Sieve Methods

Denis Xavier Charles

Preface

Sieve methods have had a long and fruitful history. The sieve of Eratosthenes (around 3rd century B.C.) was a device to generate prime numbers. Later Legendre used it in his studies of the prime number counting function $\pi(x)$. Sieve methods bloomed and became a topic of intense investigation after the pioneering work of Viggo Brun (see [**Bru16**],[**Bru19**], [**Bru22**]). Using his formulation of the sieve Brun proved, that the sum

$$\sum_{p,\ p+2 \text{ both prime}} \frac{1}{p}$$

converges. This was the first result of its kind, regarding the *Twin-prime* problem. A slew of sieve methods were developed over the years — Selberg's upper bound sieve, Rosser's Sieve, the Large Sieve, the Asymptotic sieve, to name a few. Many beautiful results have been proved using these sieves. The Brun-Titchmarsh theorem and the extremely powerful result of Bombieri are two important examples. Chen's theorem [**Che73**], namely that there are infinitely many primes p such that $p + 2$ is a product of at most two primes, is another indication of the power of sieve methods.

Sieve methods are of importance even in applied fields of number theory such as Algorithmic Number Theory, and Cryptography. There are many direct applications, for example finding all the prime numbers below a certain bound, or constructing numbers free of large prime factors. There are indirect applications too, for example the running time of several factoring algorithms depends directly on the distribution of smooth numbers in short intervals. The so called *undeniable signature schemes* require prime numbers of the form $2p + 1$ such that p is also prime. Sieve methods can yield valuable clues about these distributions and hence allow us to bound the running times of these algorithms.

In this treatise we survey the major sieve methods and their important applications in number theory. We apply sieves to study the distribution of *square-free numbers*, *smooth numbers*, and prime numbers. The first chapter is a discussion of the basic sieve formulation of Legendre. We show that the distribution of square-free numbers can be deduced using a square-free sieve[1]. We give an account of improvements in the error term of this distribution, using known results regarding the Riemann Zeta function.

The second chapter deals with Brun's Combinatorial sieve as presented in the modern language of [**HR74**]. We apply the general sieve to give a simpler proof of a theorem of Rademacher [**Rad24**]. The bound obtained by this simpler proof is slightly inferior, but still sufficient for applications such as the result of Erdős, Chowla and Briggs on the number of mutually orthogonal Latin squares. The formulation of Brun's sieve in [**HR74**] also includes a proof of the important *Buchstab identity*. We use it to derive some bounds on the distribution of smooth numbers ([**Hal70**]).

The third chapter deals with the development and the applications of Selberg's upper bound method. The proof by van Lint and Richert [**vLR65**] of the Brun-Titchmarsh theorem is given as the chief application. Hooley's improvement of bounds on prime factors in a problem studied by Chebyschev is also outlined here. The last chapter is a study of the Large Sieve. We give an outline of a proof of Bombieri's central theorem

[1] This is not a new proof - it is implicit in the work of Erdős [**Erd60**]

on the error term in the distribution of primes. A new application of the Bombieri theorem is shown; we prove that there are infinitely many primes p such that $p + 2$ is a *square-free* number with at most 7 prime factors.

Acknowledgements: I would like to thank my advisor Dr. Ken Regan, for allowing me to work on a topic of my own interest. His support, encouragement and advice has been invaluable for my work. I thank him for proofreading the entire document and his constructive comments. A special word of thanks to Dr. Jin-Yi for helping me with character sums. I thank him for answering my queries in such a way that I gained a new insight into the problem. I thank Dr. Alan Selman for his encouragement and advice. I am deeply grateful to Professors Eric Bach, Tom Cusick, Kevin Ford, and Andrew Granville for promptly answering my queries. Their suggestions, pointers, and ideas were invaluable for this work. I am indebted to the National Science foundation for the monetary support for this work, under my advisor's grant CCR 98-20140.

I thank my parents for their love, encouragement and prayers. I thank Pavan, Maurice, and Samik for pretending to be interested in sieves, and for reviewing the proofs. A special word of thanks to all my friends for anchoring me in sanity through this summer.

Denis Charles.

To
Truth and Purity

Contents

Preface 3

Chapter 0. Notation and preliminaries 9
- 0.1. Standard Nomenclature 9
- 0.2. Conventions 9
- 0.3. Preliminaries 9

Chapter 1. The sieve of Eratosthenes 13
- 1.1. Introduction 13
- 1.2. Sieve of Eratosthenes-Legendre 13
- 1.3. Smooth numbers 15
- 1.4. Density of squarefree numbers 15
- 1.5. The error term in the distribution of Squarefree numbers 18
- 1.6. Pairs of squarefree numbers 22
- 1.7. The smallest squarefree number in an arithmetic progression 26
- 1.8. The Sieve Problem 28

Chapter 2. The Combinatorial Sieve 31
- 2.1. Brun's Pure Sieve 31
- 2.2. Brun's Sieve 36
- 2.3. Orthogonal Latin Squares and the Euler Conjecture 45
- 2.4. A Theorem of Schinzel 50
- 2.5. Smooth Numbers 55
- 2.6. On the number of integers prime to a given number 57

Chapter 3. Selberg's Sieve 59
- 3.1. The Selberg upper-bound method 59
- 3.2. The Brun-Titchmarsh Theorem 66
- 3.3. Prelude to a theorem of Hooley 71
- 3.4. A theorem of Hooley 74

Chapter 4. The Large Sieve 81
- 4.1. Bounds on exponential sums 81
- 4.2. The Large Sieve 87
- 4.3. The Brun-Titchmarsh Theorem revisited 91
- 4.4. Bombieri's Theorem 93
- 4.5. Prime and Squarefree pairs 96

Bibliography 101

CHAPTER 0

Notation and preliminaries

0.1. Standard Nomenclature

The largest integer not exceeding x is denoted $\lfloor x \rfloor$. We write $a \backslash b$ for two integers a, b $a \neq 0$ if a divides b. The Möbius function is denoted by $\mu(n)$ and defined as:

$$\mu(n) = \begin{cases} (-1)^k & \text{if } n = p_1 \cdots p_k, \text{ for } 1 \leq i < j \leq k : p_i \neq p_j, \\ 0 & \text{otherwise.} \end{cases}$$

The prime counting function is $\pi(x)$ defined as the cardinality of the set $P = \{p \leq x \mid p \text{ a prime}\}$, while $\pi(x; q, a)$ will denote the cardinality of $\{p \leq x \mid p \equiv a \mod q\}$. We denote the von-Mangoldt function by $\Lambda(n)$:

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^k \text{ for a prime } p, \\ 0 & \text{otherwise,} \end{cases}$$

and its cumulation by $\psi(x) = \sum_{n \leq x} \Lambda(n)$. If $n = p_1^{e_1} \cdots p_k^{e_k}$ is the prime factorization of n then $\nu(n) = k$ denotes the number of distinct primes in the factorization. We write $\varphi(n)$ for Euler's totient function:

$$\varphi(n) = n \prod_{p \backslash n} \left(1 - \frac{1}{p}\right).$$

0.2. Conventions

The letter p will always denote a prime number. Consequently, $\sum_{n \leq p \leq m} f(p)$ will denote a sum over the *prime* numbers in the range of summation. $\mathcal{A}$ will stand for a general integer sequence to be sifted, and P for the sifting set of primes. We employ the standard O and o-notation. We use the Vinogradov notation $\ll$ to mean that inequality holds with some constant, i.e., $f(n) \ll g(n) \Rightarrow \exists c > 0 : f(n) \leq c g(n)$. If $\gcd(a, b) = 1$ for two integers a and b, then we also write $a \perp b$.

0.3. Preliminaries

THEOREM 0.3.1. *Let $n \geq 1$ be an integer. Then*

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

Proof : Since divisors that are not squarefree drop out of the sum by the definition of μ, we may without loss of generality assume that n is squarefree.

Let $n = p_1 p_2 \cdots p_l$, then any divisor d of n has the form $p_1^{e_1} p_2^{e_2} \cdots p_l^{e_l}$ with $e_i \in \{0,1\}$ for $1 \le i \le l$. Using this we can split up the sum we wish to evaluate:

$$\begin{aligned}
\sum_{d \backslash n} \mu(d) &= \sum_{\substack{p_1^{e_1} p_2^{e_2} \cdots p_l^{e_l} \\ e_1 + \cdots + e_l = \text{ even}}} 1 - \sum_{\substack{p_1^{e_1} p_2^{e_2} \cdots p_l^{e_l} \\ e_1 + \cdots + e_l = \text{ odd}}} 1 \\
&= \binom{n}{0} - \binom{n}{1} + \binom{n}{2} + \cdots + (-1)^n \binom{n}{n} \\
&= (1-1)^n \\
&= 0.
\end{aligned}$$

There is another way we could have evaluated the sum. Let $T(l)$ be the number of 0-1 strings of length l that have odd number of 1s in them. Consider the last position of such a string. If it is a 1, then we must fill the rest of the positions with an even number of 1s which can be done in $2^{l-1} - T(l-1)$ ways. If the last position is a 0, then the rest of the string must have an odd number of 1s which can be done in $T(l-1)$ ways. We have argued that $T(l)$ satisfies the following recurrence:

$$\begin{aligned}
T(l) &= T(l-1) + (2^{l-1} - T(l-1)) \\
&= 2^{l-1}.
\end{aligned}$$

Thus the number of sequences with odd number of 1s and the number of them with even number of 1s is the same, and so the above sum is zero. □

THEOREM 0.3.2. *(Möbius Inversion) If*

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

then

$$g(n) = \sum_{d \backslash n} \mu(d) f\left(\frac{n}{d}\right).$$

Proof :

$$\begin{aligned}
\sum_{d \backslash n} \mu(d) f\left(\frac{n}{d}\right) &= \sum_{d \backslash n} \mu(d) \sum_{l \backslash (n/d)} g(l) \\
&= \sum_{l \backslash n} g(l) \left(\sum_{d \backslash (n/l)} \mu(d) \right) \\
&= \sum_{l = n} g(l) && \text{by Theorem 0.3.1} \\
&= g(n).
\end{aligned}$$

□

THEOREM 0.3.3. *If*

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

then

$$g(n) = \sum_{d \backslash n} \mu\left(\frac{n}{d}\right) f(d).$$

Proof :

$$\begin{aligned}
\sum_{d\backslash n} \mu\Big(\frac{n}{d}\Big) f(d) &= \sum_{d\backslash n} \mu\Big(\frac{n}{d}\Big) \sum_{l\backslash d} g(l) \\
&= \sum_{l\backslash n} g(l) \sum_{d\backslash n/l} \mu\Big(\frac{n}{dl}\Big) \\
&= \sum_{l=n} g(l) && \text{by Theorem 0.3.1} \\
&= g(n).
\end{aligned}$$

□

THEOREM 0.3.4.

$$\begin{aligned}
\sum_{d\backslash n} \frac{\mu(d)}{d} &= \prod_{p\backslash n} \left(1 - \frac{1}{p}\right) \\
&= \prod_{p\backslash n} \left(1 + \frac{\mu(p)}{p}\right).
\end{aligned}$$

Proof : We know that $\sum_{d\backslash n} \varphi(d) = n$. Using Möbius inversion on this we get:

$$\begin{aligned}
n \prod_{p\backslash n} \left(1 - \frac{1}{p}\right) = \varphi(n) &= \sum_{d\backslash n} \mu(d) \frac{n}{d} \\
&= n \sum_{d\backslash n} \frac{\mu(d)}{d}.
\end{aligned}$$

□

REMARK 0.3.5. The proof of Theorem 0.3.4 actually works for any multiplicative function of the divisors of n in the denominator, provided it is zero at non-squarefree divisors. We could have also proved Theorem 0.3.1 using the identity:

$$\sum_{d\backslash n} \mu(d) = \prod_{p\backslash n} \big(1 + \mu(p)\big).$$

CHAPTER 1

The sieve of Eratosthenes

1.1. Introduction

The sieve of Eratosthenes is a simple effective procedure for finding all the primes up to a certain bound x. Take a list of the numbers $2, 3, \cdots, \lfloor x \rfloor$. Call 2 a prime, and start by crossing out all the multiples of 2. Because 3 is uncrossed at this stage 3 must be prime. Cross out the multiples of 3 since they are composite, and then pick the next number that is still uncrossed and repeat. If after a stage the next uncrossed number exceeds $\sqrt{x}$ then stop. At this stage all the numbers that are not crossed out are prime.

Legendre realized that this procedure can be captured succinctly in a theoretical analog of the sifting process, and used this in his study of the function $\pi(x) = \left|\{p \leq x \mid p \text{ a prime}\}\right|$.

In this chapter we will try to apply this basic technique to study some simple problems. First we shall look at the sieve applied to the problem of estimating $\pi(x)$. Although the method would lead to an exact formula for $\pi(x) - \pi(\sqrt{x})$ this does not give useful estimates for $\pi(x)$ owing to a huge error term. However we can adapt the basic method to study other sequences of numbers, for example the *squarefree* numbers, meaning numbers that are products of distinct primes. The basic sieve we develop will be more successful in dealing with squarefree numbers, essentially because they are denser than the primes. We will be able to give interesting bounds on the density of these numbers in arithmetic progressions and in pairs $(n, n+2)$. We shall also find a bound on the smallest squarefree number in an arithmetic progression. Finally we shall give the general setup of a sieve problem and re-formulate the classical sieve of Eratosthenes-Legendre in this framework.

1.2. Sieve of Eratosthenes-Legendre

Let $P_z = \prod_{p<z} p$. The sieve of Eratosthenes deletes from the list of numbers all those numbers that are not relatively prime to P_z, except the primes dividing P_z itself. We are interested in finding bounds on the cardinality of the set $S = \{n \mid n \leq x, n \perp P_z\}$. We define

$$s(n) = \begin{cases} 1, & \text{if } n \in S \\ 0 & \text{otherwise.} \end{cases}$$

This is the characteristic function of the set S. Using the properties of the Möbius function (see Chapter 0), we can write an explicit expression for $s(n)$.

$$s(n) = \sum_{d \backslash \gcd(n, P_z)} \mu(d).$$

We will call such a function $s(n)$ the *sifting function.*

Then

$$\begin{aligned}
|S| &= \sum_{n\le x} s(n) \\
&= \sum_{n\le x} \sum_{d\backslash \gcd(n,P_z)} \mu(d) \\
&= \sum_{d\backslash P_z} \mu(d)\Bigg(\sum_{\substack{n\le x\\ d\backslash n}} 1\Bigg) \\
&= \sum_{d\backslash P_z} \mu(d)\left\lfloor \frac{x}{d}\right\rfloor \\
&= \sum_{d\backslash P_z} \mu(d)\left(\frac{x}{d} + \left\lfloor \frac{x}{d}\right\rfloor - \frac{x}{d}\right) \\
&= \sum_{d\backslash P_z} \mu(d)\frac{x}{d} + \sum_{d\backslash P_z} \mu(d)\left(\left\lfloor \frac{x}{d}\right\rfloor - \frac{x}{d}\right).
\end{aligned}$$

Since each term in the second sum has absolute value at most 1, we obtain

$$\begin{aligned}
|S| &\le x\sum_{d\backslash P_z} \frac{\mu(d)}{d} + 2^{\pi(z)} \\
&= x\prod_{p\backslash P_z}\left(1-\frac{1}{p}\right) + 2^{\pi(z)}.
\end{aligned}$$

Now a theorem of Mertens states that

$$\prod_{p<z}\left(1-\frac{1}{p}\right) \sim \frac{e^{\gamma}}{\ln z},$$

and this yields the estimate:

$$|S| \le x\frac{e^{\gamma}}{\ln z} + 2^{\pi(z)}$$

provided $z\to\infty$ as $x\to\infty$.

The usefulness of the above scheme is restricted by the huge error term $2^{\pi(z)}$. For $z = O(\ln x)$ for example we get

$$\pi(x) - \pi(\ln x) = O\left(\frac{x}{\ln\ln x}\right),$$

and since $\pi(x)\le x$ we get the estimate

$$\pi(x) = O\left(\frac{x}{\ln\ln x}\right).$$

This is markedly inferior to the truth $\pi(x)\sim\frac{x}{\ln x}$.

Note that if $z=\sqrt{x}$ then $|S|$ measures $\pi(x)-\pi(\sqrt{x})$, for which we have derived the following exact formula:

$$\pi(x)-\pi(\sqrt{x})+1 = x\prod_{p<\sqrt{x}}\left(1-\frac{1}{p}\right) + \sum_{d\backslash P_{\sqrt{x}}} \mu(d)\left(\left\lfloor\frac{x}{d}\right\rfloor - \frac{x}{d}\right).$$

1.3. Smooth numbers

DEFINITION 1.3.1. A number n will be called k-*smooth* if

$$\forall p : (p \backslash n) \Rightarrow (p < k).$$

Let $\Psi(x,k) = |\{n \leq x \mid n \text{ is } k\text{-smooth}\}|$ i.e., the number of k-smooth numbers up to a bound x.

We can use our sieve argument to try to find a bound on $\Psi(x,k)$. The weakness of this simple sieve will be apparent in the bound it gives us.

PROPOSITION 1.3.2.

$$\Psi(x,k) = O\left(x\frac{\ln k}{\ln x} + 2^{\pi(x)-\pi(k)}\right).$$

Proof : Since a number is k-smooth only if all its prime divisors are below k, we can find the k-smooth numbers below a bound x, by using as our sifting set $P = \{p \mid k < p \leq x\}$. Let $P_{k,x} = \prod_{p \in P} p$.
Let $S = \{n \mid n \text{ is } k\text{-smooth}\}$, and this time define

$$s(n) = \begin{cases} 1 & \text{if } n \in S \text{ or } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Now rewriting $s(n)$ using the Möbius function, we obtain

$$s(n) = \sum_{d \backslash \gcd(n, P_{k,x})} \mu(d).$$

Setting $S(n) = |S|$, we apply Mertens' Theorem at the end to conclude:

$$\begin{aligned}
S(n) &= \sum_{n \leq x} s(n) \\
&= \sum_{n \leq x} \sum_{d \backslash \gcd(n, P_{k,x})} \mu(d) \\
&= \sum_{d \backslash P_{k,x}} \mu(d) \left(\sum_{\substack{n \leq x \\ d \backslash n}} 1 \right) \\
&= \sum_{d \backslash P_{k,x}} \mu(d) \left\lfloor \frac{x}{d} \right\rfloor \\
&= x \prod_{k < p \leq x} \left(1 - \frac{1}{p}\right) + O(2^{\pi(x)-\pi(k)}) \\
&= O\left(x\frac{\ln k}{\ln x} + 2^{\pi(x)-\pi(k)}\right).
\end{aligned}$$

□

The bound is clearly very poor. However we can improve this bound using more advanced sieve techniques. In [**Warl90**], a much better bound is given under some conditions on the sifting primes.

1.4. Density of squarefree numbers

The basic method of the sieve of Eratosthenes-Legendre can be adapted to prove a more interesting result. Let $S = \{n \mid n \leq x, n \text{ is squarefree}\}$, and let $\kappa(x) = |S|$. To obtain S as a result of a sifting process, all we need to do is take primes $p < \sqrt{x}$ and cross of multiples of p^2 from the list. We shall show that a variant of the function $s(n)$ introduced earlier works in this case.

THEOREM 1.4.1.

$$\kappa(x) = \frac{6}{\pi^2}x + O(\sqrt{x}).$$

Proof : The sifting function for this set is now

$$s(n) = |\mu(n)|,$$

and $\kappa(x) = \sum_{n\le x} s(n) = \sum_{n\le x} |\mu(n)|$. Now we reach an impasse, because there does not seem to be any easy way of evaluating this sum. The trick is to look for another expression for the sifting function.

Claim: $s(n) = \sum_{d^2\backslash n} \mu(d)$.
Proof of Claim: Any number n can be represented as $n = m^2 w$, where w is squarefree and m is the largest square divisor of n. If $n = p_1^{e_1}p_2^{e_2}\cdots p_l^{e_l}$ with $e_i = 2q_i + r_i, 0 \le r_i \le 1$, then $m = \prod_i p_i^{q_i}$ satisfies the expression. We shall write $\Box(n)$ to stand for the largest square divisor of n. Now

$$\sum_{d^2\backslash n} \mu(d) = \sum_{d\backslash\Box(n)} \mu(d),$$

and this sum is 0 unless $\Box(n) = 1$ in which case it is also 1. This proves the claim.

Setting $m = \sqrt{x}$, we obtain:

$$\begin{aligned}
\kappa(n) &= \sum_{n\le x} s(n)\\
&= \sum_{n\le x}\sum_{d^2\backslash n} \mu(d)\\
&= \sum_{d\le m} \mu(d) \sum_{\substack{n\le x\\ d^2\backslash n}} 1\\
&= \sum_{d\le m} \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor\\
&= x\sum_{d\le m} \frac{\mu(d)}{d^2} + \sum_{d\le m} \mu(d)\left(\left\lfloor \frac{x}{d^2}\right\rfloor - \frac{x}{d^2}\right)\\
&= x\sum_{d\le m} \frac{\mu(d)}{d^2} + O(m).
\end{aligned}$$

Using the fact that

$$\prod_p \left(1 - \frac{1}{p^2}\right) = \sum_{n\ge 1} \frac{\mu(n)}{n^2}$$

we get

$$\begin{aligned}
\kappa(n) &= x\prod_p \left(1 - \frac{1}{p^2}\right) - \sum_{d>m} \frac{\mu(d)}{d^2} + O(m)\\
&= x\prod_p \left(1 - \frac{1}{p^2}\right) + O(m).
\end{aligned}$$

Also

$$\prod_p \left(1 - \frac{1}{p^2}\right) = \frac{1}{\zeta(2)},$$

so that we finally get

$$\kappa(n) = x\frac{1}{\zeta(2)} + O(\sqrt{x}).$$

Euler showed that $\zeta(2) = \frac{\pi^2}{6}$, and using this in the above expression we have

$$\kappa(n) = \frac{6}{\pi^2}x + O(\sqrt{x}).$$

□

Another natural question to ask is: what is the density of squarefree numbers in an arithmetic progression? We shall give a partial answer to that question in the next theorem. Let $\kappa(x; a, l) = |\{n \leq x \mid n \text{ is squarefree}, n \equiv a \mod l\}|$.

THEOREM 1.4.2. *Let* $q > 2$ *be a prime, and let* a *be a positive integer relatively prime to* q*. Then there is a constant* $c > 0$ *depending only on* q *such that*

$$\kappa(x; a, q) \geq cx + O(\sqrt{x}).$$

Proof : Using the same idea as in the previous proof we have:

$$\kappa(x; a, q) = \sum_{\substack{n \equiv_q a \\ n \leq x}} \sum_{d^2 \backslash n} \mu(d) \tag{1.1}$$

$$= \sum_{d \leq m} \left(\sum_{\substack{d^2 \backslash n \\ n \equiv_q a \\ n \leq x}} 1 \right) \text{ where } m \text{ is } \lfloor \sqrt{x} \rfloor. \tag{1.2}$$

The quantity we need to bound is defined by

$$N(x; d, a, q) = \sum_{\substack{d^2 \backslash n \\ n \equiv_q a \\ n \leq x}} 1$$

This is essentially the number of solutions in k to the congruence

$$kd^2 \equiv a \mod q.$$

There are two cases:
[$\mathbf{d} \perp \mathbf{q}$] In this case there is a unique solution k such that

$$k \equiv a(d^{-2}) \mod q.$$

However, if $k \in \{0, 1, \cdots, q-1\}$ is such a solution then for $e \geq 1$, $k + eq$ is also a solution. Now $(k + eq)d^2 = n \leq x$, so

$$\begin{aligned}(k + eq) &\leq \frac{x}{d^2} \\ e &\leq \frac{x}{d^2 q} - \frac{k}{q} \\ e &\leq \left\lfloor \frac{x}{d^2 q} \right\rfloor \text{ as } k < q.\end{aligned}$$

[$\mathbf{d} \not\perp \mathbf{q}$] In this case there are no solutions to the congruence as $a > 0$.

Thus $N(x; d, a, q) = \lfloor \frac{x}{d^2 q} \rfloor$ if $d \perp q$, and 0 otherwise. Substituting in (1.2) we get

$$\begin{aligned}
\kappa(x; a, q) &= \sum_{d \le m} \mu(d) \left\lfloor \frac{x}{d^2 q} \right\rfloor - \sum_{\substack{d \le m \\ d \not\perp q}} \mu(d) \left\lfloor \frac{x}{d^2 q} \right\rfloor \\
&= \frac{x}{q} \left(\sum_{d \le m} \frac{\mu(d)}{d^2} - \sum_{d \not\perp q} \frac{\mu(d)}{d^2} \right) + O(m)
\end{aligned}$$

$$\begin{aligned}
\sum_{d \not\perp q} \frac{\mu(d)}{d^2} &\le \sum_{d \not\perp q} \frac{1}{d^2} \\
&= \sum_{q \backslash d, d \le x} \frac{1}{d^2} \\
&= \sum_{k \le (x/q)} \frac{1}{k^2 q^2} \\
&= \frac{1}{q^2} \sum_{k \le (x/q)} \frac{1}{k^2} \\
&\le \frac{1}{q^2} \sum_{k \ge 1} \frac{1}{k^2} \\
&\le \frac{\pi^2}{6q^2}
\end{aligned}$$

Thus we get

$$\kappa(x; a, q) \ge x \left(\frac{1}{q\zeta(2)} - \frac{\zeta(2)}{q^2} \right) + O(\sqrt{x}).$$

and hence $\kappa(x; a, q) \ge cx + O(\sqrt{x})$. $\square$

1.5. The error term in the distribution of Squarefree numbers

We proved in the previous section that $\kappa(x) - \frac{6}{\pi^2}x = O(\sqrt{x})$, and it turns out to be extremely difficult to improve on this bound. In this section we briefly digress form the topic of sieves to show a strengthening of the error term if one assumes the *Riemann Hypothesis* (henceforth called RH). First we shall strengthen the error term (unconditionally) using a theorem of *Walfisz*.

THEOREM 1.5.1 ([**Wal63**] Satz §5.5.3).

$$\sum_{n \le x} \mu(n) = Bx \exp\{-A \log^{\frac{3}{5}} x \log\log^{-\frac{1}{5}} x\}$$

for some positive constants A *and* B.

We simplify the proof in [**Wal63**] of the following theorem:

THEOREM 1.5.2 ([**Wal63**] Satz §5.6.1).

$$\kappa(x) = \frac{6}{\pi^2}x + O\left(\sqrt{x} \exp\{-c \log^{\frac{3}{5}} x \log\log^{-\frac{1}{5}} x\}\right)$$

for some positive constant $c > 0$.

Proof :

$$\begin{aligned}\kappa(x) &= \sum_{1\leq n\leq x}\sum_{d^2\backslash n}\mu(d)\\ &= \sum_{d^2 m\leq x}\mu(d)\\ &= \sum_{d^2\leq x}\mu(d)\left\lfloor\frac{x}{d^2}\right\rfloor.\end{aligned}$$

Let $S_2(x,y)=\sum_{d\leq y}\mu(d)\delta\left(\frac{x}{d^2}\right)$, where $\delta(z)=z-\lfloor z\rfloor-\frac{1}{2}$ and $M(y)=\sum_{n\leq y}\mu(n)$. Then

$$\kappa(x)=x\sum_{d^2\leq x}\frac{\mu^2(d)}{d^2}-S_2(x,\sqrt{x})-\frac{1}{2}M(\sqrt{x}).$$

In [**MV81**] (see p.255) the following bound is proved:

$$S_2(x,y)=O(x^{\frac{2}{7}}+y^{\frac{1}{2}}x^{\frac{1}{7}+\epsilon}),$$

and this implies that $S(x,\sqrt{x})=O(x^{\frac{11}{28}})$.
Now consider:

$$\begin{aligned}\sum_{d>y}\frac{\mu(d)}{d^2} &= 2\sum_{d>y}\mu(d)\int_d^\infty\frac{1}{z^3}dz\\ &= 2\int_y^\infty\frac{dz}{z^3}\left\{\sum_{y<n<z}\mu(n)\right\}\end{aligned}$$

(interchanging of the sum and the integral is valid since both of them are convergent)

$$\begin{aligned} &= 2\int_y^\infty\frac{M(z)dz}{z^3}-2M(y)\int_y^\infty\frac{dz}{z^3}\\ &= O\left\{M(y)\int_y^\infty\frac{dz}{z^3}\right\}-o(1)\\ &= O\left\{\frac{M(y)}{y^2}\right\}.\end{aligned}$$

Hence

$$\sum_{d>\sqrt{x}}\frac{\mu(d)}{d^2}=O\left\{\frac{\exp\{c\log^{\frac{3}{5}}x\log\log^{-\frac{1}{5}}x\}}{\sqrt{x}}\right\}$$

and also

$$\sum_{d\leq\sqrt{x}}\frac{\mu(d)}{d^2}=\frac{1}{\zeta(2)}+O\left\{\frac{\exp\{c\log^{\frac{3}{5}}x\log\log^{-\frac{1}{5}}x\}}{\sqrt{x}}\right\}.$$

The theorem follows from these estimates. □

Corollary 1.5.3. *The number of squarefree numbers in the interval* $[x,\cdots,x+\sqrt{x}]$ *is asymptotic to* $\frac{6\sqrt{x}}{\pi^2}$.
The corresponding problem for primes seems to be far more difficult, see [**HB88**].

It turns out that if the *Riemann Hypothesis* holds then $M(y)=O(\sqrt{y})$, and using this in the above proof we get the following theorem:

THEOREM 1.5.4. *Assuming the Riemann Hypothesis,*

$$\kappa(x) = \frac{6}{\pi^2}x + O(x^{\frac{11}{28}}). \quad \square$$

It turns out that if we assume the Riemann Hypothesis we can do better even without the strong bound on $S_2(x, y)$. We begin as we did before,

$$\begin{aligned}
\kappa(x) &= \sum_{1 \le d \le x} \mu(d) \Big\{ \sum_{\substack{1 \le n \le x \\ d^2 \backslash n}} 1 \Big\} \\
&= \sum_{d^2 n \le x} \mu(d) \\
&= \sum_{\substack{d^2 n \le x \\ d \le y}} \mu(d) + \sum_{\substack{d^2 n \le x \\ d > y}} \mu(d) \\
&= \Sigma_1 + \Sigma_2 \text{ (say)}.
\end{aligned}$$

Now (as in the proof of the previous theorem)

$$\begin{aligned}
\Sigma_1 &= \sum_{d \le y} \mu(d) \left\lfloor \frac{x}{d^2} \right\rfloor \\
&= \sum_{d \le y} \mu(d) \left(\frac{x}{d^2} - \left(\frac{x}{d^2} - \left\lfloor \frac{x}{d^2} \right\rfloor - \frac{1}{2} \right) \right) - \frac{1}{2} \sum_{d \le y} \mu(d).
\end{aligned}$$

Let as before

$$S_2(x, y) = \sum_{d \le y} \mu(d) \delta\left(\frac{x}{d^2} \right)$$

and $M(y) = \sum_{d \le y} \mu(d)$, where $\delta(z) = z - \lfloor z \rfloor - \frac{1}{2}$, so that

$$\Sigma_1 = x \sum_{d \le y} \frac{\mu(d)}{d^2} - S_2(x, y) - \frac{1}{2} M(y).$$

Let

$$f_y(s) = \frac{1}{\zeta(s)} - \sum_{d \le y} \frac{\mu(d)}{d^s}.$$

We adopt the standard convention of referring to the real part of s as σ and the imaginary part as t. If $\sigma > 1$ then we have

$$f_y(s) = \sum_{d > y} \frac{\mu(d)}{d^s},$$

since in this case we also have

$$\frac{1}{\zeta(s)} = \sum_{1 \le d} \frac{\mu(d)}{d^s}.$$

Consider

$$\begin{aligned}
\zeta(s) f_y(2s) &= \Big\{ \sum_{1 \le n} \frac{1}{n^s} \Big\} \Big\{ \sum_{d > y} \frac{\mu(d)}{d^{2s}} \Big\} \\
&= \sum_{1 \le n} \frac{1}{n^s} \Big(\sum_{\substack{d > y \\ d^2 \backslash n}} \mu(d) \Big).
\end{aligned}$$

If we look at the restricted version of this sum, namely,

$$\sum_{1\le n\le x} \frac{1}{n^s}\Bigg(\sum_{\substack{d>y\\ d^2\backslash n}} \mu(d)\Bigg),$$

then as $s \to 0$ this sum equals Σ_2. Thus we need a way of evaluating this sum when $s \to 0$. The following result (Lemma (3.12) [**Tit86**] p60) will help us do just that.

LEMMA 1.5.5. [**Tit86**] *Let $\langle a_n \rangle$ be a sequence of real numbers, such that as $\sigma \to 1$ from above,*

$$\sum_{n\ge 1} \frac{|a_n|}{n^\sigma} = O\bigg(\frac{1}{(\sigma-1)^\alpha}\bigg),$$

for some $\alpha \ge 1$. Let $\psi(n)$ be an upper bound for $|a_n|$, and define:

$$f(s) = \sum_{n\ge 1} \frac{a_n}{n^s},$$

for $\sigma > 1$. If $c > 0, \sigma \ge 0, \sigma + c > 1$, x is not an integer, and N is the nearest integer to x, then for all $T > 0$:

$$\sum_{n<x} \frac{a_n}{n^s} = \frac{1}{2\pi i}\int_{c-iT}^{c+iT} f(s+w)\frac{x^w}{w}dw + O\bigg(\frac{x^c}{T(\sigma+c-1)^\alpha}\bigg) + O\bigg(\frac{\psi(2x)x^{1-\sigma}\log x}{T}\bigg) + O\bigg(\frac{\psi(N)x^{1-\sigma}}{T|x-N|}\bigg). \quad \square$$

Applying this lemma to the series

$$\sum_{1\le n\le x} \frac{1}{n^s}\Bigg(\sum_{\substack{d>y\\ d^2\backslash n}} \mu(d)\Bigg)$$

with $c = 1 + \frac{1}{\log x}$ and $T = x$ gives remainder terms of $O(x^\epsilon)$, since $\psi(z) = O(\sqrt{z})$. Making the change of variable $w \leftarrow s$ taking the s in the lemma to be 0, and setting $x_0 = \lfloor x \rfloor + \frac{1}{2}$ so that x_0 is not an integer, we obtain

$$\Sigma_2 = \frac{1}{2\pi i}\int_{c-ix}^{c+ix} \zeta(s) f_y(2s)\frac{x_0^s}{s}ds + O(x^\epsilon).$$

Now consider splitting the integral into four regions:

$$\int_{c-ix}^{c+ix} + \int_{c+ix}^{\frac{1}{2}+ix} + \int_{\frac{1}{2}+ix}^{\frac{1}{2}-ix} + \int_{\frac{1}{2}-ix}^{c-ix}$$

(where the integrand is the same as above). Since the integrand has a simple pole at $s = 1$, with residue $2\pi i f_y(2)x_0$, we have

$$\int_{c-ix}^{c+ix} + \int_{c+ix}^{\frac{1}{2}+ix} + \int_{\frac{1}{2}+ix}^{\frac{1}{2}-ix} + \int_{\frac{1}{2}-ix}^{c-ix} = 2\pi i f_y(2)x_0$$

and so

$$\Sigma_2 = f_y(2)x_0 + \frac{1}{2\pi i}\int_{\mathcal{C}} \zeta(s) f_y(2s)\frac{x_0^s}{s}ds + O(x^\epsilon),$$

where $\mathcal{C}$ is the path made up of the line segments

$$\begin{aligned} c - ix &\longrightarrow \frac{1}{2} - ix \\ \frac{1}{2} - ix &\longrightarrow \frac{1}{2} + ix \\ \frac{1}{2} + ix &\longrightarrow c + ix. \end{aligned}$$

By Theorem (14.2) on p.337 of [**Tit86**], RH implies that $\frac{1}{\zeta(s)} = O(|t|^{\epsilon})$. Also

THEOREM 1.5.6 ([**Tit86**] (14.25A)). *Assume RH. For* s *with* $\sigma > \frac{1}{2}$,

$$\sum_{n<x} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)} + O(T^{1-\epsilon}x^2) + O(T^{\epsilon}x^{\frac{1}{2}-\sigma+\delta}). \quad \square$$

Using this we can take T large so that

$$f_y(s) = O(y^{\frac{1}{2}-\sigma+\delta'}) \tag{1.3}$$

under RH.

Also by Theorem (14.25C) [**Tit86**], RH implies $M(z) = O(z^{\frac{1}{2}+\epsilon})$. Using all this information we can bound

$$\int_{\mathcal{C}} \zeta(s) f_y(2s) \frac{x_0^s}{s} ds$$

on the contour $\mathcal{C}$: we have $f_y(2s) = O(y^{\frac{1}{2}-1+\epsilon}) = O(y^{-\frac{1}{2}+\epsilon})$ and $\zeta(s) = \frac{1}{s-1} + O(t^{\epsilon})$, and since $x^s = x^{\sigma+it} = e^{(\sigma+it)\log x} = e^{\sigma \log x + it \log x}$, we have $|x^s| = x^{\sigma}$. Thus the integral in (1.3) is:

$$O(x^{\frac{1}{2}+\epsilon}y^{-\frac{1}{2}+\epsilon}).$$

Combining all these estimates we get the following bounds:

THEOREM 1.5.7 ([**MV81**]). *Assuming the Riemann Hypothesis, for any* $y > 0$

$$\kappa(x) = \frac{x}{\zeta(2)} - S_2(x,y) + O\big(x^{\frac{1}{2}+\epsilon}y^{-\frac{1}{2}+\epsilon} + y^{\frac{1}{2}+\epsilon}\big). \quad \square$$

COROLLARY 1.5.8. *Assuming the Riemann Hypothesis,*

$$\kappa(x) = \frac{x}{\zeta(2)} + O\big(x^{\frac{1}{3}+\delta}\big).$$

Proof : Clearly we have $S_2(x,y) = O(y)$, now setting $y = x^{\frac{1}{3}}$ in the above theorem we get the result. $\square$

In the same article [**MV81**] Montgomery and Vaughan went on to estimate the sums involved more precisely to show that $\kappa(x) = \frac{1}{\zeta(2)}x + O(x^{\frac{9}{28}+\epsilon})$. Subsequently the exponent of the error term was reduced to $\frac{7}{22}$ by various authors (see [**BakPin85**]).

1.6. Pairs of squarefree numbers

The famous twin prime problem asks whether there are infinitely many primes p such that $p+2$ is also prime. Although this problem is still open, the analogous question for the squarefree numbers can be settled rather easily using the methods we have seen so far. For a more general version of this result see [**Mir49**].

Let $\kappa_2(x) = \big|\{n(n+2) \mid \mu(n)^2 = \mu(n+2)^2 = 1, n \le x\}\big|$.

THEOREM 1.6.1.

$$\kappa_2(x) = \prod_p \left(1 - \frac{2}{p^2}\right)x + O(x^{\frac{2}{3}} \ln^{\frac{4}{3}} x).$$

Proof : Let $s(n) = \sum_{d^2 \backslash n} \mu(d)$. Using this we have

$$\begin{aligned}\kappa_2(x) &= \sum_{n\le x} s(n)s(n+2) \\ &= \sum_{n\le x} \Big(\sum_{a^2\backslash n} \mu(a)\Big)\Big(\sum_{b^2\backslash n} \mu(b)\Big).\end{aligned}$$

If $a^2\backslash n$ and $b^2\backslash(n+2)$, then writing $n = k_1a^2$ and $n+2 = k_2b^2$ we have $k_1'a^2 + k_2b^2 = 2$ ($k_1' = -k_1$). This says that $\gcd(a^2, b^2)$ divides 2, so $\gcd(a, b)$ must be 1, i.e. $a \perp b$. Now interchanging the sum we get

$$\kappa_2(x) = \sum_{\substack{k_1a^2 - k_2b^2 = 2\\ k_2b^2 \le x\\ a\perp b}} \mu(a)\mu(b).$$

The rest of the proof is now to bound the above sum, and to this end we split up the sum into two parts:

$$\kappa_2(x) = \sum_{ab\le y} \mu(a)\mu(b)N(x; a^2, b^2, 2) + \sum_{\substack{ab>y\\ k_1a^2-k_2b^2=2, k_2b^2\le x}} \mu(a)\mu(b).$$

Here $N(x; a^2, b^2, 2)$ is a count of the number of solutions to the equation

$$k_1a^2 - k_2b^2 = 2,\ k_2b^2 \le x.$$

It is clear that $N(x; a^2, b^2, 2) = 0$ if $\gcd(a^2, b^2)$ does not divide 2, and otherwise

$$\begin{aligned} N(x; a^2, b^2, 2) &= \frac{x}{\operatorname{lcm}(\ a^2, b^2)} + O(1) \\ &= \frac{x}{(ab)^2} + O(1), \end{aligned}$$

since $a \perp b$.

Using this we have

$$\begin{aligned} \sum_{\substack{ab\le y\\ a\perp b}} \mu(a)\mu(b)N(x; a^2, b^2, 2) &\le \sum_{\substack{ab\le y\\ a\perp b}} \mu(a)\mu(b)\left(\frac{x}{(ab)^2} + O(1)\right) \\ &= x\sum_{ab\le y} \frac{\mu(ab)}{(ab)^2} + \sum_{ab\le y} \mu(a)\mu(b), \end{aligned}$$

since the terms with $a \not\perp b$ are killed by the Möbius function.

Thus

$$\begin{aligned} \sum_{ab\le y} \mu(a)\mu(b) &\le \sum_{ab\le y} 1 \\ &= \left\lfloor \frac{y}{1} \right\rfloor + \left\lfloor \frac{y}{2} \right\rfloor + \cdots + \left\lfloor \frac{y}{y} \right\rfloor \end{aligned}$$
$$= O\left(y \sum_{1\le k\le y} \frac{1}{k}\right) = O(y \ln y).$$

Now the sum

$$\sum_{ab\le y} \frac{\mu(ab)}{(ab)^2}$$

can be evaluated by looking at the terms with $\nu(ab) = k$. Write $a = p_1^{\epsilon_1}p_2^{\epsilon_2}\cdots p_k^{\epsilon_k}$ and $b = p_1^{\delta_1}p_2^{\delta_2}\cdots p_k^{\delta_k}$. Since $a \perp b$ we should have $(\forall i : 1 \le i \le k)\ \epsilon_i + \delta_i = 1$, so there are $2^{\nu(ab)}$ terms whose denominator is $(ab)^2$. Hence

$$\sum_{ab\le y}\frac{\mu(ab)}{(ab)^2}=\sum_{n\le y}\frac{\mu(n)2^{\nu(n)}}{n^2}$$
$$=\prod_{p\le y}\left(1-\frac{2}{p^2}\right).$$

So

$$\sum_{ab\le y}\frac{\mu(ab)}{(ab)^2}=\prod_{p}\left(1-\frac{2}{p^2}\right)-\sum_{n>y}\frac{\mu(n)2^{\nu(n)}}{n^2}.$$

We need a bound on the sum on the right hand side of the above equation.
Now

$$\sum_{ab>y}\frac{1}{(ab)^2}=\sum_{b<y,ab>y}\frac{1}{(ab)^2}+\sum_{a>y,b>y}\frac{1}{(ab)^2}.$$

The second sum converges so we need to bound on the first part of the sum. Now:

$$\sum_{b<y,ab>y}\frac{1}{(ab)^2}\le\sum_{1\le b\le y}\frac{1}{b^2}\left(\sum_{a>\frac{y}{b}}\frac{1}{a^2}\right)$$

$$\sum_{a>\frac{y}{b}}\frac{1}{a^2}\le\int_{\frac{y}{b}}^{\infty}\frac{1}{a^2}da=\frac{b}{y}$$

so we have

$$\sum_{b<y,ab>y}\frac{1}{(ab)^2}\le\frac{1}{y}\sum_{1\le b\le y}\frac{1}{b}$$
$$=\frac{1}{y}\ln y.$$

We finally get

$$x\sum_{ab\le y}\frac{\mu(ab)}{(ab)^2}=x\prod_{p}\left(1-\frac{1}{p^2}\right)+O\left(\frac{x}{y}\ln y\right).$$

Now we have to bound the sum

$$\sum_{\substack{ab>y\\k_1a^2-k_2b^2=2,k_2b^2\le x}}\mu(a)\mu(b).$$

We re-express this sum as follows:

$$\sum_{\substack{ab>y\\a^2c-b^2d=2\\b^2d\le x}}\mu(a)\mu(b)\le\sum_{\substack{a^2c-b^2d=2\\b^2d\le x\\ab>y}}1.$$

Since $a^2c=2+b^2d$, $a^2c\le 2+x$, and this gives us $c\le\frac{(x+2)}{a^2}$. Since $d\le\frac{x}{b^2}$ and $y<ab$ we have either $cd\le\frac{x(x+2)}{a^2b^2}$ or $cd\le\frac{x(x+2)}{y^2}$. This gives

$$\sum_{\substack{a^2c-b^2d=2\\b^2d\le x,ab>y}}1\le\sum_{cd<\frac{x(x+2)}{y^2}}M(x;c,d,2),$$

where $M(x;c,d,2)$ is the number of solutions of

$$ca^2 - db^2 = 2, db^2 \leq x. \tag{1.4}$$

The above equation implies that

$$\begin{aligned}\left(\frac{2c^{-1}}{p}\right) &\equiv 1 \mod p, \text{ for all } p \backslash d,\\ \left(\frac{2d^{-1}}{p}\right) &\equiv 1 \mod p, \text{ for all } p \backslash c.\end{aligned}$$

Estermann studied these congruences and for the case cd not a square he proved [**Est31**]:

$$M(x;c,d,2) = O(\ln x),$$

in fact that $M(x;c,d,2) \leq 4(\ln(x+2)+1)$.
If cd is a square then since the equation (1.4) implies $c \perp d$ we can set $c = l^2$, $d = m^2$ to obtain:

$$\begin{aligned}M(x;c,d,2) &= \sum_{l^2a^2-m^2b^2=2} 1\\ &\leq \sum_{r^2-s^2=2} 1\\ &= 0.\end{aligned}$$

In any case we have $M(x;c,d,2) = O(\ln x)$, and using this we have:

$$\sum_{cd<\frac{x(x+2)}{y^2}} M(x;c,d,2) \leq \ln x \sum_{cd<\frac{x(x+2)}{y^2}} 1.$$

For any positive constant K we have:

$$\begin{aligned}\sum_{cd<K} 1 &= \sum_{c<K} \frac{K}{c}\\ &\leq K \ln K,\end{aligned}$$

so

$$\begin{aligned}\sum_{cd<\frac{x(x+2)}{y^2}} M(x;c,d,2) &\leq \ln^2 x \frac{x(x+2)}{y^2}\\ &= O\left(\frac{x^2}{y^2}\ln^2 x\right).\end{aligned}$$

Setting $y = x^{\frac{2}{3}} \ln^{\frac{1}{3}} x$ we have

$$\begin{aligned}\sum_{ab>y} \mu(ab) &\leq \sum_{cd<\frac{x(x+2)}{2}} M(x;c,d,2)\\ &\leq x^{\frac{2}{3}} \ln^{\frac{4}{3}} x,\end{aligned}$$

and also

$$x\sum_{ab\leq y} \frac{\mu(ab)}{(ab)^2} = x\prod_p \left(1 - \frac{1}{p^2}\right) + O\left(x^{\frac{1}{3}}\left(\ln^{\frac{2}{3}} x + o(1)\right)\right).$$

The theorem follows from these two bounds. □

1.7. The smallest squarefree number in an arithmetic progression

The simple methods that we have seen so far are surprisingly powerful and provide a quick bound on the smallest squarefree number in an arithmetic progression. The following result is from [**Erd60**] and is one of the early uses of a squarefree sieve.

THEOREM 1.7.1. *Let* $a \perp D$, $1 \leq a < D$. *Then the smallest squarefree number in the arithmetic progression* $\langle a + kD \ : \ k \geq 0 \rangle$ *is*

$$O\left(\frac{D^{\frac{3}{2}}}{\ln D}\right).$$

Proof : Let $\mathcal{A} = \langle a + kD \ : \ k \geq 0 \rangle$ be the sequence. The first step would be to sift $\mathcal{A}$ by all squares of primes below a certain limit z. This will leave out only those numbers that could have a large prime as their square divisor. We will finally bound the number of such integers below x and show that there are still some numbers left over — and that will prove the theorem.
Let $P_z = \prod_{p<z} p$. The result of the sifting of the sequence $\mathcal{A}$ by P_z is:

$$\begin{aligned} S(\mathcal{A}; P_z, x) &= \sum_{\substack{n \in \mathcal{A} \\ n \leq x}} \sum_{\substack{d^2 \backslash n \\ d \backslash P_z}} \mu(d) \\ &= \sum_{d \backslash P_z} \mu(d) \left(\sum_{\substack{n \in \mathcal{A} \\ n \leq x, d^2 \backslash n}} 1 \right). \end{aligned}$$

Now

$$\sum_{\substack{n \in \mathcal{A} \\ n \leq x, d^2 \backslash n}} 1$$

is exactly the number of solutions to the following pair of congruences:

$$\begin{aligned} n &\equiv 0 \mod d^2 \\ n &\equiv a \mod D. \end{aligned}$$

Suppose $d \perp D$. Then there is exactly one solution in the interval $\operatorname{lcm}(D, d^2) = Dd^2$, so the total number of solutions in $1 \leq n \leq x$ is at most

$$\frac{x}{Dd^2} + 1.$$

If $\gcd(d, D) = \delta$ then $n = k\delta$ by the first congruence and $n - a = k'\delta$ by the second congruence. This yields $a = (k - k')\delta$ and so $\gcd(a, D) \neq 1$. This is a contradiction, so if $d \not\perp D$ there are no solutions to the congruence. Let $\overline{k} = \lfloor \frac{(x-a)}{D} \rfloor$, which is the maximum value of k for $a + kD$ to be in $\mathcal{A}$. Then

$$\begin{aligned}
S(\mathcal{A}; P_z, x) &= \sum_{d \backslash P_z, d \perp D} \mu(d)\left(\frac{x}{Dd^2} + 1\right) \\
&= \frac{x}{D} \sum_{d \backslash P_z, d \perp D} \frac{\mu(d)}{d^2} \\
&= \overline{k}\Bigg(\sum_{\substack{d \backslash P_z \\ d \perp D}} \frac{\mu(d)}{d^2} + o(1)\Bigg) \\
&= \overline{k}\Bigg(\prod_{p \backslash P_z, p \nmid d}\left(1 - \frac{1}{p^2}\right) + o(1)\Bigg) \\
&\geq \overline{k}\Bigg(\prod_{p}\left(1 - \frac{1}{p^2}\right) + o(1)\Bigg) \\
&= \overline{k}\left(\frac{6}{\pi^2} + o(1)\right).
\end{aligned}$$

Taking $\overline{k}$ to be $\frac{c\sqrt{D}}{\ln D}$ we have

$$S(\mathcal{A}; P_z, \overline{k}) \geq \frac{6}{\pi^2}\frac{c\sqrt{D}}{\ln D}.$$

The number of integers $a + kD$ in $\mathcal{A}$ for which $k < \frac{c\sqrt{D}}{\ln D}$ and also

$$\begin{aligned}
n &\equiv 0 \mod p^2 \\
n &\equiv a \mod D
\end{aligned}$$

is at most $\frac{c\sqrt{D}}{p^2 \ln D} + 1$.

Let N stand for the number of integers $k < \frac{c\sqrt{D}}{\ln D}$ in $\mathcal{A}$ for which $a + kD \not\equiv 0 \mod p^2$ for all $p \leq \sqrt{cD}$. Then

$$N \geq \frac{6}{\pi^2}\frac{c\sqrt{D}}{\ln D} - \frac{c\sqrt{D}}{\ln D}\left(\sum_{p \geq z}\frac{1}{p^2}\right) - \sum_{p \geq z, p \leq \sqrt{cD}} 1 \tag{1.5}$$

$$\geq \frac{6}{\pi^2}\frac{c\sqrt{D}}{\ln D} - \frac{c\sqrt{D}}{\ln D}\frac{1}{z} - \pi(\sqrt{cD}), \tag{1.6}$$

and so for large enough c and L

$$N > \frac{1}{2}\frac{c\sqrt{D}}{\ln D}. \tag{1.7}$$

We have used the fact that $\pi(x) < \frac{2x}{\ln x}$ for large enough x.

Now we are left with the numbers that are either squarefree or divisible by a prime $p > \sqrt{cD}$. For these numbers $a + kD$ either

$$a + kD \equiv 0 \mod p^2, k < \frac{c\sqrt{D}}{\ln D} \text{ and } p > \sqrt{cD}$$

or

$$a + kD = \alpha p^2 \text{ with } \alpha < \frac{\sqrt{D}}{\ln D}.$$

Supposing $p > D^{\frac{1}{2}+\epsilon}$, we would have $\alpha < D^{\frac{1}{2}-\epsilon}$ if D is large enough, so we also have $p < D$.

Thus $a + kD = \alpha p^2$ yields a congruence $a \equiv \alpha p^2 \mod D$. Let us fix an α; then clearly the number of such prime solutions is less than the number of solutions for the congruence $x^2 \equiv a\alpha^{-1} \mod D$, $0 < x < D$. If $a\alpha^{-1}$ is a quadratic residue modulo D, then by the Chinese Remainder Theorem there are at most $2^{\nu(D)}$

such solutions to this congruence. Since $\nu(n) = o(\ln n)$, we can write $2^{\nu(D)} = o(D^{\frac{\epsilon}{2}})$. If $p > D^{\frac{1}{2}+\epsilon}$ then there are only $D^{\frac{1}{2}-\epsilon}$ choices for α, so on the whole there are only $o(D^{\frac{1}{2}-\frac{\epsilon}{2}})$ such solutions.

Let us consider the solutions for $\sqrt{cD} < p < D^{\frac{1}{2}+\epsilon}$. We have

$$p^2 \equiv a\alpha^{-1} \mod D,\ \alpha < \frac{\sqrt{D}}{\ln D},\ \sqrt{cD} < p < D^{\frac{1}{2}+\epsilon}.$$

Let c_α be the number of solutions of this congruence for a fixed α. These solutions give rise to $\sum \binom{c_\alpha}{2}$ solutions to the congruence

$$p^2 \equiv q^2,\ p, q < D^{\frac{1}{2}+\epsilon}. \tag{1.8}$$

Since (1.8) implies $(p-q)(p+q) \equiv 0 \mod D$, the number of such solutions is at most the number of solutions to $uv \equiv 0 \mod D$, $u < 2D^{\frac{1}{2}+\epsilon}, v < 2D^{\frac{1}{2}+\epsilon}$. This gives us

$$uv = \beta D, 1 \leq \beta < 4D^{2\epsilon}. \tag{1.9}$$

Also for a fixed β the number of such solutions is less than the number of factors of the number βD, which is $o((\beta D)^\epsilon)$, so the number of solutions of (1.9) is $o((\beta D)^\epsilon)4D^{2\epsilon} = o(D^{4\epsilon})$. This gives

$$\sum \binom{c_\alpha}{2} = o(D^{4\epsilon})$$

and hence

$$\sum_{c_\alpha > 1} c_\alpha = o(D^{4\epsilon}).$$

Since $\alpha < \frac{\sqrt{D}}{\ln D}$, $\sum c_\alpha \leq \frac{\sqrt{D}}{\ln D} + o(D^{4\epsilon})$. Thus the number of integers $0 \leq k < \frac{c\sqrt{D}}{\ln D}$ for which

$$a + kD \equiv 0 \mod p^2$$

for some $p > \sqrt{cD}$ is at most $\frac{D^{\frac{1}{2}}}{\ln D} + o(D^{\frac{1}{2}-\frac{\epsilon}{2}})$. So the number of integers k, $0 \leq k < \frac{c\sqrt{D}}{\ln D}$, for which $a + kD$ is squarefree is

$$\frac{1}{2}\frac{c\sqrt{D}}{\ln D} - \frac{\sqrt{D}}{\ln D} - o(D^{\frac{1-\epsilon}{2}}) > 0$$

for large enough c. $\square$

1.8. The Sieve Problem

Now that we have seen some examples of sieve techniques at work, we can formulate the sieve problem in a generic setting so that the essential quantities are clearly visible. The notation we shall adopt is that of the seminal book by Halberstam and Richert [**HR74**].

1.8.1. Notation.

(1) $\mathcal{A}, \mathcal{B}, \cdots$ will stand for integer sequences.
(2) $\mathcal{A}_d = \langle a \in \mathcal{A} \ :\ a \equiv 0 \mod d \rangle$.
(3) $\mathcal{A}^z = \langle a \in \mathcal{A} \ :\ a \leq z \rangle$.
(4) If $\mathcal{A}$ is a finite sequence then $|\mathcal{A}|$ will denote the length of the sequence.
(5) $\mathcal{P} = \langle p_i \ :\ p_i$ is the i-th prime$\rangle$.
(6) $P_z = \prod_{p \in \mathcal{P}^z} p$.

(7) $S(\mathcal{A};\mathcal{P}^z,x)$ will be the number of elements in $\mathcal{A}^x$ that survive the sifting process by the sequence $\mathcal{P}^z$. In general the sifting is determined by a sifting function $\sigma:\mathcal{A}\to\{0,1\}$ which determines whether a number survives the sifting, but usually we will only be considering simple sifting functions like

$$\sigma(n)=1\Leftrightarrow\left(n\perp\prod_{p\in\mathcal{S}^z}p\right)$$

(8) If $\mathcal{A}$ is a finite sequence then $\omega(p)$ is defined such that $\frac{\omega(p)}{p}x$ is a good approximation to $|\mathcal{A}_p^x|$. If d is any squarefree integer we can generalize this notation by defining $\omega(d)=\prod_{p\backslash d}\omega(d)$.
(9) Define $R_d(x)=|\mathcal{A}_d^x|-\frac{\omega(p)}{p}x$, i.e. the remainder term in our estimate of $|\mathcal{A}_d^x|$.
(10) Define

$$W(z)=\prod_{p\backslash P_z}\left(1-\frac{\omega(p)}{p}\right).$$

1.8.2. The Sieve of Eratosthenes-Legendre revisited. The generic sieve problem is to estimate $S(\mathcal{A};\mathcal{P}^z,x)$. Needless to say solving the problem as stated in this generality is too great a task. This treatise will only be concerned with restricted versions of the sieve problem which nevertheless yield interesting and non-trivial results. The case of great importance is when $\mathcal{S}^z=\mathcal{P}^z$ and $\mathcal{A}$ is some subsequence of positive integers.

The sieve of Eratosthenes-Legendre can be recast in this framework as follows.
Let $\mathcal{A}$ be the sequence to be sifted, and let $\omega(d)$ and R_d be the modulo counting function and the remainder function for the sequence, respectively. Let $\mathcal{P}^z$ be the sifting sequence; then the sifting function is

$$\sigma(n)=\begin{cases}1 & \text{if, } n\perp P_z\\ 0 & \text{otherwise.}\end{cases}$$

We can rewrite $\sigma(n)$ as

$$\sigma(n)=\sum_{d\backslash\gcd(n,P_z)}\mu(d).$$

Thus we have

$$
\begin{aligned}
S(\mathcal{A},\mathcal{P}^z,x) &= \sum_{n\in\mathcal{A},n\le x} \sigma(n) \\
&= \sum_{n\in\mathcal{A}^x} \sum_{d\backslash \gcd(n,P_z)} \mu(d) \\
&= \sum_{d\backslash P_z} \mu(d)\Bigg(\sum_{\substack{n\in\mathcal{A}^x\\ d\backslash n}} 1\Bigg) \\
&= \sum_{d\backslash P_z} \mu(d)|\mathcal{A}_d^x| \\
&= \sum_{d\backslash P_z} \mu(d)\left(\frac{\omega(d)}{d}x + R_d(x)\right) \\
&= x\sum_{d\backslash P_z} \frac{\mu(d)\omega(d)}{d} + \sum_{d\backslash P_z} \mu(d)R_d(x) \\
&= x\prod_{p\backslash P_z}\left(1-\frac{\omega(p)}{p}\right) + \sum_{d\backslash P_z} \mu(d)R_d(x) \\
&= xW(z) + \sum_{d\backslash P_z} \mu(d)R_d(x) \\
&= xW(z) + \theta\sum_{d\backslash P_z} R_d(x) \text{ where } |\theta|\le 1.
\end{aligned}
$$

If we assume that $|R_d(x)| \le \omega(d)$ and suppose that $\omega(p) \le C_0$, where C_0 is some constant, then $\omega(d) \le C_0^{\nu(d)}$. So

$$
\begin{aligned}
\sum_{d\backslash P_z} R_d(x) &\le \sum_{d\backslash P_z} C_0^{\nu(d)} \\
&= \prod_{p\backslash P_z} (1+C_0) \\
&= (1+C_0)^{\pi(z)}.
\end{aligned}
$$

Thus we have proved the following theorem.

THEOREM 1.8.1. *For all sufficiently large x and $z < x$, there is a θ with $|\theta| \le 1$ (θ depending on z), such that*

$$S(\mathcal{A};\mathcal{P}^z,x) = xW(z) + \theta\sum_{d\backslash P_z} R_d(x).$$

If we have $|R_d(x)| \le \omega(d)$ and $\omega(p) \le C_0$ then

$$S(\mathcal{A};\mathcal{P}^z,x) = xW(z) + O\Big((1+C_0)^{\pi(z)}\Big).$$

It is very clear that the effectiveness of the basic sieve is limited by the fact that the remainder term is a sum over all the divisors of P_z. Beginning with the next chapter we shall systematically try to reduce this term.

CHAPTER 2

The Combinatorial Sieve

In this chapter we begin by exploring the ideas of Viggo Brun, who first showed how we can improve on the Legendre method if we relax our requirement of asymptotic results but instead look for inequalities. After developing Brun's sieve in general we shall look at applications that bring out the surprising power of the technique. We follow the presentation in *Halberstam & Richert* [**HR74**] rather closely since its form is well suited for our applications. However our development will be targeted only to the Brun's sieve.

2.1. Brun's Pure Sieve

Let $\mathcal{A}^x$ be a finite sequence of integers and let $\mathcal{S}^z$ be the sifting primes. In the previous chapter the sifting function was:

$$\sigma(n) = \sum_{d \backslash \gcd(n, P_z)} \mu(d).$$

Let us see what can be done if instead we have a pair of functions $\chi_1(d)$ and $\chi_2(d)$ such that

$$\sigma_2(n) \equiv \sum_{d \backslash n} \mu(d)\chi_2(d) \leq \sum_{d \backslash n} \mu(d) \leq \sum_{d \backslash n} \mu(d)\chi_1(d) \equiv \sigma_1(n).$$

Since

$$\begin{aligned} S(\mathcal{A}; P^z, x) &= \sum_{d \backslash P_z} \mu(d)|\mathcal{A}_d| \\ &= |\mathcal{A}| - \sum_{p \backslash P_z} |\mathcal{A}_p| + \sum_{pq \backslash P_z} |\mathcal{A}_{pq}| + \cdots \end{aligned}$$

we expect that truncating the series after an even (odd) number of sums will give us a lower (upper) bound. Brun's pure sieve is an application of this well-known idea.

Using the notation developed in the last chapter we have

$$\sum_{n \in \mathcal{A}} \sum_{\substack{d \backslash n \\ d \backslash P_z}} \mu(d)\chi_2(d) \leq S(\mathcal{A}, \mathcal{P}^z, x) \leq \sum_{n \in \mathcal{A}} \sum_{\substack{d \backslash n \\ d \backslash P_z}} \mu(d)\chi_1(d).$$

Let us first look at the upper bound:

$$\begin{aligned} \sum_{n \in \mathcal{A}} \sum_{\substack{d \backslash n \\ d \backslash P_z}} \mu(d)\chi_1(d) &= \sum_{d \backslash P_z} \mu(d)\chi_1(d)|A_d^x| \\ &= \sum_{d \backslash P_z} \mu(d)\chi_1(d)\left(\frac{\omega(d)x}{d} + |R_d(x)|\right) \\ &= x\sum_{d \backslash P_z} \mu(d)\chi_1(d)\frac{\omega(d)}{d} + \sum_{d \backslash P_z} \mu(d)\chi_1(d)|R_d(x)|. \end{aligned}$$

Let $\sigma_1(n) = \sum_{d\backslash n} \mu(d)\chi_1(d)$; then by Möbius inversion we get

$$\mu(d)\chi_1(d) = \sum_{\delta\backslash d} \mu\Big(\frac{d}{\delta}\Big)\sigma(\delta).$$

Substituting this in the above expression we get

$$\begin{aligned} x\sum_{d\backslash P_z} \mu(d)\chi_1(d)\frac{\omega(d)}{d} &= x\sum_{d\backslash P_z} \frac{\omega(d)}{d}\Big(\sum_{\delta\backslash d} \mu\Big(\frac{d}{\delta}\Big)\sigma_1(\delta)\Big) \\ &= x\sum_{\delta\backslash P_z} \frac{\sigma_1(\delta)\omega(\delta)}{\delta}\Big(\sum_{t\backslash (P_z/\delta)} \mu(t)\frac{\omega(t)}{t}\Big) \end{aligned}$$

(since $\omega(d)$ is a multiplicative function)

$$\begin{aligned} &= x\sum_{\delta\backslash P_z} \sigma_1(\delta)\frac{\omega(\delta)}{\delta}\prod_{p\backslash (P_z/\delta)}\Big(1-\frac{\omega(p)}{p}\Big) \\ &= xW(z)\sum_{\delta\backslash P_z} \sigma_1(\delta)\frac{\omega(\delta)}{\delta\prod_{p\backslash\delta}\Big(1-\frac{\omega(p)}{p}\Big)} \\ &= xW(z)\sum_{\delta\backslash P_z} \sigma_1(\delta)g(\delta) \\ &= xW(z)\Big(1+\sum_{1<\delta\backslash P_z} \sigma_1(\delta)g(\delta)\Big), \end{aligned}$$

where $g(d)$ abbreviates $\frac{\omega(d)}{d\prod_{p\backslash d}\left(1-\frac{\omega(p)}{p}\right)}$.
The remainder term is clearly at most

$$\sum_{d\backslash P_z} \mu(d)\chi_1(d)|R_d(x)| \le \sum_{d\backslash P_z} |\chi_1(d)||R_d(x)|.$$

A similar argument works for the lower bound too. Thus we have:

(2.10)
$$xW(z)\Big(1+\sum_{1<\delta\backslash P_z}\sigma_2(\delta)g(\delta)\Big) - \sum_{d\backslash P_z}|\chi_2(d)||R_d(x)| \le S(\mathcal{A},\mathcal{P}^z,x)$$

(2.11)
$$\le xW(z)\Big(1+\sum_{1<\delta\backslash P_z}\sigma_1(\delta)g(\delta)\Big) + \sum_{d\backslash P_z}|\chi_1(d)||R_d(x)|.$$

Our aim will be to minimize $|\sum_{1<\delta\backslash P_z}\sigma_i(\delta)g(\delta)|$ for $i = 1, 2$ such that the remainder term $\sum_{d\backslash P_z}|\chi_i(d)||R_d|$ is small.

A whole class of estimates can be obtained by restricting the functions χ_i to be the characteristic sequences of two divisor sets D^+ and D^- of P_z. The resulting sieves are called *Combinatorial Sieves*.
Let us consider the following functions:

$$\chi^{(r)}(d) = \begin{cases} 1 & \text{if } \nu(d) \le r, \text{ and } \mu^2(d) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

These functions restrict the divisor sets over which we take the sum. In particular the restriction is on the number of distinct prime factors of the divisors.
We will require the following lemma.

Lemma 2.1.1.

$$\sum_{0\le i\le k}(-1)^i\binom{n}{i}=(-1)^k\binom{n-1}{k}.$$

Proof : The proof is by induction on k.
For $k=0$ we have

$$(-1)^0\binom{n}{0}=\binom{n-1}{0}+\binom{n-1}{-1}=\binom{n-1}{0}.$$

Now

$$\begin{aligned}\sum_{0\le i\le (k+1)}(-1)^i\binom{n}{i}&=\sum_{0\le i\le k}(-1)^i\binom{n}{i}+(-1)^{k+1}\binom{n}{k+1}\\&=(-1)^k\binom{n-1}{k}+(-1)^{k+1}\binom{n}{k+1}\\&=(-1)^k\binom{n-1}{k}+(-1)^{k+1}\left(\binom{n-1}{k}+\binom{n-1}{k+1}\right)\\&=(-1)^{k+1}\binom{n-1}{k+1}.\end{aligned}$$

□

Lemma 2.1.2. *Let n be a positive integer and s a non-negative integer. Then*

$$\sum_{d\backslash n}\mu(d)\chi^{(2s+1)}(d)\le\sum_{d\backslash n}\mu(d)\le\sum_{d\backslash n}\mu(d)\chi^{(2s)}(d).$$

Proof : When $n=1$ all the sums are equal so we can assume $n>1$. Then

$$\begin{aligned}\sum_{d\backslash n}\mu(d)\chi^{(r)}(d)&=\sum_{1\le k\le r}(-1)^k\binom{\nu(n)}{k}\\&=(-1)^r\binom{\nu(n)-1}{r}.\end{aligned}$$

by Lemma (2.1.1). □

Now let us try to bound the terms involved in (2.11). Let $\sigma^{(r)}(n)=\sum_{d\backslash n}\mu(d)\chi^{(r)}(d)$, so that we have

$$\begin{aligned}\sigma^{(r)}(n)&=\sum_{\substack{d\backslash n\\ \nu(d)\le r}}\mu(d)\\&=(-1)^r\binom{\nu(n)-1}{r}\end{aligned}$$

and hence $|\sigma^{(r)}(n)|=\binom{\nu(n)-1}{r}\le\binom{\nu(n)}{r}$.

Then we have

$$\begin{aligned}\Big|\sum_{1<d\backslash P_z}\sigma^{(r)}(d)g(d)\Big| &\le \sum_{1<d\backslash P_z}\binom{\nu(d)}{r}g(d)\\ &= \sum_{r\le m\le \nu(P_z)=\pi(z)}\binom{m}{r}\sum_{\substack{1<d\backslash P_z\\ \nu(d)=m}}g(d)\\ &\le \sum_{m\le r}\binom{m}{r}\frac{1}{m!}\Big(\sum_{p<z}g(p)\Big)^m\\ &= \frac{1}{r!}\Big(\sum_{p<z}g(p)\Big)^r\exp\sum_{p<z}g(p).\end{aligned}$$

Suppose we make the assumption $|R_d(x)|\le\omega(d)$; then we can also bound the remainder term as follows:

$$\sum_{d\backslash P_z}|\chi^{(r)}(d)||R_d(x)|\le\sum_{d\backslash P_z,\nu(d)\le r}\omega(d)\le\Big(1+\sum_{p<z}\omega(p)\Big)^r.$$

Since

$$\sum_{\substack{d\backslash P_z\\ \nu(d)\le 2s+1}}\mu(d)|\mathcal{A}_d| = \sum_{\substack{d\backslash P_z\\ \nu(d)\le 2s}}\mu(d)|\mathcal{A}_d| - \sum_{\substack{d\backslash P_z\\ \nu(d)=2s+1}}|\mathcal{A}_d|\le S(\mathcal{A};P^z,x)\le\sum_{\substack{d\backslash P_z\\ \nu(d)\le 2s}}\mu(d)|\mathcal{A}_d|$$

we can always write

$$S(\mathcal{A};P^z,x)=\sum_{\substack{d\backslash P_z\\ \nu(d)\le r}}\mu(d)|\mathcal{A}_d|+\theta\sum_{\substack{d\backslash P_z\\ \nu(d)=r+1}}|\mathcal{A}_d|, |\theta|\le 1.$$

Putting all these together we have:

$$S(\mathcal{A};\mathcal{P}^z,x)=xW(z)\left(1+\theta\frac{1}{r!}\Big(\sum_{p<z}g(p)\Big)^r\exp\Big\{\sum_{p<z}g(p)\Big\}\right)+\theta'\Big(1+\sum_{p<z}\omega(p)\Big)^r$$

for some positive integer r, and with $|\theta|\le 1, |\theta'|\le 1$.
Thus we have proved:

Theorem 2.1.3 (Brun's Pure Sieve). *Let*

$$g(d)=\frac{\omega(d)}{d\prod_{p\backslash d}\big(1-\frac{\omega(p)}{p}\big)}$$

be well defined for all d *with* $\mu(d)\neq 0$, *and suppose* $|R_d(x)|\le\omega(d)$. *Then for every non-negative integer* r *there exist* θ,θ' *with* $|\theta|\le 1, |\theta'|\le 1$ *such that*

$$S(\mathcal{A};\mathcal{P}^z,x)=xW(z)\left(1+\theta\frac{1}{r!}\Big(\sum_{p<z}g(p)\Big)^r\exp\Big\{\sum_{p<z}g(p)\Big\}\right)+\theta'\Big(1+\sum_{p<z}\omega(p)\Big)^r. \quad\square$$

We can apply this theorem to derive a much better bound on $\pi(x)$ that we obtained eariler.
We consider the sequence $\mathbb{N}^x$, and in this case since $\mathbb{N}^x_p=\{n\le x\mid n\equiv 0 \mod p\}$ we have $|\mathbb{N}^x_p|=\lfloor\frac{x}{p}\rfloor=\frac{1}{p}x+\delta', |\delta'|<1$. So we can take $\omega(p)=1$, and the condition $|R_d(x)|\le\omega(d)$ is also satisified.
Also

$$g(p)=\frac{1}{p\big(1-\frac{1}{p}\big)}\le\frac{2}{p},$$

and this gives us

$$\sum_{p<z} g(p) \leq 2 \sum_{p<z} \frac{1}{p} \\ < 2(\ln \ln z + 1).$$

We use the trivial estimate $1 + \sum_{p<z} \omega(p) \leq z$. In this case we have

$$W(z) = \prod_{p<z} \left(1 - \frac{1}{p}\right) \\ \sim \frac{e^{-\gamma}}{\ln z}.$$

We begin with the following observations. First $\frac{1}{r!} \leq \left(\frac{e}{r}\right)^r$ by the Stirling approximation. Next if we set z such that $\sum_{p<z} g(p) \leq \lambda r$, then the result of the theorem simplifies to

$$S(\mathcal{A}; \mathcal{P}^z, x) = xW(z)\left(1 + \theta(e^{1+\lambda}\lambda)^r\right) + \theta' z^r.$$

Defining

$$r = \left\lfloor \frac{2(\ln \ln z + 1)}{\lambda} \right\rfloor + 1$$

gives us $\sum_{p<z} g(p) \leq \lambda r$.
We restrict z so that

$$\ln z = \frac{\ln x}{\gamma \ln \ln x},$$

and set

$$\lambda = \frac{\xi \ln z(\ln \ln z + 1)}{\ln x}$$

so that for a large enough x and appropriate settings of ξ, γ we get $\lambda e^{1+\lambda} \leq 1$. For this setting of z and r we have $z^r = o\big(x^{1-\epsilon}\big)$ for some $\epsilon > 0$. Thus the theorem gives

$$\pi(x) = O\left(\frac{x \ln \ln x}{\ln x}\left\{1 + \theta e^{-c \ln \ln x}\right\}\right) + o(x^{1-\epsilon}).$$

This approximation is significantly better than our first and shows the improvement that can be made using this simple idea. Next we will look at the twin primes problem, which was Brun's primary application of his pure sieve.
In this case we take the sequence to be

$$\mathcal{A} = |\{n(n+2) \mid n \leq x\}|.$$

Let $p > 2$; then $\mathcal{A}_p = \{n(n+2) \mid n \leq x, n(n+2) \equiv 0 \mod p\}$. Now $n(n+2) \equiv 0 \mod p$ only if $n \equiv 0$ or $n + 2 \equiv 0$ since p is an odd prime. Clearly 0 and $p - 2$ are two solutions in the interval $0, \cdots, p-1$. So we can take $\omega(p) = 2$, $p > 2$. For $p = 2$ we have $\omega(p) = 1$. By the Chinese Remainder Theorem we have $|R_d(x)| \leq \omega(d)$. We take the sifting primes to be $P = \{p \mid p > 2\}$. Since $S(\mathcal{A}; P^z, x)$ counts all the twin-prime pairs above z, $S(\mathcal{A}; P^z, x) + 2z$ is an upper bound on the number of twin-primes below x. Then:

$$
\begin{aligned}
W(z) &= \prod_{2<p<z} \left(1 - \frac{2}{p}\right) \\
&\leq \prod_{2<p<z} \left(1 - \frac{2}{p}\right)^2 \\
&= O\left(\frac{1}{\ln^2 z}\right).
\end{aligned}
$$

Carrying out the rest of the analysis again using $\ln z = \frac{\ln x}{\gamma \ln \ln x}$ we get the following theorem.

THEOREM 2.1.4. *Let* $\pi_2(x) = |\{p \leq x \mid p + 2 = q\}|$

$$\pi_2(x) = O\left(x \left(\frac{\ln \ln x}{\ln x}\right)^2\right).$$

The above theorem can be put in a more impressive form.

THEOREM 2.1.5.

$$\sum_{\substack{p \\ p+2=q}} \frac{1}{p} \text{ converges.}$$

Proof :

$$
\begin{aligned}
\sum_{\substack{p \\ p+2=q}} \frac{1}{p} &= \sum_n \frac{\pi_2(n) - \pi_2(n-1)}{n} \\
&= \sum_n \pi_2(n) \left(\frac{1}{n} - \frac{1}{n+1}\right) \\
&= \sum_n \pi_2(n) \frac{1}{n(n+1)} \\
&\leq B \sum_n \frac{n(\ln \ln n)^2}{n(n+1) \ln^2 n} \\
&= B \sum_n \frac{1}{n} \left(\frac{\ln \ln n}{\ln n}\right)^2 \\
&= O(1).
\end{aligned}
$$

The last step follows via

$$\sum_{n \leq x} \frac{1}{n} \left(\frac{\ln \ln n}{\ln n}\right)^2 \leq \left(\frac{2}{\ln x} + \frac{2 \ln \ln x}{\ln x} + \frac{(\ln \ln x)^2}{\ln x}\right)(1 + o(1))$$

using approximation by integration and taking the limit $x \to \infty$. □

2.2. Brun's Sieve

The second idea of Brun was to limit the remainder term by restricting the size of primes making up the divisors. This simple idea results in a sieve of remarkable power which can be used to prove rather sharp bounds on $S(\mathcal{A}, P^z, x)$. Since we are modifying the divisor sets in a non-trivial fashion we would like to have some simple conditions on the characteristic functions χ of the divisor sets, such that χ still yields good lower or upper bounds. Our first task is to find such a set of conditions.

We begin with the following observation.

PROPOSITION 2.2.1.

$$S(\mathcal{A}, P^z, x) = \sum_{d \backslash P_z} \mu(d)\chi(d)|\mathcal{A}_d| - \sum_{1<d \backslash P_z} \sigma(d) S(\mathcal{A}_d; P^z_{(d)}, z)$$

where $P^z_{(d)} = \prod_{\substack{p \in P^z \\ p \nmid d}} p$.

Proof :

$$\begin{aligned}
\sum_{d \backslash P_z} \mu(d)\chi(d)|\mathcal{A}^x_d| &= \sum_{d \backslash P_z} |\mathcal{A}^x_d| \sum_{\delta \backslash d} \mu\left(\frac{d}{\delta}\right)\sigma(\delta) \\
&= \sum_{\delta \backslash P_z} \sigma(\delta) \sum_{t \backslash P_z/\delta} \mu(t)|\mathcal{A}_{\delta t}| \\
&= \sum_{t \backslash P_z} \mu(t)|\mathcal{A}_t| + \sum_{1<\delta \backslash P_z} \sigma(\delta) \sum_{t \backslash P_z/\delta} \mu(t)|\mathcal{A}_{\delta t}| \\
&= S(\mathcal{A}, P^z, x) + \sum_{1<\delta \backslash P_z} \sigma(\delta) \sum_{t \backslash P_z/\delta} \mu(t)|\mathcal{A}_{\delta t}| \\
&= S(\mathcal{A}, P^z, x) + \sum_{1<d \backslash P_z} \sigma(d) S(\mathcal{A}_d; P^z_{(d)}, z),
\end{aligned}$$

where we have used the Möbius inversion on the expression for $\sigma(d)$ as in the previous section. □

We will use the above proposition to compare $\sum_{d \backslash P_z} \mu(d)|\mathcal{A}_d|$ with $\sum_{d \backslash P_z} \mu(d)\chi(d)|\mathcal{A}_d|$.

Now

$$\begin{aligned}
\sigma(d) &= \sum_{l \backslash d} \mu(l)\chi(l) \\
&= \sum_{l \backslash d/p} \mu(l)\chi(l) + \sum_{l \backslash d/p} \mu(lp)\chi(lp) \\
&= \sum_{l \backslash d/p} \mu(l)\chi(l) - \sum_{l \backslash d/p} \mu(l)\chi(lp) \\
&= \sum_{l \backslash d/p} \mu(l)\big(\chi(l) - \chi(lp)\big).
\end{aligned}$$

Let $q(d)$ be the smallest prime divisor of d. Now using the above expression we can write

$$\begin{aligned}
\sum_{1<d \backslash P_z} \sigma(d) S(\mathcal{A}_d; P^z_{(d)}, x) &= \sum_{\delta \backslash P_z} \sum_{\substack{p \backslash P_z \\ p<q(\delta)}} \sigma(p\delta) S(\mathcal{A}_{p\delta}; P^z_{(p\delta)}, x) \\
&= \sum_{\delta \backslash P_z} \sum_{\substack{p \backslash P_z \\ p<q(\delta)}} S(\mathcal{A}_{p\delta}; P^z_{(p\delta)}, x) \sum_{l \backslash \delta} \mu(l) \Big(\chi(l) - \chi(pl)\Big) \\
&= \sum_{l \backslash P_z} \sum_{\substack{p \backslash P_z \\ p<q(l)}} \mu(l) \Big(\chi(l) - \chi(pl)\Big) \sum_{\substack{t \backslash P_z/l \\ p<q(t)}} S(\mathcal{A}_{plt}; P^z_{(plt),x}) \\
&= \sum_{l \backslash P_z} \sum_{\substack{p \backslash P_z \\ p<q(l)}} \mu(l) \Big(\chi(l) - \chi(pl)\Big) S(\mathcal{A}_{pl}; P^p_{(pl)}, x).
\end{aligned}$$

Using this in the above proposition,

$$
\begin{aligned}
S(\mathcal{A};P^z,x) &= \sum_{d\backslash P_z}\mu(d)\chi(d)|\mathcal{A}_d| - \sum_{d\backslash P_z}\sum_{\substack{p\backslash P_z\\ p<q(d)}}\mu(d)\Big(\chi(d)-\chi(pd)\Big)S(A_{pd};P^p_{(pd)},x)\\
&= \sum_{d\backslash P_z}\mu(d)\chi(d)|\mathcal{A}_d| - \sum_{d\backslash P_z}\sum_{\substack{p\backslash P_z\\ p<q(d)}}\mu(d)\Big(\chi(d)-\chi(pd)\Big)S(A_{pd};P^p,x)
\end{aligned}
$$

since $P^p_{(pd)} = P^p$.
Suppose we have $\chi(1)=1$ and $\chi(d)=0$ for $d>1$. Then

$$S(\mathcal{A};P^z,x) = |A| - \sum_{p<z,p\in P} S(A_p;P^p,x).$$

Now let χ_1,χ_2 be the characteristic functions of the divisor sets that we wish to use to get upper and lower bounds respectively. If we arrange

$$(-1)^{i-1}\mu(d)\Big(\chi_i(d)-\chi_i(pd)\Big)\geq 0$$

whenever $pd\backslash P_z$ and $p<q(d)$ for $i=1,2$, then

$$\sum_{d\backslash P_z}\mu(d)\chi_2(d)|\mathcal{A}_d| \leq S(\mathcal{A};P^z,x) \leq \sum_{d\backslash P_z}\mu(d)\chi_1(d)|\mathcal{A}_d|.$$

The above inequality is valid (needless to say) only if the sums involving χ_i are positive. This gives us a set of sufficient conditions for our functions χ_i to be well behaved.

If $pd\backslash P_z$ and $p<q(d)$ then the conditions can be satisfied in only one of the following ways:

(1) $\chi_i(d)=\chi_i(pd)$
(2) $\chi_i(d)=1,\chi_i(pd)=0$ and $\mu(d)=(-1)^{i-1}$
(3) $\chi_i(d)=0,\chi_i(pd)=1$ and $\mu(d)=(-1)^i$.

We can avoid the last possibility by requiring that the functions χ_i be divisor closed, i.e. that

$$\chi_i(d)=1 \Rightarrow \Big(\forall\delta\backslash d:\chi_i(\delta)=1\Big).$$

So the functions χ_i for $i=1,2$ should have the following properties:

(1) If $d\backslash P_z$, then either $\chi_i(d)=0$ or $\chi_i(d)=1$;
(2) $\chi_i(1)=1$ (this is required for the derivation in Proposition (2.2.1));
(3) $\chi_i(d)=1 \Rightarrow \Big(\forall\delta\backslash d:\chi_i(\delta)=1\Big)$;
(4) $\chi_i(d)=1,\mu(d)=(-1)^i \Rightarrow \chi_i(pd)=1$ for all $pd\backslash P_z$, where $p<q(d)$.

Suppose we restrict $\chi^{(r)}$ (which was the divisor selection function of the previous section) to also limit the number of prime factors that come from a certain interval. Suppose at most δ_1 divisors can come from the interval $z_1<p<z$. Then the remainder term obeys

$$\sum_{d\backslash P_z}\chi^{(r)}(d)|R_d| \leq \Big(1+\sum_{p<z}\omega(p)\Big)^{\delta_1}\Big(1+\sum_{p<z_1}\omega(p)\Big)^{r-1-\delta_1}.$$

This allows a more accurate estimation of the remainder term. The full Brun Sieve uses n such intervals to minimize the remainder term.

The first step is to compare $\sum_{d\backslash P_z} \mu(d)\frac{\omega(d)}{d}$ with $\sum_{d\backslash P_z} \mu(d)\chi_i(d)\frac{\omega(d)}{d}$. By writing $\chi_i(d) = 1 + \mho_i(d)$, we can split the sum

$$\sum_{d\backslash P_z} \mu(d)\chi_i(d)\frac{\omega(d)}{d} = \sum_{d\backslash P_z} \mu(d)\chi_i(d)\frac{\omega(d)}{d} + \sum_{d\backslash P_z} \mu(d)\mho_i(d)\frac{\omega(d)}{d}.$$

Let $d = p_1 \cdots p_r$; then

$$\begin{aligned} 1 - \chi_i(d) &= \chi_i(p_2 \cdots p_r) - \chi_i(p_1 \cdots p_r) \\ &\quad + \chi_i(p_3 \cdots p_r) - \chi_i(p_2 \cdots p_r) \\ &\quad + \cdots \\ &\quad + \chi_i(1) - \chi_i(p_r). \end{aligned}$$

If we write $P_{(p^+,z)} = \prod_{p<q<z, q\in P} q$ then we can write the above as:

$$1 - \chi_i(d) = \sum_{p\backslash d} \Big\{ \chi_i(\gcd(d, P_{(p^+,z)})) - \chi_i(\gcd(d, P_{p,z})) \Big\}.$$

This gives us

$$\sum_{d\backslash P_z} \mu(d)\chi_i(d)\frac{\omega(d)}{d} = W(z) + \sum_{d\backslash P_z} \sum_{p\backslash d} \mu\left(\frac{d}{p}\right) \Big\{ \chi_i(\gcd(d, P_{(p^+,z)})) - \chi_i(\gcd(d, P_{(p,z)})) \Big\} \frac{\omega(d)}{d}.$$

Let $d = \delta p t$, where $\delta\backslash P_p$ and $t\backslash P_{(p^+,z)}$. Rewriting the above expression we get:

$$\begin{aligned} \sum_{d\backslash P_z} \mu(d)\chi_i(d)\frac{\omega(d)}{d} &= W(z) + \sum_{p<z} \frac{\omega(p)}{p} \sum_{\delta\backslash P_p} \mu(\delta)\frac{\omega(\delta)}{\delta} \sum_{t\backslash P_{(p^+,z)}} \mu(t)\frac{\big(\chi_i(t) - \chi_i(pt)\big)}{t}\omega(t) \\ &= W(z) + (-1)^{i-1} \sum_{p<z} \frac{\omega(p)}{p} W(p) \sum_{t\backslash P_{(p^+,z)}} \frac{\big(\chi_i(t)(1 - \chi_i(pt))\big)}{t}\omega(t), \end{aligned}$$

where we have used $\chi_i(t) - \chi_i(pt) = (-1)^{i-1}\mu(t)\chi_i(t)(1 - \chi_i(pt))$ if $pt\backslash P_z$ and $p < q(t)$. To verify this, if $\chi_i(t) = \chi_i(pt)$, then both sides are 0, and this is the case if $\chi_i(pt) = 1$ (since the χ_i are divisor closed). Now if $\chi_i(t) = 1$ and $\chi_i(pt) = 0$, then from the properties of χ_i listed above, we have that $\mu(t) = (-1)^{i-1}$, and so the relation holds.
So finally we get

$$\sum_{d\backslash P_z} \mu(d)\chi_i(d)\frac{\omega(d)}{d} = W(z)\Big\{ 1 + (-1)^{i-1} \sum_{p<z} \frac{\omega(p)}{p} \frac{W(p)}{W(z)} \sum_{t\backslash P_{(p^+,z)}} \frac{\chi_i(t)\big(1 - \chi_i(pt)\big)}{t}\omega(t) \Big\}.$$

This identity holds in general for every combinatorial sieve with χ_i satisfying the properties listed above, provided $W(z)$ and $W(p)$ are well defined. This will happen if $g(d)$ stays bounded.

Construction of the Divisor sets: Let r be a positive integer and let z_i for $1 \le i \le r$ be real numbers. We will divide the interval $[2 \cdots z]$ into r intervals as follows: let

$$2 = z_r < z_{r-1} < \cdots < z_1 < z_0 = z.$$

Let $d\backslash P_z$ and $\beta_n = \gcd(d, P_{(z_n,z)})$ for $1 \le n \le r$. Let us set $\chi_i(d) = 1$ if for all $1 \le n \le r$ we have $\nu(\beta_n) \le A + Cn$, where A and C will be picked to make χ_i an acceptable function. For the current choice χ_i is already divisor closed, so the only property we need to check is:

$$\chi_i(t) = 1, \mu(t) = (-1)^i \Rightarrow \Big(\forall pt\backslash P_z, p < q(t) : \chi_i(pt) = 1 \Big).$$

Let $z_m \le p < z_{m-1}$. Since $\chi_i(t) = 1$ we should have $\nu(\beta_m) \le A + Cm$. If $\nu(\beta_m) < A + Cm$ then $\chi_i(pt) = 1$. Now if $\nu(\beta_m) = A + Cm$, then we also have $\mu(t) = (-1)^i$. By definition $\mu(t) = (-1)^{\nu(t)}$, since $\nu(t) = A + Cm$, we have $\mu(t) = (-1)^{A+Cm}$. This suggests that we set $A = B - i$. Then we have that

$(-1)^{B+Cm} = 1$ or $B + Cm$ should be even. If we make $B + Cm$ an odd number, then the assumption that $\chi_i(t) = 1$ and $\mu(t) = (-1)^i$ results in a contradiction. Consequently, $\nu(\beta_m) = A + Cm$ cannot happen if $\chi_i(t) = 1$. For some integer b we set $B = 2b - 1$ and $C = 2$. This suggests using $\nu(\beta_n) \le 2b - 1 + i + 2n$ to be the condition on the number of factors of d in the interval $[z_n, \cdots, z)$. Summarizing, the characteristic functions of the divisor sets will be (for $i = 1, 2$)

$$\chi_i(d) = \begin{cases} 1 & \text{if } \forall m : 1 \le m \le r,\ \nu(\beta_m) \le 2b - i - 1 + 2m, \\ 0 & \text{otherwise.} \end{cases}$$

The construction was such that the above function is the characteristic function of an acceptable divisor set.

Derivation of the Sieve bounds: Now

$$\begin{aligned} &\sum_{1 \le n \le r} \sum_{z_n \le p < z_{n-1}} \frac{\omega(p) W(p)}{p W(z)} \sum_{t \backslash P_{(p^+, z)}} \frac{\chi_i(t)\big(1 - \chi_i(pt)\big)}{t} \omega(t) \\ \le &\sum_{1 \le n \le r} \frac{W(z_n)}{W(z)} \sum_{z \le p < z_{n-1}} \frac{\omega(p)}{p} \sum_{t \backslash P_{(p^+, z)}} \frac{\chi_i(t)(1 - \chi_i(pt))}{t} \omega(t). \end{aligned}$$

We have used the fact that $W(p) \le W(z_n)$ if $z_n \le p < z_{n-1}$. Now for each t which makes a contribution we have $\chi_i(pt) = 0$ and $\chi_i(t) = 1$. So we must have $\nu(t) = 2b - i + 2n - 1$ for $z_n \le p < z_{n-1}$. Hence this sum is at most

$$\sum_{1 \le n \le r} \frac{W(z_n)}{W(z)} \sum_{\substack{d \backslash P_{(z_n, z)} \\ \nu(d) = 2b - i + 2n}} \frac{\omega(d)}{d},$$

and so

$$\sum_{p < z} \frac{\omega(p) W(p)}{p W(z)} \sum_{t \backslash P_{(p^+, z)}} \frac{\chi_i(t)\big(1 - \chi_i(pt)\big)}{t} \omega(t) \le \sum_{1 \le n \le r} \frac{W(z_n)}{W(z)} \frac{1}{(2b - i + 2n)!} \Bigg(\sum_{z_n \le p < z} \frac{\omega(p)}{p} \Bigg)^{(2b - i + 2n)}.$$

Now to simplify this sum further we have to make some assumptions about the function $\omega(p)$; instead of assuming $\omega(p) = O(1)$ we shall use the more general assumption:

$$\sum_{w \le p < z} \frac{\omega(p) \ln p}{p} \le \kappa \ln\left(\frac{z}{w}\right) + \eta, \text{ for } 2 \le w \le z. \tag{2.12}$$

If indeed we had $\omega(p) = 1$, then we have

$$\sum_{w \le p < z} \frac{\ln p}{p} \le \ln\left(\frac{z}{w}\right) + 1, \text{ for } 2 \le w \le z.$$

So the assumption we have made is an assumption on the average distribution of $\omega(p)$. Such an assumption usually holds, and is much easier to verify in more complicated situations.

A question we can ask is: Does the above assumption imply a bound for the sum

$$\sum_{w \le p < z} \frac{\omega(p)}{p}?$$

Let

$$S(k) \equiv \sum_{w \le p < k} \frac{\omega(p) \ln p}{p}.$$

Since $S(k) - S(k-1) = \frac{\omega(k)\ln k}{k}$ if k is prime we have

$$\begin{aligned}\sum_{w \le p < z} \frac{\omega(p)}{p} &= \sum_{w \le k < z} \frac{S(k) - S(k-1)}{\ln k} \\ &= \sum_{w \le k < z-1} S(k)\left(\frac{1}{\ln k} - \frac{1}{\ln k + 1}\right) \\ &= \sum_{w \le k < z-1} S(k)\left(\frac{\ln(k+1) - \ln k}{\ln k \ln k + 1}\right)\end{aligned}$$

Now

$$\ln(k+1) = \ln k + \ln\left(1 + \frac{1}{k}\right),$$

and since $1 + x \le e^x$ we have $\ln\left(1 + \frac{1}{k}\right) \le \frac{1}{k}$.

Then

$$\sum_{w \le p < z} \frac{\omega(p)}{p} \le \sum_{w \le k < z-1} \frac{S(k)}{k \ln^2 k} \tag{2.13}$$

$$\le \sum_{w \le k < z-1} \frac{\kappa \ln\left(\frac{k}{w}\right) + \eta}{k \ln^2 k} \tag{2.14}$$

$$\le \kappa \sum_{w \le k < z-1} \frac{1}{k \ln k} - \ln w \sum_{w \le k < z-1} \frac{1}{k \ln^2 k} + \eta \sum_{w \le k < z-1} \frac{1}{k \ln^2 k} \tag{2.15}$$

$$\le \kappa \ln\left(\frac{\ln z}{\ln w}\right) + \frac{\eta}{\ln w}. \tag{2.16}$$

Here we have used

$$\int_w^z \frac{1}{x \ln x} dx = -\ln \ln w + \ln \ln z,$$

and

$$\int_w^z \frac{1}{x \ln^2 x} dx = \frac{1}{\ln w} - \frac{1}{\ln z}.$$

Now returning to our original problem we need bounds on

$$\frac{W(z_n)}{W(z)} = \prod_{z_n \le p < z} \frac{1}{\left(1 - \frac{\omega(p)}{p}\right)},$$

and so

$$\ln \frac{W(z_n)}{W(z)} \approx \sum_{z_n \le p < z} \frac{\omega(p)}{p}.$$

Our assumption (2.12) yields a bound on

$$\sum_{z_n \le p < z} \frac{\omega(p)}{p}$$

by (2.16). Thus we expect that a bound of the form

$$\frac{W(z_n)}{W(z)} \le e^{\gamma\left(n\lambda + \frac{c}{\ln z}\right)}$$

can be enforced with some— constants γ, λ and c. This can be achieved for example with a double-exponential fall-off of z_n with respect to z, in fact this is what we shall do later.
If a bound for $\frac{W(z_n)}{W(z)}$ of the above form exists, then this also gives us (as we might expect)

$$\sum_{z_n \le p < z} \frac{\omega(p)}{p} \le \sum_{z_n \le p < z} \ln\left(\frac{1}{1 - \frac{\omega(p)}{p}}\right)$$
$$\le \ln \frac{W(z_n)}{W(z)} < \gamma\left(n\lambda + \frac{c}{\ln z}\right).$$

Let $f = \frac{c}{\ln z}$, and suppose we can enforce $\gamma = 2$ (this helps in the simplification to follow). Then

$$\sum_{1 \le n \le r} \frac{W(z_n)}{W(z)} \sum_{\substack{d \backslash P_{(z_n, z)} \\ \nu(d) = 2b - i + 2n}} \frac{\omega(d)}{d} \le \sum_{1 \le n \le r} e^{2n\lambda + 2f} \frac{(2n\lambda + 2f)^{2b-i+2n}}{(2b - i + 2n)!}$$
$$\le \sum_{1 \le n \le r} e^{2f} \left(e^{\lambda}\right)^{2n} \frac{(2n)^{2b-i+2n}}{(2n)!(2n)^{2b-i}} \left(1 + \frac{f}{n}\right)^{2b-i+2n}$$

(since $(2b - i + 2n)! \ge (2n)!(2n)^{2b-i}$)

$$= \sum_{1 \le n \le r} e^{2f} (\lambda e^{\lambda})^{2n} \frac{(2ne^{-1})^{2n} e^{2n}}{(2n)!} (\lambda^{2b-i}) \left(1 + \frac{f}{n\lambda}\right)^{2b-i} \left(1 + \frac{f}{n\lambda}\right)^{2n}$$
$$= e^{2f} (\lambda + f)^{2b-i} \sum_{1 \le n \le r} \frac{(2ne^{-1})^{2n}}{(2n)!} (\lambda e^{1+\lambda})^{2n} \left(1 + \frac{f}{n\lambda}\right)^{2n}$$

since $\frac{(ne^{-1})^n}{n!}$ is decreasing, and $\left(1 + \frac{f}{n\lambda}\right)^{2n} \le e^{\frac{2f}{\lambda}}$. Also assuming $\lambda e^{1+\lambda} \le 1$;

$$\sum_{1 \le n \le r} \frac{W(z_n)}{W(z)} \sum_{\substack{d \backslash P_{(z_n, z)} \\ \nu(d) = 2b - i + 2n}} \frac{\omega(d)}{d} \le e^{2f} (\lambda + f)^{2b-i} 2e^{-2} e^{\frac{2f}{\lambda}} \sum_{1 \le n} \left(\lambda e^{1+\lambda}\right)$$
$$= \frac{2\lambda^{2b-i+2} e^{2\lambda}}{1 - (\lambda e^{1+\lambda})^2} \left(1 + \frac{c}{\lambda}\right)^{2b-i} e^{2f(1 + \frac{1}{\lambda})}$$
$$\le \frac{2\lambda^{2b-i+2} e^{2\lambda}}{1 - (\lambda e^{1+\lambda})^2} e^{(2b-i+4)\frac{f}{\lambda}}.$$

Thus

$$\sum_{d \backslash P_z} \mu(d) \chi_i(d) \frac{\omega(d)}{d} = W(z) \left(1 + 2\theta \frac{\lambda^{2b-i+2} e^{2\lambda}}{1 - (\lambda e^{1+\lambda})^2} e^{(2b-i+4)\frac{f}{\lambda}}\right) \text{ for } i = 1, 2.$$

Now we have to bound the remainder term, which is significantly easier. Let us assume that $\omega(p) \le A$ for some constant $A > 0$. Then

$$\sum_{d \backslash P_z} \chi_i(d) |R_d| \le \sum_{d \backslash P_z} \chi_i(d) \omega(d)$$
$$\le \left(1 + \sum_{p < z} \omega(p)\right)^{2b-i+1} \prod_{1 \le n \le r-1} \left(1 + \sum_{p < z_n} \omega(p)\right)^2$$
$$\le (1 + A(2\mathrm{li}\, z + 3))^{2b-i+1} \prod_{1 \le n \le r-1} (1 + A(2\mathrm{li}\, z_n + 3))^2 \text{ for } i = 1, 2.$$

Selection of the intervals: We select the numbers z_n with an exponential fall-off in the logarithm. Let $\Lambda > 0$ be a real number. Define

$$\ln z_n = e^{-n\Lambda} \ln z \text{ for } n = 1, \cdots, r-1;$$

and set $z_r = 2$.
Here r is selected such that

$$\ln z_{r-1} = e^{-(r-1)\Lambda} \ln z > \ln 2,$$

and

$$e^{-r\Lambda} \ln z \le \ln 2,$$

so we have

$$e^{(r-1)\Lambda} < \frac{\ln z}{\ln 2} \le e^{r\Lambda}.$$

Thus for a suitable constant B the remainder term becomes

$$\begin{aligned}\sum_{d \backslash P_z} \chi_i(d)|R_d| &\le \left(\frac{Bz}{\ln z}\right)^{2b-i+1} \prod_{1 \le n < r} \left(\frac{Bz_n e^{n\Lambda}}{\ln z}\right) \\ &= \left(\frac{Bz}{\ln z}\right)^{2b-i+1} \prod_{1 \le n \le r-1} \left(\frac{Be^{\frac{1}{2}r\Lambda}}{\ln z}\right)^{r-1} \prod_{1 \le n \le r-1} z_n^2.\end{aligned}$$

Now

$$\frac{Be^{\frac{1}{2}r\Lambda}}{\ln z} \le \frac{Be^{\Lambda/2}}{\ln z} \sqrt{\frac{\ln z}{\ln 2}} < 1,$$

and also

$$\prod_{1 \le n \le r-1} z_n^2 = \exp\left(2 \ln z \sum_{1 \le n \le r-1} e^{-n\Lambda}\right) \le z^{\frac{2}{e^\Lambda - 1}}.$$

Thus

$$\sum_{d \backslash P_z} \chi_i(d)|R_d| = O\big(z^{2b-i+1+\frac{2}{e^\Lambda - 1}}\big) \text{ for } i = 1, 2.$$

We still have to check that $\frac{W(z_n)}{W(z)} \le e^{2(n\lambda+f)}$. By our assumptions about the sum $\sum_{w \le p < z} \frac{\omega(p) \ln p}{p}$ we have

$$\begin{aligned}\frac{W(z_n)}{W(z)} &\le \exp\left(n\Lambda\kappa + \frac{2ce^{n\Lambda}}{\ln z}\right) \\ &= e^{2c} \exp\left(n\left(\Lambda\kappa + \frac{2c}{\ln z} \frac{e^{n\Lambda} - 1}{n}\right)\right), n = 1, \cdots, r.\end{aligned}$$

If $1 \le \frac{1}{1 - \frac{\omega(p)}{p}} \le A$, then

$$c = \frac{\eta}{2}\left(1 + A\kappa + \frac{\eta A}{\ln 2}\right).$$

Since $\Lambda > 0$ we have

$$\frac{e^{n\Lambda} - 1}{n} \le \frac{e^{r\Lambda} - 1}{r},$$

and this is at most

$$\Lambda\frac{e^{r\Lambda}}{r\Lambda}\le\Lambda\frac{e^{\Lambda}}{\ln 2}\frac{\ln z}{\ln(\ln z/\ln 2)}.$$

So we get

$$\frac{W(z_n)}{W(z)}\le e^{2c}\exp\left(n\Lambda\kappa\left(1+\frac{2ce^{\Lambda}}{\kappa\ln 2}\frac{1}{\ln(\ln z/\ln 2)}\right)\right)\text{ for } n=1,\cdots,r.$$

To meet our conditions on $\frac{W(z_n)}{W(z)}$ we take

$$\Lambda=\frac{2\lambda}{\kappa}\frac{1}{1+\epsilon}$$
$$\epsilon=\frac{1}{\delta e^{\frac{1}{\kappa}}},$$

and so

$$\begin{aligned}e^{\frac{2\lambda}{\kappa}}-e^{\Lambda}&\le\left(\frac{2\lambda}{\kappa}-\Lambda\right)e^{\frac{2\lambda}{\kappa}}\\&\le\epsilon^{\Lambda}e^{\frac{1}{\kappa}}.\end{aligned}$$

Since $e^{\Lambda}-1\ge\Lambda$ we have

$$\frac{e^{\frac{2\lambda}{\kappa}}-1}{e^{\Lambda}-1}\le 1+\frac{\epsilon\Lambda e^{\frac{1}{k}}}{e^{\Lambda-1}}\le 1+\epsilon e^{\frac{1}{\kappa}}=1+\frac{1}{\delta}.$$

With $\xi=1+\frac{1}{\delta}$ we obtain

$$\sum_{d\backslash P_z}\chi_i(d)|R_d|=O\left(z^{2b-i+1+\frac{2\xi}{\left(e^{\frac{2\lambda}{\kappa}}-1\right)}}\right)\text{ for } i=1,2.$$

Thus we have proved the following theorem.

THEOREM 2.2.2. *Assume that*

$$1\le\frac{1}{1-\frac{\omega(p)}{p}}\le A,$$

$$\sum_{w\le p<z}\frac{\omega(p)\ln p}{p}\le\kappa\ln\left(\frac{\ln z}{\ln w}\right)+\frac{\eta}{\ln w},$$

and

$$|R_d|\le\omega(d).$$

Let λ *be such that* $0<\lambda e^{1+\lambda}<1$. *Then*

$$S(\mathcal{A};P^z,x)\le xW(z)\left\{1+2\frac{\lambda^{2b+1}e^{2\lambda}}{1-(\lambda e^{1+\lambda})^2}\exp\left((2b+3)\frac{c}{\lambda\ln z}\right)\right\}+O\left(z^{2b-1+\frac{2\xi}{e^{\frac{2\lambda}{\kappa}}-1}}\right),\tag{2.17}$$

and

$$S(\mathcal{A};P^z,x)\ge xW(z)\left\{1-2\frac{\lambda^{2b}e^{2\lambda}}{1-(\lambda e^{1+\lambda})^2}\exp\left((2b+2)\frac{c}{\lambda\ln z}\right)\right\}+O\left(z^{2b-1+\frac{2\xi}{e^{\frac{2\lambda}{\kappa}}-1}}\right),\tag{2.18}$$

where

$$c=\frac{\eta}{2}\left(1+A\left(\kappa+\frac{\eta}{\ln 2}\right)\right)$$

and $\xi=1+\epsilon$ *for* $0<\epsilon<1$. □

Application to the Twin Primes problem : We set $\mathcal{A} = \{n(n+2) \mid n \leq x\}$. In this case we have $\omega(2) = 1$ and $\omega(p) = 2$. Further, all the conditions of Theorem (2.2.2) hold, and the lower bound is seen to be positive. Thus (2.18) tends to infinity with x, ([**HR74**], p.63) for $z = x^{\frac{1}{u}}$ with $u < 8$. This implies that every divisor of a number in the sifted set is $\geq x^{\frac{1}{u}}$ so each number in the sifted set can have at most $u < 8$ factors[1]. Thus we have the following theorem.

THEOREM 2.2.3. *There are infinitely many* n *such that* $\nu(n(n+2)) \leq 7$. □

We will look at some interesting applications of Brun's sieve in the following sections.

2.3. Orthogonal Latin Squares and the Euler Conjecture

DEFINITION 2.3.1. A Latin square of order n is an $n \times n$ matrix with entries in $S = \{1, \cdots, n\}$ such that every row and column is a permutation of the set S.

DEFINITION 2.3.2. Two Latin squares A and B or order n are said to be *mutually orthogonal* if the n^2 pairs (a_{ij}, b_{ij}) are distinct.

Here is a Latin square of order 3:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix},$$

and here is a latin square that is orthogonal to it:

$$B = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}.$$

Euler conjectured that there are no mutually orthogonal Latin squares of order n, where $n \equiv 2 \mod 4$. The conjecture was disproved for the case $n = 10$, and later Bose, Parker and Shrikande [**BPS60**] showed that for every higher $n > 6$ the conjecture was false. Let $\perp(n)$ be the number of orthogonal latin squares of order n. Chowla, Erdős and Straus [**CES60**] building on this and some previous results, established that $\perp(n) > \frac{1}{3}n^{\frac{1}{91}}$ for large enough n. The proof involves an interesting use of the Brun Sieve, and we shall give an account of this. The exponent $\frac{1}{91}$ is far from optimal and has been subsequently improved.

The starting point for the proof is the following pair of results:

THEOREM 2.3.3. [**BPS60**] *If* $k \leq \perp(m) + 1$ *and* $1 < u < m$ *then*

$$\perp(km + u) \geq \min\{\perp(k), \perp(k+1), \perp(m) + 1, \perp(u) + 1\} - 1.$$

THEOREM 2.3.4 (MacNiesh). *(1)* $\perp(ab) \geq \min\{\perp(a), \perp(b)\}$;

(2) $\perp(q) = q - 1$ *if* q *is a power of a prime.*

First we shall prove the following:

THEOREM 2.3.5.

$$\lim_{n \to \infty} \perp(n) = \infty.$$

Proof : The idea is to have a lower bound on each of the quantities involved in Theorem (2.3.3), and then use the theorem with $km + u = n$.

Let x be a large positive integer. If

$$k + 1 = \prod_{p \leq x} p^x,$$

then by Theorem (2.3.4) we have $\perp(k+1) \geq 2^x - 1 \geq x$. Also since $k \equiv 1 \mod p$ for $p \leq x$ all the prime factors k are larger than x, so applying Theorem (2.3.4) again we have $\perp(k) \geq x$.

[1] For a similar derivation see Theorem (2.3.6).

Now we select m in two pieces m_1 and m_2. The first piece is set to be

$$m_1 = k^k \prod_{\substack{q \not| n \\ q \le x}} q^k.$$

Note that m_1 is bounded in terms of x alone. Now if n is large enough the interval

$$\left[\frac{n}{(k+1)m_1} \cdots \frac{n-1}{km_1}\right]$$

contains an integer m_2 such that $m_2 \equiv 1 \mod k!$, simply because the length of the interval becomes larger than $k!$.

Now set $m = m_1 m_2$ then $\perp(m) \geq \min\{\perp(m_1), \perp(m_2)\} \geq \min\{2^k - 1, k\} \geq k$. Thus we have $\perp(m) + 1 \geq k$ to satisfy the condition of Theorem (2.3.3). Set $u = n - km$; we need to bound $\perp(u)$, but first we need to verify that $1 < u < m$. We have $\frac{n}{(k+1)m_1} < m_2 < \frac{(n-1)}{km_1}$ or $\frac{n}{(k+1)} < m < \frac{n-1}{k}$. This yields $km + 1 < n$ and $km + m > n$, which implies that $1 < u < m$. Let $p \leq x$ then

$$km \not\equiv n \mod p.$$

This is because k has prime factors only above x, m_1 has a small prime factor only if it does not divide n, and m_2 has prime factors only above $k \geq x$. Thus $km + u \not\equiv n \mod p$ for $p < x$, and so no prime smaller than x divides u. Thus we get $\perp(u) \geq x$. Now applying Theorem (2.3.3) we have $\perp(km + u) = \perp(n) \geq x$. □

Note that this has already disproved Euler's conjecture. It is clear that our method of proof relied on our ability to produce some numbers with large prime factors and some congruence properties, this indicates that a sieve argument might help. The necessary machinery from sieves is encapsulated in the following theorem:

THEOREM 2.3.6. [**Rad24**] *Let $p_1, \cdots, p_r$ be primes, and let $a_i < p_i, b_i < p_i$ be non-negative integers for $1 \leq i \leq r$. Let $D > 1$ be an integer with $\gcd(D, p_i) = 1$ for each i, $1 \leq i \leq r$, and Λ is an integer, $0 < \Lambda < D$ such that $\gcd(\Lambda, D) = 1$. Let*

$$P(D, x; p_1, a_1, b_1; p_2, a_2, b_2; \cdots; p_r, a_r, b_r) \\ = \left|\{n \leq x \mid n \equiv \Lambda (\mathrm{mod}\ D), (\forall i : 1 \leq i \leq r) : n \not\equiv a_i (\mathrm{mod}\ p_i), n \not\equiv b_i (\mathrm{mod}\ p_i)\}\right|.$$

If $p_1 < p_2 < \cdots < p_r$ and $p_i > 2$, then

$$P(D, x; p_1, a_1, b_1; \cdots; p_r, a_r, b_r) > \frac{Cx}{D \ln^2 p_r} - C' p_r^{7.938},$$

where C and C' are positive constants.

REMARK 2.3.7. The original theorem has 7.9 instead of our slightly worse 7.938, but this can be improved using a more detailed analysis of our proof.

Proof : The quantity $S(\mathcal{A}; P^z, x)$ is the number of integers in $\mathcal{A}$ that are $\not\equiv 0$ modulo p_i for each $p_i \in P, p_i \leq z$. In this case we have two constraints for each prime p_i. But we can collapse these two constraints into one as follows. The constraint for the prime i is that $n \not\equiv a_i$, $n \not\equiv b_i$ modulo p_i. So the constraint fails iff

$$(n - a_i)(n - b_i) \equiv 0 \mod p_i.$$

Let $\mathcal{A} = \{n \leq x \mid n \equiv \Lambda \mod D\}$, $\mathcal{A}_{p_i} = \{n \leq x \mid (n - a_i)(n - b_i) \equiv 0 \mod p_i\}$, and if $d = p_{i_1} \cdots p_{i_k}$ then $\mathcal{A}_d = \{n \leq x \mid \prod_{1 \leq j \leq k} (n - a_{i_j})(n - b_{i_j}) \equiv 0 \mod d\}$. Suppose $|\mathcal{A}_{p_i}| = \frac{\omega(p_i)}{p_i} x + R_{p_i}$; then we see that if d is squarefree then $|\mathcal{A}_d| = \frac{\omega(d)}{d} x + R_d$, where $\omega(d)$ is defined multiplicatively. Thus we are interested in the estimate:

$$\begin{aligned} P(D, x; p_1, a_1, b_1; \cdots; p_r, a_r, b_r) &= |\mathcal{A}| - \sum |\mathcal{A}_{p_i}| + \sum |\mathcal{A}_{p_i p_j}| - \cdots \\ &= \sum_{d \backslash p_1 \cdots p_r} \mu(d) |\mathcal{A}_d|, \end{aligned}$$

which is just the sieve estimate.

The congruence $(n - a_i)(n - b_i) \equiv 0 \mod p_i$ has at most 2 solutions modulo p_i so $\omega(p_i) = 2$ for each i. We will try to apply Brun's Sieve to this problem.

We just need to verify that the conditions of the proof of Theorem (2.2.2) are valid. First

$$\frac{1}{1 - \frac{\omega(p)}{p}} \leq 3$$

so $A = 3$. Next

$$\sum_{\substack{w \leq p < z \\ p \in \{p_1, \cdots, p_r\}}} \frac{\omega(p) \ln p}{p} \leq 2 \sum_{w \leq p < z} \frac{\ln p}{p} \leq \left(\ln \frac{z}{w} + 1 \right),$$

from which we have $\kappa = 2$, and $\eta = 2$. $R_d \leq \omega(d)$ also holds. Thus by the lower bound we have (with $b = 2$):

$$S(\mathcal{A}; P = \{p_1, \cdots, p_r\}, z) \geq |A^x| W(z) \left\{ 1 - 2 \frac{(\lambda e^\lambda)^2}{1 - (\lambda e^\lambda)^2} \exp\left(\frac{4c}{\lambda \ln z} \right) \right\} + O\left(z^{1 + \frac{2\xi}{e^{\lambda} - 1}} \right).$$

So all we need to show is that there is a λ such that

$$1 + \frac{2 + 2\epsilon}{e^{\lambda - 1}} \leq u \leq 7.938$$

and

$$1 - \frac{2(\lambda e^\lambda)^2}{1 - (\lambda e^{1+\lambda})^2} > 0.$$

Then the second condition implies

$$\lambda e^\lambda < \frac{1}{\sqrt{2 + e^2}} \approx 0.3263540699 \cdots$$

and the first implies

$$\frac{2 + 2\xi}{6.938} + 1 \leq e^\lambda.$$

Now set $\xi = 10^{-9}$, so we must have $\lambda \geq \log 1.288267513692707$. This value of λ also satisfies the other constraint.

Now we take $z = p_r$, and using $|A^x| = \frac{x}{D} + \theta$, $|\theta| < 1$,

$$S(\mathcal{A}; P^{p_r}, x) \geq \frac{Cx}{D} \prod_{1 \leq i \leq r} \left(1 - \frac{2}{p_i} \right) + O(p_r^{7.938}),$$

and also

$$\prod_i \left(1 - \frac{2}{p_i} \right) \leq \prod_{p \leq p_r} \left(1 - \frac{2}{p_i} \right).$$

Now in

$$\ln \prod_{p \leq p_r} \left(1 - \frac{2}{p_i} \right) = -2 \sum_{p \leq p_r} \frac{1}{p_i} - 2 \sum_{p \leq p_r} \sum_{m > 1} \frac{1}{m p^m}$$

the second sum converges, so we have

$$\prod_{p \leq p_r}\left(1-\frac{2}{p_i}\right)=\frac{1}{\ln^2 p_r}+o\left(\frac{1}{\ln^2 p_r}\right).$$

The theorem follows. □

Now we have the following simple lemma:

LEMMA 2.3.8. *For all* $c, 0 < c < 1$*, the number of integers* $y \leq x$ *that are divisible by a prime factor* $p > n^c$ *of* n*, is at most* $\frac{x}{cn^c}$*.*

Proof : At most $\frac{x}{p}$ integers $y \leq x$ are divisble by p and so the total number of such integers is given by:

$$\begin{aligned}\sum_{\substack{p\backslash n\\ p>n^c}}\frac{x}{p} &\leq \frac{x}{n^c}\sum_{\substack{p\backslash n\\ p>n^c}} 1\\ &\leq \frac{x}{cn^c}.\end{aligned}$$

The last part follows because, there can be at most $1/c$ prime factors of a number n that are greater than n^c. □

THEOREM 2.3.9. [**CES60**] *There is an* $n_0 > 0$ *such that for all* $n > n_0$*,* $\perp(n) > \frac{1}{3}n^{\frac{1}{91}}$*.*

Proof : The idea as before is to apply Theorem (2.3.3) to suitable k, m and u for a given n such that $n = km + u$. For this to yield a lower bound on $\perp(n)$ we need lower bounds on $\perp(k), \perp(k+1), \perp(m)$ and $\perp(u)$.

We begin with the selection of k: we need k as well as $k+1$ to have no small prime factors. This is exactly the sort of problem handled by the theorem we have just proved. It turns out that the constraints on k depend on the parity of n.

Case 1. (n even). Consider the constraints:

$$\begin{aligned}&k \equiv -1 \mod 2^{\lfloor \frac{1}{91}\lg n\rfloor}\\ &k \not\equiv 0 \text{ or } -1 \text{ for } p \leq n^{\frac{1}{10}}\end{aligned}$$

and $k < n^{\frac{1}{10}}$. The first congruence restricts k to lie in an arithmetic progression with difference $2^{\lfloor \frac{1}{91}\lg n\rfloor} < c_1 n^{\frac{1}{91}}$. The second incongruence implies that both k and $k+1$ are free of small prime factors, apart from the large power of 2 dividing $k+1$.

Now applying Theorem (2.3.6) there are at least:

$$\begin{aligned}\frac{Cn^{\frac{1}{10}}}{c_1\frac{1}{90^2}n^{\frac{1}{91}}\log^2 n}-C'n^{\frac{79.38}{10}\frac{1}{90}} &= c_2\frac{n^{\frac{81}{910}}}{\log^2 n}-C'n^{\frac{79.38}{900}}\\ &> c_3\frac{n^{\frac{81}{910}}}{\log^2 n}\end{aligned}$$

values of k satisfying the constraints.

By Lemma (2.3.8) the number of integers below $n^{\frac{1}{10}}$ that have a prime factor greater than $n^{\frac{1}{90}}$ in common with n is at most $90n^{\frac{8}{90}}$. Thus from the bound for the values of k we have that there is a k such that $\gcd(k, n) = 1$. Just by our selection of k we have that k has no small prime factors and though $k+1$ has 2 as a prime factor we still have that $k+1 \equiv 0 \mod 2^{\lfloor \frac{1}{91}\lg n\rfloor}$ and all the other factors are bigger than $n^{\frac{1}{90}}$ so using Theorem (2.3.4)

$$\begin{aligned}\perp(k) &> n^{\frac{1}{90}}-1 > \frac{1}{3}n^{\frac{1}{91}}\\ \perp(k+1) &> \min\left\{\frac{1}{2}n^{\frac{1}{91}}, n^{\frac{1}{90}}\right\}-1 > \frac{1}{3}n^{\frac{1}{91}}.\end{aligned}$$

We now set $n = n_1 + n_2k$ where $0 < n_1 < k$. We cannot directly use n_1 and n_2 in our application of Theorem (2.3.3), since we have no bounds for $\perp(n_1)$ and $\perp(n_2)$. Though we have freedom in our choice of m we are still forced to pick k as our quotient in the division of n by m to write $n = km + u$. This suggests picking a u subject to certain conditions and then set $m = \frac{n-u}{k}$. Again this immediately restricts us to look at numbers that are congruent to n_1 modulo k. Let $u = n_1 + u_1k$ where u_1 is picked according to the following conditions:

$$\begin{aligned} u_1 &\not\equiv n_1 \mod 2, \\ u_1 &\not\equiv -\frac{n_1}{k} \mod p, p \not| k, \\ u_1 &\not\equiv n_2 \mod p, \text{ for } 3 \le p \le k, \end{aligned}$$

and $u_1 < n^{\frac{159}{200}}$. The first incongruence forces u_1k to be of opposite parity from n_1 and always fixes u to be odd. In this setup we will set $m = \frac{n-u}{k} = n_2 - u_1$. We want m to be free of small prime factors to guarantee a good lower bound for $\perp(m)$ and this is taken care by the third incongruence. Meanwhile, the second incongruence arranges for u itself to have no small prime factors. The limit on u_1 is forced on us because of the limitations of Theorem (2.3.6).

The restrictions of the incongruences modulo the primes $2, 3$, and 5 can be handled by restricting u_1 to belong to an arithmetic progression with difference 30. To apply Theorem (2.3.6) we need $\gcd(u_1, 30) = 1$. If we had $\gcd(u_1, 30) > 1$, then we can set $u_1' = \frac{u_1}{\gcd(u_1,30)}$ and apply Theorem (2.3.6). Thus there are at least

$$\frac{Cn^{\frac{159}{200}}}{30\log^2 k} - C'k^{\frac{79.38}{10}} > c_4\frac{n^{\frac{159}{200}}}{\log^2 n} - C'n^{\frac{79.38}{100}} > 0$$

choices for u_1, if n is large enough. Now u is not divisible by any prime $p \le k$. First suppose that $p \not| k$, then this contradicts the incongruence $n_1 \not\equiv -u_1k \mod p$. Next, if $p \backslash k$, then $p \backslash n_1$ which implies $p \backslash n$ a contradiction to $\gcd(k, n) = 1$.
Thus $\perp(u) \ge k$, but k is not divisible by any prime $\le n^{\frac{1}{90}}$, so $\perp(u) \ge n^{\frac{1}{90}} > \frac{1}{3}n^{\frac{1}{91}}$. Now as promised we set $m = \frac{n-u}{k}$, we need to verify that $m > u > 1$ to apply Theorem (2.3.3), and observe that

$$\begin{aligned} m &> \frac{n}{n^{\frac{1}{10}}} - (1 + n^{\frac{159}{200}}) > \frac{1}{2}n^{\frac{9}{10}} \\ &> n^{\frac{1}{10}} + (1 + n^{\frac{159}{200}}) \\ &> u > 1, \end{aligned}$$

for large enough n.
Furthermore, all prime factors of m exceed k by our choice of u, and hence:

$$\perp(m) \ge k > \frac{1}{3}n^{\frac{1}{91}}.$$

Finally putting all these together and applying Theorem (2.3.3) we get: $\perp(n) > \frac{1}{3}n^{\frac{1}{91}}$ for large enough even numbers n.

Case 2. (n odd). We apply Theorem (2.3.6) to $k + 1$ instead with the following constraints:

$$\begin{aligned} &k + 1 \equiv 1 \mod 2^{\lfloor \frac{1}{91} \lg n \rfloor} \\ &k + 1 \not\equiv 0 \text{ or } 1 \mod p, p \le n^{\frac{1}{90}} \\ &k + 1 \le n^{\frac{1}{10}}. \end{aligned}$$

Now the argument proceeds with the role of k and $k+1$ interchanged, and the second set of constraints becomes:

$$\begin{aligned} u_1 &\not\equiv n_2 \mod 2, \\ u_1 &\not\equiv -\frac{n_1}{k} \mod p, p \le k, p \not| k, \\ u_1 &\not\equiv n_2 \mod p, p \le k, \end{aligned}$$

and $u_1 < n^{\frac{159}{200}}$. So here both n and m are odd. The argument then proceeds similarly.
□

Better estimates for $\perp(n)$ are known— for example in [**Wil74**] a bound $\perp(n) \ge n^{\frac{1}{17}} - 2$ is proved (for large enough n). The current best estimate seems to be $\perp(n) \ge n^{\frac{1}{14.8}}$ [**Be83**].

2.4. A Theorem of Schinzel

In this section we will give an application involving a variation of Theorem 2.3.6, where we look at some constant number of constraints. The proof is an interesting use of Brun's sieve.

THEOREM 2.4.1. [**Sch66**] *For all positive integers* h *and* $N \ge 3$ *there is an integer* D *such that:*

(1) $1 \le D \le (\log N)^{20h}$*;*
(2) $\gcd(iD+1, N) = 1$*, for* $1 \le i \le h$.

Proof : For $h = 1$ we can take $D = q - 1$, where q is the least prime not dividing N. Since $\sum_{p \le D} \log p \le \log N$, we have from [**RS62**] Theorem 10, that either $D \le 100$ or $0.84D \le \log N$. Since $D \le N$ we have $D \le (\log N)^{20}$, for all $N \ge 3$.

If $N \le (\log N)^{20h}$, then $D = N$ satisfies the conditions of the theorem, so we can assume $N > (\log N)^{20h}$, with $h \ge 2$. Now

$$N > (\log N)^{20h} \Rightarrow \log N > 20h \log\log N.$$

If $\log N < 110h$, then $N < e^{110h}$ and

$$\begin{aligned} (\log N)^{20h} \ge (110h)^{20h} &= e^{\log 110h} 20h \\ &\ge e^{\log 110 + \log h} 20h \\ &= e^{94.0069h + 20h \log h} \\ &\ge e^{114.0096h}, \end{aligned}$$

which is a contradiction to $N > (\log N)^{20h}$. Hence we must have $\log N \ge 110h$, and $\log\log N \ge \log 110 + \log h \ge 5.3936$, or $\log\log N > 5$.

Let $H = \prod_{p \le 10h} p$, and we let $p_1, \cdots, p_r$ be the primes $p_i > 10h$ such that $p_i \backslash N$. Let $p_1 < p_2 < \cdots < p_r$. Let $P(H, x; p_1, \cdots, p_r)$ be the number of integers $n \le x$ such that

$$n \equiv 0 \mod H,$$

and

$$(\forall i\, \forall j) : 1 \le i \le h, 1 \le j \le r \ : \ in + 1 \not\equiv 0 \mod p_j.$$

Since $p_i > 10h$ for all the values of i in the incongruences, i is invertible. Thus, the above constraints are equivalent to a system of h incongruences per prime (we had 2 such constraints in Theorem 2.3.6). Thus we have a system of incongruences:

$$x \not\equiv a_{ij} \mod p_j,$$

for some a_{ij}.

Here we are in a special situation of the Sieve problem. The number of primes with respect to which we sift the sequence is very small, namely we sift only by the prime factors of N, of which there can be at most $\log N$. Hence we shall re-do the analysis of the Brun sieve and thereby get a better estimate.
Let $\mathcal{A} = \{n \leq x \mid n \equiv 0 \mod H\}$, $P = \prod_{1 \leq i \leq r} p_i$ and let

$$\mathcal{A}_{p_j} = \left\{ n \in A \mid \prod_{1 \leq i \leq h} (n - a_{ij}) \equiv 0 \mod p_j \right\}.$$

We extend the notation to $\mathcal{A}_d$ for d a divisor of P.
We have that

$$P(H; x; p_1, \cdots, p_r) = \sum_{d \backslash P} \mu(d) |A_d|.$$

For $|\mathcal{A}_{p_j}|$, we can select $\omega(p) = \frac{hx}{Hp_j}$, and $R_{p_j} \leq h$ since for each congruence there is an error of at most 1 in the approximation. The denominator H can be taken out of our analysis if we set $x \leftarrow \frac{x}{H}$. We also have that $R_d \leq \omega(d)$.

Hence

$$W(k) = \prod_{1 \leq i \leq k} \left(1 - \frac{h}{p_i}\right).$$

From our earlier work in section (2), we have

$$P(H; x, p_1, \cdots, p_r) > \frac{xW(p_r)}{H}(1 + \Theta) + R,$$

where

$$\Theta = 1 - \sum_{i \leq r} \frac{\omega(p_i)}{p_i} \frac{W(p_i)}{W(p_r)} \sum_{t \backslash P_{(i \cdots r)}} \frac{\chi(t)(1 - \chi(pt))}{t} \omega(t)$$

and $P_{(i \cdots r)} = \prod_{i < k \leq r} p_k$. We let $1 \leq r_t \leq r_{t-1} \leq \cdots \leq r_0 = r$, be a sequence of integers. These correspond to the real numbers z_i, but here we select the indices of the primes instead. We use the function $\chi \equiv \chi_2$ (in the proof of Theorem 2.2.2), with $P_{(r_i \cdots r)}$ instead of $P_{(z_n, z)}$ in the definition.

We will show that in this case we can select the intervals (r_i) such that $\Theta < 1$.

Following the same argument as in Section 2 (with $b = 1$), we arrive at the following upper bound for Θ:

$$\sum_{1 \leq n \leq t} \frac{W(r_n)}{W(r)} \frac{1}{(2n+1)!} \left(\sum_{r_n \leq i \leq r} \frac{\omega(p_i)}{p_i} \right)^{2n+1}.$$

We will show later that we can pick r_i such that

$$\frac{W(r_n)}{W(r)} = \frac{1}{\prod_{r_n \leq i \leq r} (1 - \frac{h}{p_i})} \leq e^{n\gamma},$$

where $\gamma = \log 1.3$. As before

$$\sum_{r_n \leq i \leq r} \frac{\omega(p_i)}{p_i} \leq \log \left(\frac{W(r_n)}{W(r)} \right) \leq n\gamma.$$

So the bound for Θ is

$$\begin{aligned}\sum_{1\le n\le t}\frac{e^{n\gamma}}{(2n+1)!}(n\gamma)^{2n+1} &= \sum_{1\le n\le t}\frac{(ne^{-1})^{2n+1}}{(2n+1)!}e^{2n+1}\gamma^{2n+1}e^{n\gamma}\\ &\le \frac{1}{e^3(3!)}\sum_{1\le n\le t}(\gamma e^{1+\gamma})^{2n+1}\end{aligned}$$

(since $\frac{(ne^{-1})^{2n+1}}{(2n+1)!}$ is decreasing)

$$\begin{aligned}&\le \frac{1}{e^3(3!)}\gamma e^{1+\gamma}\left(\sum_{1\le n<\infty}(\gamma e^{1+\gamma})^{2n}\right)\\ &= \frac{1}{e^3(3!)}\gamma e^{1+\gamma}\frac{1}{1-(\gamma e^{1+\gamma})^2}.\end{aligned}$$

The last step follows because $\gamma e^{1+\gamma} < 1$. The final expression is $\approx 0.05478 < 1$. Thus $\Theta < 1$.

Let us define the intervals by selecting r_i (for $1 \le i \le t$), as the *least* index such that

$$\pi_i = \prod_{r_i<k\le r_{i-1}}\left(1-\frac{h}{p_k}\right) \ge \frac{1}{1.3}.$$

Since $p_i > 10h$ this is always possible. This automatically satisfies the requirements set earlier on γ. Select t such that

$$\pi_t = \prod_{1\le k\le r_{t-1}}\left(1-\frac{h}{p_k}\right) \ge \frac{1}{1.3}.$$

Since $p_i > 10h$ we have

$$1-\frac{h}{p_i} > 1-\frac{h}{10h} = \frac{9}{10}$$

so

$$\begin{aligned}\frac{9}{10}\pi_i &= \left(1-\frac{h}{10h}\right)\pi_i\\ &< \left(1-\frac{h}{p_{r_i}}\right)\pi_i,\end{aligned}$$

which by the definition of r_i is such that

$$< \frac{1}{1.3}.$$

Thus

$$\pi_i \le \frac{10}{9}\frac{1}{1.3} = \frac{1}{1.17} < \frac{8}{9}.$$

We will show that

$$\log\prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right) > \frac{-h\log\log N}{e\log eh} > -0.2h\log\log N.$$

Using the series expansion of $\log(1+x)$ we see that

$$\begin{aligned}\log \prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right)+\log \prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right)^{-h} &\ge -\sum_{1\le i\le r}\sum_{2\le m}\frac{1}{m}\left(\frac{h}{p_i}\right)^m\\ &\ge \sum_{1\le i\le r}\frac{1}{2}\sum_{1\le m}\left(\frac{h}{p_i}\right)^m\\ &= -\frac{1}{2}\sum_{1\le i\le r}\left(\frac{h}{p_i}\right)^2\left(\frac{1}{1-\frac{h}{p_i}}\right).\end{aligned}$$

We need a good bound on $\sum_i \frac{1}{p_i^2}$. We have by [**RS62**] (p.87), that

$$\sum_{x<p}\frac{1}{p^n}\le \frac{1.02n}{x^{n-1}\ln x}.$$

Using this with $n=2$ and $x=10h$, (all the primes $p_i>10h$ by our choice) we have

$$\sum_{1\le i\le r}\frac{1}{p_i^2}\le \frac{2.04}{10h\log 10h}.$$

Thus

$$\begin{aligned}-\frac{1}{2}\sum_{1\le i\le r}\frac{1}{1-\frac{h}{p_i}}\left(\frac{h}{p_i}\right)^2 &\ge -\frac{5}{9}h^2\sum_{1\le i\le r}\frac{1}{p_i^2}\\ &\ge -\frac{0.2h}{\log 10h}.\end{aligned}$$

Now if we can bound from above

$$\log \prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right)^{-h},$$

then we can obtain a lower bound on $\log\prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right)$.

Let $N'=\frac{N}{\gcd(H,N)}$. We have

$$\frac{A}{\varphi(A)}\frac{1}{\prod_{1\le i\le r}\left(1-\frac{1}{p_i}\right)}=\frac{AN'}{\varphi(AN')}.$$

By [**RS62**] Theorem (15): For $n\ge 3$

$$\frac{n}{\varphi(n)}<e^\gamma\log\log n+\frac{5}{2\log\log n},$$

where γ is the Euler constant. Also by [**RS62**] Theorem (9): $\log H<11h<0.1\log N$. Using this we have

$$\begin{aligned}\frac{HN'}{\varphi(HN')} &< e^\gamma\log\log HN'+\frac{2.51}{\log\log HN'}\\ &< e^\gamma\log\log N'+\frac{e^\gamma}{10}+\frac{2.51}{5}\\ &< e^\gamma(\log\log N+0.4),\end{aligned}$$

as $HN'\ge N$, $\log\log N>5$ by our conditions, and also $N'\le N$.

Now by [**RS62**], where a lower bound of $\frac{e^{-\gamma}}{\log x}\left(1-\frac{1}{\log^2 x}\right)$ for $\prod_{p\le x}\frac{1}{1-\frac{1}{p}}$ is given, we have:

$$\frac{H}{\varphi(H)} > e^{\gamma}\log 10h\left(1-\frac{1}{2\log^2 10h}\right) > e^{\gamma}(\log h + 2.1).$$

Since $\log\log N > \log 10h$,

$$\prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right)^{-1} < \frac{1}{e^{\gamma}(\log h+2.1)}\left\{\ e^{\gamma}\log\log N + 0.4\right\}$$

yielding

$$-h\log\prod_{1\le i\le r}\left(1-\frac{1}{p_i}\right) < h\left(\log(\log\log N+0.4) - \log(\log h+2.1) + \frac{0.2}{\log 10h}\right)$$

and finally

$$\log\prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right) > -h\big(\log\log\log N - \log\log eh\big).$$

Using $\log x - \log a = 1 + \log\left(\frac{x}{ae}\right) \le \frac{x}{ae}$, we have

$$\log\prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right) > \frac{-h\log\log N}{e\log eh}.$$

Since $\pi_i \le \frac{1}{1.17}$, we obtain

$$(t-1)\log 1.17 \le \log\prod_{1\le i\le r}\left(1-\frac{h}{p_i}\right)^{-1} \le \frac{h\log\log N}{e\log eh} < \frac{h\log\log N}{e\log(h+1)}.$$

This yields

$$\begin{aligned}(2t+1)\log(h+1) &< 3\log(h+1) + \frac{2h\log\log N}{e\log 1.17}\\ &< 3\log(h+1) + 4.7h\log\log N.\end{aligned}$$

Now $p_i > i\log i$, by [**RS62**] (Corollary to Theorem 3). Hence

$$\begin{aligned}\log\pi_i &= \sum_{r_n<i\le r_{n-1}}\log\left(1-\frac{h}{p_i}\right)\\ &> -\frac{10}{9}\sum_{r_n<i\le r_{n-1}}\frac{h}{p_s}\\ &> -\frac{10h}{9}\int_{r_n}^{r_{n-1}}\frac{dt}{t\log t}\\ &= -\frac{10h}{9}\log\frac{\log r_{n-1}}{\log r_n}.\end{aligned}$$

Since $\pi_i \le \frac{1}{1.17}$, we have

$$\begin{aligned}\frac{\log r_n}{\log r_{n-1}} &< \left(\frac{1}{1.17}\right)^{\frac{9}{10h}}\\ &< \left(1+\frac{9}{10h}\log 1.17\right)^{-1}\\ &\le (1+0.141h^{-1})^{-1},\end{aligned}$$

and so

$$\frac{\log r_n}{\log r} < (1 + 0.141h^{-1})^{-n}$$

for $1 \le n \le t-1$. Further

$$\log N \ge \sum_{1\le i\le r} \log p_i > r\log 10h \ge r\log 20,$$

so $\log r < \log\log N - 1$.

Now for the remainder term:

$$R = \sum_{d\backslash P} \chi(d)|R_d| \le \big(1 + \sum_{1\le i\le r} \omega(p_i)\big) \prod_{1\le i\le t-1} \big(1 + \sum_{j\le r_i} \omega(p_j)\big)^2$$

(since $\omega(p) = h$)

$$\le (1+hr) \prod_{1\le i\le t-1} (1+hr_i)^2.$$

Thus

$$\begin{aligned}
\log R &\le \log(1+h) + \log r + 2(t-1)\log(h+1) + 2\sum_{1\le i\le t-1} \log r_i \\
&= (2t-1)\log(h+1) + \log r + 2\sum_{1\le i\le t-1} \log r_i \\
&< 3\log(h+1) + 4.7h\log\log N + (\log\log N - 1)\left(2\sum_{0\le n}(1+0.141h^{-1})^{-n} - 1\right) \\
&< 3\log(h+1) + 4.7h\log\log N + (\log\log N - 1)(14.2h+1) \\
&< 19.4h\log\log N - 11h - 1.
\end{aligned}$$

Since $\log H < 11h$, we have

$$\log R < 19.4h\log\log N - \log H - 1,$$

and

$$\log\left(\frac{c(\log N)^{20h}}{H} \prod_{1\le i\le r}\left(1 - \frac{h}{p_i}\right)\right) > \log R,$$

where $c = 1 - 0.05478$. Thus $P(H, (\log N)^{20h}, p_1, \cdots, p_r) > 0$. Thus there is an integer D satisfying the conditions of the theorem. $\square$

2.5. Smooth Numbers

Here we illustrate the surprising power of the indentity proved in Proposition 2.2.1.

Let $P_z^x = \{p \mid z \le p < x\}$. Then setting $\chi(1) = 1$ and $\chi(d) = 1$ for $d > 1$ in Proposition 2.2.1, we have that for $2 \le z_1 \le z$:

$$S(\mathcal{A}^x; P_{z_1}^x) = S(\mathcal{A}^x; P_z^x) - \sum_{z_1\le p<z} S(\mathcal{A}_p; P_p^x).$$

Recall that $S(\mathcal{A}^x; P_z^x) = \Psi(x,z)$ the number z-smooth integers below x, also $S(\mathcal{A}_p; P_p^x) = \Psi(\frac{x}{p}, p)$.

Hence we have, for $2 \le z_1 \le z$ that

$$\Psi(x,z) = \Psi(x,z_1) + \sum_{z_1 \le p < z} \Psi\left(\frac{x}{p}, p\right). \tag{2.19}$$

As an application we show the following theorem.

THEOREM 2.5.1 ([**Hal70**]). *Let $y = x^{\frac{1}{\theta}}$ where $1 < \theta \le 2$. Then*

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

Proof :

Applying the identity (2.19) with $z = x$ and $z_1 = y$, we have

$$\Psi(x,y) = \Psi(x,x) - \sum_{y \le p < x} \Psi\left(\frac{x}{p}, p\right). \tag{2.20}$$

Now $\Psi(x,x) = \lfloor x \rfloor$. Since $1 < \theta \le 2$, $p \ge \sqrt{x}$, we have that $\frac{x}{p} \le \sqrt{x} \le p$. Consequently, $\Psi\left(\frac{x}{p}, p\right) = \left\lfloor \frac{x}{p} \right\rfloor$. Substituting in (2.20), we have

$$\begin{aligned}
\Psi(x,y) &= \lfloor x \rfloor - \sum_{y \le p < x} \left\lfloor \frac{x}{p} \right\rfloor \\
&= x - x \sum_{y \le p < x} \frac{1}{p} + O(\pi(x)) \\
&= x\left\{1 - \log\log x + \log\log y + O\left(\frac{1}{\log x}\right)\right\}.
\end{aligned}$$

Now $x \ge y^{\theta}$, so $\log x \ge \theta \log y$, and also $\log\log x \ge \log\theta + \log\log y$ this yields

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

□

The recurrence formula can be used to convert upper bounds to other useful lower bounds, and can also be used iteratively. Here is a simple example.

Let us try to evaluate $\Psi(x, x^{\frac{1}{\delta}})$ for $2 < \delta < e$ using the recurrence formula

$$\Psi(x, x^{\frac{1}{\delta}}) = \Psi(x, x^{\frac{1}{2}}) - \sum_{x^{\frac{1}{\delta}} \le p \le x^{\frac{1}{2}}} \Psi\left(\frac{x}{p}, p\right).$$

Applying the trivial bound $\Psi\left(\frac{x}{p}, p\right) \le \frac{x}{p}$

$$\begin{aligned}
\sum_{x^{\frac{1}{\delta}} \le p \le x^{\frac{1}{2}}} \Psi\left(\frac{x}{p}, p\right) &\le x \sum_{x^{\frac{1}{\delta}} \le p \le x^{\frac{1}{2}}} \frac{1}{p} \\
&= x(\log\delta - \log 2).
\end{aligned}$$

Now applying the theorem with $\theta = 2$, we have

$$\Psi(x, x^{\frac{1}{2}}) = x\left(1 - \log 2 + O\left(\frac{1}{\log x}\right)\right).$$

Thus we obtain

$$\Psi(x, x^{\frac{1}{\delta}}) \ge x\left(1 - \log\delta + O\left(\frac{1}{\log x}\right)\right).$$

Of course, in this case we could have directly derived this result as in the theorem, but this just is an illustration of the usage of Buchstab's identity. In estimating $\psi(x,y)$ we could try to use Brun's sieve as in section (1.3). It is clear however, that to obtain a good estimate we need to take $\ln z < \epsilon \ln x$, but this would make the error term very large, since that depends on the size of the interval $x - z$.

2.6. On the number of integers prime to a given number

Let $k > 1$ be an integer and $x > 1$ a real number, here we will find bounds for the sum:

$$\sum_{\substack{n \le x \\ \gcd(n,k)=1}} 1.$$

It is clear that in every interval $\mod k$ there are $\varphi(k)$ such integers. However, it is not clear how uniform the distribution of these numbers are inside the interval.

The sequence to be sifted is $\mathcal{A} = \{n \mid 1 \le n \le x\}$, and the sifting primes are $P = \{p \mid p \backslash k\}$. We assume $x \ge k$. In this case we can take $|\mathcal{A}_d| = \frac{x}{d} + R_d$, where $\omega(d) = 1$ and $R_d \le 1$. Now, $\frac{1}{1-\frac{1}{p}} \le 2$. Hence $A = 2$, we also have

$$\sum_{\substack{w \le p \le z \\ p \in P}} \frac{\omega(p) \ln p}{p} \le \ln\left(\frac{\ln z}{\ln w}\right) + \frac{1}{\ln w}$$

thus $\kappa = \eta = 1$.
To apply the lower bound estimate of the Brun sieve (with $b = 1$), we need to find λ such that

$$1 - \frac{2(\lambda e^{\lambda})^2}{1 + (\lambda e^{1+\lambda})^2} > 0$$

and

$$1 + \frac{2.01}{e^{2\lambda} - 1} < \gamma,$$

where we have used $\xi = 1.005$. It turns out that we can take $\gamma < 5$, and satisfy both the constraints for $\lambda = 0.204$. This gives

$$S(\mathcal{A}; P, z) \ge x W(z)\big(1 - o(1)\big) + O\big(z^{4.85}\big).$$

Taking $z = x^{\frac{1}{5}}$, we obtain

$$S(\mathcal{A}; P, z) \ge c \prod_{p \backslash k} \left(1 - \frac{1}{p}\right) x + O\big(x^{0.97}\big).$$

Now to get the actual estimate $\sum_{n \le x, n \perp k} 1$ we need to account for the numbers that might have been included in this estimate which are not really prime to k. Clearly, by our choice of the limit for z, each number which is over-counted must share a factor p with k that is larger than $x^{\frac{1}{5}}$. Let us assume that the largest prime factor of k is $< x^{\frac{1}{5}}$.
Thus we have:

$$\sum_{\substack{n \le x \\ \gcd(n,k)=1}} 1 \ge \frac{c\varphi(k)}{k} x + O\big(x^{0.97}\big),$$

where $c < 1$.
For the upper bound we can take the same value of λ as for the lower bound but this forces us to take $z = x^{\frac{1}{6}}$ in this case and we get

$$\sum_{\substack{n \le x \\ \gcd(n,k)=1}} 1 \le \frac{c'\varphi(k)}{k} x + O\big(x^{0.975}\big),$$

where $c' < 4$.
In summary we have proved:

THEOREM 2.6.1. *Let* $x > 0$ *and* k *a positive integer whose largest prime factor* p *is less than* $x^{\frac{1}{5}}$. *Then*

$$\frac{c\varphi(k)}{k}x + O\left(x^{0.97}\right) \leq \sum_{\substack{n \leq x \\ \gcd(n,k)=1}} 1 \leq \frac{c'\varphi(k)}{k}x + O\left(x^{0.975}\right),$$

where $c < 1$ *and* $c' < 4$ *are constants.* $\square$

CHAPTER 3

Selberg's Sieve

Around 1946 Atle Selberg introduced a new method for finding upper bounds to the sieve estimate [**Sel47**]. The method usually gives much better bounds than the Brun's sieve. To obtain lower bounds one can couple the Selberg sieve with the Buchstab identities. After developing the basic ideas of this sieve technique, we shall look at the most important application of this method - to derive inequalities of the Brun-Titchmarsh type.

3.1. The Selberg upper-bound method

Selberg's method of estimating the sum

$$S(\mathcal{A}; P^z, x) = \sum_{a\in\mathcal{A}} \Bigg(\sum_{d \backslash \gcd(a,P_z)} \mu(d) \Bigg)$$

relies on finding a sequence of numbers λ_d such that $\lambda_1 = 1$ and using the inequality:

$$S(\mathcal{A}; P^z, x) \le \sum_{a\in\mathcal{A}} \Bigg(\sum_{d \backslash \gcd(a,P_z)} \lambda_d \Bigg)^2.$$

This allows us complete freedom in our choice of the numbers λ_d for $d > 1$, and the idea of this method is to select the λ_d such that the sum is minimized. Note that setting $\lambda_1 = 1$ and $\lambda_d = 0$ for $d > 1$, leads to the trivial estimate $S(\mathcal{A}; P^z, x) \le |\mathcal{A}^x|$. Selberg's method relies on choices of λ_d that mimic the cancellation occuring in the sum $\sum_{d\backslash n} \mu(d)$. Such choices lead to better estimates when we interchange the sum.

Now

$$\sum_{a\in\mathcal{A}^x} \Bigg(\sum_{d \backslash \gcd(a,P_z)} \lambda_d \Bigg)^2 = \sum_{\substack{d_i \backslash P_z \\ i=1,2}} \lambda_{d_1}\lambda_{d_2} \Bigg(\sum_{\substack{a\in\mathcal{A}^x \\ a\equiv 0 \mod D}} 1 \Bigg),$$

where $D = \mathrm{lcm}(d_1, d_2)$. By our conventions about the sequence $\mathcal{A}$, we have

$$\sum_{\substack{a\in\mathcal{A}^x \\ a\equiv 0 \mod D}} 1 = |\mathcal{A}_D^x| = \frac{\omega(D)}{D} x + R_D.$$

This yields,

$$\begin{aligned} \sum_{\substack{d_i \backslash P_z \\ i=1,2}} \lambda_{d_1}\lambda_{d_2} |\mathcal{A}_D^x|, &= x \sum_{\substack{d_i \backslash P_z \\ i=1,2}} \lambda_{d_1}\lambda_{d_2} \frac{\omega(D)}{D} + \sum_{\substack{d_i \backslash P_z \\ i=1,2}} \lambda_{d_1}\lambda_{d_2} |R_D| \\ &= x\Sigma_1 + \Sigma_2. \end{aligned}$$

The problem of selecting λ_d already seems difficult. We can make the assumption that $\lambda_d = 0$ for $d > z$ and hope that since the second sum σ_2 contains only z^2 terms we can concentrate on minimizing the leading sum σ_1. Our first effort will be directed towards this.

Minimization of $\sum_1$: Using the fact that $\omega(d)$ is a multiplicative function, we have

$$\frac{\omega(D)}{D}=\frac{\omega(d_1)\omega(d_2)}{\omega(\gcd(d_1,d_2))}\frac{\gcd(d_1,d_2)}{d_1d_2},$$

so

$$\Sigma_1=\sum_{d_i\backslash P_z}\lambda_{d_1}\lambda_{d_2}\frac{\omega(d_1)}{d_1}\frac{\omega(d_2)}{d_2}\frac{\gcd(d_1,d_2)}{d_1d_2}.$$

Let $f(d)=\frac{\omega(d)}{d}$, so that the sum becomes

$$\Sigma_1=\sum_{d_i\backslash P_z}\lambda_{d_1}\lambda_{d_2}\frac{f(d_1)f(d_2)}{f(d)},\tag{3.21}$$

where $d=\gcd(d_1,d_2)$.

We need to get rid of the term in the denominator, and to this end we introduce the function

$$J(r)=\frac{1}{f(r)}\prod_{p\backslash r}\big(1-f(p)\big).$$

Let $r=ps$, and consider:

$$\begin{aligned}\sum_{\delta\backslash ps}J(\delta)&=\sum_{\delta\backslash s}\Big\{J(p\delta)+J(\delta)\Big\}\\&=\sum_{\delta\backslash s}\Bigg(\frac{1}{f(p\delta)}\prod_{q\backslash p\delta}\big(1-f(q)\big)+\frac{1}{f(\delta)}\prod_{q\backslash\delta}\big(1-f(q)\big)\Bigg)\\&=\sum_{\delta\backslash s}J(\delta)\Big\{\frac{1}{f(p)}\big(1-f(p)\big)+1\Big\}\\&=\frac{1}{f(p)}\sum_{\delta\backslash s}J(\delta),\end{aligned}$$

together with

$$\sum_{\delta\backslash p}J(\delta)=J(p)+J(1)=\frac{1}{f(p)}.$$

Thus we have

$$\frac{1}{f(d)}=\sum_{\delta\backslash d}J(d).$$

Substituting this for $\frac{1}{f(d)}$ in (3.21) we get,

$$\begin{aligned}\sum_{d_i\backslash P_z}\lambda_{d_1}\lambda_{d_2}\frac{f(d_1)f(d_2)}{f(d)}&=\sum_{d_i\backslash P_z}\lambda_{d_1}\lambda_{d_2}f(d_1)f(d_2)\sum_{\delta\backslash d_1,\delta\backslash d_2}J(d)\\&=\sum_{\substack{r\le z\\ r\backslash P_z}}J(r)\Big\{\sum_{\substack{r\backslash d\\ d\le z}}\lambda_d f(d)\Big\}^2.\end{aligned}$$

Let $\xi_r=\sum_{\substack{r\backslash d\\ d\le z}}\lambda_d f(d)$, so that

$$\Sigma_1=\sum_{\substack{r\le z\\ r\backslash P_z}}J(r)\xi_r^2.$$

This is what we need to minimize subject to the restriction $\lambda_1 = 1$. We wish to write this constraint as a constraint among the variables ξ_i, which would allow us to convert the minimization problem to one entirely involving the variables ξ_i.
The idea is to use Möbius inversion to pick out λ_1, and this is not difficult:

$$\begin{aligned}\sum_{r\le z}\mu(r)\xi_r &= \sum_{r\le z}\mu(r)\sum_{\substack{r\backslash d\\ d\le z}}\lambda_d f(d)\\ &= \sum_{d\le z} f(d)\lambda_d\left(\sum_{r\backslash d}\mu(d)\right)\\ &= \lambda_1 f(1)\\ &= \lambda_1\\ &= 1.\end{aligned}$$

Thus we need to minimize $\sum_{r\le z} J(r)\xi_r^2$, subject to the constraint $\sum_{r\le z}\mu(r)\xi_r = 1$. Let $F = \sum_r J(r)\xi_r^2 - \Delta$ for some real Δ. Since $\sum_{r\le z}\mu(r)\xi_r = 1$, we have $F = \sum_r J(r)\xi_r^2 - \Delta\sum_r \mu(r)\xi_r$. Minimizing F is the same as minimizing the function $\sum_{r\le z} J(r)\xi_r^2$. Let us try to complete the square term in the first sum in F. This suggests setting $\Delta \leftarrow 2\omega$, so

$$\begin{aligned}\sum_{r\le z} J(r)\xi_r^2 - 2\omega\sum_{r\le z}\mu(r)\xi_r &= \sum_{r\le z} J(r)\left\{\xi_r^2 - \frac{2\omega\mu(r)\xi_r}{J(r)}\right\}\\ &= \sum_{r\le z} J(r)\left\{\xi_r^2 - \frac{2\omega\mu(r)\xi_r}{J(r)} + \left(\frac{\omega\mu(r)}{J(r)}\right)^2\right\} - \sum_{r\le z}\frac{\omega^2\mu(r)^2}{J(r)}\\ &= \sum_{r\le z} J(r)\left\{\xi_r - \frac{\omega\mu(r)}{J(r)}\right\}^2 - \omega^2\sum_{r\le z}\frac{\mu^2(r)}{J(r)}.\end{aligned}$$

Thus at the minimum value of F we should have $\xi_r = \frac{\omega\mu(r)}{J(r)}$, and the minimum value of F would be $-\omega^2\sum_{r\le z}\frac{\mu^2(r)}{J(r)}$. To find the value of ω, we can substitute ξ_r into the constraint $\sum_{r\le z}\mu(r)\xi_r = 1$, and this gives us immediately that

$$\omega = \frac{1}{\sum_{r\le z}\frac{\mu(r)^2}{J(r)}}.$$

So

$$\begin{aligned}\min\sum_{r\le z} J(r)\xi_r^2 &= \sum_{r\le z}\omega^2\frac{\mu(r)^2}{J(r)}\\ &= \omega^2\sum_{r\le z}\frac{\mu(r)^2}{J(r)}\\ &= \frac{\omega^2}{\omega}\\ &= \omega\\ &= \frac{1}{\sum_{r\le z}\frac{\mu(r)^2}{J(r)}}.\end{aligned}$$

By our definition of the function $g(d)$ we have $g(r) = \frac{1}{J(r)}$, so

$$\sum_{r\le z}\frac{\mu(r)^2}{J(r)} = \sum_{r\le z}\mu(r)^2 g(r).$$

Set

$$G(z) = \sum_{r \le z} \mu^2(r) g(r).$$

Then the minimum value of Σ_1 is $\frac{x}{G(z)}$.

Evaluation of Σ_2: To estimate the remainder term Σ_2, we need an estimate on the size of the λ_d. We had earlier used Möbius inversion to extract λ_1 from a combination of the ξ_r, and we can repeat the process to get λ_δ for any δ.
Now by definition

$$\xi_r = \sum_{\substack{r \backslash d \\ d \le z}} \lambda_d f(d).$$

Let $r = \gamma\delta$, so that

$$\begin{aligned}
\xi_{\gamma\delta} &= \sum_{\substack{\gamma\delta \backslash d \\ d \le z}} \lambda_d f(d) \\
&= \sum_{\substack{\gamma \backslash \frac{d}{\delta} \\ d \le z}} \lambda_d f(d) \\
&= \sum_{\substack{\gamma \backslash \nu, \nu \le \frac{d}{\delta} \\ \gcd(\nu,\delta)=1}} \lambda_{\delta\nu} f(\delta\nu).
\end{aligned}$$

Since we want to extract the term with $\gamma = 1$, we calculate:

$$\begin{aligned}
\sum_{\substack{\gamma \le \frac{z}{\delta} \\ \gamma \perp \delta}} \mu(\gamma) \xi_{\gamma\delta} &= \sum_{\substack{\gamma \le \frac{z}{\delta} \\ \gamma \perp \delta}} \mu(\gamma) \sum_{\substack{\gamma \backslash \nu, \nu \le \frac{z}{\delta} \\ \nu \perp \delta}} \lambda_{\delta\nu} f(\delta\nu) \\
&= \sum_{\nu \le \frac{z}{\delta}, \nu \perp \delta} \lambda_{\delta\nu} f(\delta\nu) \left\{ \sum_{\gamma \backslash \nu} \mu(k) \right\} \\
&= \lambda_\delta f(\delta).
\end{aligned}$$

Thus

$$\lambda_\delta = \frac{1}{f(\delta)} \sum_{\substack{\gamma \le \frac{z}{\delta} \\ \gamma \perp \delta}} \mu(\gamma) \xi_{\gamma\delta},$$

and substituting for $\xi_{\gamma\delta}$ gives

$$\begin{aligned}
\lambda_\delta &= \frac{\omega}{f(\delta)} \sum_{\substack{\gamma \le \frac{z}{\delta} \\ \gamma \perp \delta}} \frac{\mu(\gamma\delta)\mu(\gamma)}{J(\gamma\delta)} \\
&= \frac{\omega\mu(\delta)}{f(\delta) J(\delta)} \sum_{\substack{\gamma \le \frac{z}{\delta} \\ \gamma \perp \delta}} \frac{\mu(\gamma)^2}{J(\delta)}.
\end{aligned}$$

Let

$$G_d(y) = \sum_{\delta < y, \delta \perp d} \mu^2(\delta) g(\delta).$$

Then

$$\lambda_\delta = \frac{\omega\mu(\delta)}{f(\delta)J(\delta)} G_\delta\left(\frac{z}{\delta}\right). \tag{3.22}$$

We will show that $|\lambda_d| \leq 1$. Observe that

$$\begin{aligned}
G(z) &= \sum_{l\backslash d} \sum_{\substack{m\leq z\\ \gcd(m,d)=l}} \mu(m)^2 g(m)\\
&= \sum_{l\backslash d} \sum_{\substack{h<\frac{z}{l}\\ \gcd(h,l)=1\\ \gcd(h,\frac{d}{l})=1}} \mu(lh)^2 h(lh)\\
&= \sum_{l\backslash d} \mu(l)^2 g(l) G_d\left(\frac{z}{l}\right)\\
&\geq \left(\sum_{l\backslash d} \mu(l)^2 g(l)\right) G_d\left(\frac{z}{d}\right)
\end{aligned}$$

and

$$\begin{aligned}
\sum_{l\backslash d} \mu(l)^2 g(l) &= \prod_{p\backslash d}(1+g(p))\\
&= \prod_{p\backslash d} \frac{p}{p-\omega(p)}\\
&= \frac{1}{\prod_{p\backslash d}(1-\frac{\omega(p)}{p})},
\end{aligned}$$

and so

$$G_d\left(\frac{z}{d}\right) \leq \prod_{p\backslash d}\left(1-\frac{\omega(p)}{p}\right) G(z). \tag{3.23}$$

Now substituting for $J(\delta)$ in (3.22), we get:

$$\lambda_d = \frac{\mu(d)}{\prod_{p\backslash d}(1-\frac{\omega(p)}{p})} \frac{G_d(z/d)}{G(z)}. \tag{3.24}$$

Thus by (3.23) and (3.24), we have $|\lambda_d| \leq 1$.

Now

$$\Sigma_2 \leq \sum_{\substack{d_i<z\\ d_i\backslash P_z}} \left|R_{\mathrm{lcm}(d_1,d_2)}\right|.$$

Fix a d; we can estimate the number of integers d_1, d_2 for which $d = \mathrm{lcm}(d_1, d_2)$. Now d as well as d_1 and d_2 are squarefree. If $d_1 = \prod_i p_i^{e_i}$ and $d_2 = \prod_i p_i^{f_i}$, then $d = \prod_i p_i^{\max\{e_i,f_i\}}$. Suppose $p\backslash d$, then $p\backslash d_1$ or $p\backslash d_2$ or p divides both of them. So the number of integers which can give rise to d as their lcm is exactly $3^{\nu(d)}$.

Using this and the fact that $d < z^2$, we get

$$\Sigma_2 \leq \sum_{d<z^2} 3^{\nu(d)} |R_d|.$$

If we also have the remainder condition $|R_d| \leq \omega(d)$, then we can simplify further:

$$\begin{aligned}
\sum_{d<z^2} 3^{\nu(d)}|R_d| &\le \sum_{\substack{d<z^2\\ d\backslash P_z}} 3^{\nu(d)}\omega(d)\\
&\le z^2 \sum_{d\backslash P_z} \frac{3^{\nu(d)\omega(d)}}{d}\\
&= z^2 \prod_{p<z, p\in P}\left(1+\frac{3\omega(p)}{p}\right)\\
&\le z^2 \prod_{p<z}\left(1+\frac{\omega(p)}{p}\right)^3\\
&\le \frac{z^2}{W^3(z)}.
\end{aligned}$$

Thus we have proved:

Theorem 3.1.1. *If* $|R_d| \le \omega(d)$, *then*

$$S(\mathcal{A}; P^z, x) \le \frac{x}{G(z)} + \frac{z^2}{W^3(z)},$$

where

$$G(z) = \sum_{r\le z} \mu^2(r) g(r).$$

The second term can also be upper bounded by

$$\sum_{\substack{d<z^2\\ d\backslash P_z}} 3^{\nu(d)}|R_d|,$$

which is also upper bounded by

$$\sum_{\substack{d<z^2\\ \Gamma(d)\subseteq P}} \mu^2(d) 3^{\nu(d)}|R_d|.$$

Here $\Gamma(d)$ stands for the set of prime divisors of d. □

We will apply the Selberg method to the simple but important case where $\omega(d) = 1$ and $|R_d| \le 1$.

Theorem 3.1.2. *Suppose* $\omega(d) = 1$ *and* $|R_d| \le 1$. *If* d *is squarefree and* $p \notin P \Rightarrow p \perp d$ *then*

$$S(\mathcal{A}; P, z) \le \frac{x}{\prod_{\substack{p<z\\ p\notin P}}\left(1-\frac{1}{p}\right)\log z} + z^2.$$

Proof : Recall that

$$g(d) = \frac{\omega(d)}{d\prod_{p\backslash d}\left(1-\frac{\omega(p)}{p}\right)}$$

where $d\backslash P_z$. In this case we have $\omega(d) = 1$, so we have

$$g(d) = \frac{1}{\varphi(d)}.$$

Let $k = \prod_{\substack{p<z \\ p \notin P}} p$. Then by definition of $G(z)$ in this case we get

$$G(z) = \sum_{\substack{d<z \\ d \perp k}} \frac{\mu^2(d)}{\varphi(d)}.$$

Let

$$S_k(z) = \sum_{\substack{d<z \\ d \perp k}} \frac{\mu^2(d)}{\varphi(d)}.$$

Then

$$\begin{aligned}
S_1(z) &= \sum_{d<z} \frac{\mu^2(d)}{\varphi(d)} \\
&= \sum_{l \backslash k} \sum_{\substack{d<z \\ \gcd(d,k)=l}} \frac{\mu^2(d)}{\varphi(d)} \\
&= \sum_{l \backslash k} \sum_{\substack{h<\frac{z}{l} \\ \gcd(h,k/l)=1 \\ \gcd(h,l)=1}} \frac{\mu^2(lh)}{\varphi(lh)} \\
&= \sum_{l \backslash k} \frac{\mu^2(l)}{\varphi(l)} \sum_{\substack{h<\frac{z}{l} \\ h \perp k}} \frac{\mu^2(h)}{\varphi(h)} \\
&= \sum_{l \backslash k} \frac{\mu^2(l)}{\varphi(l)} S_k\left(\frac{z}{l}\right) \\
&\leq \sum_{l \backslash k} \frac{\mu^2(l)}{\varphi(l)} S_k(z),
\end{aligned}$$

because $S_k(z)$ is an increasing function of z.

Now

$$\begin{aligned}
\sum_{l \backslash k} \frac{\mu^2(l)}{\varphi(l)} &= \prod_{p \backslash k} \left(1 + \frac{1}{p-1}\right) \\
&= \frac{1}{\prod_{p \backslash k}(1 - \frac{1}{p})} \\
&= \frac{k}{\varphi(k)},
\end{aligned}$$

and so

$$S_k(z) \geq \frac{\varphi(k)}{k} S_1(z).$$

To apply Theorem 3.1.1 we need a good lower bound on $G(z)$. Since $G(z) = S_k(z)$, the above derivation says that we can translate a lower bound on $S_1(z)$ to a lower bound on $S_k(z)$.

We have

$$S_1(z) = \sum_{d<z} \frac{\mu^2(d)}{d} \frac{1}{\prod_{p\backslash d}\left(1-\frac{1}{p}\right)}$$
$$= \sum_{d<x} \frac{\mu^2(d)}{d} \prod_{p\backslash d}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right).$$

If we set $\blacksquare(n)$ to be the largest squarefree divisor of n, then

$$S_1(z) = \sum_{\blacksquare(n)<z} \frac{1}{n}$$
$$\geq \sum_{n<z} \frac{1}{n}$$
$$\geq \log z.$$

So $S_k(z) \geq \frac{\varphi(k)}{k} \log z$. We know from the proof of Theorem (3.1.1) that the remainder term is at most

$$\sum_{\substack{d_i \backslash P_z \\ d_i<z}} |R_{\mathrm{lcm}(d_1,d_2)}| \leq \left(\sum_{d<z} \mu^2(d)\right)^2$$
$$< z^2.$$

Thus

$$S(\mathcal{A}; P^z, x) \leq \frac{1}{\prod_{\substack{p<z \\ p\notin P}}(1-\frac{1}{p}) \log z} x + z^2.$$

□

3.2. The Brun-Titchmarsh Theorem

The prime number theorem for arithmetic progressions states that

$$\pi(x; l, k) = \frac{\mathrm{li} x}{\varphi(k)} + O\left(x e^{-A\sqrt{\log x}}\right)$$

uniformly for $k \leq (\log x)^B$, where B is any positive constant and A is a positive constant depending only on B. This is a very narrow range of values of k. It turns out that if we assume the Extended Riemann Hypothesis, then

$$\pi(x; l, k) = \frac{\mathrm{li} x}{\varphi(k)} + O\left(\sqrt{x} \log x\right)$$

uniformly for $k \leq \frac{\sqrt{x}}{\log^2 x}$. By a careful analysis of the Selberg sieve (especially the remainder term) van Lint and Richert [**vLR65**] showed a good upper bound for $\pi(x; l, k)$ valid for any $k < x$. In this section we shall look at the proof of this result (see Theorem 3.2.5). In a later chapter we shall improve on this result using the so called *Large sieve*.

Let $k, l > 0$ be relatively prime integers, and let $x, y > 1$ be reals with $y \leq x$. We will concentrate on the sequence

$$\mathcal{A} = \{n \mid x - y < n \leq x,\ n \equiv l \mod k\}.$$

For K a multiple of k, we take as the sifting primes

$$P_K = \{p \mid p \not| K\}.$$

First we shall prove a form of the Selberg sieve, where we have a better estimate of the remainder term. We define

$$S_K(z) = \sum_{\substack{1 \le n \le z \\ n \perp K}} \frac{\mu^2(n)}{\varphi(n)}$$

as in the proof of Theorem (3.1.2), and

$$H_K(z) = \sum_{\substack{1 \le n \le x \\ n \perp K}} \mu^2(n) \frac{\sigma(n)}{\varphi(n)}$$

with $\sigma(n) = \sum_{d \backslash n} d$.

LEMMA 3.2.1.

$$S(\mathcal{A}; P_K^z, x, y) \le \frac{y}{kS_K(z)} + \frac{H_K^2(z)}{S_K^2(z)}.$$

Proof : The cardinality of the set

$$\mathcal{A}_D = \{n \mid x - y < n \le x, n \equiv l \mod k, n \equiv 0 \mod D\}$$

is $\frac{y}{kD} + R_D$. Following the proof of the Selberg sieve and using the analysis in Theorem (3.1.2) we get the first term to be

$$\frac{y}{kS_K(z)}.$$

Now the remainder term is (using $|R_d| \le 1$) at most

$$\sum_{\substack{d_i \backslash P_K \\ i=1,2}} |\lambda_{d_1} \lambda_{d_2}| = \left(\sum_{d \backslash P_K} |\lambda_d| \right)^2.$$

In the notation of this proof we have

$$\lambda_d = \mu(d) \frac{d}{\varphi(d)} \frac{S_{Kd}\left(\frac{z}{d}\right)}{S_K(z)}$$

so

$$\begin{aligned}
\sum_{d \backslash P_K} |\lambda_d| &= \sum_{\substack{1 \le d \le z \\ d \perp K}} \frac{\mu^2(d) d}{\varphi(d)} \frac{1}{S_K(z)} \sum_{\substack{1 \le m \le z/d \\ m \perp Kd}} \frac{\mu^2(m)}{\varphi(m)} \\
&= \frac{1}{S_K(z)} \sum_{\substack{1 \le d \le z \\ d \perp K}} \frac{\mu^2(d)}{\varphi(d)} \left(\sum_{\substack{1 \le m \le z/d \\ m \perp kd}} \frac{\mu^2(m)}{\varphi(m)} \right) \\
&= \frac{1}{S_K(z)} \sum_{\substack{1 \le d \le z \\ d \perp K}} \sum_{\substack{1 \le m \le z/d \\ m \perp kd}} \frac{\mu^2(md)}{\varphi(md)} d \\
&= \frac{1}{S_K(z)} \sum_{\substack{1 \le n \le z \\ n \perp K}} \frac{\mu^2(n)}{\varphi(n)} \sum_{d \backslash n} d \\
&= \frac{H_K(z)}{S_K(z)}.
\end{aligned}$$

Hence the remainder term is at most $\frac{H_K^2(z)}{S_K^2(z)}$, and the lemma follows. □

Our aim now is to find a good upper bound on $H_K^2(z)$. One idea is to use Cauchy's inequality on this sum, and this suggests that we first find a concrete upper bound for the sum $\sum_{n\leq x, n\perp K} 1$, which we have seen in the last chapter.
Using Theorem (3.1.2) we have

THEOREM 3.2.2. *If $1 \leq k < y \leq x$ and P is a set of primes p with $k \perp p$, then we have for any $z \geq 2$ that*

$$\left|\{n \mid x-y < n \leq x, n \equiv l \mod k, n \perp P_z\}\right| \leq \frac{y}{\prod_{\substack{p<z\\ p\notin P}} k \log z} + z^2.$$

LEMMA 3.2.3. *Let $p(k)$ be the largest prime divisor of k. For $x \geq e^6$ and $p(k) \leq x$ we have*

$$\sum_{\substack{n\leq x\\ n\perp K}} 1 < \frac{7\varphi(k)}{k} x.$$

Proof : Take $k = 1, y = x$ and $P = \{p \mid p \not| k\}$ in Theorem (3.2.2). For $z \leq x$ we have

$$\Phi_K(x) = \left|\{n \ : \ n \leq x, \gcd(n, \prod_{p<z, p\perp K} p) = 1\}\right| \leq \frac{x}{\prod_{\substack{p<z\\ p\backslash K}} (1-\frac{1}{p}) \log z} + z^2.$$

Thus

$$\frac{k}{\varphi(k)} \frac{\Phi_K(x)}{x} \leq \frac{1}{\prod_{p\leq x}(1-\frac{1}{p})} \left(\frac{1}{\log z} + \frac{z^2}{x}\right),$$

and using

$$\prod_{p\leq x} \left(1-\frac{1}{p}\right)^{-1} \leq e^{\gamma} \log x \left\{1 + \frac{1}{2\log^2 x}\right\}$$

and setting $z = x^{\frac{1}{3}}$, we get

$$\frac{k}{\varphi(k)} \frac{\Phi_K(x)}{x} \leq e^{\gamma} \log x \left(1 + \frac{1}{2\log^2 x}\right) \left(\frac{3}{\log x} + \frac{1}{x^{\frac{1}{3}}}\right).$$

The right hand side is decreasing, and for $x = e^6$ is < 7. □

LEMMA 3.2.4. *For $z > 10^3$, h even,*

$$\frac{H_h^2(z)}{S_h^2(z)} < 22.5 \frac{h}{\varphi(h)} \frac{z^2}{\log^2 z}.$$

Proof : Let

$$J_h(z) = \sum_{\substack{1\leq n\leq z\\ n\perp h}} \mu^2(n) \frac{\sigma^2(n)}{\varphi^2(n)},$$

and as above let $\Phi_h(z) = \sum_{\substack{1\leq n\leq z\\ n\perp h}} 1$. Now

$$H_k(z) = \sum_{\substack{1\leq n\leq z\\ n\perp k}} \mu^2(n) \frac{\sigma(n)}{\varphi(n)}.$$

Cauchy's inequality states that

$$\left(\sum_{1\leq n\leq N} a_n b_n\right)^2 \leq \left(\sum_{1\leq n\leq N} a_n^2\right)\left(\sum_{1\leq n\leq N} b_n^2\right).$$

Using this with $b_n = 1$, $a_n = \mu^2(n)\frac{\sigma(n)}{\varphi(n)}$ and observing that $\mu^4(n) = \mu^2(n)$, we have

$$H_h^2(z) \leq \Phi_h(z) J_h(z).$$

Let n be an integer and $p \perp n$; then

$$\begin{aligned}\sigma(np) &= \sum_{d\backslash np} d \\ &= \sum_{d\backslash n} d + p\sum_{d\backslash n} d \\ &= \sigma(n)(1+p),\end{aligned}$$

and also $\varphi(np) = \varphi(n)\varphi(p)$. If n is squarefree, then

$$\begin{aligned}\frac{\sigma^2(np)}{\varphi^2(np)} &= \frac{\sigma^2(n)}{\varphi^2(n)}\left(\frac{(1+p)^2}{\varphi^2(p)}\right) \\ &= \frac{\sigma^2(n)}{\varphi^2(n)}\left\{\frac{\varphi^2(p)+4p}{\varphi^2(p)}\right\} \\ &= \frac{\sigma^2(n)}{\varphi^2(n)}\left\{1+\frac{4p}{\varphi^2(p)}\right\}.\end{aligned}$$

By induction we have

$$\begin{aligned}\frac{\sigma^2(n)}{\varphi^2(n)} &= \prod_{p\backslash n}\left(1+\frac{4p}{\varphi^2(p)}\right) \\ &= \sum_{d\backslash n}\frac{4^{\nu(d)}d}{\varphi^2(d)}, \mu^2(n)=1.\end{aligned}$$

Since $2\backslash h$ we have $J_h(z) \leq J_2(z)$ and

$$\begin{aligned}J_2(z) &= \sum_{\substack{1\leq n\leq z\\ n\perp 2}} \mu^2(n)\sum_{d\backslash n}\frac{4^{\nu(d)}d}{\varphi^2(d)} \\ &= \sum_{\substack{1\leq d\leq z\\ d\perp 2}} \mu^2(d)\frac{4^{\nu(d)}d}{\varphi^2(d)}\sum_{\substack{1\leq m\leq z/d\\ m\perp 2d}}\mu^2(m) \\ &\leq z\sum_{\substack{1\leq d\leq z\\ d\perp 2}}\frac{\mu^2(d)4^{\nu(d)}}{\varphi^2(d)} \\ &\leq z\prod_{p>2}\left(1+\frac{4}{(p-1)^2}\right) \\ &< \frac{16}{5}z.\end{aligned}$$

In the proof of Theorem (3.1.2) we had proved $S_h(x) \geq \frac{\varphi(h)}{h}\log x$; now using this and Lemma (3.2.3) we have:

$$\begin{aligned}\frac{H_h^2(z)}{S_h^2(z)} &\leq \frac{7\frac{\varphi(h)}{h}z\frac{16}{5}z}{\frac{\varphi^2(h)}{h^2}\log^2 z} \\ &= 22.5\frac{z^2}{\log^2 z}\frac{h}{\varphi(h)}.\end{aligned}$$

□

THEOREM 3.2.5. *If* x *and* y *are real numbers and* k *and* l *are integers satisfying* $1 \le k < y \le x$ *with* $k \perp l$, *then*

$$\pi(x;k,l) - \pi(x-y;k,l) < \frac{3y}{\varphi(k)\log\left(\frac{y}{k}\right)} \tag{3.25}$$

and

$$\pi(x;k,l) - \pi(x-y;k,l) < \frac{y}{\varphi(k)\log\sqrt{\frac{y}{k}}}\left(1 + \frac{4}{\log\sqrt{\frac{y}{k}}}\right). \tag{3.26}$$

Proof : Let $\Delta(x,y,k,l) = \pi(x;k,l) - \pi(x-y;k,l)$ and $h = \frac{2k}{\gcd(2,k)}$. Then there is an l_1 such that $\Delta(x,y,k,l) \le \Delta(x,y,h,l_1) + 1$. For if k is even, then $h = k$, and we can take $l_1 = l$. If k is odd, then the parity of $mk + l$ changes alternately. In this case, we can set l_1 to be the solution to $l_1 \equiv 1 \mod 2$ and $l_1 \equiv l \mod k$. So at worst we miss one prime in the even subsequence.
By what we have proved so far, the sifting of the sequence $\mathcal{A}$ by P_z yields the following upper bound:

$$\Delta(x,y,k,l) \le \Delta(x,y,h,l_1) + 1 \tag{3.27}$$

$$\le \frac{y}{\varphi(h)S_1(z)} + \frac{H_h^2(z)}{S_h^2(z)} + \pi(z,h,l_1) + 1 \tag{3.28}$$

$$\le \frac{y}{\varphi(k)S_1(z)} + \frac{H_h^2(z)}{S_h^2(z)} + \pi(z,h,l_1) + 1 \text{ for any } z > 1. \tag{3.29}$$

We begin with a trivial estimate

$$\Delta(x,y,h,l_1) \le \sum_{\substack{x-y<n\le x\\ n\equiv l_1 \bmod h}} 1 \le \frac{y}{h} + 1.$$

So $\Delta(x,y,k,l) \le \frac{y}{h} + 2$. Let $u = \sqrt{\frac{y}{k}}$. Since $\varphi(k) = \varphi(h) \le \frac{1}{2}h$, we have

$$\frac{\Delta(x,y,k,l)}{y} \le \frac{1}{k} + \frac{2}{y}.$$

Using $y = u^2k$ we obtain

$$\frac{\varphi(k)\Delta(x,y,k,l)}{y} \le \frac{\varphi(k)}{k} + \frac{2\varphi(k)}{y} \le \frac{1}{2} + \frac{2\varphi(k)}{u^2k} \le \frac{1}{2} + \frac{2}{u^2}.$$

Thus

$$Q = \frac{\log\sqrt{\frac{y}{k}}\varphi(k)}{y}\Delta(x,y,k,l) \le \log u\left(\frac{1}{2} + \frac{2}{u^2}\right) < \frac{3}{2} \text{ for } 1 < u \le e^{2.9}.$$

Now

$$\pi(z,h,l_1) + 1 \le \sum_{1\le n\le z, k\perp 2} \mu^2(k) \le \frac{z-1}{2} \text{ for } z \ge 9.$$

The remainder term is at most

$$\left(\sum_{\substack{d<z\\ \gcd(d,h)=1}} \mu^2(d)\right)^2 \le \left(\sum_{\substack{d<z\\ \gcd(d,2)=1}} \mu^2(d)\right)^2, \text{ since } 2\backslash h$$

$$\le \left(\frac{z-1}{2}\right)^2 \text{ if } z \ge 9.$$

By (3.27), and the above bounds we have:

$$\begin{aligned} Q &\leq \log u \left\{ \frac{1}{\log z} + \frac{1}{u^2}\left(\left(\frac{z-1}{2} \right)^2 + \frac{z-1}{2} \right) \right\} \\ &< \log u \left\{ \frac{1}{\log z} + \frac{z^2}{4u^2} \right\} \text{ if } z \geq 9. \end{aligned}$$

Define ω by

$$u = \frac{\omega}{\sqrt{2}} e^{\omega},$$

and set $z = e^{\omega}$ so that

$$Q \leq \frac{\log(\frac{\omega}{\sqrt{2}}) + \omega}{\omega} \left\{ 1 + \frac{1}{2\omega} \right\} \text{ for } \omega \geq \log 9.$$

For $\omega \geq \sqrt{2}e > \log 9$ this function is decreasing, and for $\omega = \sqrt{2}e$ it is $< \frac{3}{2}$. This proves (3.25).

Now (3.26) is a consequence of (3.25) for $u \leq e^8$. If $e^8 < u < e^{10}$, then using the above bound for Q, we obtain

$$Q \leq \frac{\log(\frac{\omega}{\sqrt{2}}) + \omega}{\omega} \left\{ 1 + \frac{1}{2\omega} \right\}.$$

If $u > e^8$, then $\omega < 6.4$ and this gives $Q < 1.4 < 1 + \frac{4}{\log u}$. This shows (3.26) for $u < e^{10}$.
Now using (3.27) and setting $\log z = \log u - 2$, we get

$$\log \sqrt{\frac{y}{k}}(Q-1) \leq \log u \left\{ \frac{\log u}{\log z} - 1 + 48 \frac{\log u}{u^2} \frac{z^2}{\log^2 z} + \frac{\log u}{u^2} z \right\}, \quad = \log u \left\{ \frac{2}{\log u - 2} + \frac{48}{e^4} \frac{\log u}{(\log u - 2)^2} + \frac{\log u}{e^2 u} \right\},$$

which is a decreasing function in u. In particular it is < 4 if $u \geq e^{10}$. This proves (3.26). $\square$

3.3. Prelude to a theorem of Hooley

In this section we will look at a variation of a problem of Chebyschev that we shall see in the next section. The problem is to prove a lower bound on the largest prime divisor of

$$\prod_{p \leq x} (p^2 - 1) = \prod_{p \leq x} (p+1) \prod_{p \leq x} (p-1).$$

We will prove the following theorem of Motohashi [**Mot70**].

THEOREM 3.3.1. *Let* P_x *be the largest prime divisor of*

$$\prod_{p \leq x} (p^2 - 1).$$

Then $P_x > x^{\theta}$ *for any* $\theta < 1 - \frac{1}{2e^{\frac{1}{4}}}$.

Proof : In this proof q will also stand for primes, and sums or products over q will represent sums or products over primes in the range.

Consider the product $\Xi = \prod_{p \leq x} (p^2 - 1)$. Taking log on both sides, we have

$$\begin{aligned} \log \Xi &= \log \prod_{p \leq x} p^2 \left(1 - \frac{1}{p^2} \right) \\ &= 2 \sum_{p \leq x} \log p - O\left(\sum_{p \leq x} \frac{1}{p^2} \right) \\ &= 2x + O(x e^{-c\sqrt{\log x}}) - O(1). \end{aligned}$$

Let $\pi(x,k)$ be the number of primes below x such that $p^2-1 \equiv 0 \mod k$. We have that $p^2-1=(p+1)(p-1)$ and for $p>2$ we have $\gcd(p+1,p-1)=2$. If $k=q^a$, $q\neq 2$, then $p^2-1\equiv 0 \mod k$ implies that either $p+1\equiv 0 \mod k$ or $p-1\equiv 0 \mod k$. In this case we have $\pi(x,q^a)=\pi(x;-1,q^a)+\pi(x;+1,q^a)$. Furthermore, $\pi(x,2)=\pi(x)$, and $\pi(x,4)=\pi(x)$. For $a>2$, we have $\pi(x,2^a)=\pi(x;-1,2^{a-1})+\pi(x;+1,2^{a-1})$.
Using the function $\pi(x,q^a)$, we can write Ξ as

$$\prod_{q^a<x} q^{\pi(x,q^a)}.$$

For if q^a divides Ξ, then it is counted exactly a times in this product. Taking logarithms we have

$$\sum_{q^a<x}\pi(x,q^a)\log q = 2x + O(xe^{-c\sqrt{\log x}}).$$

We split up the sum as follows:

$$\begin{aligned}\sum_{q^a<x}\pi(x,q^a)\log q &= \sum_{\substack{q\le\frac{\sqrt{x}}{\log^B x}\\ a=1}} + \sum_{\substack{\frac{\sqrt{x}}{\log^B x}<q\le x^\theta\\ a=1}} + \sum_{\substack{x^\theta<q<x\\ a=1}} + \sum_{\substack{q^a<x\\ a\ge 2}}\\ &= \Sigma_1+\Sigma_2+\Sigma_3+\Sigma_4,\end{aligned}$$

where B is a positive real number. We wish to show that Σ_3 is non-zero for the value of θ claimed. Since we already have an asymptotic formula for the sum, to obtain a lower bound for Σ_3 we need upper bounds for the remaining sums. We have $\pi(x,k)\sim\frac{2\mathrm{li}x}{\varphi(k)}$.

(Σ_1) Bombieri's Theorem— which we shall prove in Chapter 4, can be used directly to bound this sum we get:

$$\begin{aligned}\Sigma_1 &= \frac{2x}{\log x}\sum_{q\le\frac{\sqrt{x}}{\log^B x}}\frac{\log q}{(q-1)} + O\left(\frac{x}{\log x}\right)\\ &= x + O\left(\frac{x\log\log x}{\log x}\right).\end{aligned}$$

(Σ_2) We have from the Brun-Titchmarsh Theorem (3.2.5) that

$$\pi(x,q)\le 4\frac{x}{(q-1)\log\left(\frac{x}{q}\right)}\left\{1+\frac{8}{\log\left(\frac{x}{q}\right)}\right\}.$$

Hence

$$\Sigma_2 \le 4x\left\{\sum_{\frac{\sqrt{x}}{\log^B x}<q\le x^\theta}\frac{\log q}{(q-1)\log\left(\frac{x}{q}\right)} + O\left(\frac{1}{(\log^2 x)}\sum_{q\le x}\frac{\log q}{q}\right)\right\},$$

and using $\sum_{p\le x}\frac{\log p}{p}\sim\log x$, we have

$$\Sigma_2 = 4x\left\{\sum_{\frac{\sqrt{x}}{\log^B x}<q\le x^\theta}\frac{\log q}{q\log\left(\frac{x}{q}\right)}\right\} + O\left(\frac{x}{\log x}\right).$$

Writing $\vartheta(x)$ for $\sum_{p\le x}\log p$, we have by partial summation:

$$\begin{aligned}\sum_{y<p\le z}\frac{\log p}{q\log\left(\frac{x}{q}\right)} &= \sum_{y<k\le z}\frac{\vartheta(k)-\vartheta(k-1)}{k\log\left(\frac{x}{k}\right)}\\ &= \sum_{y<k\le z}\vartheta(k)\left\{\frac{1}{k\log\left(\frac{x}{k}\right)} - \frac{1}{(k+1)\log\left(\frac{x}{k+1}\right)}\right\}.\end{aligned}$$

This sum boils down to

$$\sum_{y<k\le z} \frac{\vartheta(k)}{k(k+1)\log\left(\frac{x}{k}\right)},$$

and using $\vartheta(x) < x\left(1+\frac{1}{2\log x}\right)$, we get

$$\sum_{y<p\le z} \frac{\log p}{q\log\left(\frac{x}{q}\right)} \le \sum_{y<k\le z} \frac{1}{k\log\left(\frac{x}{k}\right)}.$$

Now we can bound this sum using integration to get

$$\sum_{y<p\le z} \frac{\log p}{q\log\left(\frac{x}{q}\right)} = \log\left(\log\frac{x}{z}\right) - \log\left(\log\frac{x}{y}\right) + o(1).$$

Thus

$$\sum_{\frac{\sqrt{x}}{\log^B x}<q\le x^\theta} \frac{\log q}{q\log\left(\frac{x}{q}\right)} = -\log 2(1-\theta) + o(1),$$

and so

$$\Sigma_2 \le -4\log 2(1-\theta)x + o(x).$$

(Σ_4) We split up Σ_4 into two parts,

$$\begin{aligned}\Sigma_4 &= \sum_{\substack{q^a\le x^{\frac{2}{3}}\\ a\ge 2}} + \sum_{\substack{x^{\frac{2}{3}}<q^a<x\\ a\ge 2}}\\ &= \Sigma_{41} + \Sigma_{42}, \text{ say.}\end{aligned}$$

Using the Brun-Titchmarsh theorem:

$$\begin{aligned}\Sigma_{41} &= O\left\{\sum_{q\le\sqrt{x}} \log q \frac{x}{\log x} \sum_{a\ge 2} \frac{1}{\varphi(q^a)}\right\}\\ &= O\left\{\frac{x}{\log x} \sum_{q\le\sqrt{x}} \frac{\log q}{q^2}\right\}\\ &= O\left(\frac{x}{\log x}\right)\end{aligned}$$

and

$$\begin{aligned}\Sigma_{42} &= O\left\{\sum_{q\le\sqrt{x}} \log q \sum_{x^{\frac{2}{3}}<q^a<x} \frac{x}{q^a}\right\}\\ &= O\left\{x^{\frac{1}{3}} \sum_{q\le\sqrt{x}} \log q \frac{\log x}{\log q}\right\}\\ &= O(x^{\frac{5}{6}}).\end{aligned}$$

Thus

$$\Sigma_4 = O\left(\frac{x}{\log x}\right).$$

From the bounds we have derived we get:

$$\Sigma_3 > (1 + 4\log 2(1-\theta))x + o(x).$$

Hence if $1 + 4\log 2(1-\theta) > 0$ i.e., if

$$1 - \frac{1}{2e^{\frac{1}{4}}} > \theta,$$

then there is a prime factor exceeding x^θ. $\square$

Among known improvements to this result, the best one is that the largest prime factor exceeds x^θ for $\theta = 0.677$ (see [**BakHar95**], [**BakHar98**], and also [**Ho73**]).

3.4. A theorem of Hooley

Chebyhev proved that if P_x is the largest prime factor of $\prod_{n\le x}(n^2+1)$, then $\frac{P_x}{x} \to \infty$. Hooley [**Ho67**] (see also [**Ho76**]) improved the previous best known result of

$$\frac{P_x}{x} > (\log x)^{A_1 \log\log\log x}$$

by Erdős [**Erd52**] to $P_x > x^{\frac{11}{10}}$ using the Selberg sieve. In this section we shall outline the proof given by Hooley in [**Ho76**]. The exponent $\frac{11}{10}$ has since been improved to $\theta < 1.202\cdots$, where θ is the solution to $2 - \theta - 2\log(2-\theta) = \frac{5}{4}$, by Deshouillers and Iwaniec [**DI83**] (see also [**Dar96**]).

THEOREM 3.4.1 ([**Ho76**]). *The largest prime factor of*

$$\prod_{n\le x}(n^2+1)$$

exceeds $x^{\frac{11}{10}}$ for all large enough values of x.

Proof : Let P_x be the largest prime factor of $\prod_{n\le x}(n^2+1)$, and set $N_x(l) = \left|\{n \le x \mid n^2 \equiv -1 \mod l\}\right|$. We begin by finding a lower bound for $\sum_{x\le p\le P_x} N_x(p)\log p$, as in the proof of Theorem (3.3.1). We have

$$\prod_{n\le x}(n^2+1) = \prod_{\substack{p\le P_x\\ p^\alpha < x^2+1}} p^{N_x(p^\alpha)}.$$

Taking logs,

$$\begin{aligned}\log\prod_{n\le x}(n^2+1) &= \log\prod_{n\le x} n^2\left(1+\frac{1}{n^2}\right)\\ &> \log\left(\lfloor x\rfloor !\right)^2\\ &= 2x\log x + O(x)\end{aligned}$$

by Stirling's theorem, and so

$$\sum_{\substack{p\le P_x\\ p^\alpha < x^2+1}} N_x(p^\alpha)\log p > 2x\log x + O(x).$$

Now

$$\begin{aligned}\sum_{\substack{p\le P_x\\ p^\alpha < x^2+1}} N_x(p^\alpha)\log p &= \sum_{x\le p\le P_x} N_x(p)\log p + \sum_{p\le x} N_x(p)\log p + \sum_{\substack{p\le P_x\\ \alpha>1}} N_x(p^\alpha)\log p\\ &= \Sigma_A + \Sigma_B + \Sigma_C.\end{aligned}$$

As before we proceed to upper-bound Σ_B and Σ_C, thereby obtaining a lower bound for Σ_A. Now

$$\begin{aligned} N_x(l) &= \sum_{\substack{n^2+1\equiv 0 \mod l \\ n\le x}} 1 \\ &= \sum_{\substack{\nu^2+1\equiv 0 \mod l \\ 0<\nu\le l}} \sum_{\substack{n\equiv \nu \mod l \\ n\le x}} 1. \end{aligned}$$

Let $\rho(l)$ be the number of solution to the congruence $\nu^2+1\equiv 0 \mod l$. Then since

$$\sum_{\substack{n\equiv \nu \mod l \\ n\le x}} 1 - \frac{x}{l} = O(1),$$

we have

$$N_x(l) = \frac{x\rho(l)}{l} + O(\rho(l)).$$

Now $\rho(2)=1$, and since the congruence $\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \mod p$ has no solutions for $p\equiv 3 \mod 4$, and has exactly two solutions for $p\equiv 1 \mod 4$. We conclude

$$\rho(p) = \begin{cases} 2 & \text{if } p\equiv 1 \mod 4, \\ 0 & \text{if } p\equiv 1 \mod 4. \end{cases}$$

The needed bounds are given by:

$$\begin{aligned} \Sigma_B &= x\sum_{p\le x} \frac{\rho(p)\log p}{p} + O\big(\sum_{p\le x} \rho(p)\log p\big) \\ &= 2x \sum_{\substack{p\le x \\ p\equiv 1 \mod 4}} \frac{\log p}{p} + O(x) + O\big(\sum_{p\le x} \log p\big), \\ &= x\log x + O(x). \end{aligned}$$

using $\sum_{\substack{p\le x \\ p\equiv l \mod k}} \frac{\log p}{p} = \frac{1}{\varphi(k)}\log x + O(1)$,

$$\begin{aligned} \Sigma_C &= O\Bigg(\sum_{p\le\sqrt{x^2+1}} \log p \sum_{2\le\alpha} \left\{\frac{x}{p^\alpha}+1\right\}\Bigg) \\ &= O\left(x\sum_p \frac{\log p}{p(1-\frac{1}{p})}\right) \\ &= O\left(x\sum_p \frac{\log p}{p(p-1)}\right) \\ &= O(x) \end{aligned}$$

since the sum converges.

Thus we get $\Sigma_A > x\log x + O(x)$. Our next task is to upper-bound the sum $T_x(y) = \sum_{x<p\le y} N_x(p)\log p$, which in conjunction with the above lower bound will yield a lower bound for y. It turns out that to estimate $T_x(y)$ effectively, we need to split up the sum into two parts and evaluate each of them separately. To this end let $X = x^{\frac{1}{11}}$, and assume that $x^{\frac{12}{11}} < y < x^2$. Then

$$\begin{aligned} T_x(y) &= \sum_{x<p\le xX} N_x(p)\log p + \sum_{xX<p\le y} N_x(p)\log p \\ &= T_x(xX) + T'_x(y). \end{aligned}$$

To evaluate $T_x(xX)$, we let $V_x(v) = \sum_{v<p\leq ev} N_x(p)$. Then

$$\begin{aligned} T_x(xX) &= \sum_{0\leq\alpha<\log X} \sum_{xe^{\alpha}<p\leq xe^{\alpha+1}} N_x(p)\log p \\ &\leq \sum_{0\leq\alpha<\log X} \log(xe^{\alpha+1})V_x(xe^{\alpha}). \end{aligned}$$

Now for the sum $T_x'(y)$, using the definition of $N_x(l)$, we have:

$$\begin{aligned} T_x'(y) &= \sum_{\substack{xX<p\leq y \\ pm=n^2+1 \\ n\leq x}} \log p \\ &= \sum_{m>\frac{x^2}{y\log^8 x}} \log p + \sum_{m\leq\frac{x^2}{y\log^8 x}} \log p \\ &= T_x''(y) + T_x'''(y)(say). \end{aligned}$$

Now the conditions of the summation $T_x'''(y)$ yield $m \leq \frac{x^2}{y\log^8 x}$, and so $n < \sqrt{(pm)} \leq \sqrt{\left(\frac{yx^2}{y\log^8 x}\right)} = \frac{x}{\log^4 x}$. Since $m \leq n$, we have $m \leq \frac{x}{\log^4 x}$. Using this we have

$$\begin{aligned} T_x'''(y) &= 2\log x \sum_{\substack{lm=n^2+1 \\ m,n\leq\frac{x}{\log^4 x}}} 1 \\ &= 2\log x \sum_{m\leq\frac{x}{\log^4 x}} N_{\frac{x}{\log^4 x}}(m). \end{aligned}$$

Now if $m = \prod_i p_i^{h_i}$, then $\rho(m) = \prod_i \rho(p^{h_i})$, and each of the individual terms is a constant. So $\rho(m) \leq 2^{\nu(m)}$, and this itself is upper bounded by $d(m)$, i.e. the number of divisors of m. Therefore:

$$\begin{aligned} T_x'''(y) &\leq \frac{2x}{\log^3 x} \sum_{m\leq\frac{x}{\log^4 x}} \frac{\rho(m)}{m} + O\big(\log x \sum_{m\leq\frac{x}{\log^4 x}} \rho(m)\big) \\ &= O\left\{\frac{x}{\log^3 x} \sum_{m\leq\frac{x}{\log^4 x}} \frac{\rho(m)}{m}\right\} \\ &= O\left\{\frac{x}{\log^3 x} \sum_{m\leq x} \frac{d(m)}{m}\right\}. \end{aligned}$$

Now consider

$$\begin{aligned} \left(\sum_{1\leq n\leq x} \frac{1}{n}\right)\left(\sum_{1\leq m\leq x} \frac{1}{m}\right) &= \sum_{\substack{1\leq n\leq x^2 \\ n \text{ is } x-\text{smooth}}} \frac{1}{n}\Big(\sum_{\substack{u,v\leq x \\ uv=n}} 1\Big) \\ &\geq \left(\sum_{1\leq n\leq x} \frac{d(n)}{n}\right). \end{aligned}$$

This yields $\sum_{1\leq n\leq x} \frac{d(n)}{n} = O(\log^2 x)$, and so

$$T_x'''(y) = O\left(\frac{x}{\log x}\right).$$

In $T''_x(y)$, we have $m > \frac{x^2}{y\log^8 x}$ and $pm \le x^2+1$, so $m \le \frac{x^2+1}{p}$. Furthermore $p > xX$, and so $m \le \frac{x}{X}\left(1+\frac{1}{x^2}\right) \le \frac{ex}{X}$. Thus we have

$$T''_x(y) \le \sum_{\substack{\frac{x^2}{y\log^8 x} < m \le \frac{ex}{X} \\ pm=n^2+1 \\ n\le x, p\ge xX}} \log\frac{ex^2}{m}.$$

Let

$$W_x(w) = \sum_{\substack{w<m\le ew \\ pm=n^2+1 \\ n\le x, p\ge x}} 1.$$

Then

$$T''_x(y) \le \sum_{0\le\alpha<\log Y} \log\left(xXe^{\alpha+1}\right) W_x\left(\frac{xe^{-\alpha}}{X}\right),$$

where $Y = \frac{ey\log^8 x}{xX}$. Finally,

$$T'_x(y) \le \sum_{0\le\alpha<\log Y} \log(xXe^{\alpha+1}) W_x\left(\frac{xe^{-\alpha}}{x}\right) + O\left(\frac{x}{\log x}\right).$$

We will format the sums involved for application of the Selberg sieve. Let λ be a squarefree number, and define

$$\Upsilon(u;\lambda) = \sum_{u<\lambda k\le eu} N_x(\lambda k).$$

We impose the conditions $x^{\frac{4}{5}} < u < x^{\frac{4}{3}}$ and $\lambda < \min\{\frac{u^{\frac{5}{4}}}{x}, \frac{x}{u^{\frac{3}{4}}}\}$. By a rather ingenious and elaborate argument Hooley showed that

$$\Upsilon(u;\lambda) = \frac{3x\rho(\lambda)}{2\pi\lambda}\frac{1}{\prod_{p\setminus\lambda}\left(1+\frac{1}{p}\right)} + O\left(x^{\frac{1}{2}+\epsilon}u^{\frac{3}{8}}\lambda^{-\frac{1}{2}}\right)$$

(see [**Ho76**] §2.3 - §2.6). Since the argument is not central to our application of the sieve, we exclude the derivation of this bound here.

Application of the Sieve: Let $x \le v < x^{\frac{12}{11}}$, so that v satisfies the conditions on u imposed by our bounds on $\Upsilon(u;\lambda)$. Let d denote a squarefree number, and let λ_d be the Selberg coefficients. Then

$$\begin{aligned} V_x(v) &\le \sum_{v<l\le ev} N_x(l)\left(\sum_{d\setminus l}\lambda_d^2\right) \\ &= \sum_{d_1,d_2\le z}\lambda_{d_1}\lambda_{d_2} \sum_{\substack{v<l\le ev \\ l\equiv 0 \mod \operatorname{lcm}(d_1,d_2)}} N_x(l) \\ &= \sum_{d_1,d_2\le z}\lambda_{d_1}\lambda_{d_2}\Upsilon(v;\operatorname{lcm}(d_1,d_2)\ ) \end{aligned}$$

(since $\operatorname{lcm}(d_1,d_2)\ < xv^{-\frac{3}{4}}$)

$$\le \frac{3x}{2\pi}\sum_{d_1,d_2\le z}\frac{\lambda_{d_1}\lambda_{d_2}\omega(\operatorname{lcm}(d_1,d_2))}{\operatorname{lcm}(d_1,d_2)} + O\left(x^{\frac{1}{2}+\epsilon}v^{\frac{3}{8}}\sum_{d_1,d_2\le z}\frac{|\lambda_{d_1}||\lambda_{d_2}|}{\sqrt{\operatorname{lcm}(d_1,d_2)}}\right).$$

Here

$$\omega(d) = \frac{\rho(d)}{\prod_{p\backslash d}(1+\frac{1}{p})},$$

which is clearly multiplicative. So we can apply Selberg's sieve without modification, except that the remainder term is more clearly specified in this case. Thus by Theorem (3.1.1), we have

$$V_x(\nu) \leq \frac{3x}{2\pi G(z)} + R,$$

where R is the remainder term. Now

$$G(z) = \sum_{d<z} \mu^2(d) g(d),$$

and

$$\begin{aligned} g(p) &= \frac{\omega(p)}{p\left(1-\frac{\omega(p)}{p}\right)} \\ &= \frac{2\left(1+\frac{1}{p}\right)^{-1}}{p\left(1-\frac{2}{p}\left(1+\frac{1}{p}\right)^{-1}\right)} \\ &= \frac{2}{p\left(1-\frac{1}{p}\right)}. \end{aligned}$$

Thus

$$\begin{aligned} g(d) &= \frac{\rho(d)}{d\prod_{\substack{p\backslash d \\ p\neq 2, p\equiv 1 \mod 4}}\left(1-\frac{1}{p}\right)} \\ &= \sum_{d'} \frac{\rho(dd')}{dd'}, \end{aligned}$$

where d' indicates any number whose prime factors divide d. Also $\rho(2^\alpha) = 0$, if $\alpha > 1$, and we have

$$\begin{aligned} \sum_{d\leq z} \mu^2(d) g(d) &= \sum_{d\leq z}\sum_{d'} \frac{\rho(dd')}{dd'} \\ &\geq \sum_{m\leq z} \frac{\rho(m)}{m} \\ &\geq \frac{3(1-\eta_1)}{2\pi} \log z, \end{aligned}$$

where $\eta_1 < 1$ can be chosen very small.
Here we have used

$$\sum_{m\leq z} \rho(m) = \frac{3z}{2\pi} + O(z^{\frac{3}{4}})$$

(which is proved in [**Ho76**] p. 32) and partial summation.

Also the remainder term can be bounded as follows:

$$\begin{aligned}
R &= O\left\{x^{\frac{1}{2}+\epsilon}v^{\frac{3}{8}} \sum_{d_1,d_2\le z} \frac{1}{\sqrt{\mathrm{lcm}(d_1,d_2)}}\right\} \\
&= O\left(x^{\frac{1}{2}+\epsilon}v^{\frac{3}{8}} \sum_{d\le z} \sum_{\substack{l_1,l_2\le \frac{z}{d} \\ l_1\perp l_2}} \frac{1}{\sqrt{dl_1l_2}}\right) \\
&= O\left(x^{\frac{1}{2}+\epsilon}v^{\frac{3}{8}} \sum_{d\le z} \frac{z}{\sqrt{d}} \sum_{\substack{l_1,l_2\le z/d \\ l_1\perp l_2}} \frac{1}{\sqrt{l_1l_2}}\right) \\
&= O\left(x^{\frac{1}{2}+\epsilon}v^{\frac{3}{8}} \sum_{d\le z} \frac{z}{d^{\frac{3}{2}}}\right) \\
&= O\left(x^{\frac{1}{2}+\epsilon}v^{\frac{3}{8}}z\right).
\end{aligned}$$

Selecting $z = x^{\frac{1}{2}-\eta}v^{-\frac{3}{8}}$, we get

$$V_x(v) < \frac{(1+\eta_2)x}{\log\sqrt{x}v^{-\frac{3}{8}}},$$

where η_2 can be made arbitrarily small.

Similarly

$$\begin{aligned}
W_x(w) &\le \sum_{w<m\le ew} \left(\sum_{d\backslash l} \lambda_d\right)^2 \\
&= \sum_{d_1,d_2\le z} \lambda_{d_1}\lambda_{d_2} \sum_{\substack{w<m\le ew \\ mr\times \mathrm{lcm}(d_1,d_2)=n^2+1 \\ n\le x}} 1 \\
&= \sum_{d_1,d_2\le z} \lambda_{d_1}\lambda_{d_2} \Upsilon\big(\mathrm{lcm}(d_1,d_2)\,w; \mathrm{lcm}(d_1,d_2)\,\big).
\end{aligned}$$

Carrying through the sieve estimate, we get with $z = x^{\frac{2}{7}-\eta}w^{-\frac{3}{14}}$ that

$$W_x(w) < \frac{(1+\eta_2)x}{\log x^{\frac{2}{7}}w^{-\frac{3}{14}}}.$$

Let $y = x^{\frac{11}{10}}$ and $\gamma = \log x$. Using the above estimates, we find that

$$\begin{aligned}
T_x(xX) &< x(1+\eta_2) \sum_{0\le\alpha<\log x} \frac{\alpha+\gamma+1}{\frac{1}{8}\gamma - \frac{3}{8}\alpha} \\
&< 0.8902x\log x,
\end{aligned}$$

where we have used integration to upper-bound the sum. Similarly we find

$$T'x(y) < 0.1081x\log x$$

for large enough x. Thus we get

$$T_x(x^{\frac{11}{10}}) < 0.9983x\log x,$$

and so the largest prime factor of $\prod_{n\le x}(n^2+1)$ exceeds $x^{\frac{11}{10}}$, for all large enough values of x. $\square$

CHAPTER 4

The Large Sieve

The Selberg sieve does not give good bounds if we sieve out a large number of residue classes modulo each prime in the sifting set. The large sieve was designed to handle this problem, (hence the name). The bounds are derived by relating the properties of the integer sequence to the behavior of certain exponential sums.

4.1. Bounds on exponential sums

Define $e(t) = e^{2\pi i t}$. We have $e\left(\frac{n}{q}\right) = e\left(\frac{m}{q}\right)$ if $n \equiv m \mod q$. The following property of the exponential function resembles that of the Möbius function, and is useful to study the distribution of a sequence of integers in residue classes modulo some number.

PROPOSITION 4.1.1.

$$\sum_{1 \le a \le q} e\left(\frac{an}{q}\right) = \begin{cases} q, & \textit{if } n \equiv 0 \mod q \\ 0, & \textit{otherwise.} \end{cases}$$

Proof : If $n \equiv 0 \mod q$, then $e\left(\frac{an}{q}\right) = 1$ for each a. So $\sum_{1 \le a \le q} e\left(\frac{an}{q}\right) = q$. If $n \not\equiv 0 \mod q$, then

$$\begin{aligned} \sum_{1 \le a \le q} e\left(\frac{an}{q}\right) &= \sum_{0 \le a \le q-1} e\left(\frac{an}{q}\right) \\ &= \frac{e\left(\frac{qn}{q}\right) - 1}{e\left(\frac{n}{q}\right) - 1} \\ &= 0. \end{aligned}$$

□

Let $a_1, \cdots, a_z$ be a sequence of integers, and define

$$Z(q,h) = \left|\{i \mid 1 \le i \le z, a_i \equiv h \mod q\}\right|$$

and

$$S(x) = \sum_{1 \le i \le z} e(a_i x).$$

Now for all integers a we have

$$S\left(\frac{a}{q}\right) = \sum_{1 \le h \le q} Z(q,h) e\left(\frac{ah}{q}\right). \tag{4.30}$$

Suppose all the integers in the sequence are distributed evenly among the residue classes modulo q; then using Proposition 4.1.1 we have

$$\begin{aligned} S\left(\frac{a}{q}\right) &= Z(q,h) \sum_{1 \le h \le q} e\left(\frac{ah}{q}\right) \\ &= 0, \text{ if } a \not\equiv 0 \mod q. \end{aligned}$$

If on the other hand all the integers a_i belong to a single residue class modulo q, then $\left|S\left(\frac{a}{q}\right)\right| = z$ for all integers a. Hence the distribution of the integers among the residue classes is related to $|S\left(\frac{a}{q}\right)|$. In fact, we can express $Z(q,h)$ in terms of $S\left(\frac{a}{q}\right)$ as follows:

$$S\left(\frac{a}{q}\right)e\left(\frac{-h'a}{q}\right) = \sum_{1\le h\le q} Z(q,h)e\left(\frac{ah}{q}\right)e\left(\frac{-h'a}{q}\right),$$

and therefore

$$\begin{aligned}\sum_{1\le a\le q} S\left(\frac{a}{q}\right)e\left(\frac{-h'a}{q}\right) &= \sum_{1\le a\le q} Z(q,h)\sum_{1\le h\le q} e\left(\frac{a(h-h')}{q}\right)\\ &= Z(q,h')q.\end{aligned}$$

Hence

$$qZ(q,h) = \sum_{1\le a\le q} S\left(\frac{a}{q}\right)e\left(\frac{-h'a}{q}\right). \tag{4.31}$$

It turns out that useful upper bounds can be obtained for the sum

$$\sum_{p\le x}\sum_{1\le a\le p-1}\left|S\left(\frac{a}{p}\right)\right|$$

that are largely independent of the integer sequence used to define $S(x)$.

We first prove a result that shows how the above sums are related to the distribution of the integer sequence in the residue classes.

LEMMA 4.1.2. *For all integers $q \ge 2$,*

$$\sum_{1\le a\le q-1}\left|S\left(\frac{a}{q}\right)\right|^2 = q\sum_{1\le h\le q}\left(Z(q,h)-\frac{z}{q}\right)^2.$$

Proof :

$$\begin{aligned}\sum_{1\le a\le q-1}\left|S\left(\frac{a}{q}\right)\right|^2 &= \sum_{1\le a\le q-1}\left(\sum_{1\le h\le q} Z(q,h)e\left(\frac{ah}{q}\right)\right)\left(\sum_{1\le k\le q} Z(q,h)e\left(\frac{ka}{q}\right)\right)\\ &= \sum_{1\le a\le q-1}\sum_{1\le h,k\le q} Z(q,h)Z(q,k)e\left(\frac{a(h-k)}{q}\right)\\ &= \sum_{1\le h,k\le q} Z(q,h)Z(q,k)\left(\sum_{1\le a\le q-1} e\left(\frac{a(h-k)}{q}\right)\right).\end{aligned}$$

It is easy to see that

$$\sum_{1\le a\le q-1} e\left(\frac{a(h-k)}{q}\right) = \begin{cases} q-1, & \text{if } h\equiv k \mod q\\ -1, & \text{otherwise.}\end{cases}$$

Thus

$$\begin{aligned}
\sum_{1\le a\le q-1}\left|S\left(\frac{a}{q}\right)\right|^2 &= q\sum_{1\le h\le q} Z(q,h)^2 - \sum_{1\le h,k\le q} Z(q,h)Z(q,k)\\
&= q\sum_{1\le h\le q} Z(q,h)^2 - \Big(\sum_{1\le h\le q} Z(q,h)\Big)^2\\
&= q\Big(\sum_{1\le h\le q} Z(q,h)^2\Big) - z^2\\
&= q\sum_{1\le h\le q}\Big(Z(q,h)^2 - \frac{2zZ(q,h)}{q} + \frac{z^2}{q^2}\Big)\\
&= q\sum_{1\le h\le q}\Big(Z(q,h) - \frac{z}{q}\Big)^2.
\end{aligned}$$

□

We will look at exponential sums of the form

$$S(x) = \sum_{-K\le n\le K} a_n e(nx),$$

where K is a positive integer and $a_n \in \mathbb{C}$.

Notation : We write $\|t\|$ to mean the distance from t to the nearest integer, i.e., $\|t\| = \min_n |t-n| = \big|\lfloor t+\frac{1}{2}\rfloor - t\big|$.

THEOREM 4.1.3 ([**Gal67**]). *If* $S(x) = \sum_{-K\le n\le K} a_n e(nx)$ *and* $x_1,\cdots,x_R$ *are real numbers such that*

$$\|x_r - x_s\| \ge \delta > 0 \text{ for } r\ne s,$$

then

$$\sum_{1\le r\le R} |S(x_r)|^2 \le (\delta^{-1} + 2\pi K)\sum_{-K\le n\le K} |a_n|^2.$$

Proof : For any u we can write

$$S^2(x_r) = S^2(u) + 2\int_u^{x_r} S'(t)S(t)dt.$$

Using this we have

$$|S^2(x_r)| \le |S^2(u)| + 2\left|\int_u^{x_r} |S'(t)S(t)|dt\right|.$$

We now integrate over the interval $I_t = \big(x_r - \frac{\delta}{2}, x_r + \frac{\delta}{2}\big)$, to get

$$\delta|S(x_r)|^2 \le \int_{I_r} |S(u)|^2 du + 2\int_{I_r}\left|\int_u^{x_r} |S'(t)S(t)|dt\right|du.$$

Then

$$\begin{aligned}
\int_{I_r}\left|\int_u^{x_r} |S'(t)S(t)|dt\right|du &= \int_{x_r}^{x_r+\frac{\delta}{2}}\Big(\int_{x_r}^{u} |S'(t)S(t)|dt\Big)du + \int_{x_r-\frac{\delta}{2}}^{x_r}\Big(\int_u^{x_r} |S'(t)S(t)|dt\Big)du\\
&= \int_{x_r}^{x_r+\frac{\delta}{2}} |S'(t)S(t)|\Big(x_r + \frac{\delta}{2} - t\Big)dt + \int_{x_r-\frac{\delta}{2}}^{x_r} |S'(t)S(t)|\Big(t - x_r + \frac{\delta}{2}\Big)dt\\
&\le \frac{\delta}{2}\int_{I_r} |S'(t)S(t)|dt.
\end{aligned}$$

Thus

$$\delta |S(x_r)|^2 \le \int_{I_r} |S(u)|^2 du + \delta \int_{I_r} |S'(t)S(t)| dt.$$

By our condition on the numbers x_r the intervals I_r are disjoint modulo 1 meaning that if $r \neq s$, then no point of I_r differs by an integer from another point in I_s. Since S is periodic with period 1 and is non-negative, the value of its integral overl I_r is upper bounded by its integral over $[0, 1]$. Thus summing over r:

$$\delta \sum_{1 \le r \le R} |S(x_r)|^2 \le \int_0^1 |S(t)|^2 dt + \delta \int_0^1 |S'(t)S(t)| dt.$$

Let us analyze the first integral. The exponential function satisfies

$$\int_0^1 e(nx) dx = \begin{cases} 1 \text{ if } n = 0, \\ 0 \text{ otherwise.} \end{cases}$$

We have

$$\begin{aligned} \int_0^1 |S(x)|^2 dx &= \int_0^1 S(x)\overline{S(x)} dx \\ &= \int_0^1 \sum_{-K \le m, n \le K} a_n \overline{a_m} e((n-m)x) dx \\ &= \sum_{-K \le n \le K} |a_n|^2. \end{aligned}$$

Thus the first integral is $\sum_{-K \le n \le K} |a_n|^2$. The second satisfies:

$$\int_0^1 |S'(t)S(t)| dt \le \left(\int_0^1 |S(t)|^2 dt \right)^{\frac{1}{2}} \left(\int_0^1 |S'(t)|^2 dt \right)^{\frac{1}{2}}$$

and on substituting $S'(t)$ by $\sum_{-K \le n \le K} 2\pi i a_n n e(nt)$, the right-hand side becomes

$$\begin{aligned} &= \Big(\sum_{-K \le n \le K} |a_n|^2 \Big)^{\frac{1}{2}} \Big(\sum_{-K \le n \le K} |2\pi n a_n|^2 \Big) \\ &\le 2\pi K \sum_{-K \le n \le K} |a_n|^2. \end{aligned}$$

Thus

$$\delta \sum_{1 \le r \le R} |S(x_r)|^2 \le (1 + \delta 2\pi K) \sum_{-K \le n \le K} |a_n^2|.$$

□

There is a stronger bound on the sum $\sum_{1 \le r \le R} |S(x_r)|^2$ due to Montgomery. To prove this we require the following result.

Theorem 4.1.4. *Let $\Phi_1, \cdots, \Phi_R$ and ξ be arbitrary vectors in an inner product space V over the complex numbers. Then*

$$\sum_{1 \le r \le R} |(\xi, \Phi_r)|^2 \le A ||\xi||^2,$$

where

$$A = \max_r \sum_{1 \le s \le R} |(\Phi_r, \Phi_s)|.$$

THEOREM 4.1.5. *Let* $S(x)$ *be as above, and* $x_1, \cdots, x_r$ *be real numbers with*

$$\|x_r - x_s\| \geq \delta > 0 \text{ for } r \neq s.$$

Then

$$\sum_{1\leq r\leq R} |S(x_r)|^2 \leq (2K + 3\delta^{-1}) \sum_{-K\leq k\leq K} |a_k|^2.$$

Proof : If $R = 1$ we have

$$|S(x)|^2 \leq N \sum_{M+1\leq n\leq M+N} |a_n|^2$$

by Cauchy's inequality. Hence we may assume $R \geq 2$ so $\delta \leq \frac{1}{2}$. We apply Theorem (4.1.4) with the inner product defined to be $(\phi, \psi) = \sum_k \phi_k \overline{\psi_k}$.
Take $\xi = \{a_k b_k^{-\frac{1}{2}}\}_{-K\leq k\leq K}$ and $\phi_r = \{b_k^{\frac{1}{2}} e(-kx_r)\}_{-\infty<k<\infty}$, where b_k will be defined later to be positive for $-K \leq k \leq K$, and non-negative for other k.
Using Theorem (4.1.4) we have

$$\sum_{1\leq r\leq R} |S(x_r)|^2 \leq A \sum_{-K\leq k\leq K} |a_k|^2 b_k^{-1},$$

where $A = \max_r \sum_{1\leq s\leq 1} |B(x_r - x_s|$ and $B(x) = \sum_{-\infty<k<\infty} b_k e(kx)$. To finish the proof it suffices to pick b_k such that $b_k \geq 1$ for $-K \leq k \leq K$ such that

$$\sum_{1\leq s\leq R} |B(x_r - x_s)| \leq 2K + 3\delta^{-1} \text{ for all } r.$$

If we took $b_k = 1$ for $-K \leq k \leq K$ and $b_k = 0$ otherwise, we would get the inferior estimate

$$\sum_{1\leq s\leq R} |B(x_r - x_s)| \leq 2K + O(\delta^{-1} \log \delta^{-1}).$$

Instead, take b_k to be

$$b_k = \begin{cases} 1 & \text{if } |k| \leq K, \\ 1 - \frac{(|k|-K)}{L} & \text{if } K \leq |k| \leq K + L, \\ 0 & \text{if } |k| \geq K + L, \end{cases}$$

where L will be selected later.
Using the indentity

$$\sum_{|j|\leq J} (J - |j|) e(jx) = \left| \sum_{1\leq j\leq J} e(jx) \right|^2$$
$$= \left(\frac{\sin \pi J x}{\sin \pi x} \right)^2,$$

we can write

$$B(x) = \frac{1}{L \sin^2 \pi x} \big((\sin \pi (K + L) x)^2 - (\sin \pi K x)^2 \big).$$

Hence $B(0) = 2K + L$, and

$$|B(x)| \leq \frac{1}{L(\sin^2 \pi x)} \leq \frac{1}{4L\|\alpha\|^2},$$

so that

$$\sum_{1\leq s\leq R} |B(x_r - x_s)| \leq 2K + L + 2 \sum_{1\leq h} \frac{1}{4Lh^2\delta^2}.$$

Since $\sum_{1\le h}\frac{1}{h^2}=\frac{\pi^2}{6}<2$, we have

$$\sum_{1\le s\le R}|B(x_r-x_s)|\le 2K+L+\frac{1}{L\delta^2}$$
$$\le 2K+\frac{3}{\delta}.$$

upon taking L to be the least integer $\ge\delta^{-1}$. $\square$

Consider the sum $S(x)=\sum_{M+1\le N\le M+N}a_n e(nx)$. The value of M is irrelevant to the magnitude of this sum since for any K we can set

$$T(x)=\sum_{K+1\le n\le K+N}a_{M-K+n}e(nx)$$
$$=e\big((K-M)x\big)S(x)$$

and then $|T(x)|=|S(x)|$. Thus the above theorem can be rephrased as follows.

Theorem 4.1.6. *Let*

$$S(x)=\sum_{M+1\le n\le M+N}a_n e(nx)$$

where M *and* N *are integers,* $N>0$*. Let* $x_1,\cdots,x_R$ *be distinct real numbers modulo* 1 *and* $\delta>0$ *is such that*

$$\|x_r-x_s\|\ge\delta,\ \textit{for } r\ne s.$$

Then for arbitrary a_n

$$\sum_{1\le r\le R}|S(x_r)|^2\le(N+3\delta^{-1})\sum_{M+1\le n\le M+N}|a_n|^2.\quad\square$$

We state (without proof) another version of the large sieve inequalities due to Montgomery and Vaughan [**MV73**] (Theorem 1).

Theorem 4.1.7 ([**MV73**]). *Let*

$$S(x)=\sum_{M+1\le n\le M+N}a_n e(nx),$$

let $x_1,\cdots,x_R$ *be real numbers, and set*

$$\delta=\min_{r\ne s}\|x_r-x_s\|.$$

Then

$$\sum_{1\le r\le R}|S(x_r)|^2\le(N+\delta^{-1})\sum_{M+1\le n\le M+N}|a_n|^2.$$

Moreover, if

$$\delta_r=\min_{\substack{s\\ s\ne r}}\|x_r-x_s\|$$

for all r*, then*

$$\sum_{1\le r\le R}(N+\tfrac{3}{2}\delta_r^{-1})^{-1}|S(x_r)|^2\le\sum_{M+1\le n\le M+N}|a_n|^2.\quad\square$$

4.2. The Large Sieve

In this section we will use the bounds derived in the previous section to study the distribution of integer sequences in residue classes modulo primes.

Let a_n be a sequence of complex numbers defined for $M+1 \le n \le M+N$ (where M, N are integers and $N > 0$). Define

$$Z(q,h) = \sum_{\substack{M+1 \le n \le M+N \\ n \equiv h \mod q}} a_n$$

and

$$Z(1,1) = Z = \sum_{M+1 \le n \le M+N} a_n.$$

Lemma 4.2.1. *Let*

$$S(x) = \sum_{M+1 \le n \le M+N} a_n e(nx).$$

If q *is a positive integer, then*

$$\sum_{1 \le a \le q} \left| S\left(\frac{a}{q}\right) \right|^2 = q \sum_{1 \le h \le q} \left| \sum_{d \backslash q} \frac{\mu(d)}{d} Z\left(\frac{q}{d}, h\right) \right|^2.$$

Proof : For an integer a we have (using (4.30))

$$S\left(\frac{a}{q}\right) = \sum_{1 \le h \le q} Z(q,h) e\left(\frac{ah}{q}\right).$$

By (4.31)

$$\begin{aligned} qZ(q,h) &= \sum_{1 \le a \le q} S\left(\frac{a}{q}\right) e\left(\frac{-ah}{q}\right) \\ &= \sum_{d \backslash q} \sum_{\substack{1 \le b \le \frac{q}{d} \\ \gcd(b, \frac{q}{d}) = 1}} S\left(\frac{bd}{q}\right) e\left(\frac{-bdh}{q}\right). \end{aligned}$$

Let

$$T(q,h) = \sum_{\substack{1 \le a \le q \\ a \perp q}} S\left(\frac{a}{q}\right) e\left(-\frac{ah}{q}\right),$$

so that

$$qZ(q,h) = \sum_{d \backslash q} T\left(\frac{q}{d}, h\right).$$

Applying Möbius inversion to this we get

$$T(q,h) = d \sum_{d \backslash q} \frac{\mu(d)}{d} Z\left(\frac{q}{d}, h\right).$$

Hence

$$|T(q,h)|^2 = q^2 \left| \sum_{d \backslash q} \frac{\mu(d)}{d} Z\left(\frac{q}{d}, h\right) \right|^2,$$

and therefore

$$\frac{1}{q}\sum_{1\le h\le q}|T(q,h)|^2 = q\sum_{1\le h\le q}\left|\sum_{d\backslash q}\frac{\mu(d)}{d}Z\left(\frac{q}{d},h\right)\right|^2.$$

Now

$$\begin{aligned}
q\sum_{1\le h\le q}\left|\sum_{d\backslash q}\frac{\mu(d)}{d}Z\left(\frac{q}{d},h\right)\right|^2 &= \frac{1}{q}\sum_{1\le h\le q}|T(q,h)|^2 \\
&= \frac{1}{q}\sum_{1\le h\le q}\sum_{\substack{1\le a,b\le q\\ a\perp q,b\perp q}} S\left(\frac{a}{q}\right)\overline{S\left(\frac{b}{q}\right)}e\left(\frac{(b-a)h}{q}\right) \\
&= \frac{1}{q}\sum_{\substack{1\le a,b\le q\\ a\perp q,b\perp q}} S\left(\frac{a}{q}\right)\overline{S\left(\frac{b}{q}\right)}\sum_{1\le h\le q} e\left(\frac{(b-a)h}{q}\right) \\
&= \sum_{\substack{1\le a\le q\\ a\perp q}}\left|S\left(\frac{a}{q}\right)\right|^2.
\end{aligned}$$

□

THEOREM 4.2.2. [**Mon68**] *Let $Z(q,h)$ and Z be defined as before, and let $x \ge 1$. For each prime $p \le x$ let $H(p)$ be the union of $\omega(p)$ distinct residue classes modulo p. Let a_n be complex numbers that satisfy*

$$a_n = 0 \text{ if } n \in H(p) \text{ for some } p \le x.$$

Then for each $q \le x$,

$$\mu^2(q)|Z|^2\prod_{p\backslash q}\frac{\omega(p)}{p-\omega(p)} \le q\sum_{1\le h\le q}\left|\sum_{d\backslash q}\frac{\mu(d)}{d}Z\left(\frac{q}{d},h\right)\right|^2.$$

Proof : This is clearly true if $\mu(q) = 0$, so we may assume $q \le x$ is a fixed squarefree integer. If $d\backslash q$, we define

$$K(d) = \left\{h \mid 1 \le h \le q \text{ and if } p\backslash d, \text{ then } h \in H(p), \text{ while if } p\backslash\frac{q}{d}, \text{ then } h \notin H(p)\right\}.$$

Defining $h_1 \equiv h_2$ if there is a d such that $\{h_1,h_2\} \subseteq K(d)$ yields an equivalence relation. Thus $K(d)$ when going through all the divisors of q gives a partition of $\{1,\cdots,q\}$. Now for each h we can write q uniquely as

$$q = \Bigg(\prod_{\substack{p:h\in H(p)\\ p\backslash q}} p\Bigg)\Bigg(\prod_{\substack{p:h\notin H(p)\\ p\backslash q}} p\Bigg).$$

Thus we can write any sum of the form

$$\sum_{1\le h\le q} f(h)$$

as

$$\sum_{d\backslash q}\sum_{h\in K(d)} f(h).$$

Fix a, δ where $\delta \backslash q$. Observe that

$$\left|\sum_{d\backslash q} \mu\left(\frac{q}{d}\right) d \sum_{h\in K(\delta)} Z(d,h)\right|^2 = \left|\sum_{d\backslash q} \frac{\mu(d)q}{d} \sum_{h\in K(\delta)} Z\left(\frac{q}{d},h\right)\right|^2 \tag{4.32}$$

$$= \left|\sum_{h\in K(\delta)\ \sum_{d\backslash q}} \frac{\mu(d)d}{q} Z\left(\frac{q}{d},h\right)\right|^2 \tag{4.33}$$

by changing the variable of summation from d to $\frac{q}{d}$.

Using the Cauchy-Schwarz inequality

$$\left|\sum_{h\in K(\delta)\ \sum_{d\backslash q}} \frac{\mu(d)d}{q} Z\left(\frac{q}{d},h\right)\right|^2 \leq \left(\sum_{h\in K(\delta)} 1\right)\left(\sum_{h\in K(\delta)}\left|\sum_{d\backslash q} \frac{\mu(d)d}{q} Z\left(\frac{q}{d},h\right)\right|^2\right). \tag{4.34}$$

Now consider

$$\left|\sum_{d\backslash q} \mu\left(\frac{q}{d}\right) d \sum_{h\in K(\delta)} Z(d,h)\right|^2.$$

Supposing $\gcd(\delta, d) > 1$, we can select a prime p such that $p\backslash \gcd(\delta, d)$. Then $Z(d,h)$ is a sum of a_n with $n \equiv h \mod d$, since $p\backslash d$ we also have $n \equiv h \mod p$. But $p\backslash\delta$ and $h \in K(\delta)$ implies that $n \in H(p)$ by the definition of $K(\delta)$. Thus by hypothesis $a_n = 0$ whenever $n \equiv h \mod d$ and $h \in K(\delta)$. Hence the inner sum of

$$\left|\sum_{d\backslash q} \mu\left(\frac{q}{d}\right) d \sum_{h\in K(\delta)} Z(d,h)\right|^2$$

vanishes when $\gcd(\delta, d) > 1$. Thus we obtain,

$$\sum_{d\backslash q} \mu\left(\frac{q}{d}\right) d \sum_{h\in K(\delta)} Z(d,h) = \sum_{d\backslash\left(\frac{q}{\delta}\right)} \mu\left(\frac{q}{d}\right) d \sum_{h\in K(\delta)} Z(d,h).$$

Fix d with $d\backslash(q/\delta)$. If $k \in H(p)$, then $Z(d,k) = 0$, and hence

$$\sum_{h\in K(\delta)} Z(d,h) = \sum_{\substack{1\leq k\leq d \\ \forall p\backslash d: k\notin H(p)}} Z(d,k)\left|\{h \mid h \in K(\delta), h \equiv k \mod d.\}\right|.$$

Let $S(\delta, d, k) = \left|\{h \mid h \in K(\delta), h \equiv k \mod d\}\right|$ for k such that $k \in H(p)$ for all primes p that divide d. By the Chinese Remainder Theorem $h \equiv k \mod d$ is equivalent to $h \equiv k \mod p$ for all prime p dividing d. Also $h \in K(\delta)$ implies that $h \in H(p)$ for all primes p dividing δ, and that $h \notin H(p)$ for all primes p dividing q/δ. Summarizing, we have shown that $h \in K(\delta)$ iff the following are satisfied:

1. $p\backslash d \Rightarrow h \equiv k \mod p, h \notin H(p)$
2. $p\backslash\delta \Rightarrow h \in H(p)$ and
3. $p\backslash\left(q/d\delta\right) \Rightarrow h \notin H(p)$.

Since we have k such that $k \notin H(p)$ for all primes p dividing d, the second condition in (1) is satisfied whenever the first is satisfied. We have that if $p\backslash d$, then there are exactly one solution of (1) modulo p, $\omega(p)$ solutions of (2) modulo p, while if $p\backslash\left(q/d\delta\right)$, then there are $p - \omega(p)$ solutions to (3) modulo p.

Applying the Chinese Remainder Theorem, we have

$$\begin{aligned} S(\delta, d, k) &= \left|\{h \mid 1 \leq h \leq q, h \text{ satisfies conditions } (1), (2)\&(3)\}\right| \\ &= \prod_{p\backslash\delta} \omega(p) \prod_{p\backslash(q/d\delta)} (p - \omega(p)). \end{aligned}$$

This number is independent of k, and so

$$\begin{aligned}\sum_{h\in K(\delta)} Z(d,h) &= \sum_{\substack{1\le k\le d\\ \forall p\backslash d\,:\, k\notin H(p)}} Z(d,k)\prod_{p\backslash\delta}\omega(p)\prod_{p\backslash q/d\delta}(p-\omega(p))\\ &= \sum_{1\le k\le d} Z(d,k)\prod_{p\backslash\delta}\omega(p)\prod_{p\backslash q/d\delta}(p-\omega(p))\\ &= Z\prod_{p\backslash\delta} f(p)\prod_{p\backslash q/d\delta}(p-\omega(p)).\end{aligned}$$

From this we get

$$\sum_{d\backslash q}\mu\left(\frac{q}{d}\right)d\sum_{h\in K(\delta)} Z(d,h) = \sum_{d\backslash q/\delta}\mu\left(\frac{q}{d}\right)dZ\prod_{p\backslash\delta}\omega(p)\prod_{p\backslash q/d\delta}(p-\omega(p)) \tag{4.35}$$

$$= \mu(q)Z\prod_{p\backslash\delta}\omega(p)\prod_{p\backslash q/\delta}(p-\omega(p))\sum_{d\backslash q/\delta}\mu(d)d\prod_{p\backslash d}(p-\omega(p))^{-1} \tag{4.36}$$

$$= \mu(q)Z\prod_{p\backslash\delta} f(p)\prod_{d\backslash q/\delta}(p-\omega(p))\prod_{p\backslash q/\delta}\left(1-\frac{p}{p-\omega(p)}\right) \tag{4.37}$$

$$= \mu(\delta)Z\prod_{p\backslash\delta}\omega(p)\prod_{p\backslash q/\delta}\omega(p) \tag{4.38}$$

$$= \mu(\delta)Z\prod_{p\backslash q}\omega(p). \tag{4.39}$$

Now

$$\begin{aligned}\sum_{h\in K(\delta)} 1 &= S(\delta,1,1)\\ &= \prod_{p\backslash\delta}\omega(p)\prod_{p\backslash q/\delta}(p-\omega(p)).\end{aligned}$$

Dividing (4.32) by the above factor and using (4.35) – (4.39) we find that

$$|Z|^2\prod_{p\backslash q}\omega(p)^2\prod_{p\backslash\delta}\omega(p)^{-1}\prod_{p\backslash q/\delta}(p-\omega(p))^{-1} \le \sum_{h\in K(\delta)}\left|\sum_{d\backslash q}\frac{\mu(d)q}{d}Z\left(\frac{q}{d},h\right)\right|^2.$$

Summing over all $\delta\backslash q$, the right hand side yields

$$\sum_{1\le h\le q}\left|\sum_{d\backslash q}\frac{\mu(d)q}{d}Z\left(\frac{q}{d},h\right)\right|^2.$$

Since the $K(\delta)$ partition $\{1,\cdots,q\}$, summing the left hand side yields

$$\begin{aligned}|Z|^2\left(\prod_{p\backslash d}\right)^2\sum_{\delta\backslash q}\left(\prod_{p\backslash\delta}\omega(p)\right)^{-1}\left(\prod_{p\backslash q/\delta}(p-\omega(p))\right)^{-1} &= |Z|^2\prod_{p\backslash q}\omega(p)\sum_{\delta\backslash q}\prod_{p\backslash q/\delta}\omega(p)\prod_{p\backslash q/\delta}(p-\omega(p))^{-1}\\ &= |Z|^2\prod_{p\backslash q}\omega(p)\prod_{p\backslash q}\left(1+\frac{\omega(p)}{p-\omega(p)}\right)\\ &= q|Z|^2\prod_{p\backslash q}\frac{\omega(p)}{p-\omega(p)}.\end{aligned}$$

□

THEOREM 4.2.3. [**MV73**] *Let $\mathcal{N}$ be a set of Z integers in an interval $[M+1, M+N]$. For each prime p let $\omega(p)$ denote the number of residue classes modulo p that contain no element of $\mathcal{N}$. Then*

$$Z \leq L^{-1},$$

where

$$L = \sum_{q \leq z} \left(N + \frac{3}{2}qz\right)^{-1} \mu^2(q) \prod_{p \leq q} \frac{\omega(p)}{p - \omega(p)}$$

and z is an arbitrary positive real number.

Proof : Let x_r be the numbers $\frac{a}{q}$ where $1 \leq a \leq q$, $a \perp q$ and $q \leq z$. If $\frac{a'}{q'} \neq \frac{a}{q}$, then

$$\left\| \frac{a}{q} - \frac{a'}{q'} \right\| \geq \frac{1}{qq'} \geq \frac{1}{qz}.$$

By Theorem (4.1.7) we have

$$\sum_{q \leq z} \left(N + \frac{3}{2}qz\right)^{-1} \sum_{\substack{1 \leq a \leq q \\ a \perp q}} \left| S\left(\frac{a}{q}\right) \right|^2 \leq \sum_{M+1 \leq n \leq M+N} |a_n|^2.$$

Set $a_n = 1$ or 0 according as $n \in \mathcal{N}$ or $n \notin \mathcal{N}$. Then by Theorem (4.2.2) we get

$$Z^2 \mu^2(q) \prod_{p \backslash q} \frac{\omega(p)}{p - \omega(p)} \leq \sum_{\substack{1 \leq a \leq q \\ a \perp q}} \left| S\left(\frac{a}{q}\right) \right|^2.$$

The right hand side equals Z and this proves the theorem. □

4.3. The Brun-Titchmarsh Theorem revisited

The large sieve can be used to strengthen the Brun-Titchmarsh theorem (Theorem 3.2.5). We require the following lemma.

LEMMA 4.3.1. *Let u and v be any positive real numbers. Then*

$$\sum_{\substack{q \leq n \\ q \perp k}} (1 + vq)^{-1} \frac{\mu^2(q)}{\varphi(q)} \geq \frac{\varphi(k)}{k} \sum_{q \leq u} (1 + vq)^{-1} \frac{\mu^2(q)}{\varphi(q)}.$$

Proof : Note that

$$\frac{k}{\varphi(k)} = \sum_{r \backslash k} \frac{\mu^2(r)}{\varphi(r)}.$$

Multiplying the sum on the left by this we get

$$\sum_{\substack{q \leq n \\ q \perp k}} (1 + vq)^{-1} \sum_{r \backslash k} \frac{\mu^2(qk)}{\varphi(qk)},$$

which includes all the terms of the sum on the right. □

THEOREM 4.3.2 ([**MV73**]). *Let x and y be positive real numbers, and let k and l be relatively prime positive integers. Then*

$$\pi(x + y; k, l) - \pi(x; k, l) < \frac{2y}{\varphi(k)\left(\frac{5}{6} + \log\left(\frac{y}{k}\right)\right)}.$$

Proof : We take

$$M = \left\lfloor \frac{x-l}{k} \right\rfloor$$

and

$$N = \left\lfloor \frac{x+y-k}{k} \right\rfloor - M.$$

Let $\mathcal{N}$ be the set of those integers n for which $M < n \leq M+N$, $kn+l$ is prime, and $kn+l > z$. Then $\omega(p) = 1$ whenever $p \leq z$ and $p \not| k$. Thus by Theorem (4.2.3) we have

$$\pi(x+y;k,l) - \pi(x;k,l) \leq L^{-1} + \pi(z),$$

where

$$L = \sum_{\substack{q \leq z \\ q \perp k}} (N + \frac{3}{2}qz)^{-1} \frac{\mu^2(q)}{\varphi(q)}.$$

Taking $z = \sqrt{\frac{2}{3}N}$ and using Lemma (4.3.1), we have

$$\pi(x+y;k,l) - \pi(x;k,l) < \frac{kN}{\varphi(k)J} + \sqrt{(N)},$$

where

$$J = \sum_{q \leq z} (1 + qz^{-1})^{-1} \frac{\mu^2(q)}{\varphi(q)}.$$

From [**War27**] we have

$$\sum_{q \leq \nu} \frac{\mu^2(q)}{\varphi(q)} = \log \nu + \gamma + \sum_p \frac{\log p}{p(p-1)} + o(1)$$

as $\nu \to \infty$.
By partial summation we find that

$$J = \log z + \gamma + \sum_p \frac{\log p}{p(p-1)} = \log 2 + o(1)$$

as $z \to \infty$. Setting $z = \sqrt{\frac{2}{3}N}$ we get

$$J = \frac{1}{2}\log N + \gamma + \sum_p \frac{\log p}{p(p-1)} - \frac{1}{2}\log\frac{3}{2} - \log 2 + o(1)$$

as $N \to \infty$. Since $\gamma > 0.577$,

$$\sum_p \frac{\log p}{p(p-1)} > 0.737,$$

filling in $\log 2 < 0.694$ and $\frac{1}{2}\log\frac{3}{2} < 0.203$, we finally obtain

$$J > \frac{1}{2}\log N + 0.417,$$

for large enough N. $\square$

4.4. Bombieri's Theorem

The large sieve inequalities imply that if a sequence of integers is distributed rather densely in an interval, then it cannot be very unevenly distributed modulo the primes. In this section we will prove an important theorem that quantifies the above statement for the primes themselves.
Define

$$\psi(x) = \sum_{n \le x} \Lambda(n),$$

where $\Lambda(n)$ is von-Mangoldt's function

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

Also define

$$\psi(x; q, a) = \sum_{\substack{n \le x \\ n \equiv a \mod q}} \Lambda(n).$$

Let

$$E(x; q, a) = \psi(x; q, a) - \frac{x}{\varphi(q)}$$

for $a \perp q$, and

$$E^*(x, q) = \max_{y \le x} E(y, q).$$

We will prove Bombieri's Theorem in the following form:
THEOREM 4.4.1 ([**Dav80**]). *Let $A > 0$ be fixed, and suppose $x^{\frac{1}{2}}(\log x)^{-A} \le Q \le x^{\frac{1}{2}}$. Then*

$$\sum_{q \le Q} E^*(x, q) \ll x^{\frac{1}{2}} Q (\log x)^5.$$

Proof : If χ is a multiplicative character modulo q, and define

$$\psi(y, \chi) = \sum_{n \le y} \chi(n)\Lambda(n).$$

We begin with the identity

$$\psi(y; q, a) = \frac{1}{\varphi(q)} \sum_{\chi} \overline{\chi}(a)\psi(y, \chi),$$

where the sum is over all the characters modulo q. Let χ_0 be the principal character we then define

$$\psi'(y, \chi) = \begin{cases} \psi(y, \chi) & \text{if } \chi \ne \chi_0, \\ \psi(y, \chi_0) - y & \text{if } \chi = \chi_0. \end{cases}$$

Then we have

$$\psi(y; q, a) - \frac{y}{\varphi(q)} = \frac{1}{\varphi(q)} \sum_{\chi} \overline{\chi}(a)\psi'(y, \chi),$$

and so

$$|E(y; q, a)| \le \frac{1}{\varphi(q)} \sum_{\chi} |\psi'(y, \chi)|$$

since $|\chi(a)| \le 1$. This estimate is independent of a, so that

$$E^*(y; q) \le \frac{1}{\varphi(q)} \sum_{\chi} |\psi'(y, \chi)|.$$

If $\chi \mod q$ is a character (possibly imprimitive) that is induced by $\chi_1 \mod q_1$, where χ_1 is primitive, then $\psi'(y,\chi)$ and $\psi'(y,\chi_1)$ do not differ very much:

$$\begin{aligned}\psi(y,\chi_1)-\psi'(y,\chi) &= \sum_{\substack{p^k\le y\\ p\backslash q}} \chi_1(p^k)\log p\\ &\ll \sum_{p\backslash q}\left\lfloor\frac{\log y}{\log p}\right\rfloor \log p\\ &\ll (\log y)\sum_{p\backslash q}\log p\\ &\ll (\log qy)^2.\end{aligned}$$

Hence we can replace the sum over all characters by one over the primitive characters only. Thus

$$E(x,q) \ll (\log qx)^2 + \frac{1}{\varphi(q)}\sum_{\chi}\left|\psi'(y,\chi_1)\right|,$$

and

$$E^*(x,q) \ll (\log qx)^2 + \frac{1}{\varphi(q)}\sum_{\chi}\max_{y\le x}\left|\psi'(y,\chi_1)\right|.$$

We can combine the contributions from each of the primitive characters. Since a primitive character induces characters to moduli that are multiples of q, we have

$$E^*(x,q) \ll (\log qx)^2 + \sum_{q\le Q}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi'(y,\chi)|\left(\sum_{k\le Q/q}\frac{1}{\varphi(kq)}\right),$$

where $\sum^*$ means the sum is over primitive characters modulo q.
Since $\varphi(kq)\ge\varphi(k)\varphi(q)$ we have

$$\sum_{k\le z}\frac{1}{\varphi(kq)} \le \frac{1}{\varphi(q)}\sum_{k\le z}\frac{1}{\varphi(k)}.$$

Now

$$\sum_{k\le z}\frac{1}{\varphi(k)} \le \prod_{p\le z}\left(1+\frac{1}{(p-1)}+\frac{1}{p(p-1)}+\frac{1}{p^2(p-1)}+\cdots\right).$$

Note that

$$\frac{1}{(p-1)}\frac{1}{\left(1-\frac{1}{p}\right)} = \frac{1}{p-1}\left(1+\frac{1}{p}+\frac{1}{p^2}+\cdots\right).$$

Thus

$$\begin{aligned}\left(1+\frac{1}{(p-1)}+\frac{1}{p(p-1)}+\frac{1}{p^2(p-1)}+\cdots\right) &= 1+\frac{1}{(p-1)}\frac{1}{\left(1-\frac{1}{p}\right)}\\ &= \left(1+\frac{1}{p(p-1)}\right)\frac{1}{\left(1-\frac{1}{p}\right)}.\end{aligned}$$

Using this we have

$$\begin{aligned}\sum_{k\le z}\frac{1}{\varphi(k)} &\le \prod_{p\le z}\left(1-\frac{1}{p}\right)^{-1}\left(1+\frac{1}{p(p-1)}\right)\\ &\ll \log z,\end{aligned}$$

and so

$$\sum_{q\le Q}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi'(y,\chi)|\Big(\sum_{k\le Q/q}\frac{1}{\varphi(kq)}\Big)\ll \log x\sum_{q\le Q}\frac{1}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}\big|\psi'(y,\chi)\big|.$$

Thus it suffices to show that

$$\sum_{q\le Q}\frac{1}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi'(y,\chi)|\ll x^{\frac12}Q(\log x)^4 \tag{4.40}$$

for $x^{\frac12}(\log x)^{-A}\le Q\le x^{\frac12}$.

Using the large sieve we will show that

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi(y,\chi)|\ll\big(x+x^{\frac56}Q+x^{\frac12}Q^2\big)(\log Qx)^4 \tag{4.41}$$

for all $x\ge 1$ and $Q\ge 1$.

Now observe that

$$\sum_{U<q\le 2U}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi(y,\chi)|\ge U\sum_{U<q\le 2U}\frac{1}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi(y,\chi)|,$$

and so

$$\sum_{U\le q\le 2U}\frac{1}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi(y,\chi)|\ll\Big(\frac{x}{U}+x^{\frac56}+x^{\frac12}U\Big)(\log Ux)^4$$

by (4.41).
Summing over $U=2^k$ for $k\le\log Q$, we have

$$\sum_{Q_1<q\le Q}\frac{1}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{y\le x}|\psi(y,\chi)|\le\Big(\frac{x}{Q_1}+x^{\frac56}\log Q+x^{\frac12}Q\Big)(\log Qx)^4.$$

We have used the fact that for $\chi=\chi_0$ we have $|\psi'(y,\chi_0)|\le|\psi(y,\chi_0)|$, and $\psi'(y,\chi)=\psi(y,\chi)$ if $\chi\ne\chi_0$.

This shows (4.40) for $Q_1=\log^A x$. By the Siegel-Walfisz theorem, if χ is a primitive character modulo q, $q\le(\log x)^A$, and $y\le x$, then

$$|\psi'(y,\chi)|\ll x(\log x)^{-2A}.$$

Thus the theorem follows from (4.41).

We will now sketch the proof of (4.41) (for details see [**Dav80**]). Using the large sieve we can derive the following:

$$\sum_{q\le Q}\frac{q}{\varphi(q)}\sideset{}{^*}\sum_{\chi}\max_{u}\Big|\sum_{1\le m\le M}\sum_{\substack{1\le n\le N\\ mn\le u}}a_mb_n\chi(mn)\Big| \tag{4.42}$$

$$\ll(M+Q^2)^{\frac12}(N+Q^2)^{\frac12}\Big(\sum_{1\le m\le M}|a_m|^2\Big)^{\frac12}\Big(\sum_{1\le n\le N}|b_n|^2\Big)^{\frac12}\log 2MN. \tag{4.43}$$

If $Q^2>x$ then (4.41) follows from above with $M=1$, $a_1=1$, $b_n=\Lambda(n)$, $N=x$. Thus we may assume $Q^2\le x$. It turns out that we can write

$$\psi(y,\chi)=S_1+S_2+S_3+S_4,$$

where

$$S_1 = \sum_{n \le U} \Lambda(n)\chi(n) \ll U,$$

$$S_2 = -\sum_{t \le UV} \Bigg(\sum_{\substack{t=md \\ m \le U \\ d \le V}} \mu(d)\Lambda(m) \Bigg) \sum_{r \le y/t} \chi(rt),$$

$$S_3 \ll (\log y) \sum_{d \le V} \max_{w} \Bigg| \sum_{w \le h \le y/d} \chi(h) \Bigg|, \text{ and}$$

$$S_4 = \sum_{U < m \le y/V} \Lambda(m) \sum_{V < k \le y/m} \Bigg(\sum_{\substack{d \backslash k \\ d \le V}} \mu(d) \Bigg) \chi(mk).$$

Using (4.42) and the Pólya-Vinogradov inequality (see [**Dav80**]), we can show that

$$\sum_{q \le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi} \max_{y \le x} |S_4| \ll \big(Q^2 x^{\frac{1}{2}} + QxU^{-\frac{1}{2}} + QxV^{-\frac{1}{2}} + x\big)(\log x)^4.$$

The sum S_2 can be split into $S_2 = \sum_{t \le UV} = \sum_{t \le U} + \sum_{U < t < UV} = S_2' + S_2''$, and it can be shown that

$$\sum_{q \le Q} \sideset{}{^*}\sum_{\chi} \max_{y \le x} |S_2''| \ll (Q^2 x^{\frac{1}{2}} + QxU^{-\frac{1}{2}} + Qx^{\frac{1}{2}} U^{\frac{1}{2}} V^{\frac{1}{2}} + x)(\log x)^2$$

and

$$\sum_{q \le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi} \max_{y \le x} |S_2'| \ll (Q^{\frac{5}{2}} U + x)(\log Ux)^2.$$

Also

$$\sum_{q \le Q} \frac{q}{\varphi(q)} \sideset{}{^*}\sum_{\chi} \max_{y \le x} |S_3| \ll (Q^{\frac{5}{2}} V + x)(\log Vx)^2.$$

On combining these estimates and taking $U = V = x^{\frac{2}{3}} Q^{-1}$ for $x^{\frac{1}{3}} \le Q \le x^{\frac{1}{2}}$, we obtain (4.41) in this range. For $Q \le x^{\frac{1}{3}}$, we can take $U = x^{\frac{1}{3}}$ to complete the proof of (4.41). □

The Bombieri result can be formulated as follows:

THEOREM 4.4.2. *Let* $E(x; q, a) = \pi(x; q, a) - \frac{\mathrm{li}\, x}{\varphi(q)}$ *for* $a \perp q$, $E(x; q) = \max_{a, a \perp q} \big|E(x; q, a)\big|$, *and* $E^*(x, q) = \max_{y \le x} E(y, q)$. *Then for all* $A > 0$ *there exists* $B > 0$ *such that*

$$\sum_{q \le x^{\frac{1}{2}} (\log x)^{-B}} E^*(x, q) \ll \frac{x}{\log^{1+A} x}.$$

4.5. Prime and Squarefree pairs

We can pose the following variation of the twin prime problem: *"Are there infinitely many primes* p *such that* $p + 2$ *is squarefree?"* The answer to the question is yes, and this is an almost immediate consequence of the powerful result we have proved.

THEOREM 4.5.1. *Let*

$$\Xi(x) = \big|\{p \le x \mid \mu^2(p+2) = 1\}\big|.$$

Then

$$\Xi(x) = \mathrm{Li}\ (x) \Bigg\{ \prod_{p > 2} \Bigg(1 - \frac{1}{p(p-1)} \Bigg) + O\Bigg(\frac{\ln x}{\sqrt{x}} \Bigg) \Bigg\} + O\Bigg(\frac{x}{\ln^{1+U}(x)} \Bigg) + O(x^{\frac{3}{4}} \ln^C(x))$$

for some constants $U > 0$ *and* $C > 0$.

Proof : Let $\mathcal{A} = \{p+2 \mid p \le x\}$. We have

$$\Xi(x) = \sum_{d^2 \le x} \mu(d) \Bigg(\sum_{\substack{n \in A \\ d^2 \backslash n}} 1 \Bigg).$$

Let $\mathcal{A}_{d^2} = \{p+2 \mid p \le x, p+2 \equiv 0 \mod d^2\}$. Thus by definition $|\mathcal{A}_{d^2}| = \pi(x; d^2, -2)$. Define

$$R_{d^2} = \pi(x; d^2, -2) - \frac{\text{Li }(x)}{\varphi(d^2)}.$$

Then we have

$$\begin{aligned} \Xi(x) &= \text{Li }(x) \sum_{d^2 \le x} \frac{\mu(d)}{\varphi(d^2)} + \sum_{d^2 \le x} \mu(d) |R_{d^2}| \\ &= \Sigma_1 + \Sigma_2 \end{aligned}$$

$$\begin{aligned} \Sigma_1 &= \text{Li }(x) \Bigg(\sum_{d} \frac{\mu(d)}{\varphi(d^2)} - \sum_{d > \sqrt{x}} \frac{\mu(d)}{\varphi(d^2)} \Bigg) \\ &= \text{Li }(x) \Bigg(\prod_{p > 2} \bigg(1 - \frac{1}{p(p-1)} \bigg) - \sum_{d > \sqrt{x}} \frac{\mu(d)}{\varphi(d^2)} \Bigg), \end{aligned}$$

since $|A_4| = 0$ allows omitting the prime 2.

The second sum can be upper-bounded by:

$$\begin{aligned} \sum_{d > \sqrt{x}} \frac{1}{\varphi(d^2)} &\le \sum_{d > \sqrt{x}} \frac{2 \ln d}{d^2} \\ &= O\bigg(\frac{\ln x}{\sqrt{x}} \bigg). \end{aligned}$$

The remainder term is bounded by:

$$\begin{aligned} \Sigma_2 &\le \sum_{d^2 \le x} |R_{d^2}| \\ &= \sum_{d^2 \le \frac{\sqrt{x}}{\ln^C x}} |R_{d^2}| + \sum_{\frac{\sqrt{x}}{\ln^C x} < d^2 < x} |R_{d^2}| \\ &= O\bigg(\frac{x}{\ln^{1+U} x} \bigg) + \sum_{\frac{\sqrt{x}}{\ln^C x} < d^2 < x} |R_{d^2}| \end{aligned}$$

using Bombieri's result to bound the first sum.

For the second sum, since $|R_{d^2}| \le \lfloor \frac{x}{d^2} \rfloor \le \frac{x}{d^2}$, we have

$$\sum_{\frac{\sqrt{x}}{\ln^C x}<d^2<x} |R_{d^2}| \leq x \sum_{\frac{\sqrt{x}}{\ln^C x}<d^2} \frac{1}{d^2}$$
$$= O\left(\frac{x\ln^C x}{x^{\frac{1}{4}}}\right)$$
$$= O\left(x^{\frac{3}{4}} \ln^C x\right).$$

The theorem follows from the estimates for Σ_1 and Σ_2. □

Let $\Psi(x) = \sum_{n\leq x} \Lambda'(n)$, where

$$\Lambda'(n) = \begin{cases} \log p & \text{if } n = p^k \text{ and } \mu^2(n+2) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$\Psi(x; q, a) = \sum_{\substack{n\leq x \\ n\equiv a \mod q}} \Lambda'(n),$$

and further let $E(x; q, a) = \Psi(x; q, a) - \frac{Cx}{\varphi(q)}$, $E(x; q) = \max_{a, a\perp q} |E(x; q, a)|$, and $E^*(x, q) = \max_{y\leq x} E(y, q)$, where $C = \prod_p \left(1 - \frac{1}{p(p-1)}\right)$.

Using partial summation and the above theorem we can show that for any $U > 0$,

$$\Psi(x) = Cx + O\left(\frac{x}{\log^{1+U} x}\right)$$

and

$$\Psi(x; q, a) = \frac{Cx}{\varphi(q)} + O\left(\frac{x}{\log^{1+U} x}\right), \text{ for } a \perp q.$$

THEOREM 4.5.2. *Let $A > 0$ be fixed. Then*

$$\sum_{(\log x)^A<q\leq Q} E^*(x, q) \ll x^{\frac{1}{2}} Q(\log x)^5,$$

provided $x^{\frac{1}{2}}(\log x)^{-A} \leq Q \leq x^{\frac{1}{2}}$.

The proof is a careful verification that the proof of the Bombieri Theorem goes through except for $q < (\log x)^A$. But in this range the maximum error possible is $O\left(\frac{x}{\log^{1+U} x}\right)$ so selecting U large enough we have:

THEOREM 4.5.3. *Let $A > 0$ be fixed. Then*

$$\sum_{q\leq Q} E^*(x, q) \ll x^{\frac{1}{2}} Q(\log x)^5,$$

provided $x^{\frac{1}{2}}(\log x)^{-A} \leq Q \leq x^{\frac{1}{2}}$.

There is a version of Brun's sieve that makes use of the result on the average behaviour of error terms to yield a better estimate. In particular we have ([**HR74**] Theorem 2.1′ p. 65)

THEOREM 4.5.4. *Let the following conditions hold on the sequence $\mathcal{A}$:*

(1)

$$1 \leq \frac{1}{1 - \frac{\omega(p)}{p}} \leq A_1;$$

(2)

$$\sum_{w \le p \le z} \frac{\omega(p)\log p}{p} \le \kappa \log\frac{z}{w} + A_2, \text{ if } 2 \le w \le z;$$

(3) There is a constant A_0' *such that*

$$|R_d| \le L\left(\frac{x\log x}{d} + 1\right) A_0'^{\nu(d)};$$

(4) For every postive constant $U \ge 1$ *there is a* C_0 *such that*

$$\sum_{d < x^{\alpha}\log^{-c_0} x} \mu^2(d)|R_d| = O\left(\frac{x}{\log^{\kappa+U} x}\right).$$

Let b *be a positive integer, let* λ *be a real number satisfying* $\lambda e^{1+\lambda} < 1$, *let*

$$c_1 = \frac{A_2}{2}\left(1 + A_1\kappa + \frac{A_1 A_2}{\log 2}\right),$$

and let $u = \frac{\log x}{\log z}$.
Then

$$\begin{aligned} S(\mathcal{A};P,z) \ge xW(z)\Bigg\{ & 1 - 2\frac{\left(\lambda^b e^{\lambda}\right)^2}{1-\left(\lambda e^{1+\lambda}\right)^2}\exp\left((2b+2)\frac{c_1}{\lambda\log z}\right) \\ & + O\left(Lz^{-\alpha u + 2b - 1 + \frac{2.01}{(e^{2\lambda/\kappa}-1)}} u^{C_0+1}\log^{C_0+\kappa+1} z\right) \\ & + O(u^{-\kappa}\log^{-U} X)\Bigg\}, \end{aligned}$$

where the O*-constants may depend on* $A_0', A_1, A_2, \kappa, \alpha$ *and* U, *but not on* λ *or* b.

Using this theorem with $\mathcal{A} = \{p+2 \mid p \le x, \mu^2(p+2) = 1\}$, and taking the sifting primes to be $P = \{p : p > 2\}$, we find that the lower bound is positive (and diverges) for $u < 9$. Following the same analysis as in [**HR74**] (p.67), we can also take $u < 8$ with a slightly better treatment of the principal and secondary terms involved in the proof of the above theorem. This allows us to conclude that the lower bound diverges even with $z = x^{\frac{1}{7}}$, and thus we have:

THEOREM 4.5.5. *There are infinitely many primes* p *such that* $p+2$ *is a squarefree number with at most 7 prime factors.*

The above result is different from earlier ones because of the extra condition that $p+2$ be made up only of distinct primes.

Bibliography

[BakHar98] Baker R. C., Harman G., *Shifted primes without large prime factors*, Acta Arith., **(83)**, 331-361, (1998).

[BakHar95] Baker R. C., Harman G., *The Brun-Titchmarsh Theorem on average*, Analytic Number Theory Vol 1, (Allerton Park, IL), 39-103, Progr. Math. 138, Birkhauser Boston, Boston, (1996),

[BakPin85] Baker R. C., Pintz J., *The distribution of square-free numbers*, Acta Arith., **(46)**, 71-79, (1985).

[Be83] Beth Thomas, *Eine Bemerkung zur Abschätzung der Anzahl orthogonaler lateinischer Quadrate mittels Siebverfahren.* Abh. Math. Sem. Univ. Hamburg, **53**, 284-288, (1983).

[BPS60] Bose R. C, Shrikande S. S., Parker E. T., *Further results on the construction of mutually orthogonal latin squares and the Falsity of Euler's conjecture*, Canad. J. Math. **12**, 189-203, 1960.

[Bru16] Brun V., *Omfordelingen av primtallene i forskjellige talklasser. En vre begrnsning.* , Nyt Tiddsskr. f. Math. **(27)** B, 45-58, (1916).

[Bru19] Brun V., *Le crible d'Eratostne et le thorme de Goldbach*, C. R. Acad. Sci. Paris, **(168)**, 544-546, (1919).

[Bru22] Brun V., *Das Siev des Eratosthenes*, 5. Skand. Mat. Knogr., Helsingfors, 197-203, (1922).

[CES60] Chowla S., Erdős P., Straus E. G., *On the maximal number of pairwise orthogonal Latin suqares of a given order* , Canad. J. Math. 12, 204-208, 1960.

[Che73] Chen J., *On the representation of a large even integer as the sum of a prime and the product of at most two primes*, Sci. Sinica, **(16)**, 157-176, (1973).

[Dar96] Dartyge Cécile, *Le plus grand facteur de* n^2+1 *oú* n *est presque premier.*, Acta Arith. **(76)**, no. 3, 199-226, (1996).

[Dav80] Harold Davenport, Montgomery H. L., *Multiplicative Number Theory*, 2nd ed., Springer-Verlag, (1980).

[DI83] Deshouillers, J.-M., Iwaniec Henryk, *On the greatest prime factor of* n^2+1, Ann. Inst. Fourier (Grenoble), **(32)**, no. 4., 1-11, (1983).

[Erd52] Erdős, Pál, *On the greatest prime factor of* $\prod f(k)$ J. Lond. Math. Soc., **(27)**, 379-384, (1952).

[Erd60] Erdős, Pál, *Über die kleinste quadratfreie Zahl einer arithmetischen Reihe*, Monatsh. Math. **64**, (1960), 314-316.

[Erd49] Erdős, Pál, *On some applications of Brun's method*, Acta Univ. Szeged. Sect. Sci. Math. **13**, (1949), 57-63.

[Est31] Estermann, Theodor, *Einige Sätze über quadratfreie Zahlen*, Math. Ann. vol. **(105)**, 1931. 653-662.

[Gal67] Gallagher P. X., *The large sieve*, Mathematika, **(14)**, 14-20, (1967).

[GeL66] Gel'Fond A. O., Linnik Yu. V., *Elementary Methods in the Analytic Theory of Numbers*, MIT - Press, (1966).

[Hal70] Halberstam, H.; *On integers all of whose prime factors are small*, Proc. Lond. Math. Soc. (3), **(21)**, 102-107, 1970.

[HR74] Halberstam, H.; Richert, H.-E., *Sieve Methods*, Academic Press, 1974.

[HalRo66] Halberstam, H.; Roth, K. F., *Sequences*, Oxford University Press, (1966).

[HB84] Heath-Brown D. R., *The Square Sieve and Consecutive Square-Free Numbers.* Math. Ann. **(266)**, (1984), 251-259.

[HB88] Heath-Brown D. R., *The number of primes in a short interval.* J. Reine Angew. Math. **(389)**, 22-63, (1988).

[Ho67] Hooley C., *On the greatest prime factor of a quadratic polynomial*, Acta Math., **(17)**, 281-299, (1967).

[Ho73] Hooley C., *On the largest prime factor of* $p+a$, Mathematika, **(20)**, 135-143, (1973).

[Ho76] Hooley C., *Applications of sieve methods to the theory of numbers*, Cambridge University Press, (1976).

[Iwan82] Iwaniec Henryk, *On the Brun-Titchmarsh theorem*, J. Math. Soc. Japan, **(34)**, No. 1, 95-123, (1982).

[vLR65] van Lint J. H., Richert H.-E., *On primes in arithmetic progression*, Acta Arith., **(11)**, 209-216, (1965).

[Mir49] Mirsky, L. *On the frequency of pairs of squarefree numbers with a given difference.* Bull. Amer. Math. Soc. **(55)**, 936-939, (1949).

[Mon68] Montgomery H. L., *A note on the large sieve*, J. Lond. Math. Soc., **(43)**, 93-98, (1968).

[MV73] Montgomery H. L., Vaughan R. C., *The Large Sieve*, Mathematika, **(20)**, No. 40, 119-134, (1973).

[MV81] Montgomery H. L., Vaughan R. C., *The Distribution of Squarefree Numbers*, Recent Progress in Analytic Number Theory, Academic Press, 247-256, (1981).

[Mot70] Motohashi Yoichi, *A note on the least prime in an arithmetic progression with a prime difference*, Acta Arith., **(17)**, 283-285, (1970).

[Odl71] Odlyzko Andrew M., *Sieve Methods*, Senior Thesis, California Institute of Technology, Pasadena, California, (1971).

[Rad24] Rademacher Hans, *Beiträge zur Viggo Brunschen Methode in der Zahlentheorie*, Abbh. Math. Sem. Hamburg, **(3)**, 12-30, (1924).

[RS62] Rosser J. B., Schoenfeld L., *Approximate formulas for some functions of prime numbers* , Illinois J. Math. **(6)**, 64-89, (1962).

[Sch66] Schinzel, A., *On sums of roots of unity. (Solution of two problems of R. M. Robinson)*, Acta Arith., **(11)**, 419-432, (1966).

[SchWa58] Schinzel A., Wang Y., *A note on some properties of the functions* $\varphi(n), \sigma(n)$ *and* $\theta(n)$, Ann. Polon. Math. **(4)**, 201-213, (1958).

[Sel47] Selberg A., *On an elementary method in the theory of primes*, Norske Vid. Selsk. Forh. Trondhjem **(19)**, no.18, 64-67, (1947).

[Sel71] Selberg A., *Sieve Methods*, Proc. Symp. Pure Math. **(20)**, 311-351, (1971).

[Tit86] Titchmarsh E. C., *The Theory of the Riemann Zeta-function*, 2nd Ed., Oxford University Press, (1986).

[Wal63] Walfisz Arnold, *Weylsche Exponentialsummen in der neueren zahlentheorie*, Deutscher Verlag der Wissenschaften, Berlin, (1963).

[War27] Ward D. R., *Some series involving Euler's function*, J. Lond. Math. Soc., **(2)**, 210-214, (1927).

[Warl90] Warlimont Richard, *Sieving by large prime factors*, Monatsh. Math., **(109)**, no. 3, 247-256, (1990).

[Wil74] Wilson, Richard M., *Concerning the number of mutually orthogonal Latin squares*, Discrete Math., **(9)** 181-198, (1974).