Adding Prime Numbers

University of Cambridge
Faculty of Mathematics
Department of Pure Mathematics and Mathematical Statistics

Author	Markus Schepke BSc
College	Wolfson College
Home address	Robert-Beltz-Str. 49, 19059 Schwerin, Germany
eMail	ms946@cam.ac.uk
Adviser	Professor Ben J. Green

I declare that this essay is work done as part of the Part III Examination. I have read and understood the Statement on Plagiarism for Part III and Graduate Courses issued by the Faculty of Mathematics, and have abided by it. This essay is the result of my own work, and except where explicitly stated otherwise, only includes material undertaken since the publication of the list of essay titles, and includes nothing which was performed in collaboration. No part of this essay has been submitted, or is concurrently being submitted, for any degree, diploma or similar qualification at any university or similar institution.

Cambridge, November 20, 2011

Adding Prime Numbers

University of Cambridge
Faculty of Mathematics
Department of Pure Mathematics and Mathematical Statistics

Contents

1 Introduction

1.1 Historical Remarks and Outline

In a letter dated to 7 June 1742, the German mathematician Christian Goldbach wrote to Leonhard Euler his conjecture that every even integer can be represented as the sum of two prime numbers. Over two and a half centuries later, this still remains unproved, but throughout this period, Goldbach's conjecture has been an inspiring incitement to many mathematicians which led to the development of a branch of mathematics now famous as additive number theory. The aim of this treatise is to present proofs of two renowned results in this area: Schnirelmann's Theorem (a weaker version of Goldbach's conjecture), and Waring's problem stating that every positive integer can be represented as the sum of a bounded number of k^{th} powers.

Both proofs rely on a certain measurement of sequence of integers, called the Schnirelmann density. The notion of this density will be introduced in Chapter 2 whose main purpose is to prove that any sequence with positive density is a basis of finite order, i. e. every integer can be represented as a sum of a finite number of elements of this sequence. As it turns out, the number of different representations of an integer as such a sum plays a crucial role in these kinds of problem, and therefore the remainder of this paper will be dedicated to the study of estimates for these numbers of representations.

In Chapter 3, we will prove Chebyshev's Theorem, a classical result in the distribution of primes which is required to prove the necessary lower bound for Schnirelmann's Theorem, and also an important result in its own right. Chapter 4 will present the Selberg sieve, a powerful and universally usable tool to estimate numbers of solutions. The deduced upper bound will conclude the proof of Schnirelmann's Theorem. Since Brun's Theorem on twin primes requires essentially the same estimates we will include a proof in this section. As a further application of the Schnirelmann density, we will present a solution to Waring's problem in Chapter 5. With the aid of this approach, a much shorter and more elementary proof than David Hilbert's original one is possible.

1.2 Notations and Definitions

Throughout this elaboration, we will make heavy use of implied constants. Instead of the O notation, we will utilize the symbols $\ll$ and $\gg$ which allow more intuitive estimates.

Definition Let f and g be positive functions. We write $f \ll g$ or $g \gg f$ if there is a positive constant c such that $f(x) \leq cg(x)$ for all sufficiently large x. This constant may depend only on some other constants, but not on x.

Moreover, by $\mathbb{N}$ and $\mathbb{N}_0$ we mean the set of positive and non-negative integers, respectively. By $\mathbb{P}$ we denote the set of prime numbers, by $\#S$ the cardinality of the set S, by $\log x$ the natural logarithm, and by $\lfloor x \rfloor$ the integral part of x. For the sake of brevity, we will write (m, n) for the greatest common divisor, and $[m, n]$ for the lowest common multiple of the integers m and n.

Before we proceed to introduce the Schnirelmann density, we want to give an important estimate that will be handy throughout the paper.

Theorem 1.1 (Cauchy-Schwarz inequality) *Let $x, y \in \mathbb{R}^n$. Then*

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right) \cdot \left(\sum_{i=1}^{n} y_i^2\right),$$

where equality is attained if and only if x and y are linearly dependent.

Proof. Assume that $y \neq 0$, otherwise the statement is trivial. The inequality follows at once from

$$\begin{aligned} 2\left(\left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} x_i y_i\right)^2\right) &= 2\left(\sum_{i,j=1}^{n} x_i^2 y_j^2 - \sum_{i,j=1}^{n} x_i y_i x_j y_j\right) \\ &= \sum_{i,j=1}^{n} \left(x_i^2 y_j^2 - 2x_i y_i x_j y_j + x_j^2 y_i^2\right) \\ &= \sum_{i,j=1}^{n} \left(x_i y_j - x_j y_i\right)^2 \geq 0, \end{aligned}$$

where equality is attained if and only if $x_i = \frac{x_j}{y_j} y_i$. Since there is at least one $y_j \neq 0$, this is equivalent to linear dependence. □

2 Schnirelmann's Theorem

The aim of this chapter is to prove the first notable result on Goldbach's conjecture, found in 1930 by the Russian mathematician Lev Schnirelmann[1]:

Theorem 2.1 (Schnirelmann) *There is a bounded number h such that every integer greater than 1 can be represented as a sum of at most h primes.*

The proof is entirely elementary; yet we will need estimates that require some more work and will be proved in the following chapters. First, we will introduce a kind of "measure" now famous as the Schnirelmann density. This depiction follows Nathanson's book [Nat96, Ch. 7].

2.1 Schnirelmann Density

Definition Let $A \subseteq \mathbb{Z}$ be a set of integers, and $x \in \mathbb{R}$ a real number. By

$$A(x) := \#\{a \in A : 1 \leq a \leq x\}$$

we denote the number of elements in A not exceeding x. Then we define the *Schnirelmann density* of A by

$$\sigma(A) := \inf_{n \in \mathbb{N}} \frac{A(n)}{n}.$$

First, we want to give some basic properties of the Schnirelmann density:

Lemma 2.2 *Let $A \subseteq \mathbb{Z}$, and $x \in \mathbb{R}$.*

(i) $0 \leq \sigma(A) \leq 1$.

(ii) $A(m) \geq \sigma(A)m$ *for all* $m \in \mathbb{N}$.

(iii) Let $m \geq 1$. *If* $m \notin A$ *then*

$$\sigma(A) \leq 1 - \frac{1}{m} < 1.$$

In particular, if $1 \notin A$ *then* $\sigma(A) = 0$.

[1] Schnirelmann first published his result in Russian [Sch30], later (expanded) in German [Sch33]. A comprehensible account was given by Edmund Landau [Lan30].

(iv) *The set A contains all positive integers if and only if $\sigma(A) = 1$.*

(v) *Let $\varepsilon > 0$. If $\sigma(A) = 0$, then we can find $N > 0$ such that $A(n) < \varepsilon n$ for all $n \geq N$.*

Proof. (i) Obviously $0 \leq A(x) \leq x$, so

$$0 \leq \frac{A(x)}{x} \leq 1,$$

and hence

$$0 \leq \inf_{n\in\mathbb{N}} \frac{A(n)}{n} = \sigma(A) \leq 1.$$

(ii) By definition, we have for all $m \in \mathbb{N}$:

$$\frac{A(m)}{m} \geq \inf_{n\in\mathbb{N}} \frac{A(n)}{n} = \sigma(A),$$

thus $A(m) \geq \sigma(A)m$.

(iii) If $m \notin A$, then $A(m) \leq m - 1$, so

$$\sigma(A) = \inf_{n\in\mathbb{N}} \frac{A(n)}{n} \leq \frac{A(m)}{m} \leq \frac{m-1}{m} = 1 - \frac{1}{m} < 1.$$

The case $m = 1$ is obvious by plugging 1 into the formula.

(iv) If $m \in A$ for all $m \in \mathbb{N}$ then $A(m) = m$, and so $\sigma(A) = 1$. Conversely, if there exists $m \in \mathbb{N}$ with $m \notin A$, then $\sigma(A) < 1$ by (iii).

(v) Assume to the contrary that $A(n) \geq \varepsilon n$ for all $n \in \mathbb{N}$. Then by definition,

$$\sigma(A) = \inf_{n\in\mathbb{N}} \frac{A(n)}{n} \geq \varepsilon > 0,$$

giving the contradiction. □

Before we proceed with the proof of Schnirelmann's Theorem, we want to acquaint ourselves with the notion of Schnirelmann density through some fundamental examples.

Example 2.3 (i) Let A denote the even numbers. Since $1 \notin A$, Lemma 2.2 yields $\sigma(A) = 0$.

(ii) Let A denote the odd numbers. Obviously $\sigma(A) \leq \frac{1}{2}$ since $2 \notin A$, and also $A(2m) = m$ for all positive integers m. But on the other hand, $A(2m+1) = m+1$, so

$$\frac{A(2m+1)}{2m+1} = \frac{m+1}{2m+1} > \frac{1}{2},$$

and hence
$$\sigma(A) = \inf_{n\in\mathbb{N}} \frac{A(n)}{n} = \inf_{n\in\mathbb{N}} \left\{ \frac{n}{2n}, \frac{n+1}{2n+1} \right\} = \frac{1}{2}.$$
It may be surprising that the density of the even and the odd numbers do not coincide, but this fact illustrates the sensitivity of the Schnirelmann density concerning the first values of the set.

(*iii*) Let A denote the square numbers, i. e.

$$A = \{a^2 : a \in \mathbb{N}_0\} = \{0^2, 1^2, 2^2, 3^2, \ldots\}.$$

Then obviously $A(n^2) = n$, and so

$$\sigma(A) = \inf_{n\in\mathbb{N}} \frac{A(n)}{n} \leq \inf_{n\in\mathbb{N}} \frac{A(n^2)}{n^2} = \inf_{n\in\mathbb{N}} \frac{n}{n^2} = 0.$$

This can be generalised to any set of k^{th} powers straightforwardly.

(*iv*) Let A denote the prime numbers, expanded by 1 (otherwise $\sigma(\mathbb{P}) = 0$ is trivial). By Chebyshev's Theorem (Thm. 3.1) we know that

$$A(x) = \pi(x) + 1 \ll \frac{x}{\log x}$$

for all $x \geq 2$. This again yields

$$\sigma(A) = \inf_{n\in\mathbb{N}} \frac{A(n)}{n} \ll \frac{m/\log m}{m} = \frac{1}{\log m}$$

for all $m \in \mathbb{N}$, and hence $\sigma(A) = 0$.

(*v*) We now give a non-trivial example of a sequence with positive density. Let A_2 denote the squarefree numbers, i. e. integers such that all their prime divisors are distinct. It is a straightforward calculation that

$$\lim_{x\to\infty} \frac{A_2(x)}{x} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} \approx 0.607927\ldots$$

Kenneth Rogers showed [Rog64] with an elementary argument that

$$\sigma(A_2) = \frac{53}{88} \approx 0.602273\ldots < \frac{6}{\pi^2}.$$

Analogously, we call a number k-free if no k^{th} power of any prime divides this number. Note that we consider 1 to be k-free, giving it the odd property of being k-free and a perfect k^{th} power. Certainly, every squarefree number is k-free for any $k \geq 2$, so on denoting the sequence of k-free numbers by A_k we have

$$\frac{53}{88} = \sigma(A_2) \leq \sigma(A_3) \leq \ldots \leq \sigma(A_k) \leq \sigma(A_{k+1}).$$

With the same calculation as above, we can show that

$$\lim_{x\to\infty} \frac{A_k(x)}{x} = \frac{1}{\zeta(k)}.$$

More precisely, R. L. DUNCAN [Dun65] proved that we have the following chain of inequalities:

$$\frac{53}{88} = \sigma(A_2) < \frac{1}{\zeta(2)} < \ldots < \sigma(A_k) < \frac{1}{\zeta(k)} < \sigma(A_{k+1}) < \frac{1}{\zeta(k+1)} < \ldots$$

Now we want to examine sumsets and bases:

Definition Let $A, B \subseteq \mathbb{Z}$. Then we define their *sumset* by

$$A + B := \{a + b : a \in A,\ b \in B\}.$$

This can be generalised to h sets $A_1, \ldots, A_h \subseteq \mathbb{Z}$ by

$$A_1 + \ldots + A_h := \{a_1 + \ldots + a_h : a_i \in A_i\}.$$

If $A_i = A$ for all i, then we write

$$hA := \sum_{i=1}^{h} A.$$

We call A a *basis of order h* if hA contains all positive integers, i. e. if every positive integer can be represented as the sum of (exactly) h elements of A, and a *basis of finite order* if there is an integer h such that every positive integer can be represented as the sum of at most h elements of A.

In most cases, we will assume that $0 \in A$, so that every basis of order h is a basis of order $h + 1$ as well, and we do not need to distinguish between representations as sums of *exactly* h elements and *at most* h elements of A.

By Lemma 2.2, being a basis of order h is equivalent to $\sigma(hA) = 1$. Schnirelmann made the important observation that a set with positive density is a basis of finite order. To prove this, we need some further properties of the Schnirelmann density.

Lemma 2.4 *Let $A, B \subseteq \mathbb{Z}$ be subsets containing* 0.

(i) If n is a non-negative integer and $A(n) + B(n) \geq n$, then n is contained in $A + B$.

(ii) If $\sigma(A) + \sigma(B) \geq 1$, then n is contained in $A + B$ for all positive integers n.

(iii) If $\sigma(A) \geq \frac{1}{2}$ then A is a basis of order 2.

Proof. (*i*) If $n \in A$ then $n + 0 \in A + B$, and analogue, if $n \in B$ then $0 + n \in A + B$. So assume that $n \notin A \cup B$. We define the sets

$$A' := \{n - a : a \in A,\ 1 \le a \le n - 1\},$$

and

$$B' := \{b : b \in B,\ 1 \le b \le n - 1\}.$$

Then $\#A' = A(n)$ and $\#B' = B(n)$ as $A(n-1) = A(n)$ and $B(n-1) = B(n)$ since $n \notin A, B$. Moreover

$$A' \cup B' \subseteq \{1, 2, \ldots, n - 1\},$$

i.e. $\#(A' \cup B') \le n - 1$. But

$$\#A' + \#B' = A(n) + B(n) \ge n$$

by assumption, so $A' \cap B'$ cannot be empty. Hence $n - a = b$ for some $a \in A$ and $b \in B$, so

$$n = a + b \in A + B.$$

(*ii*) By Lemma 2.2 we have

$$A(n) + B(n) \ge \sigma(A)n + \sigma(B)n = \underbrace{(\sigma(A) + \sigma(B))}_{\ge 1} n \ge n,$$

so $n \in A + B$ by (i).

(*iii*) Let $A = B$ in (ii), then $\sigma(A) + \sigma(A) \ge 1$, so $n \in A + A$ for all $n \in \mathbb{N}$, i.e. A is a basis of order 2. □

It is worth noticing that this result already suffices to describe the additive behaviour of k-free numbers completely.

Corollary 2.5 *The set of k-free numbers amended by* 0 *is a basis of order* 2 *for all* $k \ge 2$.

Proof. Let A_k denote the set of k-free numbers together with 0. According to Ex. 2.3 (v) we have

$$\sigma(A_k) \ge \sigma(A_2) = \frac{53}{88} > \frac{1}{2}.$$

So by Lemma 2.4, A_k is a basis of order 2. □

Usualy, we are not able to apply Lemma 2.4 that conveniently. However, we are now prepared to prove a powerful inequality concerning the Schnirelmann density.

Proposition 2.6 *Let $A, B \subseteq \mathbb{Z}$ subsets containing 0. Then we have the inequality*

$$\sigma(A+B) \geq \sigma(A) + \sigma(B) - \sigma(A) \cdot \sigma(B).$$

Proof. Let $n \geq 1$ be an integer, and $k = A(n)$. We numerate and order the elements of A by

$$0 = a_0 < a_1 < a_2 < \ldots < a_k \leq n.$$

Since $0 \in B$ we have $a_i = a_i + 0 \in A + B$ for all $i = 0, 1, \ldots, k$. Now let $r_i := B(a_{i+1} - a_i - 1)$ for $i = 0, \ldots, k-1$, and numerate the elements of B by

$$1 \leq b_1 < b_2 < \ldots < b_{r_i} \leq a_{i+1} - a_i - 1.$$

Note that r_i may be zero, in which case the respective statements are simply empty. For every i, we have by construction

$$a_i < a_i + b_1 < a_i + b_2 < \ldots < a_i + b_{r_i} < a_{i+1},$$

and $a_i + b_j \in A + B$ for all $j = 1, 2, \ldots, r_i$. Let $r_k = B(n - a_k)$, and

$$1 \leq b_1 < b_2 < \ldots < b_{r_k} \leq n - a_k.$$

Then again we have

$$a_k < a_k + b_1 < a_k + b_2 < \ldots < a_k + b_{r_k} \leq n,$$

and $a_k + b_j \in A + B$ for all $j = 1, 2, \ldots, r_k$. All statements combined, we thus have found the following distinct elements of $A + B$ not exceeding n:

$$\begin{aligned} a_0 &< a_0 + b_1 < \ldots < a_0 + b_{r_0} < a_1 < a_1 + b_1 < \ldots < a_1 + b_{r_0} < a_2 < \ldots \\ &< a_k < a_k + b_1 < \ldots < a_k + b_{r_k} \leq n. \end{aligned}$$

This gives us the estimate for all $n \in \mathbb{N}$:

$$\begin{aligned} (A+B)(n) &\geq A(n) + \sum_{i=0}^{k} r_i \\ &= A(n) + \sum_{i=0}^{k-1} B(a_{i+1} - a_i - 1) + B(n - a_k) \\ &\overset{\text{L2.2}}{\geq} A(n) + \sum_{i=0}^{k-1} \sigma(B)(a_{i+1} - a_i - 1) + \sigma(B)(n - a_k) \\ &= A(n) + \sigma(B) \sum_{i=0}^{k-1} (a_{i+1} - a_i) + \sigma(B)(n - a_k) - \sigma(B)k \\ &\overset{\text{telescope}}{=} A(n) + \sigma(B)n - \sigma(B)k \end{aligned}$$

$$\begin{aligned}
&\overset{k=A(n)}{=} && A(n)+\sigma(B)n-\sigma(B)A(n)\\
&= && (1-\sigma(B))A(n)+\sigma(B)n\\
&\overset{\text{L2.2}}{\geq} && (1-\sigma(B))\sigma(A)n+\sigma(B)n\\
&= && (\sigma(A)+\sigma(B)-\sigma(A)\sigma(B))n.
\end{aligned}$$

Division by n finally yields

$$\sigma(A+B)=\inf_{n\in\mathbb{N}}\frac{(A+B)(n)}{n}\geq\sigma(A)+\sigma(B)-\sigma(A)\sigma(B).$$ □

Note that the inequality is equivalent to

$$1-\sigma(A+B)\leq(1-\sigma(A))(1-\sigma(B)).$$

We can generalise this easily:

Corollary 2.7 *Let $h\in\mathbb{N}$, and $A_1,\ldots,A_h\subseteq\mathbb{Z}$ subsets containing 0. Then we have the inequality*

$$1-\sigma(A_1+\ldots+A_h)\leq\prod_{i=1}^{h}(1-\sigma(A_i)).$$

Proof. This is a straightforward induction on h. The case $h=1$ is obvious, the case $h=2$ is Prop. 2.6. Now let $h>2$. Then

$$\begin{aligned}
1-\sigma(A_1+\ldots+A_h) &= 1-\sigma((A_1+\ldots+A_{h-1})+A_h)\\
&\overset{\text{P2.6}}{\leq} (1-\sigma(A_1+\ldots+A_{h-1}))(1-\sigma(A_h))\\
&\overset{\text{I. H.}}{\leq} \left(\prod_{i=1}^{h-1}(1-\sigma(A_i))\right)(1-\sigma(A_h)),
\end{aligned}$$

as desired. □

We have now all ingredients to prove the announced link between positive density and bases of finite order.

Theorem 2.8 *If $A\subseteq\mathbb{Z}$ is a subset containing 0 with positive Schnirelmann density then A is a basis of finite order.*

Proof. As $\sigma(A)>0$ we have $0\leq 1-\sigma(A)<1$, and so there is an integer l with

$$0\leq(1-\sigma(A))^l\leq\frac{1}{2}.$$

By Cor. 2.7, this yields

$$1-\sigma(lA)\leq(1-\sigma(A))^l\leq\frac{1}{2},$$

and so $\sigma(lA) \geq \frac{1}{2}$. By Lemma 2.4, lA is a basis of order 2, i. e. every integer can be represented as the sum of two elements of lA. Each of these elements can be represented as a sum of l elements of A, so A is a basis of order $2l < \infty$. □

Unfortunately, we know by Ex. 2.3 (iv) that $\sigma(\mathbb{P}) = 0$, so we need some more work. The following general result indicates the strategy we will use.

Proposition 2.9 *Let $A = (a_1, a_2, \ldots)$ be a sequence of integers, and $r(a)$ denote the multiplicity of a in A, i. e.*

$$r(a) := \sum_{\substack{i \geq 1 \\ a_i = a}} 1.$$

Let x be a real number. If the estimate

$$\frac{1}{x} \cdot \frac{\left(\sum_{1 \leq N \leq x} r(N)\right)^2}{\sum_{1 \leq N \leq x} r(N)^2} \geq \alpha > 0$$

holds for all $x \geq 1$, then $\sigma(A) \geq \alpha > 0$, i. e. A has positive Schnirelmann density.

Proof. By the Cauchy-Schwarz inequality (Thm. 1.1), we have

$$\begin{aligned}
\left(\sum_{N \leq x} r(N)\right)^2 &= \left(\sum_{\substack{N \in A \\ N \leq x}} 1 \cdot r(N)\right)^2 \\
&\leq \left(\sum_{\substack{N \in A \\ N \leq x}} 1^2\right) \cdot \left(\sum_{\substack{N \in A \\ N \leq x}} r(N)^2\right) \\
&= A(x) \cdot \sum_{N \leq x} r(N)^2.
\end{aligned}$$

By assumption, this yields

$$\frac{A(x)}{x} \geq \frac{1}{x} \cdot \frac{\left(\sum r(N)\right)^2}{\sum r(N)^2} \geq \alpha > 0$$

for all $x \geq 1$. Therefore

$$\sigma(A) = \inf_{n \in \mathbb{N}} \frac{A(n)}{n} \geq \alpha > 0.$$

□

The next section will deal with the question how to apply these results to Goldbach's problem. Some more sophisticated estimates will be needed; their proofs are subject of the subsequent chapters.

As a concluding remark, it is worth noticing that the inequality in Prop. 2.6 was improved by Henry B. Mann [Man42]:

Theorem 2.10 (Mann) *Let $A, B \subseteq \mathbb{Z}$ subsets containing 0. Then we have the inequality*

$$\sigma(A+B) \geq \min\{\sigma(A)+\sigma(B), 1\}.$$

This bound is sharp: Let both A and B be the set

$$A := \{a \in \mathbb{Z} : a \equiv 1 \mod m\} \cup \{0\}$$

for an integer $m \geq 2$. Then $2A = A + A$ obviously consists of 0 and all the integers congruent to 1 or 2 modulo m. Analogously to Ex. 2.3 (ii), one sees that $\sigma(A) = 1/m$, while $\sigma(2A) = 2/m$, so Mann's result cannot be improved further.

2.2 Proof of Schnirelmann's Theorem

Throughout this section, let $r(N)$ denote the number of representations of the positive integer N as the sum of two primes, i. e.

$$r(N) := \sum_{p_1+p_2=N} 1.$$

Then Goldbach's original problem can be rephrased as

$$r(2N) > 0$$

for all $N \in \mathbb{N}$, so obviously it is worth studying $r(N)$ in order to solve Goldbach's problem.

A famous result in the distribution of primes is Chebyshev's Theorem (Thm. 3.1): If $\pi(x)$ denotes the number of primes not exceeding x (cf. Chapter 3), then we have for all $x \geq 2$:

$$\frac{x}{\log x} \ll \pi(x) \ll \frac{x}{\log x}.$$

From this, we immediately obtain a lower bound for $\sum r(N)$:

Lemma 2.11 *Let N be a positive integer and $x \geq 2$ a real number. Then:*

$$\sum_{N \leq x} r(N) \gg \frac{x^2}{(\log x)^2},$$

where the implied constant is absolute.

Proof. Let p and q be primes not exceeding $x/2$. Obviously, $p + q \leq x$ is an integer represented as the sum of two primes with value at most x, so is included in the set the sum is counting. We can choose $\pi(x/2)$ primes for p and q, respectively, so we obtain

$$\sum_{N \leq x} r(N) \geq (\pi(x/2))^2 \gg \left(\frac{x/2}{\log x/2}\right)^2 \gg \frac{x^2}{(\log x)^2}$$

by Chebyshev's Theorem (Thm. 3.1). □

Therefore, Chapter 3 will be dedicated to the proof of Chebyshev's Theorem. Regarding Prop. 2.9, we see that the following estimate is needed to proof Schnirelmann's Theorem:

Lemma 2.12 *Let N be a positive integer and $x \geq 2$ a real number. Then:*

$$\sum_{N \leq x} r(N)^2 \ll \frac{x^3}{(\log x)^4},$$

where the implied constant is absolute.

In his original paper, Schnirelmann used sieve methods by Viggo Brun; however, in Chapter 4 we will present sieve methods according to Atle Selberg that generally yield stronger estimates, and allow a more concise way to prove the required upper bound in Lemma 2.12.

Assuming these results, we can now prove Schnirelmann's Theorem. We saw that $\sigma(\mathbb{P}) = 0$ (Ex. 2.3 (iv)), but courtesy of these estimates, we are able to prove that $2\mathbb{P}$ (extended by 0 and 1) has positive density.

Proposition 2.13 *The set $(\mathbb{P} + \mathbb{P}) \cup \{0, 1\}$ has positive Schnirelmann density.*

Proof. Let $A := (\mathbb{P} + \mathbb{P})$, and $A' := A \cup \{0, 1\}$. By definition, $r(a)$ is counting the multiplicity of an integer a in A. By Lemma 2.11 we have for a positive integer N, and for a real number $x \geq 2$ the estimate

$$\sum_{N \leq x} r(N) \gg \frac{x^2}{(\log x)^2},$$

and by Lemma 2.12,

$$\sum_{N \leq x} r(N)^2 \ll \frac{x^3}{(\log x)^4}.$$

Combining these, we have

$$\frac{1}{x} \cdot \frac{\left(\sum r(N)\right)^2}{\sum r(N)^2} \gg \frac{1}{x} \cdot \frac{x^4/(\log x)^4}{x^3/(\log x)^4} = 1,$$

meaning that there exist $\alpha > 0$ such that

$$\frac{1}{x} \cdot \frac{\left(\sum r(N)\right)^2}{\sum r(N)^2} \geq \alpha > 0.$$

Since $1 \in A'$, we obtain

$$\frac{1}{x} \cdot \frac{\left(\sum r'(N)\right)^2}{\sum r'(N)^2} \geq \alpha' > 0$$

for all $x \geq 1$, where $r'(N)$ denotes the multiplicity of N in A'. Using Prop. 2.9, we obtain:

$$\sigma(A') \geq \alpha' > 0. \qquad \square$$

Compiling everything, we can complete the proof of Schnirelmann's Theorem.

Proof of Theorem 2.1. We want to prove that every integer greater than 1 can be represented by the sum of a bounded number of primes. By Prop. 2.13, we know that $A := (\mathbb{P} + \mathbb{P}) \cup \{0, 1\}$ has positive density, hence by Thm. 2.8 is a basis of finite order, say h. Let $N \geq 2$. Then the non-negative integer $N - 2$ can be represented by the sum of exactly h elements of A, say l zeros, k ones, and m pairs of primes $p_i + q_i$ for $i = 1, \ldots, m$, i. e.

$$N - 2 = \underbrace{1 + \ldots + 1}_{k \text{ times}} + (p_1 + q_1) + \ldots + (p_m + q_m),$$

where $h = l + k + m$. If $k = 2r$ is even then we can write

$$N = \underbrace{2 + \ldots + 2}_{r+1 \text{ times}} + p_1 + q_1 + \ldots + p_m + q_m;$$

if $k = 2r + 1$ is odd then we have

$$N = \underbrace{2 + \ldots + 2}_{r \text{ times}} + 3 + p_1 + q_1 + \ldots + p_m + q_m.$$

In each case, we can represent N as the sum of at most

$$r + 1 + m \leq 2k + m \leq 3h$$

primes, as required. $\square$

Of course, it is quite unsatisfactory that the order of the basis depends heavily on the implied constants in the estimates. The least number h of primes needed to represent any integer as a sum is called *Schnirelmann's constant*. Schnirelmann's original proof only yields that this constant is finite; the best known value to date is 7, proved by Olivier Ramaré [Ram95] in 1995. A proof of Goldbach's conjecture would immediately imply that Schnirelmann's constant is 3, but this seems not to be within sight.

2.3 Generalisations

In his original paper, Schnirelmann proved a slightly stronger result. We say that a set $P \subseteq \mathbb{P}$ contains a positive proportion of the primes, if there is $\vartheta > 0$ such that

$$P(x) \geq \vartheta \pi(x)$$

for all sufficiently large real numbers x. Schnirelmann proved that such sets also allow representations as finite sums for all sufficiently large integers. The following simple proof is due to Melvyn B. Nathanson [Nat87].

Theorem 2.14 *Let $P \subseteq \mathbb{P}$ be a set that contains a positive proportion of the primes. Then every sufficiently large integer can be represented by a bounded number of elements of P.*

Proof. Let $r_P(N)$ denote the number of representations of N as the sum of two elements of P. Obviously, $r_P(N) \leq r(N)$, and so

$$\sum_{N \leq x} r_P(N)^2 \leq \sum_{N \leq x} r(N)^2 \ll \frac{x^3}{(\log x)^4}$$

by Lemma 2.12. Moreover, we have

$$\sum_{N \leq x} r_P(N) \geq P(x/2)^2 \geq (\vartheta\pi(x/2))^2 \gg \frac{x^2}{(\log x)^2}$$

by the same argument as in Lemma 2.11. So $(P+P) \cup \{0,1\}$ has positive density and is hence a basis of finite order, i. e. there is a number h_1 such that every positive integer can be represented as the sum of at most h_1 elements of $P \cup \{1\}$. Now pick two primes $p, q \in P$. From the Euclidean algorithm we obtain the linear combination $xp - yq = 1$ with $x, y \geq 1$, so there is an integer n_0 such that every integer $n \geq n_0$ can be represented as a linear combination

$$n = a(n)p + b(n)q$$

with non-negative coefficients $a(n)$ and $b(n)$. So let $n \geq n_0$. We have the representation as the sum of elements of $P \cup \{1\}$

$$n - n_0 = p_1 + \ldots + p_r + \underbrace{1 + \ldots + 1}_{s \text{ times}},$$

where $r + s \leq h_1$, so $n_0 \leq n_0 + s \leq n_0 + h_1$. Let

$$h_2 := \max\{a(m) + b(m) : n_0 \leq m \leq n_0 + h_1\}.$$

Then we obtain

$$\begin{aligned} n &= p_1 + \ldots + p_r + s + n_0 \\ &= p_1 + \ldots + p_r + a(s+n_0)p + b(s+n_0)q \\ &= p_1 + \ldots + p_r + \underbrace{p + \ldots + p}_{a(s+n_0) \text{ times}} + \underbrace{q + \ldots + q}_{b(s+n_0) \text{ times}}, \end{aligned}$$

a representation with at most $h := h_1 + h_2$ elements of P. □

An important special case of a set containing a positive proportion of the primes are primes in arithmetic progression:

Corollary 2.15 *Let a and m be relatively prime integers with $m \geq 2$. Then every sufficiently large integer is the sum of a bounded number of primes in the residue class a modulo m.*

Proof. Let $P := \{p \in \mathbb{P} : p \equiv a \mod m\}$. According to Dirichlet's Theorem on primes in arithmetic progressions, we have

$$\lim_{x\to\infty} \frac{P(x)}{\pi(x)/\varphi(m)} = 1,$$

where $\varphi(m)$ is Euler's totient function, and so Thm. 2.14 applies. $\square$

3 Chebyshev's Theorem

In this chapter, we want to prove Chebyshev's Theorem which was an important step towards proving the Prime Number Theorem[1], but unlike this central result of analytic number theory it can be proved in a short and elementary way. The first proof was published in 1851 by Pafnuty Chebyshev [Che51]. However, the proof has been simplified remarkably since by Paul Erdős and others; we will follow the argument as expounded in Hua's [Hua82, Ch. 5] and Nathanson's [Nat96, Ch. 6] books.

Throughout this chapter, p shall denote a prime number; all sums and products containing p shall run over all primes p as indicated. First, we need to define the decisive functions.

Definition Let $x \geq 0$ be a real number. The *prime counting function* $\pi(x)$ is defined by

$$\pi(x) := \#\{p \in \mathbb{P} : p \leq x\} = \sum_{p \leq x} 1.$$

Sometimes, it is more convenient to use the weighted sum

$$\vartheta(x) := \sum_{p \leq x} \log p,$$

which is known as the *Chebyshev function*.

The Prime Number Theorem states that $\pi(x) \sim x/\log x$, i. e.

$$\lim_{x \to \infty} \frac{\pi(x)}{x/\log x} = 1,$$

but for our purpose, it is sufficient to prove the following:

Theorem 3.1 (Chebyshev) *Let $x \geq 2$. Then:*

$$\frac{x}{\log x} \ll \pi(x) \ll \frac{x}{\log x},$$

where the implied constants are absolute.

First, we want to gather several lemmas in the next section in order to prepare the proof of the main theorem.

[1] This was first proved independently by Jacques Hadamard [Had96] and Charles de la Vallée-Poussin [dlVP96] in 1896 making heavy use of complex analysis. Elementary, but still very long and hard proofs were found by Atle Selberg [Sel49] and Paul Erdős [Erd49] in 1949.

3.1 Preliminary Notes

Lemma 3.2 *Let $n \geq 1$ be an integer.*

(i) We have an estimate for the central binomial coefficient:

$$\binom{2n}{n} < 2^{2n} \leq 2n\binom{2n}{n}.$$

(ii) Moreover, we have

$$\binom{2n+1}{n} < 2^{2n}.$$

Proof. (i) As $\binom{2n}{n}$ is the central binomial coefficient, it is at least $\binom{2n}{k}$ for $k = 0, \dots, 2n$. With this, we have

$$\begin{aligned}\binom{2n}{n} &< \sum_{k=0}^{2n}\binom{2n}{k} = 2^{2n}\\ &= 1 + \sum_{k=1}^{2n-1}\binom{2n}{k} + 1\\ &\leq 2 + (2n-1)\binom{2n}{n}\\ &\leq 2n\binom{2n}{n}.\end{aligned}$$

(ii) This can be proved in a similar fashion. First note by the symmetry of the binomial coefficient that

$$\binom{2n+1}{n} = \binom{2n+1}{n+1}.$$

Hence by the same argument as above

$$2 \cdot \binom{2n+1}{n} = \binom{2n+1}{n} + \binom{2n+1}{n+1} < \sum_{k=0}^{2n+1}\binom{2n+1}{k} = 2^{2n+1}.$$ □

Lemma 3.3 (Erdős) *Let $x \geq 2$ be a real number. By*

$$\mathcal{P}(x) := \prod_{p \leq x} p$$

we denote the so-called primorial of x. Then we have the estimate

$$\mathcal{P}(x) \leq 4^x.$$

Proof. Let $x \geq 2$ be a real number and $n := \lfloor x \rfloor \leq x$ the integral part. Then obviously

$$\prod_{p \leq x} p = \prod_{p \leq n} p \qquad \text{and} \qquad 4^n \leq 4^x,$$

so it suffices to prove the statement for integers. We proceed by induction on n. The case $n = 2$ is obvious. If n is even, then it is not prime, so the primorial does not change, i. e.

$$\prod_{p \leq n} p = \prod_{p \leq n-1} p \leq 4^{n-1} < 4^n$$

by the induction hypothesis. Assume now that $n = 2m + 1$ is odd. We split the product into two parts

$$\prod_{p \leq n} p = \left(\prod_{p \leq m+1} p \right) \cdot \left(\prod_{m+1 < p \leq 2m+1} p \right)$$

and examine them individually. For the first factor, we obtain

$$\prod_{p \leq m+1} p \leq 4^{m+1} \tag{3.1}$$

by the induction hypothesis. For the second factor, we notice that in the binomial coefficient

$$\binom{2m+1}{m} = \frac{(2m+1) \cdot (2m) \cdot (2m-1) \cdots (m+2)}{m \cdot (m-1) \cdot (m-2) \cdots 2 \cdot 1}$$

every prime $m + 2 \leq p \leq 2m + 1$ divides the numerator, but not the denominator. Thus the binomial coefficient is divisible by the product of all these primes, and hence

$$\prod_{m+1 < p \leq 2m+1} p \leq \binom{2m+1}{m} \overset{\text{L3.2}}{<} 2^{2m} = 4^m. \tag{3.2}$$

Combining (3.1) and (3.2), we finally obtain

$$\prod_{p \leq n} p = \prod_{p \leq 2m+1} p = \left(\prod_{p \leq m+1} p \right) \cdot \left(\prod_{m+1 < p \leq 2m+1} p \right) \leq 4^{m+1} \cdot 4^m = 4^{2m+1} = 4^n. \qquad \square$$

For our proof we need a formula for the order of a prime p in $n!$ due to Adrien-Marie Legendre.

Definition Let n be a positive integer. By $v_p(n)$ we denote the highest power of the prime p dividing n, i. e.

$$v_p(n) := \max\{k \geq 0 : p^k \mid n\}.$$

This allows a short representation of n in the canonical prime factorization:

$$n = \prod_{p \leq n} p^{v_p(n)}.$$

Lemma 3.4 (Legendre) *Let n be a positive integer. Then:*

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

Proof. Obviously, v_p is additive, i. e.

$$v_p(m_1 m_2) = v_p(m_1) + v_p(m_2)$$

for all integers m_1 and m_2. By counting the number of factors we obtain:

$$v_p(n!) = v_p\left(\prod_{1\le m\le n} m\right) = \sum_{1\le m\le n} v_p(m) = \sum_{1\le m\le n} \sum_{\substack{k\ge 1\\ p^k\mid m}} 1 = \sum_{k\ge 1} \sum_{\substack{1\le m\le n\\ p^k\mid m}} 1 = \sum_{k=1}^{\infty} \left\lfloor \frac{n}{p^k} \right\rfloor .$$

As upper limit for the sum, $\lfloor \log n/\log p \rfloor$ can be chosen since $\frac{n}{p^k} \ge 1$ if and only if $\log n \ge k \log p$, which is equivalent to $k \le \frac{\log n}{\log p}$. □

Now, we want to establish an important connection between $\pi(x)$ and $\vartheta(x)$.

Lemma 3.5 *Let $x \ge 2$ be a real number and $0 < \varepsilon < 1$. Then we have the estimate*

$$\pi(x) \le \frac{1}{1-\varepsilon} \frac{\vartheta(x)}{\log x} + x^{1-\varepsilon}.$$

Proof. By reducing the range of the sum we obtain

$$\begin{aligned}
\vartheta(x) &= \sum_{p\le x} \log p \\
&\ge \sum_{x^{1-\varepsilon}<p\le x} \log p \\
&\ge \sum_{x^{1-\varepsilon}<p\le x} \log x^{1-\varepsilon} \\
&= (1-\varepsilon)\log x \sum_{x^{1-\varepsilon}<p\le x} 1 \\
&= (1-\varepsilon)\log x(\pi(x) - \pi(x^{1-\varepsilon})) \\
&\ge (1-\varepsilon)\pi(x)\log x - (1-\varepsilon)x^{1-\varepsilon}\log x,
\end{aligned}$$

which yields the inequality by rearrangement. □

3.2 Proof of Chebyshev's Theorem

We are now prepared to prove Chebyshev's Theorem.

Proof of Thm. 3.1. From Lemma 3.3, we attain an upper bound for $\vartheta(x)$:

$$\vartheta(x) = \sum_{p \le x} \log p = \log \prod_{p \le x} p \overset{\text{L3.3}}{\le} \log 4^x = x \log 4. \tag{3.3}$$

We want to use this to obtain an upper bound for $\pi(x)$. For this, we first notice by simple extreme value consideration that $\log x \le x^{1/2}$ for all $x \ge 2$, which is equivalent to

$$\sqrt{x} \le \frac{x}{\log x}. \tag{3.4}$$

We combine these by Lemma 3.5 to achieve our upper bound for $\pi(x)$. So let $0 < \varepsilon < 1$ and $x \ge 2$. Then:

$$\begin{aligned}
\pi(x) &\overset{\text{L3.5}}{\le} \frac{1}{1-\varepsilon} \frac{\vartheta(x)}{\log x} + x^{1-\varepsilon} \\
&\overset{(3.3)}{\le} \frac{\log 4}{1-\varepsilon} \frac{x}{\log x} + x^{1-\varepsilon} \\
&\overset{\varepsilon = 1/2}{=} 2 \log 4 \frac{x}{\log x} + \sqrt{x} \\
&\overset{(3.4)}{\le} 2 \log 4 \frac{x}{\log x} + \frac{x}{\log x} \\
&= (2 \log 4 + 1) \frac{x}{\log x} \ll \frac{x}{\log x}.
\end{aligned}$$

For the lower bound, we first observe

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} = \prod_{p \le 2n} p^{v_p((2n)!) - 2v_p(n!)}.$$

for an arbitrary integer $n \ge 1$. By Lemma 3.4, we have

$$\begin{aligned}
v_p((2n)!) - 2v_p(n!) &= \sum_{k=1}^{\lfloor \log 2n / \log p \rfloor} \left\lfloor \frac{2n}{p^k} \right\rfloor - 2 \sum_{k=1}^{\lfloor \log n / \log p \rfloor} \left\lfloor \frac{n}{p^k} \right\rfloor \\
&= \sum_{k=1}^{\lfloor \log 2n / \log p \rfloor} \left\lfloor 2 \frac{n}{p^k} \right\rfloor - 2 \left\lfloor \frac{n}{p^k} \right\rfloor.
\end{aligned}$$

Obviously, $\lfloor 2\alpha \rfloor - 2 \lfloor \alpha \rfloor$ can only be 0 or 1 for all $\alpha \ge 0$, so we have

$$v_p((2n)!) - 2v_p(n!) \le \left\lfloor \frac{\log 2n}{\log p} \right\rfloor \le \frac{\log 2n}{\log p}, \tag{3.5}$$

and hence

$$\frac{2^{2n}}{2n} \overset{\text{L3.2}}{\leq} \binom{2n}{n} \overset{(3.5)}{\leq} \prod_{p \leq 2n} p^{\log 2n/\log p} = \prod_{p \leq 2n} 2n = (2n)^{\pi(2n)}.$$

This is equivalent to

$$\pi(2n)\log 2n \geq 2n\log 2 - \log 2n. \tag{3.6}$$

Remember that $n \geq 1$ was an arbitrary integer, and $x \geq 2$, so let $n = \lfloor x/2 \rfloor$. Then $2n \leq x < 2n+2$, and so

$$\begin{aligned}\pi(x)\log x &\geq \pi(2n)\log 2n \\ &\overset{(3.6)}{\geq} 2n\log 2 - \log 2n \\ &\geq (x-2)\log 2 - \log x \\ &= x\log 2 - \log x - 2\log 2.\end{aligned}$$

This implies

$$\frac{\pi(x)}{x/\log x} \geq \log 2 - \frac{\log x + 2\log 2}{x} > 0$$

for $x > 4$. Simple computations show that $\frac{\pi(x)}{x/\log x} > 0$ for $2 \leq x \leq 4$, so

$$\pi(x) \gg \frac{x}{\log x}$$

for all $x \geq 2$. □

4 The Selberg Sieve

The aim of this chapter is to present the Selberg sieve which was introduced in 1947 by Atle Selberg in a very concise paper [Sel89]. The argument uses only elementary methods, and yet yields powerful estimates. The main statement of this chapter is therefore:

Theorem 4.1 (Selberg) *Let $B \subset \mathbb{Z}$ be a finite set of integers, and k a positive integer with*

$$\sum_{\substack{b \in B \\ k|b}} 1 = g(k) \cdot \#B + R(k),$$

where $g(k)$ is a multiplicative function[1] with $0 < g(p) < 1$ for all primes p, and $R(k)$ a certain remainder term. Let $\mathcal{N}_B(z)$ denote the number of elements of B that are not divisible by any prime $p \le z$, i. e.

$$\mathcal{N}_B(z) := \sum_{\substack{b \in B \\ p|b \Rightarrow p > z}} 1 = \sum_{\substack{b \in B \\ (b, \mathcal{P}(z))=1}} 1.$$

Let $g_1(k)$ be a completely multiplicative function with $g_1(p) = g(p)$ for all primes p. Then we have the estimate

$$\mathcal{N}_B(z) \le \frac{\#B}{\sum_{1 \le k \le z} g_1(k)} + \sum_{1 \le k_1, k_2 \le z} |R([k_1, k_2])| \cdot \prod_{p|k_1} (1 - g_1(p))^{-1} \cdot \prod_{p|k_2} (1 - g_1(p))^{-1}.$$

This estimate is a very general statement and can be applied in a variety of situations. We will employ it to obtain an upper bound for $r(N)$, the number of representations of N as the sum of two primes. Before doing so, we need to compile some auxiliary results. Our account follows Hua's book [Hua82, Ch. 19].

4.1 Preliminary Notes

We start by giving some remarks on multiplicative functions.

[1] Note that by a *multiplicative* function f we mean the property that $f(mn) = f(m)f(n)$ for relatively prime m and n. If this holds for all integers, we call f *completely multiplicative*.

Definition Let $n = \prod p^{v_p(n)}$ be a positive integer. We say that n is *squarefree* if n contains no multiple prime factors, i. e. if $v_p(n) \leq 1$ for all primes p. With this notion, we define the *Möbius function* $\mu(n)$ by

$$\mu(n) := \begin{cases} 1, & \text{if } n = 1, \\ 0, & \text{if } n \text{ is not squarefree}, \\ (-1)^{\omega(n)}, & \text{otherwise}, \end{cases}$$

where $\omega(n)$ denotes the number of distinct prime divisors of n.

Lemma 4.2 *Let f be a multiplicative function, not identically zero. Then:*

(i) For any integers d_1 and d_2, we have

$$g((d_1, d_2))g([d_1, d_2]) = g(d_1)g(d_2).$$

(ii) For any positive integer n, we have

$$\sum_{d|n} \mu(d)f(d) = \prod_{p|n}(1 - f(p)).$$

Proof. (i) Let $d_1 = \prod p^{v_p(d_1)}$ and $d_2 = \prod p^{v_p(d_2)}$. Obviously, we have

$$\begin{aligned} (d_1, d_2) &= \prod_{p \in \mathbb{P}} p^{\min\{v_p(d_1), v_p(d_2)\}}, \qquad \text{and} \\ [d_1, d_2] &= \prod_{p \in \mathbb{P}} p^{\max\{v_p(d_1), v_p(d_2)\}}. \end{aligned}$$

This yields

$$\begin{aligned} g((d_1, d_2))g([d_1, d_2]) &= \prod_{p \in \mathbb{P}} g\left(p^{\min\{v_p(d_1), v_p(d_2)\}}\right) \prod_{p \in \mathbb{P}} g\left(p^{\max\{v_p(d_1), v_p(d_2)\}}\right) \\ &= \prod_{p \in \mathbb{P}} g\left(p^{\min\{v_p(d_1), v_p(d_2)\}}\right) g\left(p^{\max\{v_p(d_1), v_p(d_2)\}}\right) \\ &= \prod_{p \in \mathbb{P}} g\left(p^{v_p(d_1)}\right) g\left(p^{v_p(d_2)}\right) \\ &= g(d_1)g(d_2). \end{aligned}$$

(ii) Let $n = p_1^{e_1} \cdots p_r^{e_r}$ with $p_i \neq p_j$ for $i \neq j$. Then

$$\begin{aligned} \prod_{p|n}(1 - f(p)) &= 1 - f(p_1) - \ldots - f(p_r) + f(p_1 p_2) + \ldots + f(p_{r-1}p_r) - \ldots \\ &= \sum_{P \subseteq \{p_1, \ldots, p_r\}} (-1)^{\#P} f\left(\prod_{p \in P} p\right) \end{aligned}$$

$$= \sum_{d|n} \mu(d) f(d),$$

where we could include divisors d of n with multiple prime factors in the sum since for those divisors we have $\mu(d) = 0$. □

During our proof, we will need the following basic connection between a multiplicative function and its sum function due to August Ferdinand Möbius.

Theorem 4.3 (Möbius inversion) *Let $f(n)$ be a multiplicative function.*

(i) Let $n_0 \geq 1$ be an integer. If

$$g(n) = \sum_{d|n} f(d)$$

for $1 \leq n \leq n_0$ then we have for such n

$$f(n) = \sum_{d|n} \mu(d) g(n/d).$$

The converse also holds.

(ii) Let $n_0 \geq 1$ be an integer. If

$$g(d) = \sum_{\substack{1 \leq n \leq n_0 \\ d|n}} f(n)$$

for $1 \leq d \leq n_0$ then we have for such d

$$f(d) = \sum_{\substack{1 \leq n \leq n_0 \\ d|n}} \mu(n) g(n/d).$$

The converse also holds.

Proof. (i) First, we need an important fact about the sum function of the Möbius function which follows immediately from Lemma 4.2 (ii) with $f(n) \equiv 1$:

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

This yields:

$$\sum_{d|n} \mu(d) g(n/d) = \sum_{rs=n} \mu(r) g(s)$$
$$= \sum_{rs=n} \mu(r) \sum_{d|s} f(d)$$

$$
\begin{aligned}
&= \sum_{d|n} f(d) \sum_{\substack{rs=n \\ d|s}} \mu(r) \\
&= \sum_{d|n} f(d) \underbrace{\sum_{r|\frac{n}{d}} \mu(r)}_{=1 \text{ iff } n=d} = f(n).
\end{aligned}
$$

The converse is proved in exactly the same way.

(*ii*) This can be proved with the same argument as in (i). □

The following two statements yield important estimates required to prove Selberg's inequality.

Proposition 4.4 *Let $a_i > 0$ and b_i be real numbers for $i = 1, \dots, n$. Then the minimal value of the quadratic form*

$$\sum_{i=1}^{n} a_i x_i^2$$

subject to the constraint

$$\sum_{i=1}^{n} b_i x_i = 1$$

is

$$m := \left(\sum_{i=1}^{n} \frac{b_i^2}{a_i} \right)^{-1},$$

where the minimal value is attained for

$$x_i = \frac{b_i}{a_i} \cdot m, \qquad i = 1, \dots, n.$$

Proof. By the Cauchy-Schwarz inequality, we know

$$1 = \left(\sum_{i=1}^{n} b_i x_i \right)^2 = \left(\sum_{i=1}^{n} \frac{b_i}{\sqrt{a_i}} \cdot \sqrt{a_i} x_i \right)^2 \leq \left(\sum_{i=1}^{n} \frac{b_i^2}{a_i} \right) \cdot \left(\sum_{i=1}^{n} a_i x_i^2 \right).$$

The linear constraint assures that $x \neq 0$, and hence equality is achieved if and only if

$$\sqrt{a_i} x_i = t \cdot \frac{b_i}{\sqrt{a_i}}$$

for some $t \in \mathbb{R}$. Plugging this into the constraint yields

$$1 = \sum_{i=1}^{n} b_i x_i = \sum_{i=1}^{n} \frac{b_i^2}{a_i} t,$$

proving the statement. □

Proposition 4.5 *Let $g(n)$ be a completely multiplicative function with $0 \leq g(p) < 1$ for all primes p, and β_n a sequence with $\beta_n \geq 0$ for all n. Then for $z \geq 1$, and any sequence k_n:*

$$\sum_{1 \leq n \leq z} \beta_n g(n) \prod_{p | k_n} (1 - g(p))^{-1} \geq \sum_{1 \leq n \leq z} g(n) \sum_{\substack{m|n \\ p | \frac{n}{m} \Rightarrow p | k_m}} \beta_m.$$

Proof. Using the geometric series, we obtain:

$$\begin{aligned}
\sum_{1 \leq n \leq z} \beta_n g(n) \prod_{p | k_n} (1 - g(p))^{-1} &= \sum_{1 \leq n \leq z} \beta_n g(n) \prod_{p | k_n} \sum_{m \geq 0} g(p)^m \\
&= \sum_{1 \leq n \leq z} \beta_n g(n) \prod_{p | k_n} \sum_{m \geq 0} g(p^m) \\
&= \sum_{1 \leq n \leq z} \beta_n g(n) \sum_{\substack{r \geq 1 \\ p|r \Rightarrow p|k_n}} g(r) \\
&= \sum_{1 \leq n \leq z} \beta_n \sum_{\substack{r \geq 1 \\ p|r \Rightarrow p|k_n}} g(nr) \\
&= \sum_{1 \leq n \leq z} \beta_n \sum_{\substack{s \geq 1, n|s \\ p|\frac{s}{n} \Rightarrow p|k_n}} g(s) \\
&= \sum_{s \geq 1} g(s) \sum_{\substack{1 \leq n \leq z, n|s \\ p|\frac{s}{n} \Rightarrow p|k_n}} \beta_n \\
&\geq \sum_{1 \leq s \leq z} g(s) \sum_{\substack{1 \leq n \leq z, n|s \\ p|\frac{s}{n} \Rightarrow p|k_n}} \beta_n \\
&= \sum_{1 \leq s \leq z} g(s) \sum_{\substack{n|s \\ p|\frac{s}{n} \Rightarrow p|k_n}} \beta_n.
\end{aligned}$$

□

Our argument relies on solutions to certain congruences. In order to calculate their numbers, we need the following:

Proposition 4.6 *Let $f(x) \in \mathbb{Z}[x]$ be a polynomial, and m_1, m_2 relatively prime integers. Then the number of solutions to the congruence*

$$f(x) \equiv 0 \mod m_1 m_2$$

(counting only distinct congruence classes) is the product of the numbers of solutions to the congruences

$$f(x) \equiv 0 \mod m_1 \qquad \textit{and} \qquad f(x) \equiv 0 \mod m_2.$$

Proof. Obviously, each solution modulo $m_1 m_2$ is also a solution both modulo m_1 and modulo m_2. Let conversely c_1 be a solution modulo m_1 and c_2 a solution modulo m_2.

By the Chinese remainder theorem, there exists a unique c with $c \equiv c_1 \mod m_1$ and $c \equiv c_2 \mod m_2$. Since $m_1 \mid f(c)$ and $m_2 \mid f(c)$ hold, so does $m_1 m_2 \mid f(c)$, using $(m_1, m_2) = 1$. $\square$

Finally, we will need to estimate the growth of the divisor function.

Lemma 4.7 *Let $d(n) := \sum_{k|n} 1$ denote the number of divisors of n, and $z \geq 2$. Then:*

$$\sum_{1 \leq n \leq z} \frac{d(n)}{n} \gg (\log z)^2,$$

where the implied constant is absolute.

Proof. Using

$$\sum_{1 \leq n \leq z} \frac{1}{n} \gg \log z,$$

we obtain:

$$\begin{aligned}
\sum_{1 \leq n \leq z} \frac{d(n)}{n} &= \sum_{1 \leq n \leq z} \frac{1}{n} \sum_{u|n} 1 \\
&= \sum_{1 \leq u \leq z} \sum_{\substack{1 \leq n \leq z \\ u|n}} \frac{1}{n} \\
&= \sum_{1 \leq u \leq z} \sum_{\substack{1 \leq v \leq z/u \\ n = uv}} \frac{1}{n} \\
&= \sum_{1 \leq u \leq z} \sum_{1 \leq v \leq z/u} \frac{1}{uv} \\
&= \sum_{1 \leq uv \leq z} \frac{1}{uv} \\
&\geq \sum_{1 \leq u, v \leq \sqrt{z}} \frac{1}{uv} \\
&= \left(\sum_{1 \leq u \leq \sqrt{z}} \frac{1}{u} \right)^2 \\
&\gg \left(\log z^{1/2} \right)^2 \gg (\log z)^2.
\end{aligned}$$

$\square$

4.2 Deduction of Schnirelmann's Theorem from the Selberg Sieve

Armed with these tools, we can prove an upper bound for $r(N)$.

Lemma 4.8 *Let $N \geq 2$ be an integer. Then*

$$r(N) \ll \frac{N}{(\log N)^2} \sum_{k|N} \frac{\mu(k)^2}{k},$$

where the implied constant is absolute.

Proof. For $N = 2, 3$, we have $r(N) = 0$, and for $N = p_1 + p_2$ odd, we must have $p_1 = 2$ or $p_2 = 2$, so $r(N) \leq 2$. So henceforth, we will assume $N \geq 4$ even. Writing $\mathcal{S}(N)$ for the number of representations of N as the sum of two primes, where both primes exceed $\sqrt{N}$, we have

$$r(N) = \sum_{p_1+p_2=N} 1 \leq \sum_{\substack{p_1+p_2=N \\ p_1,p_2>\sqrt{N}}} 1 + \sum_{\substack{p_1+p_2=N \\ p_1\leq\sqrt{N}}} 1 + \sum_{\substack{p_1+p_2=N \\ p_2\leq\sqrt{N}}} 1 \leq \mathcal{S}(N) + 2\sqrt{N}.$$

Define

$$B := \{c(N-c) : c = 1, \ldots, N\}.$$

If $p_1 + p_2 = N$ and $p_1, p_2 > \sqrt{N}$ then $p_1(N - p_1) = p_2(N - p_2) = p_1p_2$ is not divisible by any prime not exceeding $\sqrt{N}$, so with the notation of Thm. 4.1, we have $\mathcal{S}(N) \leq \mathcal{N}_B(z)$ for all $1 < z \leq \sqrt{N}$. Our task is therefore to find an upper bound for $\mathcal{N}_B(z)$.

Let $M(k)$ denote the number of solutions to

$$x(N-x) \equiv 0 \mod k,$$

with $0 \leq x < k$. By Prop. 4.6, this is a multiplicative function. Using $M(k)$, we have for $k \geq 1$:

$$\sum_{\substack{b\in B \\ k|b}} 1 = \sum_{\substack{1\leq c\leq N \\ k|c(N-c)}} 1 = \sum_{\substack{1\leq c\leq N \\ c(N-c)\equiv 0 \mod k}} 1 \leq \left(\frac{N}{k} + 1\right) M(k) = \frac{M(k)}{k} N + M(k).$$

Moreover:

$$\sum_{\substack{b\in B \\ k|b}} 1 \geq \left\lfloor \frac{N}{k} \right\rfloor M(k) > \left(\frac{N}{k} - 1\right) M(k) = \frac{M(k)}{k} N - M(k).$$

Defining

$$g(k) := \frac{M(k)}{k},$$

we obtain

$$\sum_{\substack{b \in B \\ k \mid b}} 1 = N \cdot g(k) + R(k),$$

where $|R(k)| \leq M(k) \leq k$. Since $M(k)$ is multiplicative, so is $g(k)$. For primes p, we count solutions to $x(N-x) \equiv 0 \mod p$, which is equivalent to $x \equiv 0 \mod p$ or $N - x \equiv 0 \mod p$. Obviously, $x = 0$ is always a solution in $0 \leq x < N$. If $p \mid N$, this is the only one; for $p \nmid N$, there has to be another one, where $N \equiv x \mod p$, so we have

$$g(p) = \begin{cases} \frac{1}{p}, & p \mid N, \\ \frac{2}{p}, & p \nmid N. \end{cases}$$

Since $2 \mid N$ by assumption, we also have $g(2) = 1/2$, and so $0 < g(p) < 1$ for all primes p. Furthermore, we define the completely multiplicative function g_1 by $g_1(p) = g(p)$ for primes p, so we can apply Thm. 4.1.

Let k be a positive integer with $k = p_1^{a_1} \cdots p_r^{a_r}$ for distinct primes p_i. Then:

$$g_1(k) = \prod_{1 \leq i \leq r} g_1(p_i)^{a_i} = \prod_{1 \leq i \leq r} \frac{M(p_i)^{a_i}}{p_i^{a_i}} = \frac{1}{k} \prod_{\substack{1 \leq i \leq r \\ p_i \mid N}} M(p_i)^{a_i} \prod_{\substack{1 \leq i \leq r \\ p_i \nmid N}} M(p_i)^{a_i} = \frac{1}{k} \prod_{\substack{1 \leq i \leq r \\ p_i \nmid N}} 2^{a_i}.$$

We define

$$h(k) = h\left(p_1^{a_1} \cdots p_r^{a_r}\right) := d\left(\prod_{\substack{1 \leq i \leq r \\ p_i \nmid N}} p^{a_i}\right) = \prod_{\substack{1 \leq i \leq r \\ p_i \nmid N}} d(p^{a_i}) = \prod_{\substack{1 \leq i \leq r \\ p_i \nmid N}} (1 + a_i) = \prod_{\substack{p \mid k \\ p \nmid N}} (1 + v_p(k)).$$

This yields by the above:

$$g_1(k) \geq \frac{1}{k} \prod_{\substack{1 \leq i \leq r \\ p_i \nmid N}} 2^{a_i} \geq \frac{1}{k} \prod_{\substack{1 \leq i \leq r \\ p_i \nmid N}} (1 + a_i) = \frac{1}{k} \cdot h(k).$$

In combination with Prop. 4.5, where $\beta_n = h(n)$, and $k_n = N$ for all n, we obtain:

$$\begin{aligned} \prod_{p \mid N} (1 - g_1(p))^{-1} \sum_{1 \leq k \leq z} g_1(k) &\geq \sum_{1 \leq k \leq z} h(k) \frac{1}{k} \prod_{p \mid N} \left(1 - \frac{1}{p}\right)^{-1} \\ &\geq \sum_{1 \leq k \leq z} \frac{1}{k} \sum_{\substack{m \mid k \\ p \mid \frac{k}{m} \Rightarrow p \mid N}} h(m). \end{aligned}$$

Write now $k = p_1^{a_1} \cdots p_t^{a_t} q_1^{b_1} \cdots q_u^{b_u}$, where $p_i \mid N$, and $q_j \nmid N$. The condition $m \mid k$ in the sum means that $m = p_1^{c_1} \cdots p_t^{c_t} q_1^{d_1} \cdots q_u^{d_u}$, where $0 \leq c_i \leq a_i$, and $0 \leq d_j \leq b_j$; the second condition implies that $\frac{k}{m} = p_1^{a_1 - c_1} \cdots p_t^{a_t - c_t}$, i.e. $d_j = b_j$ for all $j = 1, \ldots, u$. Therefore, the second sum runs over all m with $m = p_1^{c_1} \cdots p_t^{c_t} q_1^{b_1} \cdots q_u^{b_u}$, where $0 \leq c_i \leq a_i$ for all $i = 1, \ldots, t$. Thus

$$h(m) = \prod_{\substack{p|m \\ p\nmid N}} (1 + v_p(m)) = (1 + b_1) \cdots (1 + b_u).$$

Using Prop. 4.7, we hence obtain

$$\begin{aligned}
\prod_{p|N} (1 - g_1(p))^{-1} \sum_{1\leq k\leq z} g_1(k) &\geq \sum_{1\leq k\leq z} \frac{1}{k} \sum_{\substack{m|k \\ p|\frac{k}{m} \Rightarrow p|N}} h(m) \\
&= \sum_{1\leq k\leq z} \frac{1}{k} \sum_{0\leq c_i\leq a_i} h\left(p_1^{c_1} \cdots p_t^{c_t} q_1^{b_1} \cdots q_u^{b_u}\right) \\
&= \sum_{1\leq k\leq z} \frac{1}{k} \sum_{0\leq c_i\leq a_i} (1 + b_1) \cdots (1 + b_u) \\
&= \sum_{1\leq k\leq z} \frac{1}{k} (1 + a_1) \cdots (1 + a_t) \cdot (1 + b_1) \cdots (1 + b_u) \\
&= \sum_{1\leq k\leq z} \frac{1}{k} d(p^{a_1}) \cdots d(p^{a_t}) \cdot d(p^{b_1}) \cdots d(p^{b_u}) \\
&= \sum_{1\leq k\leq z} \frac{d(k)}{k} \gg (\log z)^2.
\end{aligned}$$

Note that

$$\prod_{p\in\mathbb{P}} \left(1 - \frac{1}{p^2}\right) = \zeta(2)^{-1} < \infty,$$

where $\zeta(s) = \sum n^{-s}$ denotes the Riemann ζ-function. Applying Lemma 4.2 (ii) to the multiplicative function $\frac{\mu(k)}{k}$, this gives us

$$\begin{aligned}
\sum_{1\leq k\leq z} g_1(k) &\gg (\log z)^2 \cdot \prod_{p|N} (1 - g_1(p)) \\
&= (\log z)^2 \cdot \prod_{p|N} \left(1 - \frac{1}{p^2}\right) \cdot \prod_{p|N} \left(1 + \frac{1}{p}\right)^{-1} \\
&\geq (\log z)^2 \cdot \prod_{p\in\mathbb{P}} \left(1 - \frac{1}{p^2}\right) \cdot \prod_{p|N} \left(1 - \frac{\mu(p)}{p}\right)^{-1} \\
&\gg (\log z)^2 \cdot \left(\sum_{k|N} \mu(k) \frac{\mu(k)}{k}\right)^{-1}.
\end{aligned}$$

Now let k be a positive integer. This yields

$$\prod_{p|k} (1 - g_1(p))^{-1} \leq (1 - g_1(2))^{-1} \cdot (1 - g_1(3))^{-1} \cdot \prod_{5\leq p|k} (1 - g_1(p))^{-1}$$

$$\begin{aligned}
&\le \left(1-\frac{1}{2}\right)^{-1}\cdot\left(1-\frac{2}{3}\right)^{-1}\cdot\prod_{5\le p|k}\left(1-\frac{2}{p}\right)^{-1}\\
&= 2\cdot 3\cdot\prod_{5\le p|k}\frac{p}{p-2}\\
&\le 6\cdot\prod_{p|k}2\le 6\cdot\prod_{p|k}p\le 6k.
\end{aligned}$$

Applying these results to the Selberg sieve (Thm. 4.1), we obtain

$$\begin{aligned}
\mathcal{S}(N)\le\mathcal{N}_B(z)&\le\frac{N}{\sum_{1\le k\le z}g_1(k)}+\sum_{1\le k_1,k_2\le z}|R([k_1,k_2])|\prod_{p|k_1}(1-g_1(p))^{-1}\prod_{p|k_2}(1-g_1(p))^{-1}\\
&\ll\frac{N}{(\log z)^2\left(\sum_{k|N}\frac{\mu(k)^2}{k}\right)^{-1}}+\sum_{1\le k_1,k_2\le z}[k_1,k_2]\cdot k_1\cdot k_2\\
&\le\frac{N}{(\log z)^2}\sum_{k|N}\frac{\mu(k)^2}{k}+\left(\sum_{1\le k\le z}k^2\right)^2\\
&\ll\frac{N}{(\log z)^2}\sum_{k|N}\frac{\mu(k)^2}{k}+z^6.
\end{aligned}$$

Taking $z=N^{1/12}\le N^{1/2}$, this finally yields:

$$\begin{aligned}
r(N)&\le\mathcal{S}(N)+2\sqrt{N}\\
&\ll\frac{N}{(\log N^{1/12})^2}\sum_{k|N}\frac{\mu(k)^2}{k}+N^{6/12}+2N^{1/2}\\
&\ll\frac{N}{(\log N)^2}\sum_{k|N}\frac{\mu(k)^2}{k},
\end{aligned}$$

where we used the fact that

$$\frac{x}{(\log x)^2}\ge\sqrt{x}$$

for sufficiently large x (cf. proof to Thm. 3.1). □

From this upper bound, it is easy to deduce the required upper bound for $\sum r(N)^2$:

Proof of Lemma 2.12. For an arbitrary $x\ge 2$, we need to prove that

$$\sum_{1\le N\le x}r(N)^2\ll\frac{x^3}{(\log x)^4}.$$

Applying Lemma 4.8, we obtain:

$$\sum_{1\le N\le x}r(N)^2\ll\sum_{1\le N\le x}\frac{N^2}{(\log N)^4}\left(\sum_{k|N}\frac{\mu(k)^2}{k}\right)^2$$

$$
\begin{aligned}
&\leq \frac{x^2}{(\log x)^4} \sum_{1\leq N\leq x} \sum_{k_1|N} \sum_{k_2|N} \frac{\mu(k_1)^2\mu(k_1)^2}{k_1k_2} \\
&\leq \frac{x^2}{(\log x)^4} \sum_{1\leq N\leq x} \sum_{k_1,k_2|N} \frac{1}{k_1k_2} \\
&= \frac{x^2}{(\log x)^4} \sum_{1\leq k_1,k_2\leq x} \frac{1}{k_1k_2} \sum_{\substack{1\leq N\leq x\\ [k_1,k_2]|N}} 1 \\
&= \frac{x^2}{(\log x)^4} \sum_{1\leq k_1,k_2\leq x} \frac{1}{k_1k_2} \frac{x}{[k_1,k_2]} \\
&\leq \frac{x^3}{(\log x)^4} \sum_{1\leq k_1,k_2\leq x} \frac{1}{k_1k_2} \frac{1}{(k_1k_2)^{1/2}} \\
&\leq \frac{x^3}{(\log x)^4} \sum_{k_1,k_2=1}^{\infty} \frac{1}{(k_1k_2)^{3/2}} \\
&= \frac{x^3}{(\log x)^4} \left(\sum_{k=1}^{\infty} \frac{1}{k^{3/2}}\right)^2 \ll \frac{x^3}{(\log x)^4},
\end{aligned}
$$

where we used the fact that $[k_1,k_2] \geq \sqrt{k_1k_2}$. □

Hence, it remains to prove Selberg's inequality (Thm. 4.1) in order to prove Schnirelmann's Theorem.

4.3 Proof of the Selberg Sieve

We first establish the following result:

Proposition 4.9 *Let $B \subset \mathbb{Z}$ be a finite set of integers, and k a positive integer with*

$$\sum_{\substack{b\in B\\ k|b}} 1 = g(k)\cdot \#B + R(k),$$

where $g(k)$ is a multiplicative function with $0 < g(p) < 1$ for all primes p, and $R(k)$ a certain remainder term. Let $\mathcal{N}_B(z)$ denote the number of elements of B that are not divisible by any prime $p \leq z$. Then we have the estimate

$$\mathcal{N}_B(z) \leq \frac{\#B}{s} + \sum_{1\leq k_1,k_2\leq z} \lambda_{k_1}\lambda_{k_2} R([k_1,k_2]),$$

where

$$s := \sum_{1\leq k\leq z} \frac{\mu(k)^2}{f(k)},$$

$$f(k) := \sum_{d|k} \frac{\mu(d)}{g(k/d)},$$

and

$$\lambda_k := \frac{\mu(k)}{sf(k)g(k)} \sum_{\substack{1\le m\le z/k \\ (m,k)=1}} \frac{\mu(m)^2}{f(m)}$$

Proof. Let $\lambda_1, \dots, \lambda_{\lfloor z \rfloor}$ be arbitrary real numbers with $\lambda_1 = 1$ and $\lambda_i \geq 0$. Then:

$$\begin{aligned}
\mathcal{N}_B(z) &= \sum_{\substack{b\in B \\ p|b\Rightarrow p>z}} 1 \\
&= \sum_{\substack{b\in B \\ p|b\Rightarrow p>z}} \left(\sum_{\substack{1\le k\le z \\ k|b}} \lambda_k \right)^2 \\
&\le \sum_{b\in B} \sum_{\substack{1\le k_1,k_2\le z \\ k_1,k_2|b}} \lambda_{k_1}\lambda_{k_2} \\
&= \sum_{1\le k_1,k_2\le z} \lambda_{k_1}\lambda_{k_2} \sum_{\substack{b\in B \\ [k_1k_2]|b}} 1 \\
&= \sum_{1\le k_1,k_2\le z} \lambda_{k_1}\lambda_{k_2} \left(g([k_1k_2]) \cdot \#B + R([k_1k_2]) \right) \\
&\overset{\text{L4.2}}{=} \#B \underbrace{\sum_{1\le k_1,k_2\le z} \frac{\lambda_{k_1}g(k_1)\lambda_{k_2}g(k_2)}{g((k_1,k_2))}}_{=:S(z)} + \underbrace{\sum_{1\le k_1,k_2\le z} \lambda_{k_1}\lambda_{k_2}R([k_1k_2])}_{=:T(z)}.
\end{aligned}$$

We therefore need an upper bound for $S(z)$. Since $g(k)$ is multiplicative, so is $1/g(k)$. Hence we have by Möbius inversion (Thm. 4.3):

$$\frac{1}{g(k)} = \sum_{d|k} f(d).$$

This yields:

$$\begin{aligned}
S(z) &= \sum_{1\le k_1,k_2\le z} \frac{1}{g((k_1,k_2))} \lambda_{k_1}g(k_1)\lambda_{k_2}g(k_2) \\
&= \sum_{1\le k_1,k_2\le z} \sum_{d|(k_1,k_2)} f(d)\lambda_{k_1}g(k_1)\lambda_{k_2}g(k_2) \\
&= \sum_{1\le d\le z} f(d) \sum_{\substack{1\le k_1,k_2\le z \\ d|(k_1,k_2)}} \lambda_{k_1}g(k_1)\lambda_{k_2}g(k_2)
\end{aligned}$$

$$= \sum_{1\le d\le z} f(d) \sum_{\substack{1\le k_1\le z\\ d|k_1}} \lambda_{k_1} g(k_1) \sum_{\substack{1\le k_2\le z\\ d|k_2}} \lambda_{k_2} g(k_2)$$

$$= \sum_{1\le d\le z} f(d) \left(\sum_{\substack{1\le k\le z\\ d|k}} \lambda_k g(k) \right)^2 .$$

We now want to use Prop. 4.4 to calculate the minimal value of $S(z)$. For this purpose, let

$$x_d := \sum_{\substack{1\le k\le z\\ d|k}} \lambda_k g(k).$$

By Möbius inversion (Thm. 4.3), we have

$$\lambda_k g(k) = \sum_{\substack{1\le m\le z\\ k|m}} \mu(m/k) x_m = \sum_{1\le m\le \frac{z}{k}} \mu(m) x_{mk}. \tag{4.1}$$

For $k=1$, we have $\lambda_1 = 1$ and $g(1) = 1$ since $g(k)$ is multiplicative, and hence the linear constraint

$$1 = \lambda_1 g(1) = \sum_{1\le m\le z} \mu(m) x_m.$$

By Prop. 4.4 we therefore obtain, that the minimum value of $S(z)$ is

$$s^{-1} = \sum_{1\le k\le z} \frac{\mu(k)^2}{f(k)},$$

which is attained for

$$x_d = \frac{\mu(d)}{f(d)} \cdot s^{-1}.$$

Plugging this into (4.1) yields the choice for λ_k in the statement:

$$\begin{aligned}\lambda_k g(k) &= \sum_{1\le m\le \frac{z}{k}} \mu(m) \frac{\mu(mk)}{f(mk)} s^{-1}\\ &= \sum_{\substack{1\le m\le \frac{z}{k}\\ (m,k)=1}} \mu(m) \frac{\mu(mk)}{f(mk)} s^{-1}\\ &= \frac{\mu(k)}{sf(k)} \sum_{\substack{1\le m\le \frac{z}{k}\\ (m,k)=1}} \frac{\mu(m)^2}{f(m)},\end{aligned}$$

where we used that $\mu(km) = 0$ if $(m,k) > 1$. By this choice, we have $S(z) = s^{-1}$, and hence

$$\mathcal{N}_B(z) \le \#B \cdot S(z) + T(z) = \frac{\#B}{s} + \sum_{1\le k_1,k_2\le z} \lambda_{k_1}\lambda_{k_2} R([k_1k_2]). \qquad \square$$

In order to prove Selberg's inequality (Thm. 4.1), we obviously need a lower bound for s and an upper bound for λ_k.

Proof of Thm. 4.1. First, we will show that

$$\sum_{1\le k\le z} g_1(k) \le \sum_{1\le k\le z} \frac{\mu(k)^2}{f(k)} = s.$$

For a prime p, we have

$$f(p) = \sum_{d|p} \frac{\mu(d)}{g(p/d)} = \frac{\mu(1)}{g(p)} + \frac{\mu(p)}{g(1)} = \frac{1}{g(p)} - 1 = \frac{1-g(p)}{g(p)}.$$

Let k be squarefree. Then:

$$\begin{aligned}
\frac{\mu(k)^2}{f(k)} &= \mu(k)^2 \prod_{p|k} \frac{1}{f(p)} \\
&= \mu(k)^2 \prod_{p|k} \frac{g(p)}{1-g(p)} \\
&= \mu(k)^2 \frac{\prod_{p|k} g_1(p)}{\prod_{p|k} 1-g_1(p)} \\
&= \mu(k)^2 g_1(k) \prod_{p|k} (1-g_1(p))^{-1} \ge 0. \qquad (4.2)
\end{aligned}$$

This identity holds still for $k=1$ and non-squarefree k (in this case both sides simply vanish), so by Prop. 4.5 we have:

$$\begin{aligned}
\sum_{1\le k\le z} \frac{\mu(k)^2}{f(k)} &= \sum_{1\le k\le z} \mu(k)^2 g_1(k) \prod_{p|k} (1-g_1(p))^{-1} \\
&\ge \sum_{1\le k\le z} g_1(k) \sum_{\substack{m|k \\ p|\frac{k}{m} \Rightarrow p|m}} \mu(m)^2.
\end{aligned}$$

Let d_k denote the greatest squarefree divisor of k (also known as the *radical* of k). Then if $p \mid \frac{k}{d_k}$ then $p \mid k$, and so $p \mid d_k$, i. e. d_k satisfies the condition on m in the second sum, hence this sum does not vanish. This yields:

$$\begin{aligned}
\sum_{1\le k\le z} \frac{\mu(k)^2}{f(k)} &\ge \sum_{1\le k\le z} g_1(k) \sum_{\substack{m|k \\ p|\frac{k}{m} \Rightarrow p|m}} \mu(m)^2 \\
&\ge \sum_{1\le k\le z} g_1(k) \sum_{d_k|k} \mu(d_k)^2
\end{aligned}$$

$$\geq \sum_{1\leq k\leq z} g_1(k).$$

It remains to prove that the upper bound for $|\lambda_k|$. Using the fact that both $f(k)$ and $g(k)$ are non-negative, we obtain

$$\begin{aligned}
|\lambda_k| &= \left|\frac{\mu(k)}{f(k)g(k)}\right| \cdot \underbrace{\left|\sum_{\substack{1\leq m\leq z/k\\(m,k)=1}} \frac{\mu(m)^2}{f(m)}\right| \cdot \left|\sum_{1\leq m\leq z} \frac{\mu(m)^2}{f(m)}\right|^{-1}}_{\leq 1} \\
&\leq \left|\frac{\mu(k)}{f(k)g(k)}\right| = \frac{\mu(k)^2}{f(k)g_1(k)} \\
&\overset{(4.2)}{=} \mu(k)^2 \cdot \frac{g_1(k)}{g_1(k)} \cdot \prod_{p|k}(1-g_1(p))^{-1} \\
&\leq \prod_{p|k}(1-g_1(p))^{-1},
\end{aligned}$$

where again we needed the fact that $\mu(k) = 0$ if k is not squarefree. Applying these estimates to Prop. 4.9, we finally have:

$$\begin{aligned}
\mathcal{N}_B(z) &\leq \frac{\#B}{s} + \sum_{1\leq k_1,k_2\leq z} \lambda_{k_1}\lambda_{k_2} R([k_1,k_2]) \\
&\leq \frac{\#B}{s} + \sum_{1\leq k_1,k_2\leq z} |\lambda_{k_1}|\cdot|\lambda_{k_2}|\cdot|R([k_1,k_2])| \\
&\leq \frac{\#B}{\sum_{1\leq k\leq z} g_1(k)} + \sum_{1\leq k_1,k_2\leq z} |R([k_1,k_2])| \cdot \prod_{p|k_1}(1-g_1(p))^{-1} \cdot \prod_{p|k_2}(1-g_1(p))^{-1},
\end{aligned}$$

proving the required estimate. □

4.4 Applications of the Selberg Sieve to Twin Primes

In our proof of Schnirelmann's Theorem, we applied the Selberg sieve to the sequence $(n(N-n))_{1\leq n\leq N}$ in order to sift out elements with more than two prime factors. We can apply the same idea to twin primes, i. e. primes p such that $p+2$ is prime as well. For this, we examine the sequence $(n(n+2))_{1\leq n\leq N}$, and want to find members with exactly two prime factors. The famous twin prime conjecture states that there are infinitely many twin primes. This has been suspected since antiquity, but still remains unproved. For the given reasons, this conjecture is believed to be equally hard as Goldbach's problem. This section will show how to apply Selberg's sieve methods to twin primes and similar problems.

Leonard Euler was the first to prove that the harmonic series that runs only over prime values is still divergent, i. e.

$$\sum_{p\in\mathbb{P}} \frac{1}{p} = \infty.$$

The Norwegian mathematician Viggo Brun proved in 1919 that the harmonic series over the twin primes converges, using sieve methods based on the inclusion-exclusion principle. Unfortunately, this leaves the question about the infinity of the twin primes unanswered. Selberg's sieve is a more sophisticated version of Brun's sieve, and we will use the results of the previous sections to obtain stronger estimates and more general results than Brun's original ones.

Theorem 4.10 *Let N be a positive integer, $x \geq 0$ a real number, and $\pi_N(x)$ denote the number of primes p not exceeding x such that $p+N$ is prime as well. Then*

$$\pi_N(x) \ll \frac{x}{(\log x)^2} \sum_{k\mid N} \frac{\mu(k)^2}{k},$$

where the implied constant is absolute.

From this result, we can immediately deduce Brun's Theorem:

Corollary 4.11 (Brun) *The sum over the reciprocals of the twin primes converges, i. e.*

$$\lim_{N\to\infty} \sum_{\substack{2\leq p\leq N\\ p,p+2\in\mathbb{P}}} \left(\frac{1}{p} + \frac{1}{p+2}\right) = B_2 < \infty,$$

where B_2 is known as Brun's constant.

Proof. Let $p_1, p_2, \ldots$ denote the sequence of primes such that $p_i + 2$ is prime as well. Obviously

$$\sum_{\substack{2\leq p\leq x\\ p,p+2\in\mathbb{P}}} \left(\frac{1}{p} + \frac{1}{p+2}\right) \leq \sum_{\substack{2\leq p\leq x\\ p,p+2\in\mathbb{P}}} \left(\frac{1}{p} + \frac{1}{p}\right) = 2 \sum_{\substack{2\leq p\leq x\\ p,p+2\in\mathbb{P}}} \frac{1}{p},$$

so it suffices to prove that $\sum 1/p_i$ converges. By Thm. 4.10, we have

$$n = \pi_2(p_n) \ll \frac{p_n}{(\log p_n)^2} \leq \frac{p_n}{(\log n)^2},$$

and hence

$$\frac{1}{p_n} \ll \frac{1}{n(\log n)^2}.$$

Using

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(-\frac{1}{\log x}\right) = \frac{1}{x(\log x)^2}$$

this yields

$$\begin{aligned}
\sum_{n=1}^{M} \frac{1}{p_n} &= \frac{1}{3} + \sum_{n=2}^{M} \frac{1}{p_n} \\
&\ll \frac{1}{3} + \sum_{n=2}^{M} \frac{1}{n(\log n)^2} \\
&\leq \frac{1}{3} + \frac{1}{2(\log 2)^2} + \int_2^M \frac{\mathrm{d}x}{x(\log x)^2} \\
&= \frac{1}{3} + \frac{1}{2(\log 2)^2} + \frac{1}{\log 2} - \frac{1}{\log M} \\
&\xrightarrow{M\to\infty} \frac{1}{3} + \frac{1}{2(\log 2)^2} + \frac{1}{\log 2} < \infty.
\end{aligned}$$

□

It therefore remains to prove Thm. 4.10 which can be done in exactly the same way as Lemma 4.8.

Proof of Thm. 4.10. Without loss of generality, we can assume that N is even since otherwise one member of the pair $(n, n+N)$ must be even, and hence there is at most one such pair. Writing $\mathcal{S}(x)$ for the number of primes p between $\sqrt{x}$ and x such that $p+N$ is prime as well, we obtain

$$\pi_N(x) = \sum_{\substack{2\leq p\leq x \\ p,p+N\in\mathbb{P}}} 1 = \sum_{\substack{2\leq p\leq \sqrt{x} \\ p,p+N\in\mathbb{P}}} 1 + \sum_{\substack{\sqrt{x}< p\leq x \\ p,p+N\in\mathbb{P}}} 1 \leq \mathcal{S}(x) + \sqrt{x}.$$

We define the sequence

$$B := \{c(c+N) : 1 \leq c \leq x\}.$$

If p and $p+N$ are both prime with p exceeding $\sqrt{x}$ then $p(p+N)$ is not divisible by any prime not exceeding $\sqrt{x}$. Hence $\mathcal{S}(x) \leq \mathcal{N}_B(z)$ for any $2 < z \leq \sqrt{x}$, so again we need an upper bound for $\mathcal{N}_B(z)$.

Let $M(k)$ denote the number of solutions to the congruence

$$y(y+N) \equiv 0 \mod k$$

with $0 \leq y < k$. By Prop. 4.6, this is a multiplicative function with

$$\sum_{\substack{b\in B \\ k|b}} 1 = \sum_{\substack{1\leq c\leq x \\ c(c+N)\equiv 0 \mod k}} 1 = \#B \cdot g(k) + R(k),$$

where $g(k) = M(k)/k$, and $|R(k)| \leq M(k) \leq k$ (cf. proof of Lemma 4.8). For primes p, the function $M(p)$ is counting solutions to $y(y+N) \equiv 0 \mod p$ which has the solutions

$y = 0$ and $y \equiv -N \mod p$ where these two coincide if $p \mid N$. This yields

$$g(p) = \begin{cases} \frac{1}{p}, & p \mid N, \\ \frac{2}{p}, & p \nmid N. \end{cases}$$

Since $2 \mid N$ by assumption, we have $g(2) = 1/2$, and so $0 < g(p) < 1$ for all primes p. Defining the completely multiplicative function g_1 by $g_1(p) := g(p)$, we can apply Thm. 4.1, and hence obtain the same estimates as in our proof to Lemma 4.8. This yields

$$\mathcal{S}(x) \leq \mathcal{N}_B(z) \ll \frac{\#B}{(\log z)^2} \sum_{k \mid N} \frac{\mu(k)^2}{k} + z^6.$$

Taking $z = x^{1/12} \leq x^{1/2}$, we finally obtain

$$\begin{aligned} \pi_N(x) &\leq \mathcal{S}(x) + \sqrt{x} \\ &\ll \frac{\#B}{(\log x^{1/12})^2} \sum_{k \mid N} \frac{\mu(k)^2}{k} + x^{6/12} + x^{1/2} \\ &\ll \frac{x}{(\log x)^2} \sum_{k \mid N} \frac{\mu(k)^2}{k}. \end{aligned}$$ □

Further applications of the Selberg sieve include estimates for the number of primes in a certain interval, and the proof of the Brun-Titchmarsh Theorem on the number of primes in arithmetic progression, but an exhaustive account of these theorems would go far beyond the scope of this treatise.

5 Waring's Problem

It was known since the ancient Greece that every positive integer can be represented as the sum of at most four squares, but it should take until 1770 when Joseph Louis Lagrange proved his famous four-square theorem. In the same year, Edward Waring asked the question if for every integer $k \geq 2$ there is a constant $g(k)$ such that every integer can be represented as the sum of at most $g(k)$ powers of exponent k. Many special cases have been settled during the 19^{th} century, but a general solution to Waring's problem would not have been found until 1909 when David Hilbert came up with his proof [Hil09] which made heavy use of analytic means but would also yield upper bounds for $g(k)$.

An elementary proof was given by Yuri Linnik [Lin43] in 1943 applying the Schnirelmann density to this problem. An account of the original proof was presented by Aleksandr Khinchin [Khi52] reviving Linnik's combinatorial argument. Georg Johann Rieger [Rie54] refined Linnik's method to obtain upper bounds for $g(k)$ (although these estimates are weaker than those analytic means yield). In Melvyn B. Nathanson's book [Nat00], a generalisation to integer-valued polynomials is presented. We will follow Hua Loo Keng's argument [Hua59, Hua82] employing exponential sums to receive the required estimates. Donald J. Newman [New60, New97] gives a similar proof with different estimates. An account of the proof, using Hua's methods but avoiding the exponential sums, was provided by Yuri V. Nesterenko [Nes06].

Throughout this chapter, let $k \geq 2$ be a fixed exponent, and A_k be the sequence of all k^{th} powers, i. e.

$$A_k := \{n^k : n \in \mathbb{N}_0\}.$$

All implied constants may only depend on k (and c consequently). We aim to prove the following statement:

Theorem 5.1 (Waring's Problem) *The sequence A_k is a basis of finite order for all $k \geq 2$, i. e. there is a constant $g(k)$ such that every non-negative integer can be represented as the sum of at most $g(k)$ elements of A_k.*

Because of Thm. 2.8 it suffices to prove that there exists a constant $c = c(k) \geq k$ such that cA_k has positive density. Since 0 and 1 are members of A_k for all k, this implies that A_k itself is a basis of finite order. The strategy is the same as in our proof of Schnirelmann's Theorem in Chapter 2: We give upper and lower bounds for the number of representations of an integer as the sum of k^{th} powers, and by this, deduce that the set cA_k has positive density. The next section will outline the way to achieve this.

5.1 Preliminary Notes

First we define $r_c(N)$ to be the number of solutions to

$$x_1^k + \ldots + x_c^k = N,$$

with $x_i \geq 0$ for all $i = 1, \ldots, c$, i. e.

$$r_c(N) = \sum_{\substack{x_1^k + \ldots + x_c^k = N \\ x_i \geq 0}} 1.$$

The required lower bound is easy:

Lemma 5.2 *Let x be a real number. Then:*

$$\sum_{N \leq x} r_c(N) \gg x^{c/k},$$

where the implied constant depends on k only.

Proof. Let n_i for $i = 1, \ldots, c$ be a non-negative integer not exceeding $(x/c)^{1/k}$. Then certainly

$$n_1^k + \ldots + n_c^k \leq x,$$

so for every choice of n_i we receive one of the representation the sum is counting. In total, we have $\lfloor x/c \rfloor^{c/k}$ such distinct choices giving us the claimed lower bound. □

Analogously to the proof of Schnirelmann's Theorem, we now need an upper bound for $r_c(N)$. Our aim is therefore to prove the following estimate:

Lemma 5.3 *Let x be a real number, and N a non-negative integer. Then:*

$$r_c(N) \ll N^{\frac{c}{k}-1},$$

where the implied constant depends on k only.

By summation, we immediately gain the estimate

$$\sum_{N \leq x} r_c(N)^2 \ll x^{2\frac{c}{k}-1},$$

enabling us to immediately apply Prop. 2.9. However, we will use the slightly stronger result of Lemma 5.3 to deduce a solution to Waring's problem in a more insightful way.

Proof of Thm. 5.1. Assume that $\sigma(cA_k) = 0$. Then by Lemma 2.2, for all $\varepsilon > 0$ we find some x such that

$$cA_k(x) < \varepsilon x. \tag{5.1}$$

Obviously, if $N \notin cA_k$, then $r_c(N) = 0$, so we can skip those N when counting solutions. Using Lemma 5.3, we obtain:

$$\begin{aligned}
\sum_{0 \le N \le x} r_c(N) &= \sum_{\substack{0 \le N \le x \\ N \in cA_k}} r_c(N) \\
&\overset{\text{L5.3}}{\ll} \sum_{\substack{0 \le N \le x \\ N \in cA_k}} N^{\frac{c}{k}-1} \\
&\le \sum_{\substack{0 \le N \le x \\ N \in cA_k}} x^{\frac{c}{k}-1} \\
&= x^{\frac{c}{k}-1} cA_k(x) \\
&\overset{(5.1)}{<} x^{\frac{c}{k}-1} \cdot \varepsilon x = \varepsilon x^{c/k}.
\end{aligned}$$

As ε can be chosen to be arbitrarily small, this contradicts our estimate in Lemma 5.2. Thus the set cA_k must have positive density, and is hence by Thm. 2.8 a basis of finite order, say h. So every non-negative integer can be represented as the sum of h elements of cA_k. But these elements are themselves sums of k^{th} powers, so every integer non-negative integer can be represented as the sum of at most $h \cdot c < \infty$ such k^{th} powers, solving Waring's problem. $\square$

Consequently, it remains to prove Lemma 5.3 for some constant c. As it turns out, it suffices to take $c = 8^{k-1}$, so we will henceforth use this value. We will now reduce Lemma 5.3 to an estimate for exponential sums. For the sake of clarity, we define

$$e(x) := \exp(2\pi i x).$$

First, we need an easy tool we will use several times to transform counting solution into estimating integrals:

Lemma 5.4 *Let q be an integer. Then:*

$$\int_0^1 e(q\alpha)\, \mathrm{d}\alpha = \begin{cases} 1, & \text{if } q = 0, \\ 0, & \text{if } q \neq 0. \end{cases}$$

Proof. The case $q = 0$ is obvious since the integration is empty. For $q \neq 0$, the integrand $e(q\alpha)$ describes a $|q|$-fold closed circle around the origin, hence the integral vanishes by Cauchy's integral theorem. (The statement also follows immediately from the fundamental theorem of calculus.) $\square$

We will now establish the required estimate:

Theorem 5.5 *Let $P \geq 1$ be an integer, and $c = 8^{k-1}$. Then:*

$$\int_0^1 \left| \sum_{x=0}^{P} e(x^k \alpha) \right|^c \mathrm{d}\alpha \ll P^{c-k},$$

where the implied constant depends on k only.

Before we proceed to prove Thm. 5.5, we will show how to prove Lemma 5.3 with this tool.

Proof of Lemma 5.3. Let N be an arbitrary non-negative integer. Then, using the fact that $|e(x)| = 1$ for all real numbers x:

$$\begin{aligned}
r_c(N) = |r_c(N)| &= \left| \sum_{\substack{x_1^k + \ldots + x_c^k = N \\ x_i \geq 0}} 1 \right| \\
&\overset{\text{L5.4}}{=} \left| \sum_{x_1=0}^{\lfloor N^{1/k} \rfloor} \cdots \sum_{x_c=0}^{\lfloor N^{1/k} \rfloor} \int_0^1 e(\alpha(x_1^k + \ldots + x_c^k - N)) \,\mathrm{d}\alpha \right| \\
&= \left| \int_0^1 \left(\sum_{x_1=0}^{\lfloor N^{1/k} \rfloor} e(x_1^k \alpha) \right) \cdots \left(\sum_{x_c=0}^{\lfloor N^{1/k} \rfloor} e(x_c^k \alpha) \right) e(-N\alpha) \,\mathrm{d}\alpha \right| \\
&\leq \int_0^1 \left| \sum_{x=0}^{\lfloor N^{1/k} \rfloor} e(x^k \alpha) \right|^c |e(-N\alpha)| \,\mathrm{d}\alpha \\
&= \int_0^1 \left| \sum_{x=0}^{\lfloor N^{1/k} \rfloor} e(x^k \alpha) \right|^c \mathrm{d}\alpha \\
&\overset{\text{T5.5}}{\ll} \left\lfloor N^{1/k} \right\rfloor^{c-k} \leq N^{\frac{c}{k}-1}.
\end{aligned}$$

□

In order to solve Waring's problem, we therefore need to prove Thm. 5.5 which is the subject of the next section.

5.2 Linear Equations and Exponential Sums

In the proof of Thm. 5.5, we will need to estimate the number of solutions to linear equations. The required upper bounds are provided by the next proposition:

Proposition 5.6 *Let N, X, Y be integers with $X, Y \geq 1$, and $q(N)$ denote the number of integral solutions to*

$$x_1y_1 + x_2y_2 = N \tag{5.2}$$

with $|x_i| \leq X$ and $|y_i| \leq Y$. Then:

$$q(N) \leq \begin{cases} 27X^{3/2}Y^{3/2}, & \text{if } N = 0, \\ 60XY \sum\limits_{d|N} \frac{1}{d}, & \text{if } N \neq 0. \end{cases}$$

Proof. First, let $N = 0$. Obviously, for x_i and y_i there are $2X + 1$ and $2Y + 1$ choices, respectively. Consider x_1, x_2, and y_1 to be chosen. Then there is at most one choice for y_2 left, so

$$q(0) \leq (2X + 1)^2(2Y + 1) \leq (3X)^2 3Y = 27X^2Y.$$

On the other hand, if x_1, y_1, and y_2 are considered to be chosen. Then we have $q(0) \leq 27XY^2$, so altogether we have

$$q(0) \leq \min\{27X^2Y, 27XY^2\} \leq \sqrt{27X^2Y \cdot 27XY^2} = 27X^{3/2}Y^{3/2},$$

where we used that the minimum does not exceed the geometric mean.

Not let $N \neq 0$. Without loss of generality, we can assume $X \leq Y$. By $q_1(N)$, we denote the number of integral solutions to (5.2) with $(x_1, x_2) = 1$, where $|x_2| \leq |x_1| \leq X$ and $|y_i| \leq Y$. Clearly $x_1 \neq 0$, because otherwise $x_2 = 0$, giving $N = 0$ which contradicts our assumption. Now by $q_2(N, x_1, x_2)$, we denote the number of integral solutions to (5.2) with fixed x_1 and x_2. Since x_1 and x_2 are relatively prime, this is soluble for all N by the Euclidean algorithm. Given one particular solution (y_1', y_2'), all other solutions are of the form

$$y_1 = y_1' + tx_2, \qquad y_2 = y_2' + tx_1.$$

Since we required $|y_i| \leq Y$, this gives us the condition

$$|t| \leq \left|\frac{y_2' - y_2}{x_1}\right| \leq \frac{2Y}{|x_1|}.$$

So the number of possible values for t and hence the value of $q_2(N; x_1, x_2)$ is bounded by:

$$q_2(N; x_1, x_2) \leq 2\frac{2Y}{|x_1|} + 1 \leq \frac{4Y + X}{|x_1|} \leq \frac{5Y}{|x_1|}.$$

Thus we can estimate the value of $q_1(N)$ by summation over all cases:

$$\begin{aligned} q_1(N) &\leq \sum_{1 \leq |x_1| \leq X} \sum_{|x_2| \leq |x_1|} \frac{5Y}{|x_1|} \\ &\leq 5Y \sum_{1 \leq |x_1| \leq X} \frac{1}{|x_1|}(2|x_1| + 1) \end{aligned}$$

$$\begin{aligned} &\leq 5Y \sum_{1 \leq |x_1| \leq X} 3 \\ &\leq 5Y \cdot 3 \cdot 2X = 30XY. \end{aligned}$$

Dropping the condition $|x_2| \leq |x_1|$, this hence yields that there are at most $2 \cdot 30XY$ solutions to (5.2) with $(x_1, x_2) = 1$.

Now let $(x_1, x_2) = d > 1$ with $d \mid N$. (Otherwise there are not solutions to (5.2) according to the Euclidean algorithm.) We examine the equation

$$x_1' y_1 + x_2' y_2 = \frac{N}{d}.$$

Obviously, every solution to this equation yields a solution to (5.2) with $x_1 = dx_1'$ and $x_2 = dx_2'$, where the restrictions have changed to

$$|x_i'| \leq \frac{X}{d}, \qquad |y_i| \leq Y, \qquad (x_1', x_2') = 1.$$

By the above, the number of solutions to this kind of equation does not exceed $60\frac{X}{d}Y$. Summing over all divisors of N, we finally obtain

$$q(N) \leq 60XY \sum_{d \mid N} \frac{1}{d},$$

as required. $\square$

Instead of proving Thm. 5.5 directly, we will prove a stronger result which allows us to tackle the problem by induction.

Theorem 5.7 (Hua's Lemma) *Let $P \geq 1$ be an integer, $f(x) = a_k x^k + \ldots + a_1 x + a_0 \in \mathbb{Z}[x]$ a polynomial with degree k, and coefficient*

$$a_k \ll 1, \qquad a_{k-1} \ll P, \qquad \ldots, \qquad a_1 \ll P^{k-1}, \qquad a_0 \ll P^k.$$

Then we have the estimate:

$$\int_0^1 \left| \sum_{x=0}^{P} e(f(x)\alpha) \right|^{8^{k-1}} \mathrm{d}\alpha \ll P^{8^{k-1}-k}, \tag{5.3}$$

where the implied constant depends on k only.

Obviously, $f(x) = x^k$ is a polynomial satisfying the condition of the theorem, so Thm. 5.5 is an immediate consequence of the above.

Proof. We proceed by induction over k, the degree of the polynomial f. So, let $k = 2$, i. e. we have

$$f(x) = a_2 x^2 + a_1 x + a_0,$$

where

$$a_2 \ll 1, \qquad a_1 \ll P, \qquad a_0 \ll P^2.$$

Let Q denote the number of integral solutions to

$$f(x_1) + f(x_2) + f(x_3) + f(x_4) - f(y_1) - f(y_2) - f(y_3) - f(y_4) = 0 \tag{5.4}$$

with $0 \le x_i, y_j \le P$. Then we have

$$\begin{aligned} Q &= \sum_{\substack{0\le x_i,y_j\le P \\ f(x_1)+\ldots+f(x_4)-f(y_1)-\ldots-f(y_4)=0}} 1 \\ &= \sum_{0\le x_i,y_j\le P} \int_0^1 e(\alpha(f(x_1)+\ldots+f(x_4)-f(y_1)-\ldots-f(y_4)))\,\mathrm{d}\alpha \\ &= \int_0^1 \left(\sum_{0\le x\le P} e(f(x)\alpha)\right)^4 \left(\sum_{0\le x\le P} e(-f(x)\alpha)\right)^4 \mathrm{d}\alpha \\ &= \int_0^1 \left(\sum_{0\le x\le P} e(f(x)\alpha)\right)^4 \left(\overline{\sum_{0\le x\le P} e(f(x)\alpha)}\right)^4 \mathrm{d}\alpha \\ &= \int_0^1 \left|\sum_{0\le x\le P} e(f(x)\alpha)\right|^8 \mathrm{d}\alpha, \end{aligned}$$

where $\bar{z}$ denotes the complex conjugate of z. Hence, we need to find an upper bound for Q. For $i = 1, \ldots, 4$, define

$$z_i := x_i - y_i, \qquad w_i := a_2(x_i + y_i) + a_1.$$

For any solution to

$$z_1w_1 + z_2w_2 + z_3w_3 + z_4w_4 = 0, \tag{5.5}$$

we obtain by plugging in

$$z_iw_i = (x_i - y_i)(a_2(x_i + y_i) + a_1) = a_2x_i^2 + a_1x_i - a_2y_i^2 - a_1y_i = f(x_i) - f(y_i),$$

and hence a solution to (5.4). Let R denote the number of integral solutions to (5.5) with $|z_i|, |w_j| \ll P$. By $a_m \ll P^{2-m}$ and $0 \le x_i, y_j \le P$, we have $a_mx_i^m \ll P^{2-m}P^m = P^2$ and $a_my_j^m \ll P^2$, thus the restriction on z_i and w_j ensures that every solution to (5.5) in z_i and w_j yields a solution to (5.4) in $f(x_i)$ and $f(y_j)$. Our task is therefore to find an upper bound for R, giving an upper bound for Q and hence by the above for the exponential sum in (5.3). So let $q(N)$ denote the number of integral solution to

$$z_1w_1 + z_2w_2 = N$$

with $|z_i|, |w_j| \ll P$. By Prop. 5.6, we have

$$q(N) \le \begin{cases} 27P^3, & \text{if } N = 0, \\ 60P^2 \sum\limits_{d|N} \frac{1}{d}, & \text{if } N \ne 0. \end{cases}$$

To count all solution to (5.5), we need to sum up over all possible cases for N, where $|N| \ll P^2$ with the same argument as before for $|z_i|, |w_j| \ll P$. For every choice of N, we have $q(N)$ solutions to $z_1w_1 + z_2w_2 = N$ and $q(-N) = q(N)$ solutions to $z_3w_3 + z_4w_4 = -N$. This finally yields

$$\begin{aligned}
\int_0^1 \left| \sum_{0\le x\le P} e(f(x)\alpha) \right|^8 \mathrm{d}\alpha = Q \le R \\
&= \sum_{|N|\ll P^2} q(N)^2 \\
&= q(0)^2 + \sum_{0<|N|\ll P^2} q(N)^2 \\
&\ll P^6 + \sum_{0<N\ll P^2} P^4 \left(\sum_{d|N} \frac{1}{d} \right)^2 \\
&= P^6 + P^4 \sum_{0<N\ll P^2} \sum_{d_1,d_2|N} \frac{1}{d_1d_2} \\
&= P^6 + P^4 \sum_{1\le d_1,d_2\ll P^2} \frac{1}{d_1d_2} \sum_{\substack{0<N\ll P^2 \\ [d_1,d_2]|N}} 1 \\
&\ll P^6 + P^4 \sum_{1\le d_1,d_2\ll P^2} \frac{1}{d_1d_2} \cdot \frac{P^2}{[d_1,d_2]} \\
&\le P^6 + P^4 \sum_{1\le d_1,d_2\ll P^2} \frac{1}{d_1d_2} \cdot \frac{P^2}{(d_1d_2)^{1/2}} \\
&\le P^6 + P^6 \sum_{d_1,d_2=1}^{\infty} \frac{1}{(d_1d_2)^{3/2}} \ll P^6,
\end{aligned}$$

where we used the fact that $[d_1, d_2] \ge \sqrt{d_1d_2}$. This proves the induction basis.

Now let $k \ge 3$ and assume that the statement is true for $k-1$. Writing

$$\varphi(x,y) := \frac{1}{y}(f(x+y) - f(x))$$

for $y \neq 0$, we see

$$\begin{aligned}
\left| \sum_{x=0}^{P} e(f(x)\alpha) \right|^2 &= \left(\overline{\sum_{x=0}^{P} e(f(x)\alpha)} \right) \cdot \left(\sum_{y=0}^{P} e(f(y)\alpha) \right) \\
&= \left(\sum_{x=0}^{P} e(-f(x)\alpha) \right) \cdot \left(\sum_{y=0}^{P} e(f(y)\alpha) \right)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{x=0}^{P}\sum_{y=0}^{P} e(\alpha(f(y)-f(x))) \\
&= \sum_{x=0}^{P}\sum_{-x\le y\le P-x} e(\alpha(f(x+y)-f(x))) \\
&= \sum_{x=0}^{P} e(0) + \sideset{}{'}\sum_{0<|y|\le P}\ \sideset{}{'}\sum_{0\le x\le P} e(h\varphi(x,y)\alpha) \\
&= P+1+\sideset{}{'}\sum_{0<|y|\le P}\ \sideset{}{'}\sum_{0\le x\le P} e(y\varphi(x,y)\alpha),
\end{aligned}
$$

where $\sum'$ means that the summation is only over the relevant part of the set. Fixing $y\neq 0$ in $\varphi(x,y)$, we observe

$$
\begin{aligned}
\varphi(x,y) &= \frac{1}{y}(f(x+y)-f(x)) \\
&= \frac{1}{y}\left(\sum_{i=0}^{k} a_i(x+y)^i - \sum_{i=0}^{k} a_i x^i\right) \\
&= \frac{1}{y}\left(\sum_{i=1}^{k} a_i(x+y)^i - \sum_{i=1}^{k} a_i x^i\right) \\
&= \frac{1}{y}\left(\sum_{i=1}^{k} a_i \sum_{j=0}^{i}\binom{i}{j} x^{i-j}y^j - \sum_{i=1}^{k} a_i x^i\right) \\
&= \frac{1}{y}\left(\sum_{i=1}^{k} a_i x^i + \sum_{i=1}^{k} a_i \sum_{j=1}^{i-1}\binom{i}{j} x^{i-j}y^j + \sum_{i=1}^{k} a_i y^i - \sum_{i=1}^{k} a_i x^i\right) \\
&= \sum_{i=1}^{k} a_i \sum_{j=1}^{i-1}\binom{i}{j} x^{i-j}y^{j-1} + \sum_{i=1}^{k} a_i y^{i-1},
\end{aligned}
$$

and hence, $\varphi(x,y)$ is a polynomial in x with degree $k-1$ whose coefficients are integers that satisfy the conditions of the induction hypothesis. We now define

$$
a_y := \sideset{}{'}\sum_{0\le x\le P} e(y\varphi(x,y)\alpha),
$$

giving us by the above:

$$
\begin{aligned}
\left|\sum_{x=0}^{P} e(f(x)\alpha)\right|^{2\cdot 8^{k-2}} &= \left((P+1)+\sideset{}{'}\sum_{0<|y|\le P} a_y\right)^{8^{k-2}} \\
&= \sum_{i=0}^{8^{k-2}}\binom{8^{k-2}}{i}(P+1)^i\left(\sideset{}{'}\sum_{0<|y|\le P} a_y\right)^{8^{k-2}-i}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \max\left\{(P+1)^{8^{k-2}}, \left|\sum_{0<|y|\leq P}{}' a_y\right|^{8^{k-2}}\right\} \cdot \sum_{i=0}^{8^{k-2}} \binom{8^{k-2}}{i} \\
&= 2^{8^{k-2}} \cdot \max\left\{(P+1)^{8^{k-2}}, \left|\sum_{0<|y|\leq P}{}' a_y\right|^{8^{k-2}}\right\} \\
&\ll \max\left\{P^{8^{k-2}}, \left|\sum_{0<|y|\leq P}{}' a_y\right|^{8^{k-2}}\right\}.
\end{aligned}
$$

In order to prove the statement, we need to check the cases that the maximum takes on either value. So first assume

$$\left|\sum_{0<|y|\leq P}{}' a_y\right| \leq P.$$

Applying this to the exponential sum in (5.3) yields:

$$
\begin{aligned}
\int_0^1 \left|\sum_{x=0}^{P} e(f(x)\alpha)\right|^{8^{k-1}} \mathrm{d}\alpha &= \int_0^1 \left(\left|\sum_{x=0}^{P} e(f(x)\alpha)\right|^{2\cdot 8^{k-2}}\right)^4 \mathrm{d}\alpha \\
&\ll \int_0^1 \left(P^{8^{k-2}}\right)^4 \mathrm{d}\alpha \\
&= P^{4\cdot 8^{k-2}} \leq P^{8^{k-1}-k}.
\end{aligned}
$$

We now assume

$$\left|\sum_{0<|y|\leq P}{}' a_y\right| > P.$$

Applying the Cauchy-Schwarz inequality (Thm. 1.1) repeatedly, we obtain:

$$
\begin{aligned}
\left|\sum_{x=0}^{P} e(f(x)\alpha)\right|^{2\cdot 8^{k-2}} &\ll \left|\sum_{0<|y|\leq P}{}' a_y\right|^{2^{3(k-2)}} \\
&\leq \left(\sum_{0<|y|\leq P}{}' |1| \cdot |a_y|\right)^{2^{3(k-2)}} \\
&\leq \left(\left(\sum_{0<|y|\leq P}{}' 1\right) \cdot \left(\sum_{0<|y|\leq P}{}' |a_y|^2\right)\right)^{2^{3(k-2)-1}}
\end{aligned}
$$

$$
\begin{aligned}
&\le \left(\left(\sideset{}{'}\sum_{0<|y|\le P} 1 \right)^2 \cdot \left(\sideset{}{'}\sum_{0<|y|\le P} 1 \right) \cdot \left(\sideset{}{'}\sum_{0<|y|\le P} |a_y|^{2^2} \right) \right)^{2^{3(k-2)-2}} \\
&\le \left(\left(\sideset{}{'}\sum_{0<|y|\le P} 1 \right)^{2^3-1} \cdot \left(\sideset{}{'}\sum_{0<|y|\le P} |a_y|^{2^3} \right) \right)^{2^{3(k-2)-3}} \le \cdots \\
&\le \left(\sideset{}{'}\sum_{0<|y|\le P} 1 \right)^{2^{3(k-2)}-1} \cdot \left(\sideset{}{'}\sum_{0<|y|\le P} |a_y|^{2^{3(k-2)}} \right) \\
&\le (3P)^{8^{k-2}-1} \cdot \sideset{}{'}\sum_{0<|y|\le P} |a_y|^{8^{k-2}} \\
&\ll P^{8^{k-2}-1} \cdot \sideset{}{'}\sum_{0<|y|\le P} \left| \sideset{}{'}\sum_{0\le x\le P} e(y\varphi(x,y)\alpha) \right|^{8^{k-2}} \qquad (5.6)
\end{aligned}
$$

Writing the sum as a Fourier series, we have

$$
\left| \sideset{}{'}\sum_{0\le x\le P} e(y\varphi(x,y)\alpha) \right|^{8^{k-2}} = \sum_n A_n e(yn\alpha).
$$

From the restrictions for the coefficients, we obtain

$$
n \ll \max_{0\le x\le P} |\varphi(x,y)| \ll P^{k-1}.
$$

Calculating the coefficients of the Fourier series and applying the induction hypothesis, we have

$$
\begin{aligned}
|A_n| &= \left| \int_0^1 \left| \sideset{}{'}\sum_{0\le x\le P} e(y\varphi(x,y)\beta) \right|^{8^{k-2}} \cdot e(-n\beta)\,\mathrm{d}\beta \right| \\
&\le \int_0^1 \left| \sideset{}{'}\sum_{0\le x\le P} e(y\varphi(x,y)\beta) \right|^{8^{k-2}} \cdot |e(-n\beta)|\,\mathrm{d}\beta \\
&\ll \int_0^1 P^{8^{k-2}-(k-1)}\,\mathrm{d}\beta = P^{8^{k-2}-(k-1)}.
\end{aligned}
$$

Raising (5.6) to the 4^{th} power and integrating over α from 0 to 1, this finally yields:

$$
\int_0^1 \left| \sum_{x=0}^{P} e(f(x)\alpha) \right|^{8^{k-1}} \mathrm{d}\alpha
$$

$$
\begin{aligned}
&\ll \int_0^1 P^{4\cdot 8^{k-2}-4} \cdot \left(\sideset{}{'}\sum_{0<|y|\le P} \left| \sideset{}{'}\sum_{0\le x\le P} e(y\varphi(x,y)\alpha) \right|^{8^{k-2}} \right)^4 \mathrm{d}\alpha \\
&= P^{4\cdot 8^{k-2}-4} \cdot \int_0^1 \left(\sideset{}{'}\sum_{0<|y|\le P} \sum_{|n|\ll P^{k-1}} A_n e(yn\alpha) \right)^4 \mathrm{d}\alpha \\
&= P^{4\cdot 8^{k-2}-4} \cdot \sum_{\substack{0<|y_i|\le P \\ |n_j|\ll P^{k-1}}} A_{n_1}\cdots A_{n_4} \int_0^1 e(\alpha(y_1n_1+\ldots+y_4n_4))\,\mathrm{d}\alpha \\
&= P^{4\cdot 8^{k-2}-4} \cdot \sum_{\substack{0<|y_i|\le P, |n_j|\ll P^{k-1} \\ y_1n_1+\ldots+y_4n_4=0}} A_{n_1}\cdots A_{n_4} \\
&\ll P^{4\cdot 8^{k-2}-4} \cdot P^{4\cdot 8^{k-2}-4(k-1)} \cdot \sum_{\substack{0<|y_i|\le P, |n_j|\ll P^{k-1} \\ y_1n_1+\ldots+y_4n_4=0}} 1 \\
&\ll P^{4\cdot 8^{k-2}-4} \cdot P^{4\cdot 8^{k-2}-4(k-1)} \cdot P^{3k} = P^{8^{k-1}-k},
\end{aligned}
$$

where we used the estimate for the number of integral solutions to $y_1n_1+\ldots+y_4n_4=0$ with $|y_in_i| \ll P^k$ according to the argument in the induction basis. This proves the theorem, and hence solves Waring's problem. □

5.3 Concluding Remarks and Generalisations

Finally, it is worth noticing that the two main results of this paper can be combined to obtain the Waring-Goldbach problem: *Can every (sufficiently large) integer be represented as the sum of a bounded number of k^{th} powers of primes?* Some progress has been made for small exponents, yet the general question remains unanswered. Hua [Hua65] gives an account on this topic. But this is just one example of the many unsolved problems in additive number theory. The fact that methods from many areas of mathematics can be applied to these problems makes it an active and exciting field of research with many interesting results yet to be expected.

Bibliography

[Che51] Pafnuty Chebyshev, *Sur la fonction qui détermine la totalité des nombres premiers inférieurs à une limite donnée*, Mémoires présentés à l'Académie Impériale des sciences de St.-Pétersbourg par divers savants **6** (1851), 141–157.

[dlVP96] Charles de la Vallée Poussin, *Recherches analytiques de la théorie des nombres premiers*, Annales de la Société Scientifique de Bruxelles **20** (1896), 183–256.

[Dun65] R. L. Duncan, *The Schnirelmann density of the k-free integers*, Proc. Amer. Math. Soc. **16** (1965), 1090–1091. MR 0186652 (32 #4110)

[Erd49] Paul Erdős, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proceedings of the National Academy of Sciences of the United States of America **35** (1949), 374–384.

[Had96] Jacques Hadamard, *Sur la distribution des zéros de la fonction $\zeta(s)$ et ses conséquences arithmétiques*, Bulletin de la Société Mathématique de France **24** (1896), 199–220.

[Hil09] David Hilbert, *Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem)*, Mathematische Annalen **67** (1909), 281–300.

[Hua59] Loo Keng Hua, *Die Abschätzung von Exponentialsummenund ihre Anwendung in der Zahlentheorie*, Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen **13** (1959), 1–126.

[Hua65] ______, *Additive theory of prime numbers*, Translations of mathematical monographs, vol. 13, American Mathematical Society, Providence, 1965.

[Hua82] ______, *Introduction to number theory*, Springer, Berlin, 1982.

[Khi52] A.Y. Khinchin, *Three pearls of number theory*, Graylock, Rochester, 1952.

[Lan30] Edmund Landau, *Die Goldbachsche Vermutung und der Schnirelmannsche Satz*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse (1930), 255–276.

[Lin43] Juri Wladimirowitsch Linnik, *Elementarnoe rešenie problemy Waring'a po metodu šnirel'mana*, Matematičeskij Sbornik N. F. **12** (1943), 225–230.

[Man42] Henry B. Mann, *A proof of the fundamental theorem on the density of sums of sets of positive integers*, Annals of Mathematics **43** (1942), 523–527.

[Nat87] Melvyn B. Nathanson, *A generalization of the Goldbach-Shnirel'man theorem*, The American Mathematical Monthly **94** (1987), 768–771.

[Nat96] ______, *Additive number theory. the classical bases*, Springer, New York, 1996.

[Nat00] ______, *Elementary methods in number theory*, Springer, New York, 2000.

[Nes06] Yu. V. Nesterenko, *On Waring's problem (elementary methods)*, Journal of Mathematical Sciences (New York) **137** (2006), 4699–4715.

[New60] Donald J. Newman, *A simplified proof of Waring's conjecture*, Michigan Mathematical Journal **7** (1960), 291–295.

[New97] ______, *Analytic number theory*, Springer, New York, 1997.

[Ram95] Olivier Ramaré, *On šnirel'man's constant*, Annali della Scuola Normale Superiore di Pisa **22** (1995), 645–706.

[Rie54] Georg Johann Rieger, *Zu Linniks Lösung des Waringschen Problems: Abschätzung von $g(n)$*, Mathematische Zeitschrift **60** (1954), 213–234.

[Rog64] Kenneth Rogers, *The Schnirelmann density of the squarefree integers*, Proc. Amer. Math. Soc. **15** (1964), 515–516. MR 0163893 (29 #1192)

[Sch30] Lew Genrichowitsch Schnirelmann, *Ob additivnich swoistwach tschisel*, Iswestija Donskowo Politechitscheskowo Instituta (Nowotscherkask) **14** (1930), 3–27.

[Sch33] ______, *Über additive Eigenschaften von Zahlen*, Mathematische Annalen **107** (1933), 649–690.

[Sel49] Atle Selberg, *An elementary proof of the prime-number theorem*, Annals of Mathematics **50** (1949), 305–313.

[Sel89] ______, *On an elementary method in the theory of primes*, Collected Papers, vol. I, Springer, Berlin, 1989, pp. 363–366.