The Schnirelmann-Goldbach Theorem

MDNT

December 2025

Contents

1 What are Sieves?

Suppose we were tasked to estimate the number of primes in the interval $[N, 2N]$, for some large integer N. One way to approach this problem is by means of the inclusion-exclusion principle. We could first remove all the multiples of 2, remove all the multiples of 3, so a preliminary estimate will be

$$\#\{p \in [N, 2N] : p \text{ prime}\} \approx N - \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{N}{3} \right\rfloor \dots$$

But whilst we were removing the multiples of 2 and 3, we double-counted the multiples of 6, we 'removed' them twice! To correct for this, we will have to add all the multiples of 6 to our estimate, which gives us

$$\#\{p \in [N, 2N] : p \text{ prime}\} \approx N - \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor \frac{N}{3} \right\rfloor + \left\lfloor \frac{N}{6} \right\rfloor \dots$$

It would be too ambitious to aim to continue this process forever, as N is large, so we are forced to be slightly clever about when we stop this process. As a matter of formalising this argument, one would use the *Möbius function* function, defined as follows

Definition 1.1. The *Möbius function* $\mu(n)$ is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^k & \text{if } n \text{ is the product of } r \text{ distinct primes,} \\ 0 & \text{if } n \text{ is divisible by the square of a prime.} \end{cases}$$

As a sanity check, the reader is invited to check that $\mu(2) = \mu(3) = -1,$ and $\mu(6) = 1$, which agrees with our intuition. This is essentially *Brun's Pure Sieve*, developed by Viggo Brun in 1915. Using his sieve [Bru19], in 1919, Brun showed that the sum of the reciprocals of the twin primes is convergent! More specifically, he showed the following

Theorem 1.2. (Brun, 1920) Let $\pi_2(x)$ denote the number of primes p not exceeding x such that $p + 2$ is also prime. Then

$$\pi_2(x) \ll \frac{x(\log\log x)^2}{(\log x)^2}.$$

We would not explore the Brun Sieve in this article. Instead, we will go through the Selberg Sieve, a method developed by the late Fields Medalist Atle Selberg, which in many cases leads to better bounds than Brun's method. Using the sieve, we will prove the Schnirelmann-Goldbach theorem, which states that every integer greater than one is the sum of a bounded number of primes.

2 Preliminaries

We begin with introducing some notation and some results, most of which will be from Foundations or Introduction to Number Theory.

Definition 2.1. Let D be a subset of the complex numbers $\mathbb{C}$ and let $f : D \to \mathbb{C}$ be a complex-valued map defined on D. We will write

$$f(x) = O(g(x))$$

if $g : D \to \mathbb{R}^+$ and there is a positive constant A such that

$$|f(x)| \leq Ag(x)$$

for all $x \in D$.

Remark. We will also use the notation $f(x) \ll g(x)$ or $g(x) \gg f(x)$ to indicate $f(x) = O(g(x))$.

Definition 2.2. An *arithmetic function* is a complex-valued function whose domain is the set of all positive integers.

Definition 2.3. An arithmetic function $f(n)$ is *multiplicative* if

$$f(mn) = f(m)f(n),$$

for all m, n satisfying $\gcd(m, n) = 1$.

Similarly, we define

Definition 2.4. An arithmetic function $f(n)$ is *completely multiplicative* if

$$f(mn) = f(m)f(n),$$

for all $m, n \in \mathbb{N}$.

For ease of notation, we let (m, n) denote the greatest common divisor of the integers m and n, and $[m, n]$ to denote the least common multiple of the integers m and n.

Definition 2.5. A nonempty set $\mathcal{D}$ of positive integers is called *divisor-closed* if whenever $n \in \mathcal{D}$ and d divides n, then $d \in \mathcal{D}$.

Using the properties of the Möbius function, we get the following result.

Theorem 2.6 (Möbius inversion formula)**.** Let $\mathcal{D}$ be a finite divisor-closed set, and let f and g be functions defined on $\mathcal{D}$. If

$$g(n) = \sum_{\substack{d \in \mathcal{D} \\ n|d}} f(d),$$

for all $n \in \mathcal{D}$, then

$$f(n) = \sum_{\substack{d \in \mathcal{D} \\ n|d}} \mu(\frac{d}{n}) g(d)$$

for all $n \in \mathcal{D}$. A similar converse statement also holds.

Proof. Refer to Appendix A □

We introduce the following identity, which we will use in the next chapter.

Lemma 2.7. Let φ denote the Euler totient function, and μ denote the Möbius function. We have the following identity

$$\frac{d}{\varphi(d)} = \sum_{r|d} \frac{\mu(r)^2}{\varphi(r)}.$$

Proof. Observe that both sides of the expression are multiplicative; hence, it suffices to show the result for prime powers. Let $d = p^k$, then

$$\frac{p^k}{\varphi(p^k)} = \frac{p^k}{p^k - p^{k-1}} = \frac{p}{p-1} = 1 + \frac{1}{p-1} = \sum_{r|p^k} \frac{\mu(r)^2}{\varphi(r)}.$$

□

We now introduce some prime-counting functions.

Definition 2.8.

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

and

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

where p is prime.

The following theorem is due to Chebyshev.

Theorem 2.9 (Chebyshev, 1850)**.**

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

Proof. ([Coj05], p.8) We note that

$$\prod_{n<p\leq 2n} p \;\Bigg|\; \binom{2n}{n}.$$

Now since

$$\binom{2n}{n} \leq 2^{2n},$$

after taking logarithms,

$$\vartheta(2n) - \vartheta(n) \leq 2n \log 2.$$

We repeat this argument successivly to get

$$\vartheta(n) - \vartheta(\frac{n}{2}) \leq n \log 2$$
$$\vartheta(\frac{n}{2}) - \vartheta(\frac{n}{4}) \leq \frac{n}{2} \log 2$$

$$\vdots$$

Now by summing the inequalities, we get

$$\vartheta(2n) \leq 4n \log 2.$$

In other words,

$$\vartheta(x) = O(x).$$

Finally

$$\begin{aligned} x \gg \vartheta(x) &\geq \sum_{\sqrt{x}<p\leq x} \log p \\ &\geq \frac{1}{2}(\log x)(\pi(x) - \pi(\sqrt{x})) \\ &\geq \frac{1}{2}(\log x)(\pi(x)) + O(\sqrt{x}\log x). \end{aligned}$$

Now using[1]

$$\frac{x}{\log x} \gg x^{\varepsilon}, \text{for all } 0 < \varepsilon < 1,$$

we conclude that $\pi(x) = O\left(x/\log x\right)$. □

Using the result above, we have the following result.

Lemma 2.10. Let $r(N)$ denote the number of representations of the integer N as the sum of two primes. Then

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

Proof. [Nat96], pp The proof is a counting argument, a common strategy in this article. We note that if p, q are primes such that $p, q \leq x/2$, then $p + q \leq x$. Hence, by Theorem 2.9

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

□

Definition 2.11. To every $n \times n$ symmetric matrix $A = (a_{i,j})$ we associate the *quadratic form* F_A defined by

$$F_A(x_1, \ldots, x_n) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,j}x_i x_j.$$

For example, if we are considering I_n, the $n \times n$ identity matrix, then the associated quadratic form is

$$F_{I_n}(x_1, \ldots, x_n) = x_1^2 + \cdots + x_n^2.$$

The following result is an application of the Cauchy-Schwarz inequality, which we state without proof.

[1]The reader is invited to check this by considering the limit of $x^{\varepsilon}/(x/\log x)$ as $x \to \infty$.

Lemma 2.12. Let $a_1, a_2, \ldots, a_n$ be positive real numbers and $b_1, b_2 \ldots, b_n$ be any real numbers. The minimum value of the quadratic form

$$\mathcal{Q}(y_1, y_2, \ldots, y_n) = a_1 y_1^2 + \cdots + a_n y_n^2$$

subject to the linear constraint $b_1 y_1 + \ldots b_n y_n = 1$ is

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

and this value is attained iff $y_i = \frac{m b_i}{a_i}$, for all $i = 1, 2, \ldots, n$.

Proof. Refer to Appendix A. □

3 Selberg Sieve

In this section, we will go through the Selberg Sieve, introduced by Atle Selberg in a short paper of his back in 1947 [Sel84]. We will use the sieve to obtain an upper bound on the representations of an even integer as the sum of two primes. We begin with a simpler case, with the sifting of an interval.

Let P be a positive integer. Let $S(x, y; P)$ denote the number of intgers $x < n \leq x+y$ such that $(n, P) = 1$. The goal is to derive an upper bound for $S(x, y; P)$.

Selberg began by replacing the Möbius function in Bruns Pure sieve with Λ_n, a real-valued arithmetic function, subject only to the constraint that $\Lambda_1 = 1$. Hence

$$\left(\sum_{d|n} \Lambda_d \right)^2 \geq \begin{cases} 1 \text{ if } n = 1 \\ 0 \text{ if } n \geq 1. \end{cases}$$

Noting that the sum over the $\Lambda_d' s$ squared is nonnegative,

$$\begin{aligned} S(x, y; P) &\leq \sum_{x < n \leq x+y} \left(\sum_{\substack{d|P \\ d|n}} \Lambda_d \right)^2 \qquad (1) \\ &= \sum_{x < n \leq x+y} \sum_{\substack{d|P \\ d|n}} \Lambda_d \sum_{\substack{e|P \\ e|n}} \Lambda_e \\ &= \sum_{\substack{d|P \\ e|P}} \Lambda_d \Lambda_e \sum_{\substack{x < n \leq x+y \\ d|n, e|n}} 1 \\ &= \sum_{\substack{d|P \\ e|P}} \Lambda_d \Lambda_e \left(\left\lfloor \frac{x+y}{[d,e]} \right\rfloor - \left\lfloor \frac{x}{[d,e]} \right\rfloor \right) \\ &= \sum_{\substack{d|P \\ e|P}} \Lambda_d \Lambda_e \left(\frac{x+y}{[d,e]} - \frac{x}{[d,e]} + O(1) \right) \end{aligned}$$

$$= y \sum_{\substack{d|P \\ e|P}} \frac{\Lambda_d \Lambda_e}{[d,e]} + \sum_{\substack{d|P \\ e|P}} \Lambda_d \Lambda_e \cdot O(1)$$
$$= y \sum_{\substack{d|P \\ e|P}} \frac{\Lambda_d \Lambda_e}{[d,e]} + O\bigg(\Big(\sum_{d|P} |\Lambda_d| \Big)^2 \bigg),$$

where the first line follows from our constraint that $\Lambda_1 = 1$, and in the fourth equality we used $\lfloor t \rfloor = t + O(1)$. The final equality follows from the definition of big-O. The final expression we obtain has the status of 'main term' and 'error term' respectively.

This is where Selberg's genius becomes apparent. With some effort, one can rewrite the expression in the main term of the inequality as a quadratic form, and by applying Lemma 2.12, we can minimize it, which in turn yields better bounds than those obtained using the Brun Sieve.

As a slight digression, using his approach, one can improve the bound in Theorem 1.2 to

$$\pi_2(N) \ll \frac{x}{(\log x)^2}.$$

Now, further suppose that $\Lambda_n = 0$ for $n > z^2$, where z is a parameter we choose. We have the following result.

Theorem 3.1. (Selberg, 1947) Let x, y and z be real numbers such that $y > 0$ and $z \geq 1$. For any positive integer P we have

$$S(x, y; P) \leq \frac{y}{L_P(z)} + O(z^2 L_p(z)^{-2}),$$

where

$$L_P(z) = \sum_{\substack{n \leq z \\ n|P}} \frac{\mu(n)^2}{\varphi(n)},$$

and φ is the Euler totient function.

Remark. Before we begin with the lengthy proof, we provide a brief overview of the main ideas here. We will rewrite the main term into a quadratic form. By Lemma 2.6, we derive a linear constraint to our quadratic form, and hence we can apply Lemma 2.12. For the error term, we want to control Λ_d. Since we know precisely what our choices of Λ_d must be from Lemma 2.12, we have to be a little clever in bounding the error term.

Proof. ([Mon06],[Nat96]) We continue from (1). We may assume that P is square-free. Since $[d,e](d,e) = de$, and $\sum_{d|n} \varphi(d) = n$, we see that

$$\frac{1}{[d,e]} = \frac{(d,e)}{de} = \frac{1}{de} \sum_{f|d, f|e} \varphi(f).$$

[2]This is truncating our sieve, just as in the introduction.

Hence

$$\begin{aligned}\sum_{\substack{d|P\\e|P}}\frac{\Lambda_d\Lambda_e}{[d,e]} &= \sum_{\substack{d|P\\e|P}}\frac{\Lambda_d\Lambda_e}{de}\sum_{f|d,f|e}\varphi(f)\\ &= \sum_{f|P}\varphi(f)\sum_{f|d|P}\frac{\Lambda_d}{d}\sum_{f|e|P}\frac{\Lambda_e}{e}\\ &= \sum_{f|P}\varphi(f)y_f^2, \end{aligned} \tag{2}$$

where

$$y_f = \sum_{\substack{d\\f|d|P}}\frac{\Lambda_d}{d}.$$

Noting that the set of all such $d's$, ie $\mathcal{D} = \{d : d \mid P\}$ is a divisor closed set, by Theorem 2.6 and changing indexes after we have

$$\frac{\Lambda_d}{d} = \sum_{\substack{d\\d|f|P}} y_f\mu\left(\frac{f}{d}\right) \tag{3}$$

So the constraint $\Lambda_1 = 1$ is now equivalent to the linear constraint:

$$\sum_{f|P} y_f\mu(f) = 1.$$

The goal is to minimize our expression in (2). We do that by choosing specific values for the $y'_f s$.[3] By Lemma 2.12, we see that the choice of values for our $y'_k s$ are

$$y_k = \frac{\mu(f)}{\varphi(f)L_p(z)}, \quad \text{where} \quad L_p(z) = \sum_{\substack{n\le z\\n|P}}\frac{\mu(n)^2}{\varphi(n)}.$$

We are done with the main term. For the error term, noting that P is square-free,

$$\Lambda_d = \frac{d}{L_P(z)}\sum_{\substack{f\\d|f|P\\f\le z}}\frac{\mu(f)\mu(f/d)}{\varphi(f)}.$$

Letting $m = f/d$, the expression now reads

$$\Lambda_d = \frac{d\mu(d)}{L_P(z)\varphi(d)}\sum_{\substack{m|P\\(m,d)=1\\m\le z/d}}\frac{\mu(m)^2}{\varphi(m)}.$$

Therefore,

$$\sum_{d\le z}|\Lambda_d| \le \frac{1}{L_P(z)}\sum_{d\le z}\frac{d}{\varphi(d)}\sum_{m\le z/d}\frac{1}{\varphi(m)}$$

[3] We can think of this as choosing values for our $\Lambda'_d s$, which are the so called 'sieve-weights'.

$$= \frac{1}{L_P(z)} \sum_{m \leq z} \frac{1}{\varphi(m)} \sum_{d \leq z/m} \frac{d}{\varphi(d)}.$$

By Lemma 2.7, and yet another counting argument,

$$\begin{aligned}
\sum_{d \leq z/m} \frac{d}{\varphi(d)} &= \sum_{d \leq z/m} \sum_{r|d} \frac{\mu(r)^2}{\varphi(r)} \\
&= \sum_{r \leq z/m} \frac{\mu(r)^2}{\varphi(r)} \sum_{\substack{d \leq z/m \\ r|d}} 1 \\
&= \sum_{r \leq z/m} \frac{\mu(r)^2}{\varphi(r)} \left\lfloor \frac{z/m}{r} \right\rfloor \\
&\leq \frac{z}{m} \sum_{r=1}^{\infty} \frac{\mu(r)^2}{r\varphi(r)} \ll \frac{z}{m},
\end{aligned}$$

where the conclusion follows since the sum is convergent. (See Appendix A). Finally,

$$\sum_{d \leq z} |\Lambda_d| \ll \frac{z}{L_P(z)} \sum_{m \leq z} \frac{1}{m\varphi(m)} \ll \frac{z}{L_P(z)},$$

which is precisely the claim. □

Remark. To actually use this form of the result, one has to find a lower bound for the quantity $\frac{z}{L_P(z)}$. However, we will not cover it in this article; readers may refer to [Mon06] to find a full account of this last part.

We now move on to an application of the sieve. As mentioned before, we aim to derive an upper bound for the number of representations of an even integer as the sum of two primes.

Theorem 3.2. Let N be an even integer, and let $r(N)$ denote the number of representations of N as a sum of two primes. Then

$$r(N) \ll \frac{N}{(\log N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right),$$

where the implied constant is absolute.

Unfortunately, we would not be covering the entire proof of this result. We will only highlight the first few steps, which involve setting up the scene for the Selberg Sieve approach. The reader may wish to consult Nathonson/Ben Green.

Proof. (Sketch) We first define

$$\mathcal{A} := \{n(N-n) : n \leq N\}.$$

Let $z := N^{\frac{1}{3}}$. If $n \leq N$, with both n and $N-n$ being prime, then we either have (i) $n \leq z$, (ii) $N - n \leq z$ or (iii) $n(N-n)$ has no prime factors $\leq z$. Hence

$$r(N) \leq 2z + S(\mathcal{A}, z),$$

where $S(\mathcal{A}, z)$ denotes the number of elements of $\mathcal{A}$ that are not divisible by z. The main difference here is that we now aim to sift over a *set*, whereas in our previous theorem, we sifted over an *interval*. To reduce the problem of sifting a set into sifting an interval, let $\mathcal{A}_d$ denote the elements in $\mathcal{A}$ that are divisible by d, that is

$$\mathcal{A}_d = \{n \leq N : d | n(N-n)\}.$$

It now follows that

$$\begin{aligned} S(\mathcal{A}, z) &\leq \sum_{n \leq N} \Bigg(\sum_{\substack{d | n(N-n) \\ d \leq z}} \Lambda_d \Bigg)^2 \\ &= \sum_{d_1, d_2 \leq z} \Lambda_{d_1} \Lambda_{d_2} |\mathcal{A}_{[d_1, d_2]}|, \end{aligned} \tag{4}$$

after expanding out and rearranging the sum, a similar technique we used at the beginning of this chapter. Now we need to estimate the size of $\mathcal{A}_d$. We have

$$|\mathcal{A}_d| = g(d)\frac{N}{d} + R_d, \tag{5}$$

where $g(d)$ is the number of solutions $x(N-x) \equiv 0 \mod d$, and $|R_d| \leq g(d)$. Intuitively, we first split the interval of size N into smaller intervals of length d, hence the quantity $\frac{N}{d}$. Then we count the number of solutions in each interval, as each interval forms a set of complete residues modulo d, hence $g(d)\frac{N}{d}$. The R_d term comes from the fact that we cannot exactly divide $\{1, 2, \ldots, N\}$ into intervals of length d, and there can be an interval of length $< d$ left, that may or may not contain solutions, hence $R_d \leq g(d)$. Without loss of generality, we may assume d to be squarefree. Then, by the Chinese Remainder Theorem, we have

$$g(d) = \prod_{p | d} g(p). \tag{6}$$

Putting everything together by substituting 5 and 6 into 4

$$S(\mathcal{A}, z) \leq N \sum_{d_1, d_2 \leq z} \frac{\Lambda_{d_1} \Lambda_{d_2} g([d_1, d_2])}{[d_1, d_2]} + \sum_{d_1, d_2 \leq z} |\Lambda_{d_1}||\Lambda_{d_2}| g([d_1, d_2]).$$

Thus, we have successfully reduced our problem of sifting a set into sifting an interval, and we are in a position to apply our previous techniques used in Theorem 3.1. After some work (namely, minimising the quadratic form), one can show that the minimum value of our main term is

$$\frac{1}{D} = \sum_{d \leq z, d \text{ squarefree}} \frac{1}{f(d)},$$

where

$$f(k) = \sum_{\delta | k} \mu\left(\frac{k}{\delta}\right) \frac{\delta}{g(\delta)}.$$

It remains to bound the main term and error term; we omit the lengthy proof here and take the following bounds for granted. The reader is directed to Nathonson/Ben Green for the full proof. We have

$$D \gg \log^2(N) \cdot (\prod_{p|N}(1+\frac{1}{p}))^{-1},$$

and

$$E = \sum_{d_1,d_2 \leq z} |\Lambda_{d_1}||\Lambda_{d_2}| g([d_1,d_2]) \ll N^{\frac{2}{3}}.$$

Finally,

$$S(\mathcal{A},z) \ll \frac{N}{(\log N)^2} \prod_{p|N}\left(1+\frac{1}{p}\right),$$

which follows after substituting our bound for D and E, and the claim is proven. $\square$

4 Schnirelmann Density

We begin this section by defining the *Schnirelmann Density*, who first studied it in 1939 [L G39]. It will give us information, or more precisely, quantify how 'dense' the set we are considering is. This will be key in proving the Schnirelmann-Goldbach Theorem.

Definition 4.1. Let $\mathcal{A}$ be a set of infinite non-negative integers. Then the *Schnirelmann Density* of $\mathcal{A}$, denoted $\sigma(\mathcal{A})$ is given by

$$\sigma(\mathcal{A}) = \inf_{N=1,2,\ldots} \frac{1}{N} \left|\mathcal{A} \cap \{1,\ldots,N\}\right|.$$

We shall also use the following notation

$$\mathcal{A}[N] = \mathcal{A} \cap \{1,\ldots,N\}.$$

In other words, we can rewrite Definition 4.1 as

$$\sigma(\mathcal{A}) = \inf_{N=1,2,\ldots} \frac{\mathcal{A}[N]}{N}.$$

For every set $\mathcal{A}$ of integers, we deduce that

$$0 \leq \sigma(\mathcal{A}) \leq 1,$$

which follows from the definition of $\sigma(\mathcal{A})$,

Moreoever, if $\sigma(\mathcal{A}) = \alpha$, then we have

$$\mathcal{A}[N] \geq \alpha N.$$

It now follows that $\sigma(\mathcal{A}) = 1$ if and only if $\mathcal{A}$ contains every positive integer.

The Schnirelmann density looks very abstract at first sight, but it really aligns with our intuition. Suppose we were tasked to write 89712239[4] as a sum of $1's$ and $2's$. It would not require much to convince ourselves that this would be pretty straightforward. However, suppose we wanted to write the same number, but as a sum of $6's$ and $7's$, then the result is no longer immediately obvious.

The Schnirelmann density does exactly this; by definition, it is *very* sensitive to the small values of the set $\mathcal{A}$ we are considering. Indeed, if $\sigma(\mathcal{A}) > 0$, then $1 \in \mathcal{A}$.

Definition 4.2. Let $\mathcal{A}, \mathcal{B}$ be sets of integers. We define the *sumset* $\mathcal{A}+\mathcal{B}$ to be the set consisting of all integers of the form $a+b$, where $a \in \mathcal{A}$ and $b \in \mathcal{B}$. In set notation

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

Similarly, we define the *h-fold sumset of* $\mathcal{A}$ as

$$h\mathcal{A} = \underbrace{\mathcal{A}+\cdots+\mathcal{A}}_{h \text{ times}}.$$

Definition 4.3. A set $\mathcal{A}$ is called a *basis of order* h if $h\mathcal{A}$ contains every nonnegative integer. In other words, every nonnegative integer can be written as a sum of at most h (not necessarily distinct) elements of $\mathcal{A}$.

We also say that $\mathcal{A}$ is a *basis of finite order* if $\mathcal{A}$ is a basis of order h, for some $h \geq 1$. It follows that $\mathcal{A}$ is a basis of finite order if and only if $\sigma(h\mathcal{A}) = 1$ for some $h \geq 1$.

Using this newly defined object, Schnirelmann showed the following.

Theorem 4.4. (Schnirelmann, 1940) Let $\mathcal{A}$ be a set of integers such that $0 \in \mathcal{A}$ and $\sigma(\mathcal{A}) > 0$. Then $\mathcal{A}$ is a basis of finite order.

Needless to say, the rest of this chapter will be devoted to proving this result. We start gently with some elementary lemmas.

Lemma 4.5. Let $\mathcal{A}$ and $\mathcal{B}$ be sets of integers such that $0 \in \mathcal{A}, 0 \in \mathcal{B}$. If $n \geq 0$, and $\mathcal{A}[n] + \mathcal{B}[n] \geq n$, then $n \in \mathcal{A}+\mathcal{B}$.

Proof. [Nat96] If $n \in \mathcal{A}$ or $n \in \mathcal{B}$ then we are done, henceforth suppose not. We define the sets $\mathcal{A}'$ and $\mathcal{B}'$ by

$$\mathcal{A}' := \{n-a : a \in \mathcal{A}, 1 \leq a \leq n-1\} \quad \text{and} \quad \mathcal{B}' := \{b : b \in \mathcal{B}, 1 \leq b \leq n-1\}.$$

Then $|\mathcal{A}'| = \mathcal{A}[n]$, as $n \notin \mathcal{A}$. Similarly we have $|\mathcal{B}'| = \mathcal{B}[n]$. Moreover,

$$\mathcal{A}' \cup \mathcal{B}' \subseteq [1, n-1].$$

Since $|\mathcal{A}'| + |\mathcal{B}'| = \mathcal{A}[n] + \mathcal{B}[n] \geq n$, it follows that $\mathcal{A}' \cap \mathcal{B}' \neq \emptyset$. Hence $n-a = b$ for some $a \in \mathcal{A}, b \in \mathcal{B}$, and $n = a+b \in \mathcal{A}+\mathcal{B}$. □

We immediately deduce the following result.

[4]The authors' favourite prime!

Corollary 4.6. Let $\mathcal{A}$ be a set of integers such that $0 \in \mathcal{A}$ and $\sigma(\mathcal{A}) \geq 1/2$. Then $\mathcal{A}$ is a basis of order 2.

Proof. [Nat96] By definition of $\sigma(\mathcal{A})$, we have

$$\mathcal{A}[n] \geq \frac{n}{2},$$

for every integer nonnegative integer n. Hence, $\mathcal{A}[n]+\mathcal{A}[n] \geq n$, and the claim follows from our previous result. □

Theorem 4.7. (Schnirelmann, 1940) Let $\mathcal{A}$ and $\mathcal{B}$ be sets of integers such that $0 \in \mathcal{A}$ and $0 \in \mathcal{B}$. Let $\sigma(\mathcal{A}) = \alpha$ and $\sigma(\mathcal{B}) = \beta$. Then

$$\sigma(\mathcal{A} + \mathcal{B}) \geq \alpha + \beta - \alpha\beta.$$

Proof. [Nat96] We count the number of elements in $(\mathcal{A} + \mathcal{B})[N]$. Let

$$1 = a_0 < a_1 < \cdots < a_k \leq N$$

be the $k+1$ elements of $\mathcal{A}[N]$. Since $0 \in \mathcal{B}$, each of these elements are in $(\mathcal{A}+\mathcal{B})[N]$. Now for each $i = 0, 1, \ldots, k-1$, consider the set $a_i + \mathcal{B}[a_{i+1} - a_i - 1]$. We claim that these elements lie strictly in between a_i and a_{i+1}. Indeed, let $b_1, b_2, \ldots, b_j$ be elements of $\mathcal{B}[a_{i+1} - a_i - 1]$, then

$$\begin{aligned} 1 \leq b_1 < \cdots < b_j &\leq a_{i+1} - a_i - 1 \\ 1 + a_i \leq b_1 + a_i < \cdots < b_j + a_i &\leq a_{i+1} - 1 \\ a_i < b_1 + a_i < \cdots < b_j + a_i &< a_{i+1}. \end{aligned}$$

Consider the elements of $a_k + \mathcal{B}[N - a_k]$, they lie in the interval $(a_k, N]$. Putting everything together

$$\begin{aligned} (\mathcal{A} + \mathcal{B})[N] &= \mathcal{A}[N] + \sum_{i=0}^{k-1} |\mathcal{B}[a_{i+1} - a_i - 1]| + |\mathcal{B}[N - a_k]| \\ &\geq \alpha N + \beta \sum_{i=0}^{k-1} a_{i+1} - a_i - 1 + N - a_k \\ &= \alpha N + \beta(-1 - k + N) \\ &\geq \alpha N + \beta(-\alpha N + N) \\ &= \alpha N + \beta N - \alpha\beta N. \end{aligned}$$

Finally,

$$\sigma(\mathcal{A} + \mathcal{B}) = \inf_{n=1,2,\ldots} \frac{(\mathcal{A} + \mathcal{B})[N]}{N} \geq \alpha + \beta - \alpha\beta.$$

□

The final inequality can also be expressed as follows

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

We are finally ready to prove Theorem 4.4.

Proof. (of Theorem 4.4)[Nat96] Let $\sigma(\mathcal{A}) = \alpha$, by definition we know that $0 < \alpha \leq 1$. If $\alpha = 1$, then there is nothing to prove. Suppose that $\alpha < 1$, then

$$(1-\alpha)^k \leq 1/2, \qquad \text{for some positive integer } k.$$

After applying Theorem 4.7 inductively,

$$1 - \sigma(k\mathcal{A}) \leq (1-\sigma(\mathcal{A}))^k = (1-\alpha)^k \leq 1/2.$$

So we have $\sigma(k\mathcal{A}) \geq 1/2$, and by Corollary 4.6, $k\mathcal{A}$ is a finite basis of order 2. It follows that $\mathcal{A}$ is a finite basis of order $2k$. □

5 Towards the Goldbach Conjecture

In this section, we will finally reap the rewards of all the hard work we have previously done. As usual, we begin with a few lemmas.

Lemma 5.1. Let d, e be positive integers. Then

$$[d, e] \geq \sqrt{de}.$$

Proof. Let $p_1, \ldots, p_k$ be the complete list of prime factors of both d and e. It follows that

$$d = p_1^{a_1} \ldots p_k^{a_k} \quad \text{and} \quad e = p_1^{b_1} \ldots p_k^{b_k},$$

for nonnegative integers $a_1, \ldots, a_k, b_1, \ldots, b_k$. On the other hand, define $c_i := \max\{a_i, b_i\}$ for each $i = 1, \ldots, k$, then we have

$$[d, e] = p_1^{c_1} \ldots p_k^{c_k}.$$

Hence, it suffices to show that $c_i \geq (a_i + b_i)/2$, which follows by how we defined c_i. □

Lemma 5.2. Let $r(N)$ denote the number of representations of N as the sum of two primes. Then

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

Proof. [Nat96] By Theorem 3, if N is even, then

$$r(N) \ll \frac{N}{(\log N)^2} \prod_{p|N} \left(1 + \frac{1}{p}\right) \leq \frac{N}{(\log N)^2} \sum_{d|N} \frac{1}{d}.$$

The inequality also holds for odd integers, since an odd integer N can be written as a sum of two primes if and only if $N - 2$ is prime, in which case $r(N) = 2$. Hence

$$r(N)^2 \ll \frac{N^2}{(\log N)^4} \left(\sum_{d|N} \frac{1}{d}\right)^2.$$

Now summing both sides up to x, we obtain

$$\begin{aligned}
\sum_{N\le x} r(N)^2 &\ll \sum_{N\le x} \frac{N^2}{(\log N)^4} \sum_{d|N} \frac{1}{d} \sum_{e|N} \frac{1}{e} \\
&\ll \frac{x^2}{(\log N)^4} \sum_{N\le x} \sum_{d|N} \sum_{e|N} \frac{1}{de} \\
&= \frac{x^2}{(\log N)^4} \sum_{d,e\le x} \frac{1}{de} \sum_{\substack{N\le x \\ [d,e]|N}} 1 \\
&= \frac{x^2}{(\log N)^4} \sum_{d,e\le x} \frac{1}{de} \frac{x}{[d,e]} \\
&\le \frac{x^3}{(\log N)^4} \sum_{d,e\le x} \frac{1}{(de)^{3/2}} \\
&\le \frac{x^3}{(\log N)^4} \left(\sum_{d\le x} \frac{1}{d^{3/2}} \right)^2 \\
&\ll \frac{x^3}{(\log N)^4},
\end{aligned}$$

as the final sum is convergent. This concludes the proof. □

Theorem 5.3. (Goldbach-Schnirelmann) Every integer greater than one is the sum of a bounded number of primes, where the implied constant is absolute.

Proof. [Nat96] We start by showing the set

$$\tilde{\mathcal{P}} := \{0, 1\} \cup \{p + q : p, q \text{ prime}\}$$

has positive Schnirelmann density. Let $r(N)$ denote the number of representations of N as the sum of two primes. By the Cauchy-Schwarz inequality,

$$\left(\sum_{N\le x} r(N) \right)^2 \le \sum_{\substack{N\le x \\ r(N)\ge 1}} 1 \sum_{N\le x} r(N)^2 \le \tilde{\mathcal{P}}[x] \sum_{N\le x} r(N)^2.$$

By Lemma 2.10 and Lemma 5.2, we have

$$\begin{aligned}
\frac{\tilde{\mathcal{P}}[x]}{x} &\ge \frac{1}{x} \frac{\left(\sum_{N\le x} r(N) \right)^2}{\sum_{N\le x} r(N)^2} \\
&\gg \frac{1}{x} \frac{x^4}{(\log x)^4} \frac{(\log x)^4}{x^3} \\
&\gg 1.
\end{aligned}$$

So there exists a number $c_1 > 0$ such that $\tilde{\mathcal{P}}[x] \ge c_1 x$ for all $x \ge x_0$. Since $1 \in \tilde{\mathcal{P}}$, it follows that there exists a number $c_2 > 0$ such that $\tilde{\mathcal{P}}[x] \ge c_2 x$ for $1 \le x \le x_0$. Now choose $c := \min\{c_1, c_2\}$, then $\tilde{\mathcal{P}}[x] > cx$, for all $x \ge 1$, and so $\sigma(\tilde{\mathcal{P}}) > 0$. Since $\tilde{\mathcal{P}}$

has positive Schnirelmann density, by Theorem 4.4, $\tilde{\mathcal{P}}$ is a basis of finte order, say of order h for some positive integer h. Let $N \geq 2$. If $N = 2$ then we are done. Suppose not, write

$$N - 2 = \underbrace{1 + 1 + \cdots + 1}_{k \text{ times}} + (p_1 + q_1) + \cdots + (p_i + q_i),$$

with $i + k \leq h$. If k is even, then we can write N as

$$N = \underbrace{2 + 2 + \cdots + 2}_{\frac{k}{2}+1 \text{ times}} + (p_1 + q_1) + \cdots + (p_i + q_i).$$

If k is odd, then we write

$$N = \underbrace{2 + 2 + \cdots + 2}_{\frac{k-1}{2} \text{ times}} + 3 + (p_1 + q_1) + \cdots + (p_i + q_i).$$

In both cases, we have written N as the sum of at most

$$i + i + k \leq 3h$$

primes, and the claim follows. □

6 What now?

We have shown that there *exists* a bound, but what about the work done to actually compute it? We let h be the smallest number h such that every integer greater than one can be written as a sum of at most h primes. In the literature, we usually refer to the number h as *Schnirelmann's constant*. In particular, the Goldbach Conjecture implies that Schnirelmann's constant is 3. There has been substantial progress in trying to improve the upper bound for Schnirelmann's constant. In 1969, Klimov gave an upper bound of $6 \cdot 10^9$, which was improved by Klimov, Pil'tai, and Sheptitskaya to 115. In 1977, R.C Vaughan further reduced the bound to 27. A full historical account up to this point can be found at [Vau79]. The final major advancement was made by Harald Helfgott in 2013[Hel15], who solved the ternary Goldbach problem via the Hardy-Littlewood circle method, which in turn implies that Schnirelmann's constant is at most 4.

A Appendix A

Theorem A.1. Let μ denote the Möbius function. Then

$$\sum_{d|n} \mu(d) = \begin{cases} 1, & \text{if } n = 1, \\ 0, & \text{if } n \geq 2. \end{cases}$$

Proof. The claim is true for $n = 1$. Suppose $n \geq 2$, then

$$n = \prod_{i=1}^{k} p_i^{r_i},$$

where $k \geq 1$, and the $p_i's$ are distinct primes. Let $\sum'$ denote a sum over squarefree integers. Then

$$\begin{aligned} \sum_{d|n} \mu(d) &= \sum_{d|n}{}' \mu(d) \\ &= \sum_{d|p_1 \ldots p_k} \mu(d) \\ &= \sum_{d|p_1 \ldots p_k} (-1)^{\omega(d)}, \end{aligned}$$

where $\omega(d)$ is the number of distinct prime divisors n. Finally,

$$\sum_{d|p_1 \ldots p_k} (-1)^{\omega(d)} = \sum_{l=0}^{k} \binom{k}{l} (-1)^l = 0.$$

□

Theorem A.2. Let $\mathcal{D}$ be a finite divisor-closed set, and let f and g be functions defined on $\mathcal{D}$. If

$$g(n) = \sum_{\substack{d \in \mathcal{D} \\ n|d}} f(d),$$

for all $n \in \mathcal{D}$, then

$$f(n) = \sum_{\substack{d \in \mathcal{D} \\ n|d}} \mu(\frac{d}{n}) g(d)$$

for all $n \in \mathcal{D}$. Conversely, if

$$f(n) = \sum_{\substack{d \in \mathcal{D} \\ n|d}} \mu\left(\frac{d}{n}\right) g(d)$$

for all $n \in \mathcal{D}$, then

$$g(n) = \sum_{\substack{d \in \mathcal{D} \\ n|d}} f(d).$$

Proof.

$$\begin{aligned}
\sum_{\substack{d\in\mathcal{D}\\ n|d}} \mu\left(\frac{d}{n}\right) g(d) &= \sum_{\substack{d\in\mathcal{D}\\ n|d}} \mu\left(\frac{d}{n}\right) \sum_{\substack{k\in\mathcal{D}\\ k|d}} f(k) \\
&= \sum_{nh\in\mathcal{D}} \mu(h) \sum_{\substack{k\in\mathcal{D}\\ nh|k}} f(k) \\
&= \sum_{nh\in\mathcal{D}} \mu(h) \sum_{nhl\in\mathcal{D}} f(nhl) \\
&= \sum_{nr\in\mathcal{D}} f(nr) \sum_{\substack{h\in\mathcal{D}\\ h|r}} \mu(h) \\
&= \sum_{nr\in\mathcal{D}} f(nr) \sum_{h|r} \mu(h) \\
&= f(n),
\end{aligned}$$

where the final equality is obtained using the previous result. The proof in the opposite direction is similar. □

Lemma A.3. Let $a_1, a_2, \ldots, a_n$ be positive real numbers and $b_1, b_2 \ldots, b_n$ be any real numbers. The minimum value of the quadratic form

$$\mathcal{Q}(y_1, y_2, \ldots, y_n) = a_1 y_1^2 + \cdots + a_n y_n^2$$

subject to the linear constraint $b_1 y_1 + \ldots b_n y_n = 1$ is

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

and this value is attained iff $y_i = \frac{m b_i}{a_i}$, for all $i = 1, 2, \ldots, n$.

Proof. Note that

$$\begin{aligned}
1 &= \left(\sum_{i=1}^{n} b_i y_i\right)^2 \\
&= \left(\sum_{i=1}^{n} \left(\frac{b_i}{\sqrt{a_i}}\right) \sqrt{a_i} y_i\right)^2 \\
&\leq \left(\sum_{i=1}^{n} \frac{b_i^2}{a_i}\right) \left(\sum_{i=1}^{n} a_i y_i^2\right),
\end{aligned}$$

after applying the Cauchy-Schwarz inequality. Hence

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

Moreover, equality holds if and only if there exists a real number t such that for all $i = 1, \ldots, n$ we have

$$\sqrt{a_i} y_i = \frac{t b_i}{\sqrt{a_i}},$$

or equivalently

$$y_i = \frac{tb_i}{a_i}.$$

This implies that

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

and so

$$t = m,$$

which implies

$$y_i = \frac{mb_i}{a_i}.$$

Conversely, if $y_i = \frac{mb_i}{a_i}$ for each i, then $\sum_{i=1}^{n} b_i y_i = 1$ and $\mathcal{Q}(y_1, \ldots, y_n) = m$. □

Theorem A.4. The sum

$$\sum_{r=1}^{\infty} \frac{1}{r\varphi(r)}$$

is convergent.

Proof. We first show the inequality

$$\varphi(n) \geq \sqrt{\frac{n}{2}},$$

for sufficiently large n. Since φ is multiplicative, it suffices to show the inequality for prime powers. Let $n = p^k$, then the inequality is equivalent to

$$p^{k-1}(p-1) \geq \frac{p^{k/2}}{\sqrt{2}},$$

which can be reduced to

$$p^{k-2}(p-1)^2 \geq \frac{1}{2}.$$

If $k = 1$, then the inequality reads

$$(p-1)^2 \geq \frac{p}{2},$$

which holds for all $p \geq 2$. If $k \geq 2$, then there is nothing to prove. Anyhow, we have that the inequality holds, with equality at $n = 2$. The convergence of our summand now follows. Indeed,

$$\sum_{r=1}^{\infty} \frac{1}{r\varphi(r)} \leq \sum_{r=1}^{\infty} \frac{\sqrt{2}}{r\sqrt{r}},$$

and the latter series converges. □

References

[Bru19] V. Brun. “La série

$$\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\frac{1}{17}+\frac{1}{19}+\frac{1}{29}+\frac{1}{31}+\frac{1}{41}+\frac{1}{43}+\frac{1}{59}+\frac{1}{61}+\cdots$$

où les dénominateurs sont des « nombres premiers jumeaux », est convergente ou finie.” French. In: *Bull. Sci. Math., II. Sér.* 43 (1919), pp. 100–104. ISSN: 0007-4497.

[Coj05] Murty MR Cojocaru AC. *An Introduction to Sieve Methods and Their Applications*. Cambridge University Press, 2005.

[Hel15] Harald Andres Helfgott. “The ternary Goldbach problem”. In: *arXiv preprint arXiv:1501.05438* (2015). URL: https://arxiv.org/abs/1501.05438.

[L G39] L. G. Shnirel’man. “On the additive properties of numbers”. In: *Uspekhi Mat. Nauk* (1939).

[Mon06] Vaughan RC. Montgomery HL. *Multiplicative Number Theory I: Classical Theory*. Cambridge University Press, 2006.

[Nat96] Melvyn B. Nathanson. *Additive Number Theory The Classical Bases*. Springer New York, NY, 1996.

[Sel84] Atle Selberg. “On an Elementary Method in the Theory of Primes”. In: *Norske Vid. Selsk. Forh., Trondheim,* (1984).

[Vau79] RC Vaughan. “A survey of some important problems in additive number theory”. In: *Societe Math. France, Asterisque* 61 (1979), pp. 213–222.