot 7\cdot 5\cdot 3\cdot 2}\right)\Big] - \left(1 + 5 + \frac{5\cdot 4}{1\cdot 2} + \frac{5\cdot 4\cdot 3}{1\cdot 2\cdot 3}\right)$$
$$= x\big[0.2078 - 0.0121 + 0.0004\big] - 26 = 0.1961x - 26.$$

Here we have set aside all terms of the form $\frac{1}{p_a p_b p_c p_d}$ and of the form $\frac{1}{p_a p_b p_c p_d p_e}$.

FIGURE 3.2.1.3: Rui's version of Brun's Example 2

The language Rui uses to describe the terms that are set aside is that they "are added on a small scale" ([44], p. 108). What does "small scale" mean in this context? The typography of Rui's version as published in Wang [53] is of little help in making sense of this phrase, because it makes no distinction between the terms that are kept in the expansion and the discarded terms, i.e., those that are "ajoutés en petit," which might be better translated as "added in small print" or "in small typeface," although without a corresponding typographical change to indicate which terms are meant, a better translation by itself would be of little use.

But by studying Brun's original version we can, with just a little effort, come to see what he means. In particular, we can understand why he describes the discarded terms as having the form $\frac{1}{p_a p_b p_c p_d}$ and $\frac{1}{p_a p_b p_c p_d p_e}$. In the original French *we can see that* the terms in small print, when multiplied by the terms outside the parentheses in which they live, yield terms of the form $\frac{1}{p_a p_b p_c p_d}$ and $\frac{1}{p_a p_b p_c p_d p_e}$, exactly the forms that Brun is referring to.

The formula obtained by dropping the terms "in small print" from the Legendre expansion can be specified unambiguously. This formula is used in Brun's §3 along with Stirling's formula to construct a lower-bound estimate that is better than the default lower-bound estimate (see Section 3.3). Brun ends §3 by saying (p. 114, formula (12)) that the lower bound thus achieved, where the remainder term increases as a power of the number of terms used in the expansion, is better than the lower bound developed earlier (p. 113, formula (9), which shows the default lower bound), but that it is still not good enough "for our purpose." Thus, at the end of §3 Brun sets the stage for §4, where he invents an improved lower-bound formula by dropping terms "to the right of vertical lines."

3.2.1.4 Discarding Terms "to the Right of Vertical Lines"

Brun's third example (pp. 109 and 110 of Rui's translation) illustrates the third method he describes for finding a lower bound for the Legendre formula, one that he'll develop in §4 and which, along with appropriate changes for use in his double sieve, will allow him to prove his theorem on Goldbach's Conjecture in §6 of his paper.

The typography of the original French in Brun's third example also makes evident what he means by dropping terms "to the right of vertical lines." Here, as it turns out, Rui's translation and the typography he uses *are rather faithful* to the original French, even surpassing it in one regard (Rui presents this example on a single page (p. 109)), though falling short in another (the English translation does not display the rightmost vertical line that appears (albeit faintly) in the original French publication).

In Figure 3.2.1.4 below, I have re-constructed the French version of Brun's third example so that it fits in a single page, to emphasize the structure of his expansion and the visual effects of the, admittedly faint, vertical lines. Although this figure allows us to understand the *result* of dropping terms "to the right of vertical lines," the figure itself reveals little of the *algorithm* that generated these effects.

Ex. 3) $N(1, x, 2, 3, 5, 7, 11, 13, 17, 19) >$

$$
\begin{aligned}
& x\Big[1 - \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{5} - \tfrac{1}{7} - \tfrac{1}{11} - \tfrac{1}{13} - \tfrac{1}{17} - \tfrac{1}{19} \\
& + \tfrac{1}{3\cdot 2} \\
& + \tfrac{1}{5\cdot 2} + \tfrac{1}{5\cdot 3}\left(1 - \tfrac{1}{2}\right) \\
& + \tfrac{1}{7\cdot 2} + \tfrac{1}{7\cdot 3}\left(1 - \tfrac{1}{2}\right) + \tfrac{1}{7\cdot 5}\left[1 - \tfrac{1}{2} - \tfrac{1}{3} + \tfrac{1}{3\cdot 2}\right] \\
& + \tfrac{1}{11\cdot 2} + \tfrac{1}{11\cdot 3}\left(1 - \tfrac{1}{2}\right) + \tfrac{1}{11\cdot 5}\left[1 - \tfrac{1}{2} - \tfrac{1}{3} + \tfrac{1}{3\cdot 2}\right] + \tfrac{1}{11\cdot 7}\left(1 - \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{5} + \tfrac{1}{3\cdot 2} + \tfrac{1}{5\cdot 2}\right) \\
& + \tfrac{1}{13\cdot 2} + \tfrac{1}{13\cdot 3}\left(1 - \tfrac{1}{2}\right) + \tfrac{1}{13\cdot 5}\left[1 - \tfrac{1}{2} - \tfrac{1}{3} + \tfrac{1}{3\cdot 2}\right] + \tfrac{1}{13\cdot 7}\left(1 - \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{5} + \tfrac{1}{3\cdot 2} + \tfrac{1}{5\cdot 2}\right) \\
& + \tfrac{1}{17\cdot 2} + \tfrac{1}{17\cdot 3}\left(1 - \tfrac{1}{2}\right) + \tfrac{1}{17\cdot 5}\left[1 - \tfrac{1}{2} - \tfrac{1}{3} + \tfrac{1}{3\cdot 2}\right] + \tfrac{1}{17\cdot 7}\left(1 - \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{5} + \tfrac{1}{3\cdot 2} + \tfrac{1}{5\cdot 2}\right) \\
& + \tfrac{1}{19\cdot 2} + \tfrac{1}{19\cdot 3}\left(1 - \tfrac{1}{2}\right) + \tfrac{1}{19\cdot 5}\left[1 - \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{3\cdot 2}\right] + \tfrac{1}{19\cdot 7}\left(1 - \tfrac{1}{2} - \tfrac{1}{3} - \tfrac{1}{5} + \tfrac{1}{3\cdot 2} + \tfrac{1}{5\cdot 2}\right)\Big] - 72 \\
& = 0{,}163\,x - 72
\end{aligned}
$$

FIGURE 3.2.1.4: Lower bound for $N(1, x, 2, 3, 5, \ldots, 19)$: dropping terms "to the right of vertical lines"

As we saw in the previous section, it is straightforward to understand the algorithm underlying Brun's second example (see Sections 3.2.1.3 and 3.3) where he discards terms "in small print." That algorithm discards terms "in the tail" of the expansion of the main term (when that expansion has the form shown in 1.4.1.4), i.e., the Σ-terms with the largest denominators.

However, the algorithm behind Brun's third example, which discards terms "to the right of vertical lines," is more intricate. In addition to discarding terms in the tail of the expansion of the main term when it is expressed as in 1.4.1.4, this algorithm also reduces the number of terms of the expansion, both positive and negative terms in a precisely specified way (see Section 3.4.1).

Consider for instance the full summation of the form $\sum\sum_{p_1<p_2}\frac{1}{p_1 p_2}$. In this case, if there are r sifting primes total, then there are $\binom{r}{2}$ summands in that full summation. However, using Brun's algorithm for discarding terms "to the right of vertical lines," not all such summands are included in the expansion of the main term. See Section 3.4.2.

Next we briefly discuss Brun's notation from §2 and note that only some of his notation is needed to trace the main thread of his later developments in §3, §4, and §6.

3.2.2 Brun's Notation in §2 for his Generalized Sieve

The notation that Brun develops in §2 is for a more general sieve than the kind we have been considering up to this point. Our presentation has not yet discussed this notation because it hasn't so far been needed to convey the main ideas. Later, when we consider the double sieve again in Section 3.5, we will have occasion to mention some of the notation from Brun's more generalized sieve.

Brun's sieve operates on a sequence that is not necessarily a sequence of consecutive integers from 1 to x. The first element of the sifting sequence can be an arbitrary positive integer, which he calls Δ. Also, he considers generalized arithmetic sequences with common difference D, where D is greater than or equal to 1 and relatively prime to each of the sifting primes $p_1, p_2, \ldots p_r$. Further, Brun's sieve does not necessarily operate by trapping multiples of the sifting primes per se, but rather by trapping certain sequences that have a common difference of p_1, or of p_2, ..., or of p_r. So, Brun's sieve would allow us to sieve the sequence of consecutive odd integers $3, 5, \ldots x$ by the sifting sequence $2, 7, 12, 17, ..., 37$. In that case, $\Delta = 3$, and $D = 5$. Here, the sifting sequence starts with

2, and we sift by a sequence with common different 5 (a prime). In general, the sifting sequences can start with any positive integer.

Brun's original formulation for $N(1, x, p_1, p_2, \ldots, p_r)$ is different from the Legendre formula we have used in Chapter 1 for $N(1, x, p_1, p_2, \ldots, p_r)$. Brun later *derives* the Legendre formula and various lower bounds for it. Indeed, his examples 1, 2, and 3 on pp. 108-110 look very much like examples we've seen already in Chapter 1.

Later, in Section 3.5, we'll consider two main pieces of Brun's generalized notation in his §6. First, in dealing with the double sieve, he sifts by "double rows" like those we exhibited in Figure 1.9.0.2. Second, in his example on p. 127 the sequence to be sifted has common difference $D = 2$, and starts at 1, even though his general notation allows for other Δ and D values. The result is just a double sieve that operates on a sequence of consecutive odd numbers beginning at 1, very much like the simple example shown in Figure 1.9.0.3.

3.3 Discarding Terms in "Small Print"

In §3 Brun extends and generalizes computations like those from his second example on p. 108 to show that the sieve method of discarding terms "in small print" yields an improved lower bound estimate for $N(1, x, p_1, p_2, \ldots, p_r)$. Brun first establishes that the operation of this sieve – where the main-term expansion is truncated at an "m-indexed Σ-term" (where m is odd; see Section 3.3.2) – yields a formula that does indeed give a lower bound for N, the number of numbers that fall through his sieve[3]. That formula has the form $N > M - R$. Brun then uses algebra to establish recursive relationships among certain terms in the expansion of his sieve. Then he invokes Stirling's formula to show in a simple form one particular relationship among certain terms of the expansion. He uses these results to establish a lower bound A for M. Then he uses a counting argument to find an upper bound B for R. The bounds A and B, applied to the fundamental inequality $N > M - R$, give Brun his final result of this section – a lower bound for N which holds for the mild condition that $m + 2$ is greater than the sum of the reciprocals of the sifting primes $p_1, p_2, \ldots, p_r$ (see the end of Section 3.3.4 for why this condition exists).

So, in effect, "all" that Brun does is start with a particular form (3.3.2.2) of the fundamental inequality $N > M - R$, and then proceed to bound M below by A and

[3] An argument for why this sieve does in fact give a lower bound for N is given in 3.3.2.

R above by B, thus establishing a new inequality of the form $N > A - B$, which gives an improved lower bound for N (in the sense that it reduces the size of the remainder term). Thus, with $\sigma = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_r}$:

$$N(1, x, p_1, p_2, \ldots, p_r) > x\big[(1-\frac{1}{p_1})(1-\frac{1}{p_2})\ldots(1-\frac{1}{p_r}) - (\frac{e\sigma}{m+1})^{m+1})\big] - r^{m+1}. \quad (3.3.0.1)$$

3.3.1 Algorithm for Discarding Terms in "Small Print"

Recall Brun's second example from page 108, which is shown again below in Figure 3.3.1.1. Although it may not be immediately obvious how to algorithmically drop terms "in small print" from a sieve, Brun shows for his example that removing terms in small print means removing those terms of the expansion that are of the form $\dfrac{1}{p_a p_b p_c p_d}$ and $\dfrac{1}{p_a p_b p_c p_d p_e}$. Thus, *discarding terms in "small print" means removed terms from the tail end of the expansion of the main term in the Legendre expansion.*

Eg. 2) $N(1, x, 2, 3, 5, 7, 11) > x\Big[1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} - \frac{1}{7} - \frac{1}{11}$

$+ \frac{1}{3\cdot2} + \frac{1}{5\cdot2} + \frac{1}{5\cdot3}(1 - \frac{1}{2}) + \frac{1}{7\cdot2} + \frac{1}{7\cdot3}(1 - \frac{1}{2}) + \frac{1}{7\cdot5}(1 - \frac{1}{2} - \frac{1}{3})$

$+ \frac{1}{11\cdot2} + \frac{1}{11\cdot3}(1 - \frac{1}{2}) + \frac{1}{11\cdot5}(1 - \frac{1}{2} - \frac{1}{3}) + \frac{1}{11\cdot7}(1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5})\Big]$

$- 26$,

where the terms set aside are added on a small scale. One can also write

$$N(1, x, 2, 3, 5, 7, 11) > x\Big[(1 - \frac{1}{2})(1 - \frac{1}{3})(1 - \frac{1}{5})(1 - \frac{1}{7})(1 - \frac{1}{11})$$
$$- \Big(\frac{1}{7\cdot5\cdot3\cdot2} + \frac{1}{11\cdot5\cdot3\cdot2} + \frac{1}{11\cdot7\cdot3\cdot2} + \frac{1}{11\cdot7\cdot5\cdot2} + \frac{1}{11\cdot7\cdot5\cdot3}\Big)$$
$$+ \Big(\frac{1}{11\cdot7\cdot5\cdot3\cdot2}\Big)\Big] - \Big(1 + 5 + \frac{5\cdot4}{1\cdot2} + \frac{5\cdot4\cdot3}{1\cdot2\cdot3}\Big)$$
$$= x\big[0.2078 - 0.0121 + 0.0004\big] - 26 = 0.1961x - 26.$$

Here we have set aside all terms of the form $\dfrac{1}{p_a p_b p_c p_d}$ and of the form $\dfrac{1}{p_a p_b p_c p_d p_e}$.

FIGURE 3.3.1.1: Rui's version of Brun's Example 2

Thus, the truncation of the expansion of the main term is such that we keep an odd number ($m = 3$ in the example above) of "Σ" terms in the expansion and discard the rest. That is, from the full expansion of the main term

$x\Big[1 - \sum_p \frac{1}{p} + \sum\sum_{p_1<p_2} \frac{1}{p_1 p_2} - \sum\sum\sum_{p_1<p_2<p_3} \frac{1}{p_1 p_2 p_3} + \ldots + (-1)^r \sum\sum\ldots\sum_{p_1<\ldots<p_r} \frac{1}{p_1 p_2 \ldots p_r}\Big],$

we keep only $x\Big[1 - \sum_p \frac{1}{p} + \sum\sum_{p_1<p_2} \frac{1}{p_1 p_2} - \sum\sum\sum_{p_1<p_2<p_3} \frac{1}{p_1 p_2 p_3}\Big]$.

At this point, we have not shown that by truncating the expansion of the main term in this manner and adjusting the remainder term appropriately, the resulting expression of a main term minus a remainder term does give a lower bound for N. The next section addresses this issue by giving an overview of how Brun's calculations for this sieve do yield a lower bound for N. The details of Brun's calculations follow in subsequent sections.

3.3.2 Brun's Sieve in §3 Does Give a Lower Bound for N

Recall from Section 1.5 the inequality that gives the "worst" lower bound for N:

$$N(1, x, p_1, p_2, \ldots, p_r) > x \sum_{d|p_1 p_2 \ldots p_r} \frac{\mu(d)}{d} - R, \tag{3.3.2.1}$$

where $R = 2^r$, with r equal to the number of sifting primes used in the sieve. This inequality can also be written

$$N > x\Big[1 - \sum_p \frac{1}{p} + \sum\sum_{p_1<p_2} \frac{1}{p_1 p_2} - \sum\sum\sum_{p_1<p_2<p_3} \frac{1}{p_1 p_2 p_3} + \ldots + (-1)^r \sum\sum\ldots\sum_{p_1<\ldots<p_r} \frac{1}{p_1 p_2 \ldots p_r}\Big] - R.$$

Brun basically takes as his point of departure an inequality similar to the one above, but with a truncated expansion of the main term and a correspondingly adjusted remainder term R. This inequality (essentially formula (8) on p. 112) is

$$N > \Big[x - \sum_p \frac{x}{p} + \sum\sum_{p_1<p_2} \frac{x}{p_1 p_2} - \ldots - \sum\sum\ldots\sum_{p_1<p_2<\ldots<p_m} \frac{x}{p_1 p_2 \ldots p_m}\Big] - R, \tag{3.3.2.2}$$

where m is an odd number less than r. (Note that the values of R in these last two inequalities, although they refer to the number of terms in the expansion of their respective main terms, will in general not be the same.)

It is not immediately clear why, even though the fully expanded right-hand side in 3.3.2.1 gives a lower bound for N, so, too, does the truncated right-hand side given in

3.3.2.2. Because Brun arrives at this fact by an argument that is not discussed in this thesis, it is useful to come at this result through a different argument, given next.

Observe that the expansion of the main term in 3.3.2.2 is truncated after subtracting an "odd-numbered Σ-term" (since m is odd). Recall that the Principle of Inclusion-Exclusion (PIE) operates by alternatingly *overcounting* and *undercounting* the number of elements that fall through the sieve (see Example 1.4.0.1). These overcounts and undercounts are tracked, respectively, by the addition and subtraction of the Σ-terms in the expansion. So, truncating the expansion of the PIE at an "odd-numbered Σ-term" (i.e., right after *subtracting* a group of Σ-terms from the expansion), results in an undercount of the number of terms that fall through the sieve, i.e., a lower bound for N. So far, so good.

Now compare 3.3.2.2 to the inequality below, which, based on how the PIE operates, we know holds:

$$N > x - \sum_{p} \left\lfloor \frac{x}{p} \right\rfloor + \sum_{p_1<p_2}\sum \left\lfloor \frac{x}{p_1 p_2} \right\rfloor - \ldots - \sum_{p_1<p_2<\ldots<p_m}\sum \ldots \sum \left\lfloor \frac{x}{p_1 p_2 \ldots p_m} \right\rfloor . \tag{3.3.2.3}$$

Each summand on the right-hand side of 3.3.2.3 of the form $\lfloor \frac{x}{d} \rfloor$ (where d is a prime or a product of primes) can also be written in the form $\frac{x}{d} - (\frac{x}{d} - \lfloor \frac{x}{d} \rfloor)$. And, of course, there are as many expressions of the form $\frac{x}{d} - (\frac{x}{d} - \lfloor \frac{x}{d} \rfloor)$ as there are summands of the form $\lfloor \frac{x}{d} \rfloor$. Designate the number of such expressions by R.

Further, each expression of the form $(\frac{x}{d} - \lfloor \frac{x}{d} \rfloor)$ is a positive number less than 1, and since there are R such expressions, the sum of all such expressions is at most R. But, because of the alternating signs in 3.3.2.3, not all expressions $(\frac{x}{d} - \lfloor \frac{x}{d} \rfloor)$ contribute a positive amount to the right-hand side. This means that the total contribution of all the expressions of the form $(\frac{x}{d} - \lfloor \frac{x}{d} \rfloor)$ is a number, say k, whose magnitude is less than R. So, since $\lfloor \frac{x}{d} \rfloor = \frac{x}{d} - (\frac{x}{d} - \lfloor \frac{x}{d} \rfloor)$ and since k collects the total contribution of the terms $(\frac{x}{d} - \lfloor \frac{x}{d} \rfloor)$, the formula in 3.3.2.3 can be written

$$N > \left[x - \sum_{p} \frac{x}{p} + \sum_{p_1<p_2}\sum \frac{x}{p_1 p_2} - \ldots - \sum_{p_1<p_2<\ldots<p_m}\sum \ldots \sum \frac{x}{p_1 p_2 \ldots p_m} \right] + k, \tag{3.3.2.4}$$

where k (since its magnitude is less than R) is less than the number of summands in the brackets.

Now, since $|k| < R$, we still maintain the inequality if we write

$$N > \Big[x - \sum_p \frac{x}{p} + \sum_{p_1<p_2}\sum \frac{x}{p_1 p_2} - \ldots - \sum_{p_1<p_2<\ldots<p_m}\sum \ldots \sum \frac{x}{p_1 p_2 \ldots p_m}\Big] - R, \qquad (3.3.2.5)$$

where R is the number of terms inside the brackets. This last inequality is the same as that in 3.3.2.2, which is what we set out to show holds.

Note, however, that although Brun's manipulations do result in a lower-bound inequality for N that changes the size of the remainder term for the better by reducing it from 2^r to r^{m+1}, it is still possible (as shown in Section 3.3.7) that this lower bound for N is negative. Nonetheless, the technique Brun uses in §3, specifically his result that uses Stirling's formula to bound the size of the mth term of the expansion, forms a key part of Brun's complex manipulations in §4 of his paper, in which he does establish that for sufficiently large numbers x, the lower bound for $N(1, x, p_1, p_2, \ldots, p_r)$ is positive.

3.3.3 Stirling's Formula is Used to Bound Part of the Expansion

To bound the mth term in the expansion of the main term, Brun uses a recursive relationship among certain terms in the expansion of the main term; he then employs Stirling's formula to arrive at an upper bound on the m-indexed Σ-term (pp. 13-16 in the original [3] give more details than the corresponding pages 111-116 in Rui's translation [44]).

Brun defines σ as the sum of the reciprocals of the sifting primes; that is, $\sigma = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_r}$. Next, he defines $\Sigma_1 = \sum_p \frac{1}{p} = \sigma$, $\Sigma_2 = \sum_{p_1<p_2}\sum \frac{1}{p_1 p_2}$, $\Sigma_3 = \sum_{p_1<p_2<p_3}\sum\sum \frac{1}{p_1 p_2 p_3}$, etc.

With this notation, the expansion

$$x\Big[1 - \sum_p \frac{1}{p} + \sum_{p_1<p_2}\sum \frac{1}{p_1p_2} - \sum_{p_1<p_2<p_3}\sum\sum \frac{1}{p_1p_2p_3} + \ldots + (-1)^r \sum_{p_1<\ldots<p_r}\sum \ldots \sum \frac{1}{p_1p_2\ldots p_r}\Big],$$

can be written as

$$x\Big[(1 - \Sigma_1 + \Sigma_2 - \Sigma_3 + \ldots (-1)^r \Sigma_r)\Big].$$

Then, using the definitions of σ, Σ_1, and Σ_2, Brun shows that

$$\sigma \cdot \Sigma_1 = \sigma \cdot \sigma = \sigma^2 = (\tfrac{1}{p_1})^2 + (\tfrac{1}{p_2})^2 + \ldots + (\tfrac{1}{p_r})^2 + 2\Sigma_2 > 2\Sigma_2.$$

That is, $\sigma \cdot \Sigma_1 > 2\Sigma_2$. Brun continues, in a similar manner:

$$\sigma \cdot \Sigma_2 > 3\Sigma_3,$$

$$\sigma \cdot \Sigma_3 > 4\Sigma_3,$$

$$\dots$$

$$\sigma \cdot \Sigma_{m-1} > m\Sigma_m,$$

and therefore, on dividing through by m, also that

$$\Sigma_m < \frac{\sigma}{m}\Sigma_{m-1}, \tag{3.3.3.1}$$

from which it follows that

$$\Sigma_m < \frac{\sigma^m}{m!}, \tag{3.3.3.2}$$

which upon using Stirling's formula ($n! = (\frac{n}{e})^n(\sqrt{2\pi n} + \theta), -1 < \theta < 1$), yields

$$\Sigma_m < \left(\frac{e\sigma}{m}\right)^m. \tag{3.3.3.3}$$

That is, Σ_m is bounded above by $\left(\frac{e\sigma}{m}\right)^m$. Next, we'll see that this *upper* bound on Σ_m allows Brun to find a *lower* bound for the main term of the expansion $M = x\Big[(1 - \Sigma_1 + \Sigma_2 - \dots (-1)^m\Sigma_m\Big]$

3.3.4 Bounding the Main Term

Brun next reformulates the expansion of the truncated main term, beginning with an add-and-subtract argument. He expresses

$$N(1, x, p_1, p_2, \dots, p_r) > x\Big[(1 - \Sigma_1 + \Sigma_2 - \dots - \Sigma_m + \Sigma_{m+1} - \dots + \dots (-1)^r\Sigma_r)\Big] - R.$$

as

$$\begin{aligned} N(1, x, p_1, p_2, \dots, p_r) > \\ & x\Big[(1 - \Sigma_1 + \Sigma_2 - \dots - \Sigma_m + \Sigma_{m+1} - \dots + \dots (-1)^r\Sigma_r) \\ & - (\Sigma_{m+1} - \Sigma_{m+2} + ...(-1)^r\Sigma_r)\Big] - R, \end{aligned} \tag{3.3.4.1}$$

where R is the number of terms in the expansion.

Brun remarks that we know the value of the terms in the first set of parentheses of 3.3.4.1. That value is just $[(1-\frac{1}{p_1})(1-\frac{1}{p_2})\dots(1-\frac{1}{p_r})]$, using the equivalence of the expansion of the main term that we already showed in 1.4.1.5 and 1.4.1.6.

Next, Brun says, the terms of 3.3.4.1 in the second pair of parentheses are decreasing (in absolute value) when $m+2 > \sigma = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_r}$ (see below); therefore, the value of all the terms in the second set of parentheses is less than Σ_{m+1}, which, in turn, by 3.3.3.3 is less than $\left(\frac{e\sigma}{m+1}\right)^{m+1}$. Combining the above observations, which is a matter of algebra, gives Brun his result, that

$$N(1, x, p_1, p_2, \ldots, p_r) > x\Big[(1-\frac{1}{p_1})(1-\frac{1}{p_2})\dots(1-\frac{1}{p_r}) - (\frac{e\sigma}{m+1})^{m+1}\Big] - R, \quad (3.3.4.2)$$

where m is odd and less than r, $m+2 > \sigma$, and R is the number of terms in the expansion.

By subtracting $(\frac{e\sigma}{m+1})^{m+1}$ from the main term in the expansion, Brun gives a lower bound for the main term (in effect giving $M > A$). Next, he will find an upper bound for R (in effect giving $R < B$). Then he will be able to combine the information in these two inequalities to arrive at his final result in §3.

It still remains to show that the terms involving $\Sigma_j, j > m$ are decreasing in magnitude when $m+2 > \sigma = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_r}$, i.e., when $\frac{\sigma}{m+2} < 1$. This result follows from the recursive relationships Brun developed earlier.

Recall that Brun established that $\Sigma_m < \frac{\sigma}{m}\Sigma_{m-1}$. Therefore, $\Sigma_{m+2} < \frac{\sigma}{m+2}\Sigma_{m+1}$, which is sufficient to show that the terms involving $\Sigma_j, j > m$ are decreasing in magnitude when $m+2 > \sigma = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_r}$.

3.3.5 Bounding the Remainder Term

Consider again the inequality

$$N(1, x, p_1, p_2, \ldots, p_r) > x\Big[(1-\frac{1}{p_1})(1-\frac{1}{p_2})\dots(1-\frac{1}{p_r}) - (\frac{e\sigma}{m+1})^{m+1}\Big] - R. \quad (3.3.5.1)$$

We observe that R, the number terms in the resulting expansion of the main term, can be most easily determined by using the form of the main term given below:

$$N > x\Big[1 - \sum_{p}(\frac{1}{p}) + \sum_{p_1<p_2}\sum(\frac{1}{p_1p_2}) - \ldots - \sum_{p_1<p_2<\ldots<p_m}\sum\cdots\sum(\frac{1}{p_1p_2\ldots p_m}))\Big] - R, \quad (3.3.5.2)$$

Thus, $R = 1+\binom{r}{1}+\binom{r}{2}+\ldots+\binom{r}{m}$, which, as Brun indicates, is less than $1+r+r^2+\ldots+r^m$, which in turn is less than r^{m+1}. Thus, this sieve shrinks the remainder term from 2^r, which was the remainder term in the default lower bound, to r^{m+1}.

3.3.6 Combining Results Yields a New Lower Bound for N

Thus, combining the bounds we found in the previous two sections we at last obtain

$$N(1, x, p_1, p_2, \ldots, p_r) > x\big[(1-\frac{1}{p_1})(1-\frac{1}{p_2})\ldots(1-\frac{1}{p_r})-(\frac{e\sigma}{m+1})^{m+1})\big]-r^{m+1}, \quad (3.3.6.1)$$

where m is odd and such that $m+2 > \sigma = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_r}$.

But why does Brun say at the end of §3 (p. 114) that the increase of the error term in this lower bound "is still too great for our purpose"? One possibility is that he knew that his formula in 3.3.6.1 could still result in a negative lower bound for $N(1, x, p_1, p_2, \ldots, p_r)$.

3.3.7 Problem: the Improved Lower Bound Could Still be Negative

Using 3.3.6.1, we can calculate the improved lower bound for $N(1, x, 2, \ldots 19)$ and $m = 3$. We get

$$\begin{aligned} &N(1, x, 2, 3, \ldots, 19) > \\ &\qquad x\Big[(1-\frac{1}{2}) + \ldots + (1-\frac{1}{19}) - (\frac{e\cdot 1.455}{4})^4\Big] - 8^4 \qquad (3.3.7.1) \\ &\qquad \approx x[0.171 - (0.9888)^4] - 4096 \approx x[-0.785] - 4096 \end{aligned}$$

Thus, even with this improved lower bound formula, depending on the particular values x, r, and m, we could still end up with a negative lower bound. Note that in this example even if x is "sufficiently large" the lower bound will still be negative.

3.4 Discarding Terms "to the Right of Vertical Lines"

Brun's Example 3, shown above in Figure 3.2.1.4, shows what he means by discarding terms "to the right of vertical lines." As in the previous sieve where we "discard terms in small print," there does exist an algorithm to identify such terms, though Brun's complicated presentation of that algorithm is rather difficult to understand. Adding to that difficulty are the challenges in understanding the calculations Brun subsequently makes and the additional results he exploits (theorems from Stirling and Mertens) to arrive at his chief results in this section.

Notwithstanding these difficulties and challenges, though, some definite markers can be laid down to help readers find their way through Brun's arguments. In particular, the steps Brun takes in §4 are broadly similar to the steps he takes in §3: beginning with a form of the fundamental inequality $N > M - R$ for his sieve, Brun bounds M below by A and bounds R above by B, thus yielding a new lower bound $N > A - B$.

There are three main differences between the steps Brun takes in §3 and the steps he takes in §4. First, Brun uses theorems from Mertens to establish certain relationships among his sifting primes when the first sifting prime p_1 is sufficiently large. (These relationships involve the use of constants $\alpha > 0$ and $\alpha_0 > \alpha$, to which Brun later gives on p. 120 the specific values of 1.5 and 1.51, respectively.) Second, the route that Brun takes to establish a lower bound for M is much more convoluted in §4 than the corresponding route in §3. This is because Brun makes use of many more auxiliary terms and inequalities, which incorporate previous results from Mertens and Stirling in ways that are not blindingly obvious. Third, Brun's calculation of the bound on the remainder term R in §4 uses auxiliary terms and an argument that is more complicated than the counting argument he used to bound the remainder term in §3.

3.4.1 Algorithm for Discarding Terms "to the Right of Vertical Lines"

In this section we describe the basic *idea* behind the algorithm for the sieve Brun uses in §4, and we sketch the correspondence between the *notation* Brun uses for this sieve and the *idea* behind the sieve. In Section 3.4.2 we give an argument for why this sieve does give a lower bound for N.

3.4.1.1 The Basic Idea Behind the Algorithm

Brun begins §4 by presenting an algorithm that gives a lower bound for N when discarding terms "to the right of vertical lines." Brun's development of this algorithm uses notation not discussed yet in this thesis, so we need a separate argument (Section 3.4.2) for why this algorithm does in fact provide a lower bound for N. Nonetheless, here we can pick up Brun's thread in his formula 14′ on p. 115, which is shown as 3.4.1.1 below.

$$N(D,x,p_1,\ldots,p_r) > \frac{x}{D}\Big[1 - S_1 + S_2 - \ldots - S_{2n-1}\Big] - R. \tag{3.4.1.1}$$

One notes immediately the similarity of this formula with formula (8) from §3 (p. 112):

$$N(D,x,p_1,\ldots,p_r) > \frac{x}{D}\Big[1 - \Sigma_1 + \Sigma_2 - \ldots - \Sigma_m\Big] - R. \tag{3.4.1.2}$$

(In both formulas, R is the number of terms in the main-term expansion, and for our purposes, we can take $D = 1$ and assume that we're sifting a sequence of consecutive integers that start at 1.)

The similarity of 3.4.1.1 and 3.4.1.2 is twofold: the S-terms and the Σ-terms are constructed as the sums of products of reciprocals of primes, and in each formula the truncation of the expansion is done at an odd-indexed term ($2n - 1$ in the case of the S-terms, and m in the case of the Σ-terms, which, recall from Section 3.3.2, is assumed to be odd).

The key difference between the two formulas lies in how the S-terms and Σ-terms are constructed: each even-indexed S-term contain fewer summands than the corresponding even-indexed Σ-term, as can be seen by comparing the diagrams Brun uses to describe how the S-terms (p. 117) and the Σ-terms (pp. 111-112) are constructed. Example 3.4.1.1 below uses Brun's notation to describe a particular case where the number of terms in S_2 is less than the number of terms in Σ_2. The significance of the differences between S-terms and Σ-terms comes into play in the consequences Brun derives on pp. 118-120 based on certain additional assumptions and other results he provides on pp. 116-117 (see Section 3.4.4).

Later in §4, Brun uses certain initial assumptions about the distribution of the sifting primes, along with particular results by Stirling and Mertens, to find a lower bound A for the main term of the expansion and an upper bound B for the remainder term, which, taken together, allow him to show that $N > A - B$. The computations for the lower bound for N that Brun establishes in §4 are then modified as appropriate in §6 to yield a lower bound for N that allows Brun to establish his theorem on Goldbach's Conjecture.

3.4.1.2 Brun's Notation for the Algorithm

Recall from Section 3.4.1.1 this algorithm results in fewer summands being included in the S-terms that remain after we truncate the main-term expansion at an odd number (than there were in the corresponding Σ-terms resulting from the previous sieve). The notation Brun uses (in formula (14) on p. 115) to indicate which summands remain in S_2 is

$$\sum_{a \le r} \sum_{\substack{b<a \\ b<t}} \frac{1}{p_a p_b}.$$

We can make some immediate sense of this notation by recalling that r is the index of the rth sifting prime p_r and noting that the index a of p_a and the index b of p_b are to be less than r in some specified way.

In particular, when sifting by primes $2, 3, 5, 7, \ldots, p_r$, the above notation is intended to specify that the index a can range over the set $\{2, 3, 4, \ldots, r\}$, corresponding to the fact that p_a can range over the set of primes $\{3, 5, 7, 11, \ldots, p_r\}$. Thus, when $a = 2$, $p_a = 3$; when $a = 3$, $p_a = 5$; when $a = 4$; $p_a = 7$, and so forth. Similarly, the index b on p_b can range over the set $\{1, 2, 3, \ldots, t-1\}$, corresponding to the fact that p_b can range over the set $\{2, 3, \ldots, p_{t-1}\}$, where $t < a$ (which Brun assumes but does not state explicitly).

To better understand how Brun's *notation* for this algorithm corresponds to the *idea* behind the algorithm, consider the following example, which shows that Brun's notation does indeed reduce the number of terms in the S-terms remaining after the main-term expansion is truncated at an odd-indexed S-term.

Example 3.4.1.1. *Suppose we sift the consecutive numbers* $1, 2, 3, \ldots, 170$ *by the six primes* $2, 3, 5, 7, 11,$ *and* 13*, where* $p_r = p_6 = 13$*, the sixth sifting prime, and suppose that we truncate the main-term expansion at* S_5*. Show that the application of Brun's algorithm with* $t = 4$ *results in a smaller number of summands of the form* $\sum_{a \le r} \sum_{\substack{b<a \\ b<t}} \frac{1}{p_a p_b}$ *(these are the summands in* S_2*) compared to the "full" number of summands of the form* $\sum\sum_{b<a\le r} \frac{1}{p_a p_b}$ *from the previous algorithm (these are the summands in* Σ_2 *for the sieve that discarded terms "in small print").*

The "full" number of summands in Σ_2 in the expansion of the main term, according to the specification $\sum\sum_{b<a\le r} \frac{1}{p_a p_b}$ for the previous algorithm, is 15, since there are $\binom{6}{2}$ ways of choosing p_a and p_b according to this specification.

Now, apply Brun's algorithm, truncating the expansion at S_5 according to the specification $\sum_{a\le r}\sum_{\substack{b<a\\b<t}}\frac{1}{p_a p_b}$, with t equal to 4. Then there are only 12 summands of the specified form, because the index b must be less than t, which equals 4. That is, the indices for b range over the set $\{1,2,3\}$, which means that the primes p_b in the summands $\frac{1}{p_a p_b}$ range over the set $\{2,3,5\}$. Therefore, the primes 7 and 11 are not allowable values for p_b, and so the three terms $\frac{1}{13\cdot 11}$, $\frac{1}{13\cdot 7}$, and $\frac{1}{11\cdot 7}$ are not in the summands of the form $\frac{1}{p_a p_b}$ according to the specification $\sum_{a\le r}\sum_{\substack{b<a\\b<t}}\frac{1}{p_a p_b}$. Thus, three positive terms are removed from what was, in the sieve that discarded terms"in small print," the sum of fifteen terms of the form $\frac{1}{p_a p_b}$. □

Brun's notation for this sieve also specifies how the number of odd-indexed S-terms is reduced, as the next example shows, which considers summands of the form $\frac{1}{p_a p_b p_c}$ according to the specification $\sum_{a\le r}\sum_{\substack{b<a\\b<t}}\sum_{\substack{c<b\\c<t}}\frac{1}{p_a p_b p_c}$.

Example 3.4.1.2. *Suppose we sift the consecutive numbers* $1,2,3,\ldots,170$ *by the six primes* $2,3,5,7,11,$ *and* 13*, where* $p_r = p_6 = 13$*, the sixth sifting prime, and suppose that we truncate the main-term expansion at* S_5*. Show that the application of Brun's algorithm with* $t = 4$ *results in a smaller number of summands in* S_3*, i.e., summands of the form* $\sum_{a\le r}\sum_{\substack{b<a\\b<t}}\sum_{\substack{c<b\\c<t}}\frac{1}{p_a p_b p_c}$ *, compared to the "full" number of summands of the form* $\sum\sum\sum_{c<b<a\le r}\frac{1}{p_a p_b}$ *from the previous algorithm (these are the summands in* Σ_3 *for the sieve that discarded terms "in small print").*

The "full" number of summands in Σ_3 in the expansion of the main term, according to the specification $\sum\sum\sum_{c<b<a\le r}\frac{1}{p_a p_b}$ for the previous algorithm, is 20, since there are $\binom{6}{3}$ ways of choosing p_a, p_b, and p_c according to this specification.

Now, applying Brun's specification (as was done in Example 3.4.1.1), we see that $b < t = 4$ as before, and now we have also that $c < b$. Thus, the indices c range only over the set $\{1,2\}$, and therefore, the primes p_c range only over the set $\{2,3\}$. Therefore, the following summands *are* in Σ_3 in the sieve that discards terms in "small print," but *are not* in S_3: $\frac{1}{13\cdot 11\cdot 7}$, $\frac{1}{13\cdot 11\cdot 5}$, $\frac{1}{13\cdot 7\cdot 5}$, and $\frac{1}{11\cdot 7\cdot 5}$. Thus, S_3 contains only sixteen summands, whereas Σ_3 contains twenty summands. □

Consider again Brun's Example 3, as presented in Figures 3.2.1.4 and 3.4.1.2 (from p. 110), below.

Ex. 3) $N(1, x, 2, 3, 5, 7, 11, 13, 17, 19) >$

$$x\Big[1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} - \frac{1}{7} - \frac{1}{11} - \frac{1}{13} - \frac{1}{17} - \frac{1}{19}$$

$$+ \frac{1}{3\cdot 2}$$

$$+ \frac{1}{5\cdot 2} + \frac{1}{5\cdot 3}\left(1 - \frac{1}{2}\right)$$

$$+ \frac{1}{7\cdot 2} + \frac{1}{7\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{7\cdot 5}\begin{bmatrix} 1 - \frac{1}{2} - \frac{1}{3} \\ + \frac{1}{3\cdot 2} \end{bmatrix}$$

$$+ \frac{1}{11\cdot 2} + \frac{1}{11\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{11\cdot 5}\begin{bmatrix} 1 - \frac{1}{2} - \frac{1}{3} \\ + \frac{1}{3\cdot 2} \end{bmatrix} + \frac{1}{11\cdot 7}\left(\begin{array}{l} 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} \\ + \frac{1}{3\cdot 2} \\ + \frac{1}{5\cdot 2} \end{array}\right)$$

$$+ \frac{1}{13\cdot 2} + \frac{1}{13\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{13\cdot 5}\begin{bmatrix} 1 - \frac{1}{2} - \frac{1}{3} \\ + \frac{1}{3\cdot 2} \end{bmatrix} + \frac{1}{13\cdot 7}\left(\begin{array}{l} 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} \\ + \frac{1}{3\cdot 2} \\ + \frac{1}{5\cdot 2} \end{array}\right)$$

$$+ \frac{1}{17\cdot 2} + \frac{1}{17\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{17\cdot 5}\begin{bmatrix} 1 - \frac{1}{2} - \frac{1}{3} \\ + \frac{1}{3\cdot 2} \end{bmatrix} + \frac{1}{17\cdot 7}\left(\begin{array}{l} 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} \\ + \frac{1}{3\cdot 2} \\ + \frac{1}{5\cdot 2} \end{array}\right)$$

$$+ \frac{1}{19\cdot 2} + \frac{1}{19\cdot 3}\left(1 - \frac{1}{2}\right) + \frac{1}{19\cdot 5}\begin{bmatrix} 1 - \frac{1}{2} + \frac{1}{3} \\ + \frac{1}{3\cdot 2} \end{bmatrix} + \frac{1}{19\cdot 7}\left(\begin{array}{l} 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{5} \\ + \frac{1}{3\cdot 2} \\ + \frac{1}{5\cdot 2} \end{array}\right)\Big] - 72$$

$$= 0{,}163\,x - 72$$

FIGURE 3.4.1.1: Lower bound for $N(1, x, 2, 3, 5, \ldots, 19)$: dropping terms "to the right of vertical lines"

Here we have set aside the terms on the right of the vertical lines. One see that the expression is of the form

$$1 - \sum \frac{1}{p_a} + \sum \sum \frac{1}{p_a p_b} - \sum \sum \sum \frac{1}{p_a p_b p_c} + \sum \sum \sum \sum \frac{1}{p_a p_b p_c p_d} \quad ,$$

where p_a, p_b, p_c and p_d run through the following values

p_a	2	3	5	7	11	13	17	19
p_b	2	3	5	7				
p_c	2	3	5	7				
p_d	2							

in which $a > b > c > d$.

FIGURE 3.4.1.2: Specification for dropping terms "to the right of vertical lines"

There are two issues to point out with this example. First, although the truncation of the main-term stops after *adding* rather than subtracting a group of terms, it turns out that, numerically, the lower bound achieved in this example is indeed an improvement on the default lower bound. Specifically, for this example the default lower bound for N works out to be $0.171x - 256$, and Brun's new lower bound for N is $0.163x - 72$. This increase in the lower bound for N speaks to the power of this sieve, since by PIE considerations alone, we might be tempted to conclude that when we truncate the main-term expansion after *adding* some terms, the result will be an overcount of the number of elements that fall through the sieve, i.e., an upper bound for N. Yet, in this case, because the reduction in the number of the even-indexed S-terms is so great, this sieve does indeed give a lower bound for N, as Brun's calculations show.

Second, Brun indicates that the subscripts a, b, c, and d on the primes are related as follows: $a > b > c > d$. This is true enough, although when Brun presents this example on pages 109 and 110 of §2, he has not yet supplied the general notation for this sieve that he gives in §4. Thus, it is difficult to understand the significance of $a > b > c > d$ in this example.

However, with the description given above of Brun's notation and with Brun's example 3.4.1.1 at hand, we can understand the indices a, b, c, d to be ranging over the sets $\{1, 2, 3, 4, 5, 6, 7, 8\}$, $\{1, 2, 3, 4\}$, $\{1, 2, 3, 4\}$, and $\{1\}$, respectively, and the requirement that $a > b > c > d$ to be specifying the allowable values of the indices on the primes p_a, p_b, p_c, and p_d when constructing the products of the reciprocals of those primes.

3.4.2 Brun's Sieve in §4 Does Give a Lower Bound for N

Recall that Brun's second sieve discards terms "to the right of a vertical line." This means that the full expansion of N is truncated by removing certain terms from the expansion. In this section, we revert to Brun's original notation and explain how he uses subscripts on the summations of the expansion of N to discard terms "to the right of a vertical line." In particular, we show why the following inequality, which is Brun's formula (14) on p. 115 of Rui's translation, holds:

$$N(D,x,p_1,\ldots,p_r) > \frac{x}{D}\Big[1 - \sum_{a\le r}\frac{1}{p_a} + \sum_{a\le r}\sum_{\substack{b<a\\b<t}}\frac{1}{p_ap_b} - \sum_{a\le r}\sum_{\substack{b<a\\b<t}}\sum_{\substack{c<b\\c<t}}\frac{1}{p_ap_bp_c} + \sum_{a\le r}\sum_{\substack{b<a\\b<t}}\sum_{\substack{c<b\\c<t}}\sum_{\substack{d<c\\d<u}}\frac{1}{p_ap_bp_cp_d} - \ldots\Big] - R, \tag{3.4.2.1}$$

where D is the common difference between terms (and is relatively prime to the sifting primes $p_1, p_2, \ldots, p_n$), t is a whole number less than r, u is a whole number less than t, and R is the total number of terms inside the brackets. (Note that in Rui's version there is a typo in the first summation inside the brackets: he has $\frac{1}{p_r}$ instead of $\frac{1}{p_a}$.) En route to showing why formula (14) holds, we discuss Brun's formulas $(1), (2), (3), (3'), (4)$, and (13).

The key point is that when Brun speaks of "discarding terms to the right of a vertical line" he means discarding *positive* terms from the right-hand side of an *equation* for N. The result, after those terms are dropped, is an *inequality*, which gives a lower bound for N.

This point is brought out most clearly through examples that sift a sequence of consecutive integers from 1 through x, i.e., examples that are arithmetic sequences that start at 1 and that have common difference $D = 1$. This kind of example should be familiar from earlier sections of the thesis [4].

[4] In Brun's early formulas and derivations he mentions that the starting point of the sequences he considers is some positive Δ such that $\Delta \le D$ (Rui, p. 103). In other, simplified formulas he omits the starting value Δ but keeps the common difference D.

3.4.2.1 Examples Illustrating Key Formulas from Brun's §2

Brun's formula (1) (p. 102) is:

$$N(D, x, p_1, \ldots, p_r) = N(D, x, p_1, \ldots, p_{r-1}) - N(D \cdot p_r, x, p_1, \ldots, p_{r-1}), \qquad (3.4.2.2)$$

where $N(D, x, p_1, \ldots, p_r)$ is the number of numbers that fall through a sieve operating on an arithmetic sequence with common difference D relatively prime to each of the sifting primes $p_1, p_2, \ldots, p_r$ (pp. 100-102). Brun uses an implicit starting point of $\Delta \in \mathbb{Z}+$ in his notation, which has not been particularly relevant to our discussion so far.

Up to this point in this thesis, we have been using the 1 in $N(1, 37, 2, 3, 5)$ to indicate that the sequence to be sifted begins at 1, rather than that its common difference is 1. In what follows, we still use the same notation but now we let 1 be the "D" in Brun's notation, and we make the background assumption that the numbers to be sifted also start at 1. Thus, the number 1 does double duty in our notation: it indicates both the common difference D of the arithmetic sequence to be sifted and also the starting point of that sequence (which in Brun's most general notation is Δ).

3.4.2.2 Deriving Brun's Formula (2) from his Formula (1)

We use an example to show how Brun's formula (2) follows from successive applications of his formula (1). Consider a sieve that sifts the consecutive integers from 1 through 37 by the primes 2, 3, and 5. To find $N(1, 37, 2, 3, 5)$ we can use Brun's formula (1) to obtain:

$$N(1, 37, 2, 3, 5) = N(1, 37, 2, 3) - N(1 \cdot 5, 37, 2, 3). \qquad (3.4.2.3)$$

In words, the number of numbers falling through a sieve that sifts all integers from 1 through 37 by the primes 2, 3, and 5 *equals* the number of numbers falling through a sieve that sifts the integers from 1 through 37 by the primes 2 and 3 *minus* the number of numbers that fall through a sieve that sifts multiples of 5 in the sequence $1, 2, 3, \ldots, 37$ by the primes 2 and 3. Using # to indicate cardinality gives

$\#\{1, 7, 11, 13, 17, 19, 23, 29, 31, 37\} =$
$\#\{1, 5, 7, 11, 13, 17, 19, 23, 25, 29, 31, 35, 37\} - \#\{5, 25, 35\}$, or $10 = 13 - 3$.

Applying Brun's formula (1) again to expand the term $N(1, 37, 2, 3)$ in 3.4.2.3 gives

$$N(1, 37, 2, 3) = N(1, 37, 2) - N(1 \cdot 3, 37, 2). \tag{3.4.2.4}$$

In words, the number of numbers falling through a sieve that sifts the integers from 1 through 37 by 2 and 3 *equals* the number of numbers falling through a sieve that sifts the integers from 1 through 37 by 2 *minus* the number of numbers falling through a sieve that sifts multiples of 3 in the sequence $1, 2, 3, \ldots, 37$ by 2.

In other words, the number of integers from 1 through 37 that are not multiples of 2 or 3 **equals** the number of integers from 1 through 37 that are not multiples of 2 (i.e., the odd numbers from 1 through 37) **minus** the number of multiples of 3 in the sequence $1, 2, 3, \ldots, 37$ that are not multiples of 2. Using # to indicate cardinality gives
#{1,5,7,11,13,17,19,23,25,29,31,35,37} =
#{1,3,5,7,9,11,13,15,17,19,21,23,25,27,29,31,33,35,37} - #{3,9,15,21,27,33}.
That is, $13 = 19 - 6$.

We can also decompose the term $N(1, 37, 2)$ in 3.4.2.4 to obtain

$$N(1, 37, 2) = N(1, 37) - N(1 \cdot 2, 37). \tag{3.4.2.5}$$

In 3.4.2.5, $N(1, 37)$ equals the number of integers from 1 through 37 that are multiples of the common difference 1, i.e., just the number of integers from 1 through 37. And $N(1 \cdot 2, 37)$ is the number of multiples of 2 in the range $1, 2, 3, \ldots, 37$ (i.e., the even numbers from 1 through 37). Thus, 3.4.2.5 can be read as "the number of odd numbers from 1 through 37 equals 37 minus the number of even numbers from 1 through 37."

Putting together the decompositions so far, we have that
$N(1, 37, 2, 3, 5) = N(1, 37) - N(1 \cdot 2, 37) - N(1 \cdot 3, 37, 2) - N(1 \cdot 5, 37, 2, 3)$ [5], which is a specific example of Brun's formula (2) at the bottom of p. 103:
$N(D, x, p_1, \ldots, p_r) = N(D, x) - N(D \cdot p_1, x) - N(D \cdot p_2, x, p_1) - \ldots - N(D \cdot p_r, x, p_1, \ldots, p_r)$.

3.4.2.3 From Brun's Formula (2) to Formula (3)

Examples of Brun's equations (2) and (3), respectively, are:

$N(1, 37, 2, 3, 5) = N(1, 37) - N(1 \cdot 2, 37) - N(1 \cdot 3, 37, 2) - N(1 \cdot 5, 37, 2, 3)$, and

[5] We can verify this by seeing that $10 = 37 - 18 - 6 - 3$.

$$N(1,37,2,3,5) = N(1,37) - N(1\cdot 2,37) - N(1\cdot 3,37) - N(1\cdot 5,37)$$
$$+N(1\cdot 3\cdot 2,37) + N(1\cdot 5\cdot 2,37) + N(1\cdot 5\cdot 3,37,2).$$

The second equation should look familiar from earlier discussions of the PIE.

To show that these two equations give the same value for $N(1,37,2,3,5)$, we show that the right-hand sides of two equations are equal. We show that the right-hand sides of these equations are equal by showing that the terms following $N(1,37)-N(1\cdot 2,37)$ in the first equation evaluate to the same number as the terms folowing $N(1,37)-N(1\cdot 2,37)$ in the second equation.

In the first equation, the value of $N(1\cdot 3,37,2)$ is 6, which is the number of multiples of 3 in the sequence $1,2,3,4,\dots,37$ that are not multiples of 2, i.e., the cardinality of $\{3,9,15,21,27,33\}$. Also in the first equation, the value of $N(1\cdot 5,37,2,3)$ is 3, the number of multiples of 5 in the sequence $1,2,3,4,\dots,37$ that are not also multiples of 2, is 3, i.e., the cardinality of $\{5,25,35\}$. So, from the difference $N(1,37)-N(1\cdot 2,37)$ on the right-hand side of the first equation we are subtracting a total of 9.

In the second equation, the value of $N(1\cdot 3,37)$ is 12, which is the number of multiples of 3 in the sequence $1,2,3,4,\dots,37$. Also, $N(1\cdot 5,37)$ in the second equation equals 7. Finally in the second equation, $N(1\cdot 3\cdot 2,37)=6$, $N(1\cdot 5\cdot 2,37)=3$, and $N(1\cdot 5\cdot 3,37,2)=1$. Thus, in the second equation the terms following $N(1,37)-N(1\cdot 2,37)$ evaluate to $-12-7+6+3+1=-9$, which matches the amount in the first equation that is subtracted from $N(1,37)-N(1\cdot 2,37)$.

Next, we re-format the second equation (which is of the form of Brun's formula (3)), so the reader can better see how discarding positive terms to the right of an appropriate vertical line would give a lower bound for N. Thus reformatted, our particular example looks like

$$\begin{aligned} N(1,37,2,3,5) = N(1,37) & \\ - N(1\cdot 2,37) & \\ - N(1\cdot 3,37) & \\ - N(1\cdot 5,37) & \\ & + N(1\cdot 3\cdot 2,37) \\ & + N(1\cdot 5\cdot 2,37) + N(1\cdot 5\cdot 3,37,2). \end{aligned} \tag{3.4.2.6}$$

Studying this newly formatted equation, we can see why discarding terms "to the right of a vertical line" (judiciously placed) would be the same as discarding certain terms after that follow a plus sign. (We'll call all terms in this equation and in Brun's formula (3), "N-terms.")

Notice in the above equation that the N-term $N(1 \cdot 5 \cdot 3, 37, 2)$, which can be seen to lie "to the right of a vertical line," is positive. Also positive are the N-terms $N(1 \cdot 3 \cdot 2, 37)$ and $N(1 \cdot 5 \cdot 2, 37)$, which, along with the N-term $N(1 \cdot 5 \cdot 3, 37, 2)$, lie "to the right of a (different) vertical line." Thus, discarding terms from the equation that lie to the right of either of those vertical lines will yield a lower bound for $N(1, 37, 2, 3, 5)$.

The next section explains why, in general, dropping N-terms that "lie to the right of a vertical lines" from Brun's formula (3) yields a lower bound for N.

3.4.2.4 The "N-terms" in Formula (3)

Recall that the value of each N-term is the number of terms that fall through a particular sieve, which is described by that N-term. Although that number cannot be negative, it seems at first glance that it could be zero.

Brun handles this possibility in two ways. First, his definitions of the N-terms are such that each *defined* N-term is positive. Second, Brun uses circumspect language when talking about dropping terms to the right of a vertical line.

The result is that dropping terms to the right of a vertical line in formula (3) means dropping terms that are either positive or undefined (and so essentially have a value of zero in the arithmetic on the right-hand side of (3)).

In Rui's translation, Brun defines $N(D, x)$ as the "numbers [sic] of the terms between 0 and x of the progression $\Delta \quad \Delta + D \;\; \Delta + 2D \;\; \ldots \;\; \Delta + \lambda D$, where $0 < \Delta \leq D, \;\; \Delta + \lambda D \leq x < \Delta + (\lambda + 1)D$" (Rui, p. 103).

Note the last inequality, which gives bounds on x. That inequality shows that x must be at least as large as some multiple of D plus some offset. Given the context of Brun's notation in §2, it is reasonable to suppose that λ, the multiple of D, is a positive integer. It is also reasonable to suppose that the D in the definition of $N(D, x)$ on p. 103 can just as well refer to a Dp_i value such as those found in the first line of formula (3) at the top of p. 104. A further reasonable supposition is that both the D in $N(D, x)$

and $N(D,x)$ itself can also just as well refer to any of the N-terms in formula (3) or in subsequent expansions based on (3), such as the un-numbered formula between formulas (13) and (14).

Based on these three reasonable suppositions, we are led to conclude that each N-term is defined only when the value of x is at least as large as positive integral multiple of its corresponding "D-component", e.g., Dp_rp_{r-1}.

Thus, it is reasonable to say that all the N-terms on the right-hand side of (3) are either positive or undefined (in which case they "count for zero" as far as arithmetic is concerned).

Also note that Brun never makes the claim that *all* the terms to the right of a vertical line in formula (3) are positive. His language is rather more delicate.

Brun's first mention of discarding terms "to the right of a vertical line" is in §2, where he says:

> When the question is to determine a lower bound for $N(D,x,p_1,\ldots,p_r)$ we can set aside as many *positive* terms as we want in the formula (3). One can choose these terms in several different ways ... for example, the terms which lie on the right of a vertical line. (Rui, p. 104, emphasis added)

Later, in the second paragraph of §4, Brun says, "At first we set aside in the formula (3) all positive terms on the right on (sic) a vertical line" (Rui, p. 114). The result is formula (13), which gives a lower bound for N.

So, based on Brun's definitions of the N-terms and his circumspect language on pages 104 and 114 about discarding positive terms "to the right of a vertical line," the fact that some of the N-terms may count for naught simply doesn't matter in his second sieve's procedure for creating a lower bound for N.

3.4.2.5 An Example Illustrating Brun's Formula (13)

Brun's formula (13) is an inequality giving a lower bound for N:

$$\begin{aligned} N(D, x, p_1, p_2, \ldots, p_r) > N(D, x) - \sum_{a \leq r} N(D \cdot p_a, x) \\ + \sum_{a \leq r} \sum_{\substack{b<a \\ b<t}} N(D \cdot p_a \cdot p_b, x, p_1, \ldots, p_{b-1}) \end{aligned} \tag{3.4.2.7}$$

This section describes in more detail how the subscripts in formula (13) work.

Consider the subscript b in formula (13) below, and notice that $b < t$, where t is a whole number less than r. Thus, Brun's notational device operates, in effect, to drop one or more positive terms from the right-hand side of his formula (3). Let's see how.

Now expand $N(D, 1, p_1, p_2, p_3, p_4, p_5)$ according to Brun's formula (13), with $r = 5$ and $t = 4$. Since $t = 4$ and $b < a$, the subscript b can run through only the values 1, 2, and 3, which means that the primes p_b can run through only the values p_1, p_2, and p_3. Thus, Brun's formula (13) in this case yields

$$\begin{aligned} &N(D, x, p_1, \ldots, p_r) \\ &> N(D, x) \\ &-N(D \cdot p_1, x) \\ &-N(D \cdot p_2, x) \\ &-N(D \cdot p_3, x) \\ &-N(D \cdot p_4, x) \\ &-N(D \cdot p_5, x) \\ &+ N(D \cdot p_2 \cdot p_1, x) \\ &+ N(D \cdot p_3 \cdot p_1, x) + N(D \cdot p_3 \cdot p_2, x, p_1) \\ &+ N(D \cdot p_4 \cdot p_1, x) + N(D \cdot p_4 \cdot p_2, x, p_1) + N(D \cdot p_4 \cdot p_3, x, p_1, p_2) \\ &+ N(D \cdot p_5 \cdot p_1, x) + N(D \cdot p_5 \cdot p_2, x, p_1) + N(D \cdot p_5 \cdot p_3, x, p_1, p_2) \end{aligned} \tag{3.4.2.8}$$

where the sole positive term that has been dropped to the right of a vertical line is $N(D \cdot p_5 \cdot p_4, x, p_1, p_2, p_3)$, which would have been the last term included in the equality for Brun's formula (3).

So, instead of using in the expansion of the right-hand side of the above inequality all $\binom{5}{2} = 10$ terms for the different combinations of $p_a\ p_b$, as we would do for the first sieve if we didn't truncate its expansion (see the terms in the "triangular array" of pairs of terms $p_a p_b$ that Brun's gives on the bottom of p. 104 of Rui), we truncate that full group of ten terms by eliminating that one term to the right of a particular vertical line that truncates the expansion before the last (positive) term can be added.[6]

Now let's see how the numbers in the above inequality work in a specific example. Concretely, let's consider

$$\begin{aligned}
&N(1, 125, 2, 3, 5, 7, 11)\\
&\quad > N(1, 125)\\
&\quad -N(1 \cdot 2, 125)\\
&\quad -N(1 \cdot 3, 125)\\
&\quad -N(1 \cdot 5, 125)\ -N(1 \cdot 7, 125)\ -N(1 \cdot 11, 125)\\
&\quad + N(1 \cdot 3 \cdot 2, 125)\\
&\quad + N(1 \cdot 5 \cdot 2, 125) + N(1 \cdot 11 \cdot 3, 125, 2)\\
&\quad + N(1 \cdot 7 \cdot 2, 125) + N(1 \cdot 7 \cdot 3, 125, 2) + N(1 \cdot 7 \cdot 5, 125, 2, 3)\\
&\quad + N(1 \cdot 11 \cdot 2, 125) + N(1 \cdot 5 \cdot 3, 125, 2) + N(1 \cdot 11 \cdot 5, 125, 2, 3),
\end{aligned} \tag{3.4.2.9}$$

where the sole positive term that has been dropped to the right of a vertical line is $N(1 \cdot 11 \cdot 7, 125, 2, 3, 5)$.

The result is that $N(1, 125, 2, 3, 5, 7, 11)$

> 125

-64

-41

[6] Note that, as Brun himself says, the terms dropped are positive: "D'abord nous écartons dans la formule (3) tous terms *positifs* à droite d'une ligne verticale" (Brun, p. 17, emphasis added).

$$
\begin{aligned}
&-25\\
&-17\\
&-11\\
&+20\\
&+12+2\\
&+8+3+1\\
&+5+4+1.
\end{aligned}
\tag{3.4.2.10}
$$

Simplifying, we have $N(1, 125, 2, 3, 5, 7, 11) > 25$. Now, the value of the term discarded "to the right of a vertical line" is $N(1 \cdot 11 \cdot 7, 125, 2, 3, 5)$, which equals 1. So, adding 1 to 25 gives 26, which should be the exact value of $N(1, 125, 2, 3, 5, 7, 11)$. Is it? Recall that $N(1, 125, 2, 3, 5, 7, 11)$ is the number of numbers that pass through a sieve of consecutive integers from 1 to 125 when sifting by the primes 2, 3, 5, 7, and 11. That number is just $\pi(125) - \pi(\sqrt{125}) + 1$, which is 26 since $\pi(125) = 30$, and $\pi(\sqrt{125}) = 5$.

3.4.2.6 Brun's Formula (14) follows from his Formula (13)

Brun applies formula (13) twice to give this intermediate result, which is not numbered:

$$
\begin{aligned}
N(D, x, p_1, p_2, \ldots, p_r) > N(D, x) \\
&- \sum_{a \leq r} N(D \cdot p_a, x) \\
&+ \sum_{a \leq r} \sum_{\substack{b<a \\ b<t}} N(D \cdot p_a \cdot p_b, x) \\
&- \sum_{a \leq r} \sum_{\substack{b<a \\ b<t}} \sum_{\substack{c<b \\ c<t}} N(D \cdot p_a \cdot p_b \cdot p_c, x) \\
&+ \sum_{a \leq r} \sum_{\substack{b<a \\ b<t}} \sum_{\substack{c<b \\ c<t}} \sum_{\substack{d<c \\ d<u}} N(D \cdot p_a \cdot p_b \cdot p_c \cdot p_d, x, p_1, \ldots, p_{d-1})
\end{aligned}
\tag{3.4.2.11}
$$

where u is a whole number less than t (p. 115 of [44]).

From the above inequality, using the information that $N(d, x) = \frac{x}{d} + \theta$, where $-1 \leq \theta < 1$, Brun derives his formula (14):

$$N(D,x,p_1,\ldots,p_r) > \frac{x}{D}\Big[1 - \sum_{a\le r}\frac{1}{p_a} + \sum_{a\le r}\sum_{\substack{b<a\\b<t}}\frac{1}{p_ap_b} - \sum_{a\le r}\sum_{\substack{b<a\\b<t}}\sum_{\substack{c<b\\c<t}}\frac{1}{p_ap_bp_c} + \sum_{a\le r}\sum_{\substack{b<a\\b<t}}\sum_{\substack{c<b\\c<t}}\sum_{\substack{d<c\\d<u}}\frac{1}{p_ap_bp_cp_d} - \ldots\Big] - R, \tag{3.4.2.12}$$

where D is the common difference between terms (and is relatively prime to the sifting primes $p_1, p_2, \ldots, p_n$), t is a whole number less than r, u is a whole number less than t, and R is the total number of terms inside the brackets. (Note that in Rui's version there is a typo in the first summation inside the brackets: he has $\frac{1}{p_r}$ instead of $\frac{1}{p_a}$.)

To arrive at (14) from the intermediate result above, Brun essentially factors out $\frac{x}{D}$ from each term on the right-hand side of the intermediate formula, and then, since each remainder θ has absolute value less than or equal to 1, he maintains the inequality by subtracting the cumulative worst-case θ values (collected in the R, the number of terms in the expansion) from the expression inside the brackets. Finally, note that Brun uses the ellipsis *not* to indicate infinitely many additions and subtractions, but rather to indicate some finite number of PIE-like terms in the expansion. (Evidently, because of the way Brun uses subscripts, he cannot specify precisely the last term in the expansion.)

3.4.3 Overview of Brun's Steps in §4

Brun begins with inequality (14′) on p. 115, which upon substituting E_n for $1 - S_1 + S_2 - \ldots - S_{2n-1}$, becomes

$$N(D,x,p_1,\ldots,p_r) > \frac{x}{D}[E_n] - R. \tag{3.4.3.1}$$

Using results of Stirling and Mertens, Brun determines a lower bound for E_n. (Section 3.4.6 gives an overview of the steps Brun uses to derive a lower bound for E_n.) Denote that lower bound of E_n by A. Thus, $E_n > A$, and so by 3.4.3.1

$$N(D,x,p_1,\ldots,p_r) > \frac{x}{D}[A] - R. \tag{3.4.3.2}$$

Next, Brun determines an upper bound for R, the number of terms in his expansion of the main term. Call that upper bound B. Thus, $B > R$, and so by 3.4.3.2

$$N(D,x,p_1,\ldots,p_r) > \frac{x}{D}[A] - B. \tag{3.4.3.3}$$

This last result (formula (20) on p. 121) "is valid for all successive prime numbers $p_1, \ldots, p_r$ with $p_1 \geq p_e$, where p_e denotes a determinable prime number" (p. 121). In other words, the result holds when the smallest sifting prime p_1 is "sufficiently large," which implies that it holds for x sufficiently large, since by assumption of how the sieve operates, if the smallest sifting prime is "sufficiently large" so, too, must x be, since all the sifting primes are less than x. See the discussion in 3.4.4.

Once Brun has a lower bound for $N(D, x, p_1, \ldots, p_r)$ where p_1 is a sufficiently large prime he then extends his methods to determine a lower bound for $N(D, x, 2, 3, \ldots, p_r)$, where the smallest sifting prime is 2 and the largest sifting prime p_r is sufficiently large (p. 122, inequality (22)). The lower bound for $N(D, x, 2, 3, \ldots, p_r)$ allows Brun to obtain the following result near the end of §4 (p. 122):

$$N(D, x, 2, 3, \ldots, p(x^{1/6})) > \frac{1.008x}{\log\, x}, \tag{3.4.3.4}$$

which allows him to state that when sifting x consecutive numbers by the primes $2, 3, \ldots, p(x^{1/6})$ there remain more than $\frac{x}{\log\, x}$ terms, provided $x > x_0$, where x_0 "denotes a determinable number" (p. 122).

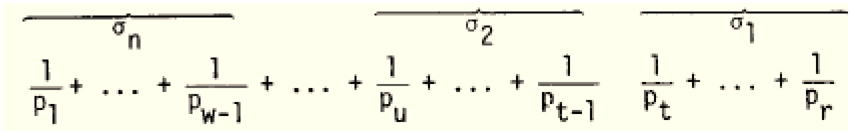
This last result, suitably adapted in §6 for his double sieve, will allow Brun to prove his theorem on Goldbach's Conjecture.

Below we provide some additional details on how Brun establishes his lower bound for E_n and his upper bound for R. But first we back up a little in §4 and look at Brun's choice of the sifting primes, his choice of a particular constant ($\alpha > 1$) that he uses in his calculations, and his use of key results from Mertens.

3.4.4 Assumptions Used by Brun for his Sieve in §4

On p. 116,

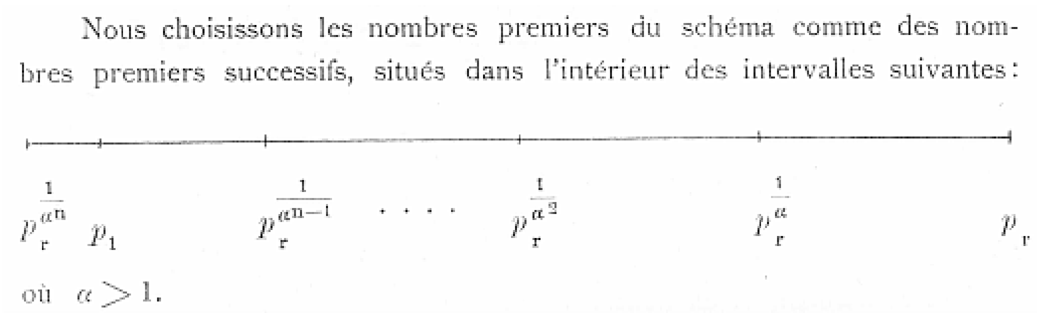
1. Brun decomposes the overall sum of the reciprocals of sifting primes (designated simply by σ in §3) into several component sums $\sigma_1, \sigma_2, \ldots \sigma_n$ (Figure 3.4.4.1);



$$\overbrace{\frac{1}{p_1} + \dots + \frac{1}{p_{w-1}}}^{\sigma_n} + \dots + \overbrace{\frac{1}{p_u} + \dots + \frac{1}{p_{t-1}}}^{\sigma_2} \; \overbrace{\frac{1}{p_t} + \dots + \frac{1}{p_r}}^{\sigma_1}$$

FIGURE 3.4.4.1: Brun's initial choice of sifting primes in §4

2. then, he sets bounds on the sizes of the sifting primes, based on a constant $\alpha > 1$. See Figure 3.4.4.2, which is taken from the French version since Rui's translation has a typographical error in the corresponding figure. (Figures 3.4.4.1 and 3.4.4.2 taken together indicate that each sum $\sigma_i, i = 1, \dots, n$ of reciprocals of sifting primes lies within an interval whose endpoints are certain fractional powers of p_r); and



Nous choisissons les nombres premiers du schéma comme des nombres premiers successifs, situés dans l'intérieur des intervalles suivantes:

$$p_r^{\frac{1}{\alpha^n}} \quad p_1 \qquad p_r^{\frac{1}{\alpha^{n-1}}} \quad \dots \quad p_r^{\frac{1}{\alpha^2}} \qquad p_r^{\frac{1}{\alpha}} \qquad p_r$$

où $\alpha > 1$.

FIGURE 3.4.4.2: The sifting primes are constrained to lie within certain intervals

3. then he gives the formulas from Mertens (3.4.4.1 and 3.4.4.2) that he will use on pp. 117-120

$$\sum_{2}^{x} \frac{1}{p} = \log \log(x) + 0.261... \; + +\theta \frac{5}{\log x}, -1 < \theta < 1, \tag{3.4.4.1}$$

and

$$\prod_{2}^{x} (1 - \frac{1}{p}) = e^{\frac{7\Theta}{\log x}} \; \frac{0.561...}{\log x}, -1 < \Theta < 1. \tag{3.4.4.2}$$

On p. 117 (Figure 3.4.4.3) Brun uses these last two formulas to show that, *for sufficiently large* p_1 each of the σ_i is bounded above and that each product π_i is bounded below. Both these bounds are used on p. 119 to help establish a lower bound for E_n and hence for the main term $M = xE_n$.

Hence we conclude

$$\sum_{x}^{x^\alpha} \frac{1}{p} = \log \alpha + \theta \, \frac{5(1+\frac{1}{\alpha})}{\log x} \quad , \qquad \prod_{x}^{x^\alpha} \left(1 - \frac{1}{p}\right) = \frac{1}{\alpha}\, e^{(1+1/\alpha)7\theta/\log x} .$$

But in that case we can choose p_1 sufficiently large for which

$$\sigma_1 = \frac{1}{p_t} + \ldots + \frac{1}{p_r} < \log \alpha_0 \ ,$$

$$\sigma_2 = \frac{1}{p_u} + \ldots + \frac{1}{p_{t-1}} < \log \alpha_0 \ , \ \ldots$$

$$\sigma_n = \frac{1}{p_1} + \ldots + \frac{1}{p_{w-1}} < \log \alpha_0 \qquad (15)$$

and

$$\pi_1 = \left(1 - \frac{1}{p_t}\right) \ldots \left(1 - \frac{1}{p_r}\right) > \frac{1}{\alpha_0} \quad ,$$

$$\pi_2 = \left(1 - \frac{1}{p_u}\right) \ldots \left(1 - \frac{1}{p_{t-1}}\right) > \frac{1}{\alpha_0} \ , \ \ldots$$

$$\pi_n = \left(1 - \frac{1}{p_1}\right) \ldots \left(1 - \frac{1}{p_{w-1}}\right) > \frac{1}{\alpha_0} \quad , \qquad (16)$$

whenever $\alpha_0 > \alpha$.

We suppose particularly $\log \alpha_0 < 1$.

FIGURE 3.4.4.3: Results that hold when p_1 (and hence x) is sufficiently large

Brun uses his constraints on the sifting primes in three principal ways. First, he uses them on p. 117 (Figure 3.4.4.3) to show that he can bound the sums σ_i above by a small number $\log \alpha_0$) and bound the products π_i below by a small number $\frac{1}{\alpha_0}$ (small, because he supposes that $\alpha_0 > \alpha > 1$ and that $\log \alpha_0 < 1$).

Second, once Brun has bounds on the σ_i and the π_i, he uses those bounds at various points in his detailed calculations on pp. 118-120 to arrive, finally, at a lower bound for N.

Third, Brun uses the constraints on the sifting primes, along with an auxiliary expression (near the bottom of p. 120) to give an upper bound for R.

Then, using his lower bound for M and his upper bound for R, Brun exploits the fundamental inequality $N > M - R$ to arrive arrive at a lower bound for N (formula (20), p. 121) that holds for all successive prime numbers $p_1, \ldots, p_r$, with $p_1 \geq p_e$, where p_e denotes a “determinable prime number.”

Note that Brun uses the phrase “sufficiently large” (p. 117 of [44]) or “suffisamment grand” (p. 19 of [3]) only once in his paper. The English phrase “sufficiently large” occurs often in English paraphrases of Brun's theorem, as in "every sufficiently large even number can be represented as the sum of two 9-primes," but Brun in his original French version of the theorem does not use the corresponding French phrase "suffisamment grand" when stating his theorem. Instead, he speaks of all even numbers x, greater than x_0, where x_0 is a determinable number ("x_0 désigne un nombre déterminable", [3], p. 32).

> One can write [every] even number x, greater than x_0, as a sum of two numbers, whose numbers of prime factors do not exceed nine. x_0 denotes a determinable number and the prime factors can be different or not.

3.4.5 The Values of Particular Determinable Numbers is Never at Issue

The exact value of any "determinable number" that Brun refers to is never at issue in his derivations. The reason is that such a number could, in principle, always be determined by tracing computations forwards from Brun's first statement concerning "sufficiently large" numbers on p. 117 (Figure 3.4.4.3). Note particularly the phrase, "But in that case, we can choose p_1 sufficiently large ..."

But since in these formulas the first sifting prime p_1 is, by the design of the sieve, less than x, this also means that x, too, has to be sufficiently large.

In his subsequent intermediate results in §4, Brun uses a chain of reasoning for other "determinable numbers" that ultimately leads back to this first sifting prime p_1 being able to be chosen sufficiently large to bound the above sums and products.

In other words, the fact that we can pick the smallest sifting prime to be sufficiently large for certain purposes - i.e., so that the bounds in Figure 3.4.4.3 hold - underlies several subsequent calculations that, in the end, give rise to the wording in his theorem on p. 131.

3.4.6 Bounding the Main Term

To establish a lower bound for E_n, Brun

1. Creates a recurrence relation among the terms in the expansion of the main term, which we can formulate as $E_{m+1} > FE_m - G$.

2. Determines, via results of Stirling and Mertens, an upper bound for the term G, which we can call H. So, $H > G$.

3. Determines a lower bound for E_{m+1}. Since $H > G$ and $E_{m+1} > FE_m - G$, we have that $E_{m+1} > FE - H$.

4. Uses results from Mertens and the fact that $E_{m+1} > FE_m - H$ to obtain his lower bound for E_n. As in Section 3.4.3, we call this lower bound by A. It is a function of the sifting primes and the constants α and α_0 that were introduced on p. 117, where he obtains results based on his use of theorems from Mertens.

3.4.7 Bounding the Remainder Term

On p. 120 Brun uses the bounds on the individual p_i, which he had set on p. 116, along with specific values for his constants α and α_0 (see pp. 116 and 117 where Brun introduces α and α_0) to show at the bottom of p. 120 that the number of terms in the expansion is

$$R < p_r \cdot p_r^{2/\alpha} \cdot \ldots \cdot p^{\frac{\alpha+1}{\alpha-1}} = p_r^5.$$

Then, with this bound on R, he concludes in inequality (20) on p.121 that

$$N(D, x, p_1, \ldots, p_r) > \frac{1.008x}{D} 0.3(1 - \frac{1}{p_1}) \ldots (1 - \frac{1}{p_r}) - (p_r^5), \tag{3.4.7.1}$$

which "is valid for all successive prime numbers $p_1, \ldots p_r$, with $p_1 \geq p_e$, where p_e denotes a determinable prime number."

Note that now, instead of saying that p_1 can be chosen sufficiently large, he says that $p_1 \geq p_e$, where p_e denotes a "determinable" prime number.

So, here is the transition in language from a number being able to be chosen "sufficiently large" to that number being specified as larger than some other "determinable" number.

These two different types of expression point to the same underlying reality: that is, there exists some number (i.e., a number that can in principle be determined) such that a given property holds for all numbers exceeding that number.

From this last inequality (3.4.7.1) Brun next extends his results to find a lower bound for $N(D, x, 2, 3, \ldots, p_r)$, that is, a lower bound for the sieve when the sifting primes start with 2 rather than a sufficiently large p_1.

3.4.8 Combining and Extending Results

Brun uses the lower bound for E_n obtained earlier (inequality (19) on p. 120) to determine the inequality (21) on p. 122:

$$N(D, x, 2, 3, \ldots, p_r) > \frac{x}{D}0.3(1 - \frac{1}{p_1}) \ldots (1 - \frac{1}{p_r}) - 2^e(p_r^5), \tag{3.4.8.1}$$

which is "valid for all $r > e$, where e denotes a determinable number, ..." (p. 122).

*Note that here and in the next two inequalities the e being referred to is **not** the e that is the base of the natural logarithm.*

This lower bound for $N(D, x, 2, 3, \ldots, p_r)$ can be further simplified by applying a formula from Mertens dealing with products like $\pi_1 = (1 - \frac{1}{p_t}) \ldots (1 - \frac{1}{p_r})$ to determine a number c such that

$$N(D, x, 2, 3, \ldots, p_r) > \frac{0.168x}{D\log p_r} - 2^e(p_r^5), \tag{3.4.8.2}$$

"for all $r > c$ where c denotes a determinable number ($c \geq e$)" (p. 122).

Brun makes use of this formula in his next example, where he chooses $D = 1$ and $p_r = p(\sqrt[6]{x})$ to state that for all $x > x_0$:

$$N(1, x, 2, 3, \ldots, p(\sqrt[6]{x})) > \frac{1.008x}{x} - 2^e x^{5/6} > \frac{x}{\log x},$$

which allows him to state, as Rui's translation has it:

> When we efface from x consecutive numbers the terms from two to two, then from three to three, etc; finally from $p(\sqrt[6]{x})$ to $p(\sqrt[6]{x})$, there remain always more than $\frac{x}{\log\, x}$ terms, provided $x > x_0$.

That is, when we sift x consecutive numbers by the primes $2, 3, \ldots, p(x^{1/6})$ there remain more than $\frac{x}{\log\, x}$ terms, provided $x > x_0$.

This last result, suitably adapted in §6 for Brun's double sieve, allows Brun to prove his theorem on Goldbach's Conjecture.

3.5 Brun Proves his Theorem on Goldbach's Conjecture

In §6 of his paper, Brun develops formulas using a method "completely analogous" (p. 128) to the one he developed formulas in §4. As a result, we can see that Brun follows the same conceptual scheme that he did in §4. That is, he begins with an initial lower bound $N > M - R$ for his double sieve (Section 3.5.2), and then bounds M below by A and R above by B (see Section 3.5.3), to give a lower bound $N > A - B$ that is large enough to allow him to prove his result.

3.5.1 Overview

In §6 when Brun derives the consequences of his initial lower bound, he takes advantage of the assumptions, results, and calculations that he made in §4, so his arguments in §6 follow very closely the arguments he makes in §4. The main changes are that for the double sieve: 1) he sets α to 1.25 and α_0 to 1.2501 instead of to 1.5 and 1.51, respectively, as he did in §4; and 2) his initial developments in §6 make the assumption that "none of the double effacements are reduced to a single one" (p. 130), i.e., that there aren't duplicates within corresponding pairs of sifting primes (see, e.g., Figure 3.5.1.1), though he subsequently shows (pp. 130-131) that even if there are duplicate effacements, he still obtains the same lower bound for N, which is large enough to allow him to prove his result on Goldbach's Conjecture.

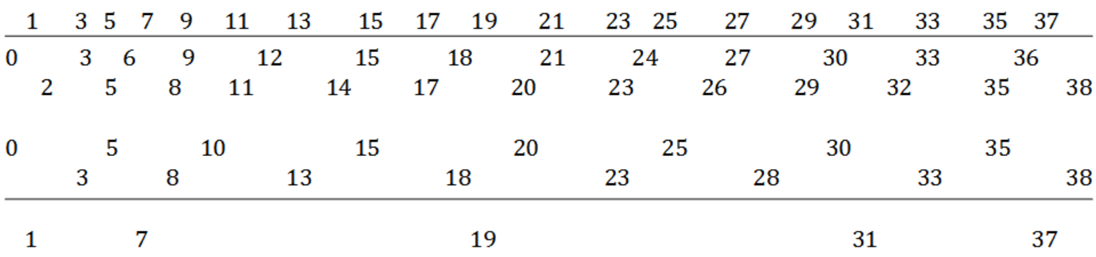


FIGURE 3.5.1.1: Brun's way of showing his double sieve

3.5.2 Initial Lower Bound for the Double Sieve

After Brun sets out the notation for his double sieve, he gives the initial lower bound for that sieve in formula (24) on p. 126 (Figure 3.5.2.1).

$$\frac{D}{x}P(D,\ x,\ p_1,\ldots,p_r) > 1 - \sum_{a\leq r}\frac{2}{p_a} + \sum_{\omega_1}\sum\frac{2^2}{p_a p_b}\left(1 - \sum_{c<b}\frac{2}{p_c}\right)$$
$$+ \sum_{\omega_1'}\sum\ \sum_{\omega_1}\sum\frac{2^4}{p_a p_b p_c p_d}\left(1 - \sum_{e<d}\frac{2}{p_e}\right) + \ldots + \frac{RD}{x}\ , \qquad (24)$$

where $\omega_1' \leq \omega_1$ etc.

FIGURE 3.5.2.1: Brun's initial lower bound for his double sieve

Note the similarity of this formula to the inequality in 3.5.2.1 below, which is the lower bound for the sieve from §4 that discards terms "to the right of vertical lines." The chief difference [7] between the two initial lower bounds is that powers of 2 now appear in the numerators of the summands in the inequality for the double sieve. This difference reflects the fact that 3.5.2.1 is dealing with Brun's *double* sieve, where there are double rows of sifting primes (as shown in Figure 3.5.1.1 and described in 1.9) rather than single rows of sifting primes single in the sieve from §4 (as in Figure 1.9.0.1 and described in Section 3.4).

$$N > x - \sum_{a\leq r}\frac{x}{p_a} + \sum_{\substack{b<a\\ b<t}}\frac{x}{p_a p_b} - \ldots - \sum_{p_a<p_b<\ldots<p_m}\sum\ldots\sum\frac{x}{p_a p_b\cdots p_m} - R. \qquad (3.5.2.1)$$

Then, as he did in §3 and §4 when deriving improved lower bounds for N, Brun carries out calculations based on certain assumptions to bound the main term M below by A and the remainder term above by B, eventually showing that the lower bound for N in his double sieve suffices to allow him to prove his result on Goldbach's Conjecture.

3.5.3 Bounding the Main Term and the Remainder

Brun uses arguments on pp. 128-129, which parallel similar arguments in §4, to establish a lower bound for E_n, when α is set to 1.25 and α_0 is set to 1.2501. This lower bound for E_n is given in formula (26) on p. 129 and in the formula 3.5.3.1 below:

$$En > 0.05(1-\frac{2}{p_1})\ldots(1-\frac{2}{p_r}). \qquad (3.5.3.1)$$

[7]Other differences between the two lower-bound formulas are due to Brun's general notation (e.g., his use of D for the common difference of the elements in the sequence to be sifted, his use of ω_1 and ω to indicate sets over which sets the subscripts $a, b, c, d, \ldots$ can range,and his final $\frac{RD}{x}$ to indicate the more general remainder term. Note that the "+" sign in front of this last fraction is another typo in Rui's translation; a "−" sign used in the original French version (p. 27).

Next, Brun introduces auxiliary formulas on p. 129 to establish an upper bound of p_r^9 for the remainder the remainder R:

In particular, he gives the following result (formula (29) on p. 130):

$$P(D, x, 3, 5, \dots, p_r) > \frac{x}{D} \cdot \frac{0.041}{(\log p_r)^2} - 3^e (p_r)^9, \tag{3.5.3.2}$$

which holds "for all $r > c$, where $c \geq e$." (Note there is a typo in Rui's version, which has e^e instead of 3^e in the last term.

The formula 3.5.3.2 is the analog, for the case of the double sieve, of formula (22) from §4:

$$N(D, x, 2, 3, \dots, p_r) > \frac{0.168x}{D \log p_r} - 2^e (p_r)^5, \tag{3.5.3.3}$$

which holds "for all $r > c$ where c denotes a determinable number ($c \geq e$)" (p. 122).

3.5.4 Combining and Extending Results

Combining the above bounds for the main term and the remainder yields formula (27) on p. 130:

$$En > 0.05(1 - \frac{2}{p1}) \dots (1 - \frac{2}{p_r}) - p_r^9, \tag{3.5.4.1}$$

which is valid "for all successive prime numbers $p_1, \dots, p_r$ whenever $p_1 \geq p_e$, where p_e denotes a determinable prime number.

The similarity of the above language and reasoning with that given for formula (20) on p. 121 is unmistakable.

Then, after giving additional results analogous to results in §4, Brun gives on p. 130 the following result (analogous to yet another formula in §4):

$$N(D, x, 3, 3, \dots, p(x^{1/10})) > \frac{0.41x}{D(\log x)^2} - 3^e x^{9/10} > \frac{0.4x}{D(\log x)^2},$$

which holds "for all $x > x_0$." This last result is the one that, with $D = 2$, eventually carries him through to the statement of his theorem, announced on p. 131.

> One can write [every] even number x, greater than x_0, as a sum of two numbers, whose numbers of prime factors do not exceed nine. x_0 denotes a determinable number and the prime factors can be different or not.

3.6 Summary

Having gone through some of the details of Brun's paper, we now have a fuller context for appreciating what he accomplished and how he did it.

Brun shows that for sufficiently large even x enough of the "right kinds" of numbers fall through his double sieve, where the "right kinds" of numbers are 9-primes such that certain pairs of which sum to the even number x.

- Brun evidently noticed that what we have called "the default lower bound" for the Legendre expansion needs to be improved, at least enough so that we can be sure it is positive.
- He manipulated the expansion of the main term in the Legendre formula and used Stirling's formula to show that under a mild condition when the truncation of the expansion stops at the mth Σ-term, the error term in the Legendre formula can be reduced from 2^r to r^{m+1}. (The mild condition is that $m+2 > \sigma = 1/p_1+\ldots+1/p_r$.)
- He then used particular versions of formulas from Mertens, along with Stirling's formula, to show via complicated combinatorial and asymptotic arguments that he could manipulate both the expansion of the main term in the Legendre formula and the error term to arrive at a positive lower bound for $N(1, x, 2, 3, \ldots, p(\sqrt[6]{x}))$ when x is sufficiently large.
- He carried over in §6 his analysis from §4 to handle the case of the double sieve, making a few slight needed modifications, which eventually allowed him to establish his theorem on Goldbach's Conjecture.

REFERENCES

[1] V. Brun. Uber das goldbachsche gesetz und die anzahl der primzahlpaare. *Archiv for Mathematik og Naturvidenskab*, 34(8):3–19, 1915.

[2] V. Brun. Le crible d'Eratosthène et le théorème de Goldbach. http://visualiseur.bnf.fr/CadresFenetre?O=NUMM-3121&M=pagination, 1919.

[3] V. Brun. Le crible d'Eratosthène et le théorème de Goldbach. Christiania Vidensk. Selsk. Skr. 1920, Nr. 3, 1920.

[4] V. Brun. Reflections on the sieve of eratosthenes. *Norske Vid. Selsk. Skr.*, 1:1–9, 1967.

[5] G. Cantor. Verification jusqu'a 1000 du theoreme empirique de goldbach. *Compte rendu de la 23eme session de l'Association francaise pour l'avancement des sciences*, 23(2):117–134, 1894.

[6] J. R. Chen. On the representation of a large even integer as the sum of a prime and the product of at most two primes. *Scientia Sinica*, 16:157–176, 1973.

[7] Alina Carmen Cojocaru and M. Ram Murty. *An Introduction to Sieve Methods and Their Applications*. CUP, Cambridge, 2005.

[8] Anne-Marie Décaillot. *Cantor et la France. Correspondance du mathématicien allemand avec les Francais a la fin du XIX siecle.* Kimé, Paris, 2008.

[9] John Derbyshire. *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics.* Joseph Henry Press, Washington, D.C., 2003.

[10] M. Desboves. Sur un theoreme de legendre et son application a la recherche de limites qui comprennent entre elles des nombres premiers. *Nouvelles annales de mathematique 1re serie*, 14:281–295, 1855.

[11] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 5.10)*, 2013. http://www.sagemath.org.

[12] D.J.Bernstein. primegen. https://cr.yp.to/primegen.html.

[13] T. Oliveira e Silva S. Herzog and S. Pardi. Empirical verification of the even goldbach conjecture and computation of prime gaps up to $4 \cdot 10^{18}$. *Mathematics of Computation*, 83:2033–2060, 2014.

[14] J. Echeverría. Empirical methods in mathematics. a case study: Goldbach's conjecture. In G. Munévar, editor, *Spanish Studies in the Philosophy of Science*. Kluwer, Boston, 1996.

[15] P. Erdős. Some recent advances and current problems in number theory. In T. L. Saaty, editor, *Lectures on Modern Mathematics, V. 3*, pages 196–244. Wiley, New York, 1965.

[16] J. Friedlander. Producing prime numbers via sieve methods. In J. Friedlander, D. R. Heath Brown, H. Iwaniec, and J Kaczorowski, editors, *Analytic Number Theory*, pages 1–49. Springer, Berlin, 2006.

[17] John Friedlander and Henryk Iwaniec. *Opera De Cribro*. AMS, Providence, RI, 2010.

[18] P.H Fuss. *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIeme siècle, précédée d'une notice sur les travaux de Léonard Euler, tant imprimés qu'inédits et publiée sous les auspices de l'Académie impériale des sciences de Saint-Pétersbourg*. Johnson Reprint Corp., New York, 1968.

[19] A. O. Gelfond, Y. V. Linnik, and A. Feinstein (trans.). *The Prime Numbers and Their Distribution*. Rand McNally, Chicago, 1965.

[20] Timothy Gowers, June Barrow-Green, and Imre Leader. *The Princeton Companion to Mathematics*. Princeton University Press, Princeton, 2008.

[21] George Greaves. *Sieves in Number Theory*. Springer, Berlin, 2001.

[22] J. Hadamard. Un travail de jean merlin sur les nombres premiers. *Bulletin de Sciences Mathematiques*, 39:121–136, 1915.

[23] H. Halberstam and H. E. Richert. *Sieve Methods*. Academic Press, 1974.

[24] G. H. Hardy and J. E. Littlewood. Some problems of 'partitio numerorum'; iii: on the expression of a number as the sum of primes. *Acta Mathematica*, 44(1):1–70, 1923.

[25] R. Haufsner. Uber das goldbach'sche gesetz, 1892.

[26] R. Haussner. Tafeln fur das goldbach'sche gesetz, 1897.

[27] H. A. Helfgott. The ternary Goldbach problem. https://arxiv.org/abs/1501.05438, 2015.

[28] D. Hilbert. Mathematical problems. *Bulletin (New Series) of the American Mathetical Society*, 37(4):407–436, 2010.

[29] H. Iwaniec. Sieve methods. *Proceedings of the International Congress of Mathematicians*, 1:357–364, 1978.

[30] H. Iwaniec and E. Kowalski. *Analytic Number Theory*. AMS, Providence, RI, 2004.

[31] G. J. O. Jameson. *The Prime Number Theorem*. Cambridge University Press, Cambridge, 2003.

[32] E. Kowalski. *The Large Sieve and its Applications: Arithmetic Geometry, Random Walks and Discrete Groups*. Cambridge University Press, Cambridge, 2008.

[33] E. Landau. Gelöste und engelloste probleme aus der theorie der primzahlverteilung und der reimannschen zetafunction. *Proceedings of the International Congress of Mathematicians*, 1:93–108, 1913.

[34] A. M. Legendre. *Essai sur la théorie des nombres*. Coucier, Paris, second edition, 1808.

[35] R. A. Mollin. An overview of sieve methods. *International Journal of Contemporary Mathematical Sciences*, 5(2):67–80, 2010.

[36] Hugh L. Montgomery and Robert C. Vaughan. *Multiplicative Number Theory I: Classical Theory*. Academic Press, Waltham, MA, 1995.

[37] W. Narkiewicz. *The Development of Prime Number Theory : From Euclid to Hardy and Littlewood*. Springer, Berlin, 2000.

[38] W. Narkiewicz. *Rational Number Theory in the 20th Century: From PNT to FLT*. Springer, Berlin, 2012.

[39] M. B. Nathanson. *Additive Number Theory: The Classical Bases*. Springer, New York, 1996.

[40] Nicomachus and M. L. D'ooge (trans.). *Introduction to Arithmetic*. Macmillan, New York, 1926.

[41] J. J. O'Connor and E. F. Robertson. Christian goldbach. http://www-history.mcs.st-and.ac.uk/Biographies/Goldbach.html.

[42] J. Pintz. Landau's problems on primes. *Journal de ThÃľorie des Nombres de Bordeaux*, 21(2):357–404, 2009.

[43] P. Pollack. *Not Always Buried Deep: A Second Course in Elementary Number Theory*. AMS, Providence, RI, 2009.

[44] Y. Rui. The sieve of eratosthenes and goldbach's conjecture. In Y. Wang, editor, *The Goldbach Conjecture*, pages 99–136. World Scientific, Singapore, 2002.

[45] I. G. Schnirelman. Uber additive eigenschaften von zahlen. *Math. Ann.*, 107:649–690, 1933.

[46] C. J. Scriba. Viggo brun in memoriam (1885-1978). *Historia Mathematica*, 7(1):1–6, 1980.

[47] N.J.A. Sloane. The on-line encyclopedia of integer sequences. http://oeis.org.

[48] R. Stanley. *Enumerative Combinatorics, 1*. Cambridge University Press, Cambridge, UK, 1997.

[49] J. Stopple. *A Primer of Analytic Number Theory*. Cambridge University Press, Cambridge, 2003.

[50] J. J. Sylvester. On the partition of an even number into two primes. *Proc. London Math. Soc.*, 4:4–6, 1871.

[51] J. J. Sylvester. On the Goldbach-Euler theorem regarding prime numbers. *Nature*, 55(1418):196–197, 1896.

[52] G. Tenenbaum and M. Mèndes France. *The Prime Numbers and Their Distribution.* AMS, Providence, RI, 2000.

[53] Yuan Wang. *The Goldbach Conjecture.* World Scientific, Singapore, second edition, 2002.

[54] E. Waring. *Meditationes Algebraicæ. Ab Edvardo Waring, M.D. Mag. Coll. Cantab. Soc. Matheseos professore Lucasiano, Regiæ societatis, et Bononiensis scientiarum academiæ socio.* Cantabrigiæ: Typis Academis excudebat J. Archdeacon. Veneunt apud J. Woodyer, Cantabrigiæ; J. Beecroft, T. Payne, et T. Cadell, Londini; et D. Prince, Oxon. M. DCC. LXX, 1770.

[55] D. Zagier. The first fifty million prime numbers. *Mathematical Intelligencer*, 1(1, supplement):7–19, 1978.