Sieving with Fourier Polynomials on Primes

—-

Large sieve, Brun-Titchmarsh Theorem and Cusps

—

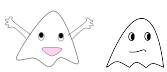
International Center for Theoretical Sciences

May 4th / May 8th 2026

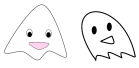
—

Olivier Ramaré

May 4, 2026



Contents



Introduction

The players

With the aim of understanding some given sequence, say $(u_n)_{n\leqslant N}$, we consider the associated *trigonometric polynomial*, which we may also call a *Fourier polynomial*, defined by

$$S(\alpha) = \sum_{n\leqslant N} u_n e(n\alpha) \quad \text{where} \quad e(x) = \exp(2i\pi x). \tag{1}$$

The sequence to which S corresponds remains understood by the context in this notation. The length N of the sequence is of utmost importance, and we may also have in some situations

$$S(\alpha) = \sum_{M<n\leqslant M+N} v_n e(n\alpha).$$

Notice that, when α is rational, say $\alpha = a/q$, then $S(a/q)$ depends only the variables

$$S(q;b) = \sum_{\substack{n\leqslant N\\ n\equiv b[q]}} u_n$$

so, in short, only on the distribution of the sequence (u_n) modulo q.

Present situation

When using the large sieve inequality for sieving purpose, whether as Linnik originaly did or in the modern version that is Montgomery's sieve, we rely on two informations:

- Some *local lower bounds* of arithmetical nature, see Theorem 1.2 or (3) below (see also Theorem 1.1 and Theorem 1.3),
- and a *global upper bound*, usually the Large Sieve inequality, see Theorem 4.2 or (4) below.

This process has proven to be very efficient and to lead to the best known results in several cases, the most famous one being surely the Brun-Titchmarsh Theorem. These lectures are largely centered on an example of crucial interest. Let us define

$$S(\alpha) = \sum_{M<p\leqslant M+N} e(p\alpha) \tag{2}$$

where p denotes a prime number and M and $N \geqslant 1$ are real numbers. We readily find that, when $q \leqslant M$, we have

$$\sum_{a \bmod^* q} |S(a/q)|^2 \geqslant \frac{\mu^2(q)}{\varphi(q)} |S(0)|^2 \tag{3}$$

and the Large Sieve inequality will tell us that

$$\sum_{q\leqslant Q}\sum_{a \bmod^* q}|S(a/q)|^2 \leqslant (N+Q^2)S(0). \tag{4}$$

When we join both, we swiftly infer that

$$(\log Q)|S(0)|^2 \leqslant \sum_{q\leqslant Q}\frac{\mu^2(q)}{\varphi(q)}|S(0)|^2 \leqslant (N+Q^2)S(0)$$

from which we deduce that the number $S(0)$ of primes in the interval $(M, M+N]$ is at most* $(2+o(1))N/\log N$. The same process applies to primes in arithmetic progressions. Here is the reference result in this area that has been coined "the Brun-Titchmarsh inequality" by Y. Linnik in his book [22] (see Lemma 1.3.1 therein), because E. Titchmarsh used in [53] a similar inequality which he proved by employing the Brun sieve.

Theorem *𝒜*

When ℓ is prime to k, $M \geqslant 0$ and $N > k$ are real numbers, we have

$$\sum_{\substack{M<p\leqslant M+N\\ p\equiv \ell[k]}} 1 \leqslant \frac{2N}{\varphi(k)\log(N/k)}.$$

This very precise form has been given by H. Montgomery & R. C. Vaughan in [29]. It is usually believed that the factor 2 that appears on the right-hand side is superfluous. It seems to have appeared for the first time in the paper [7] by I. Čulanovskiĭ. As it turns out, several different proofs of this inequality lead to the same factor 2. We shall give two such proofs, one as above in Chapter 4, and another (new) one in Chapter 6. The Selberg sieve would lead to a third proof. The hitherto asymptotically strongest upper bound has been given in [38]. See also the preprint [56] by T. Yamada.

Factor 2 in the Brun-Titchmarsh inequality and Siegel zeros

Reducing the factor 2 in the Brun-Titchmarsh inequality, see Theorem , would have a deep consequence on bounding $L(1,\chi)$ from below, when χ is a quadratic

*We selected $Q = \sqrt{N}/\log N$ and assumed that $Q \leqslant M$, an assumption that is easily lifted.

[22] Y. Linnik, 1961, "The dispersion method in binary additive problems".
[53] E. Titchmarsh, 1930, "A divisor problem."
[29] H. Montgomery and R. Vaughan, 1973, "The large sieve".
[7] I. V. Čulanovskiĭ, 1948, "Certain estimates connected with a new method of Selberg in elementary number theory".
[38] O. Ramaré and J.-C. Schlage-Puchta, 2008, "Improving on the Brun-Titchmarsh theorem".
[56] T. Yamada, 2023, *Explicit improvements of the Brun-Titchmarsh theorem for arbitrary intervals.*

Dirichlet character, and consequently on the zeros of such functions in the vicinity of $s = 1$. Here is a result due to Motohashi in [30].

Theorem $\mathscr{B}$

There exist two effective constants c_3 and c_4, such that for $k \geqslant c_4$, the following two conditions are equivalent.

(a) There exist a constant $\xi > 0$ such that for any ℓ prime to k, we have, with $X = k^{c_3}$:
$$\sum_{\substack{X<p\leqslant 2X\\ p\equiv \ell[k]}} 1 \leqslant \frac{2-\xi}{\phi(k)} \sum_{X<p\leqslant 2X} 1.$$

(b) For any real character modulo k, we have $L(1,\chi) \gg 1/\log k$.

A direct proof from the circle method viewpoint

With the idea of finding where the loss of this factor 2 occurs, we may start from the Parseval identity on $\mathbb{R}/\mathbb{Z}$, i.e. $\int_0^1 |S(\alpha)|^2 d\alpha = S(0)$. No loss occurs in this complete integral. On splitting the unit circle into Farey arcs, and working out some local lower bounds to replace (3), we shall prove the next theorem.

Theorem $\mathscr{C}$

When $M \geqslant Q$, we have $\dfrac{\log Q}{1+Q^2/V} \displaystyle\int_{-\infty}^{\infty} \Big| \sum_{\substack{|p-t|\leqslant V\\ M<p\leqslant M+N}} 1 \Big|^2 \frac{dt}{4V^2} \leqslant S(0).$

Since we readily prove that

$$\Big(1-\frac{2V}{N}\Big)\frac{1}{N}\Big|\sum_n u_n\Big|^2 \leqslant \int_{-\infty}^{\infty} \Big|\sum_{|n-t|\leqslant V} u_n\Big|^2 \frac{dt}{4V^2},$$

we may for instance select $Q^2 = V = N/\log N$ and deduce the Brun-Titchmarsh inequality, yet again with a loss of a factor 2.

This proof is a surprise as the reader will see that we have barely the feeling of loosing anything. But this is assuming that the main contributions are indeed at the points a/q for $q \leqslant Q$ and that $S(\alpha)$ has a rather sharp decrease around such points. The philosophical outcome is that this is *not* the situation.

[30] Y. Motohashi, 1979, "A note on Siegel's zeros".

Large values of the Fourier polynomial

If we assume that we have enough primes in the interval $(M, M+N]$, where are the large values of the Fourier polynomial? Let us introduce a notion for clarity.

Definition 0.1

We define the set of A-*cusps* by $\mathscr{C}(A) = \big\{\alpha \in \mathbb{R}/\mathbb{Z} : |S(\alpha)| \geqslant S(0)/A\big\}$.

As the involved trigonometric polynomial is continuous, the set $\mathscr{C}(A)$ is closed, hence compact, and is more precisely a finite union of arcs. The above set may sometimes be called *spectrum*, as in Section 3.4 or [46] by T. Sanders, but, first this word is overloaded and, second the right-hand side is often N/A rather than the one we employ above.

Let us assume that we have a $K \geqslant 1/2$ such that

$$S(0) \geqslant \frac{N}{K \log N}. \tag{5}$$

For instance, the number of primes in the interval $[N^9, N^9 + N]$ is typically $N/(9 \log N)$.

We first will prove that the number of A-cusps is well controlled.

Theorem $\mathscr{D}$

There exist positive constants C_1 and C_2 such that the following holds. Define $D(A)$ to be the maximal cardinality of a $1/N$-well spaced subset of $\mathscr{C}(A)$. We have, for $A \leqslant \sqrt{N}$,

$$\frac{C_1 A^2}{K \log(2A)} \leqslant D(A) \leqslant C_2 A^2 K \log(2A).$$

Furthermore, we have

$$C_1 A \leqslant \int_1^A \frac{D(a)}{a^2} da, \quad \int_1^A \frac{D(a)}{a^3} da \leqslant C_2 K \log(2A).$$

Both upper bounds are valid also for $A \geqslant \sqrt{N}$.

The lower bound for $D(A)$ is fact comes from the fact that we are able to produce Farey points (i.e. rational points) where $S(\alpha)$ takes large values. Let us mention this fact specifically.

[46] T. Sanders, 2011, "On Roth's theorem on progressions".

Theorem $\mathscr{E}$

The set $\mathcal{F} = \big\{(a/q) : a \bmod^* q, q \leqslant A\big\} \cap \mathscr{C}(A)$ contains more than $A^2/(7000K \log A)$ elements when $A \in [2, \sqrt{N}]$ and $N \geqslant 10^4$.

Notice that there are about $6A^2/\pi^2$ Farey points with denominator $q \leqslant A$, so that this theorem in fact shows that many cusps seem indeed to be located around the Farey points. For many A's, this proportion of cusps is even positive.

Let us spend some time on this *production* of Farey points where $S(\alpha)$ takes large values. We easily check that

$$\sum_{a \bmod^* q} |S(a/q)| \geqslant \mu^2(q) S(0)$$

provided that $q \leqslant M$, and this ensures the existence of at least *one* point a_0/q such that $|S(a_0/q)| \geqslant \mu^2(q)S(0)/\varphi(q)$. Producing only one point for every square-free q would lead to A points in $\mathscr{C}(A)$, while we need $C_1 A^2/\log(2A)$ of them. But, if all the other points a/q corresponded to very small values of $S(a/q)$, then in fact $|S(a_0/q)|$ would be much bigger and would have $S(a_0/q) \gg S(0)$. We have however a result that prevents such values a_0/q to be too numerous; such a case may happen, but it will be rare, as q varies. This follows from the next large sieve inequality.

Theorem $\mathscr{F}$

Let $\mathcal{X}$ be a δ-well spaced subset of $\mathbb{R}/\mathbb{Z}$ and $N \geqslant 1$. Let $(u_p)_{p \leqslant N}$ be a sequence of complex numbers. We have

$$\sum_{x \in \mathcal{X}} \Big| \sum_{M < p \leqslant M+N} u_p e(xp) \Big|^2 \ll \frac{N + \delta^{-1}}{\log N} \log(2|\mathcal{X}|) \sum_{M < p \leqslant M+N} |u_p|^2.$$

Generalisations

These lectures are centered on the case of primes. However a large part of this material (but not everything!) may be extended to what we loosely call "a general sieving situation". A more precise meaning is given in Chapter 1. Main examples are

- The (possibly finite) sequence of primes twins, of (sieve) dimension 2,
- The sequence of integers that are both primitive Gaussian and primitive Loeschian (i.e. integers that may be written in the form $u^2 + v^2$ with u and v coprime (these are the primitive Gaussian integers), and in the form $u^2 + uv + v^2$ with also u and v coprime (these are the primitive Loeschian

integers)), of sieve dimension 3/4. Its members below 1500 are

$$1, 17, 53, 89, 125, 197, 233, 269, 305, 377, 449, 485, 593, 629,$$
$$773, 809, 845, 1025, 1097, 1205, 1277, 1313, 1385, 1493.$$

- The sequence of odd primitive Gaussian integers n that are such that $n+4$ is also a primitive Gaussian integer. This sequence is of sieve dimension 1, it can be shown to be infinite. Its members below 600 are

$$1, 13, 25, 37, 61, 85, 97, 109, 145, 169, 181, 193, 229, 265, 277,$$
$$289, 313, 349, 373, 397, 421, 445, 457, 481, 505, 541, 565.$$

Intended schedule:

Day 1 Introduction. Local estimates in the case of primes and some smoothing considerations.

Day 2 Proof of the Large Sieve inequality from Selberg's Lemma and deduction of the Brun-Titchmarsh inequality. Link with Siegel zeros.

Day 3 Proof of Montgomery's sieve from the Parseval Identity. A comparison note with Gallagher's approach to the Large Sieve inequality.

Day 4 Large Sieve inequality for primes. Enveloping sieve for primes.

Day 5 Special Large Sieve inequality when the phase set is small. On the number of rational cusps.

1 Local estimates, modulo q

We work in this chapter on functions over $\mathbb{Z}/q\mathbb{Z}$ that have a support on a restricted "multiplicative" subset.

1.1. Three geometrical notions modulo q

Let us start with an easy notion which takes longer to define than to grasp. A subset $\mathcal{K}_q \subset \mathbb{Z}/q\mathbb{Z}$ is said to be *multiplicative** if, when the decomposition of q in prime factors reads

$$q = p_1^{e_1} p_2^{e_2} \cdots p_r^{e_r}, \quad (\forall i \neq j, \quad p_i \neq p_j),$$

and the Chinese Remainder Map is defined by

$$\begin{aligned} \sigma : \ \mathbb{Z}/q\mathbb{Z} &\to \prod_{1 \leqslant i \leqslant r} \mathbb{Z}/p_i^{e_i}\mathbb{Z} \\ x &\mapsto \big(x \bmod p_i^{e_i}\big), \end{aligned}$$

we have the property

$$\sigma^{-1}\big(\sigma(\mathcal{K}_q)\big) = \mathcal{K}_q. \tag{1.1}$$

This is often written in the shorter form

$$\mathcal{K}_q = \prod_{1 \leqslant i \leqslant r} \mathcal{K}_{p_i^{e_i}}. \tag{1.2}$$

We further need a second notion. For any divisor d of q, we define

$$\mathcal{K}_d = \mathcal{K}_q / d\mathbb{Z}. \tag{1.3}$$

We say that *Johnsen-Gallagher condition* condition holds whenever

$$\forall d | q, \forall y \in \mathcal{K}_d, \#\{x \in \mathcal{K}_q : x \equiv y [d\}] = |\mathcal{K}_q|/|\mathcal{K}_d|. \tag{1.4}$$

This is equivalent to saying that the number of preimages in $\mathcal{K}_q$ of any point y of $\mathcal{K}_d$ does not depend on y.

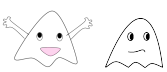
When q is square-free and $\mathcal{K}_q$ is multiplicative, this condition always holds.

We finally need to endow the vector space $\mathscr{F}(\mathcal{K}_q)$ of complex valued functions over $\mathcal{K}_q$ with a hermitian product, and this one is given by

$$[f, g]_{\mathcal{K}_q} = \frac{1}{|\mathcal{K}_q|} \sum_{c \in \mathcal{K}_q} f(c)\overline{g(c)}. \tag{1.5}$$

When $\mathcal{K}_q = \mathbb{Z}/q\mathbb{Z}$, we simply write $[f, g]_q$. Notice that if f and g are in $\mathscr{F}(\mathcal{K}_q)$ and are considered as elements of $\mathscr{F}(\mathbb{Z}/q\mathbb{Z})$ that vanish outside $\mathcal{K}_q$, then $[f, g]_{K_q} = (q/|\mathcal{K}_q|)[f, g]_q$. In particular orthogonality is preserved.

*In earlier work, I used *multiplivatively split* instead of the simpler *multiplicative*.



1.2. A local lower bound

Here is a lemma based on his ideas.

Theorem 1.1

Under the Johnsen-Gallager condition (1.4) and when (u_n) is carried by $\mathcal{K}_q$, we have

$$\sum_{a \bmod^* q} \Big|\sum_n u_n e(na/q)\Big|^2 \geqslant \prod_{p^\alpha \| q} \Big(\frac{p^\alpha}{|\mathcal{K}_{p^\alpha}|} - \frac{p^{\alpha-1}}{|\mathcal{K}_{p^{\alpha-1}}|}\Big) |\mathcal{K}_q| \sum_{b\in\mathcal{K}_q} \Big|\sum_{n\equiv b[q]} u_n\Big|^2.$$

When further $\sum_n u_n = 0$, we may divide the RHS by

$$m(q) = \max_{p^\alpha \| q} \Big(1 - \frac{|\mathcal{K}_{p^\alpha}|}{p|\mathcal{K}_{p^{\alpha-1}}|}\Big)$$

provided $m(q)$ does not vanish.

We shall give a full proof of this lower bound as well as a full description of the LHS quadratic form in the next chapter. The first and easier first inequality is what we shall use later in these lectures. When q is square-free, Huxley in Section 6 of [20] sketches the proof of such an inequality. We now provide a full proof in the case when q is square-free. The prime-case is more detailled in the next section.

Proof. Let us set $v_b = \sum_{n\equiv b[q]} u_n$. We have to study the eigenvalues of the quadratic form

$$\sum_{b,b'\in\mathcal{K}_q} v_b\, \overline{v_{b'}}\, c_q(b-b')$$

where $c_q(a)$ is the Ramanujan sum. The operator

$$U_q(\cdot;\mathcal{K}_q) : (v_b)_b \mapsto \Big(\sum_{b\in\mathcal{K}_q} v_b\, c_q(b-b')\Big)_{b'} \tag{1.6}$$

is self-adjoint since

$$\big[U_q(v;\mathcal{K}_q), w\big]_{\mathcal{K}_q} = \big[v, U_q(w;\mathcal{K}_q)\big]_{\mathcal{K}_q},$$

it may be diagonalised in an orthonormal basis.

When $q = p$, we have $c_p(b-b') = p-1$ when $b = b'$ and $c_p(b-b') = -1$ otherwise. This implies that the matrix representing our operator is a circulant matrix. Let us set for short $T = |\mathcal{K}_p|$. With

$$P(Y) = p - 1 - X - Y - Y^2 - \cdots - Y^{T-1} = p - X - \frac{1-Y^T}{1-Y}, \tag{1.7}$$

[20] M. Huxley, 1972, "Irregularity in sifted sequences".

the characteristic polynomial of our matrix is

$$P(1)P(\xi)P(\xi^2)\cdots P(\xi^{T-1}) = (p-T-X)(p-X)^{T-1}.$$

The eigenvalues are thus $p-T = p-|\mathcal{K}_p|$ with multiplicity 1 and p with multiplicity $T-1 = |\mathcal{K}_p|-1$. The operator modulo q is the tensor product of the operators modulo each prime divisors of q. The eigenvalues are thus

$$\Big(\prod_{p|d}(p-|\mathcal{K}_p|)\frac{q}{d}\Big)_{q|d} \tag{1.8}$$

with multiplicity $\prod_{p|q/d}(|\mathcal{K}_p|-1)$. The smallest one occurs when $d=q$ with multiplicity 1, and the readers will swiftly check that it is attached to the eigenvector $(1)_{b\in\mathcal{K}_q}$. Therefore, when v is orthogonal to this vector, i.e. when $\sum_n u_n = 0$, we may use the second smallest eigenvalue. The proof follows swiftly from there. □

The first inequality has an important consequence, namely:

$$\sum_{a \bmod^* q}\Big|\sum_n u_n e(na/q)\Big|^2 \geqslant \prod_{p^\alpha\| q}\Big(\frac{p^\alpha}{|\mathcal{K}_{p^\alpha}|}-\frac{p^{\alpha-1}}{|\mathcal{K}_{p^{\alpha-1}}|}\Big)\Big|\sum_n u_n\Big|^2 \tag{1.9}$$

which follows from Theorem 1.1 since

$$|\mathcal{K}_q|\sum_{b\in\mathcal{K}_q}\Big|\sum_{n\equiv b[q]}u_n\Big|^2 \geqslant \Big|\sum_n u_n\Big|^2.$$

The intermediate quantity is however more refined in case the sequence (u_n) is *not* equidistributed modulo q. This consequence of the potential imbalance may be proved by appealing to the Lagrange's Identity:

$$\sum_i |a_i|^2\sum_i |b_i|^2 = \Big(\sum_i a_i\overline{b_i}\Big)^2 + \tfrac{1}{2}\sum_{i,j}\big|a_i\overline{b_j}-a_j\overline{b_i}\big|^2, \tag{1.10}$$

see for instance the book [51] of J.M. Steele. It implies in our case that

$$|\mathcal{K}_q|\sum_{b\in\mathcal{K}_q}\Big|\sum_{n\equiv b[q]}u_n\Big|^2 = \Big|\sum_n u_n\Big|^2 + \tfrac{1}{2}\sum_{b,b'\in\mathcal{K}_q}\Big|\sum_{n\equiv b[q]}u_n - \sum_{n\equiv b'[q]}u_n\Big|^2.$$

Another path would be to rely on the Mean-Variance Identity i.e.

$$\sum_{i\leqslant n}|a_i|^2 = \sum_{i\leqslant n}\Big|a_i-\frac{1}{n}\sum_j a_j\Big|^2 + \frac{1}{n}\Big|\sum_i a_i\Big|^2. \tag{1.11}$$

We leave the details to the reader and end here this section.

[51] J. M. Steele, 2004, *The Cauchy-Schwarz master class.*

1.3. A direct proof in the prime case

Theorem 1.2

When (u_n) is supported on prime-to-q integers n and q is square-free, we have

$$\sum_{a \bmod^* q} \Big|\sum_n u_n e(na/q)\Big|^2 \geqslant \sum_{b \bmod^* q} \Big|\sum_{n\equiv b[q]} u_n\Big|^2 \geqslant \frac{1}{\varphi(q)}\Big|\sum_n u_n\Big|^2.$$

Theorem 1.1 tells us also that, when $\sum_n u_n = 0$, we may multiply this lower bound by $\min_{p|q} p$. Eq. (6.10) proposes a variant of the above inequality, while Theorem 1.3 below generalizes it. This proof extends to more general sieving situations.

Proof. We argue by recursion on the number of prime factors of q. Let us start with the prime case and use $q = p$ for clarity. We readily find that

$$\begin{aligned}\sum_{a \bmod^* p} \Big|\sum_n u_n e(na/p)\Big|^2 &= p \sum_{b \bmod^* p} \Big|\sum_{n\equiv b[p]} u_n\Big|^2 - \Big|\sum_{b \bmod^* p}\sum_{n\equiv b[p]} u_n\Big|^2 \qquad (1.12)\\ &\geqslant p \sum_{b \bmod^* p} \Big|\sum_{n\equiv b[p]} u_n\Big|^2 - (p-1)\sum_{b \bmod^* p}\Big|\sum_{n\equiv b[p]} u_n\Big|^2\\ &\geqslant \sum_{b \bmod^* p}\Big|\sum_{n\equiv b[p]} u_n\Big|^2.\end{aligned}$$

This inequality is very similar to the one employed in Montgomery's sieve. Let us now consider $q = q_1 p$ with q_1 square-free and prime to p. We find that

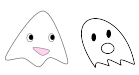
$$\begin{aligned}\sum_{a \bmod^* q} \Big|\sum_n u_n e(na/q)\Big|^2 &= \sum_{a_1 \bmod^* q_1}\sum_{c \bmod^* p}\Big|\sum_n u_n e\Big(\frac{na_1}{q_1} + \frac{nc}{p}\Big)\Big|^2\\ &\geqslant \sum_{b \bmod^* p}\sum_{a_1 \bmod^* q_1}\Big|\sum_{n\equiv b[p]} u_n e\Big(\frac{na_1}{q_1}\Big)\Big|^2\end{aligned}$$

by invoking the prime case. We may now use a recursion hypothesis and derive

$$\sum_{a \bmod^* q} \Big|\sum_n u_n e(na/q)\Big|^2 \geqslant \sum_{b \bmod^* p}\sum_{b_1 \bmod^* q_1}\Big|\sum_{\substack{n\equiv b[p]\\ n\equiv b_1[q_1]}} u_n\Big|^2.$$

The claimed inequality follows from there. □

The above proof proves in fact a more general result which reads as follows.



Theorem 1.3

When (u_n) is supported on prime-to-q_1 integers n and q_1 is square-free and prime to q_2, we have

$$\sum_{a \bmod^* q_1q_2} \Big|\sum_n u_n e\Big(\frac{na}{q_1q_2}\Big)\Big|^2 \geqslant \sum_{b \bmod^* q_1} \sum_{a_2 \bmod^* q_2} \Big|\sum_{n\equiv b[q_1]} u_n e\Big(\frac{na_2}{q_2}\Big)\Big|^2$$
$$\geqslant \frac{1}{\varphi(q_1)} \sum_{a_2 \bmod^* q_2} \Big|\sum_n u_n e\Big(\frac{na_2}{q_2}\Big)\Big|^2.$$

1.4. An even simpler proof in the prime case!

The proof of Theorem 1.2 extends to more general sieving situations. Here is an almost trivial proof, which is this time very specific to the primes.

Theorem 1.4

When (u_n) is supported on prime-to-q integers n and q is square-free, we have

$$\sum_{a \bmod^* q} \Big|\sum_n u_n e(na/q)\Big|^2 \geqslant \frac{1}{\varphi(q)}\Big|\sum_n u_n\Big|^2.$$

Notice that this inequality is generally considered to optimal since $\sum_n u_n e(na/q) = \mu(q)\sum_n u_n/\varphi(q)$ when the sequence (u_n) is evenly distributed among the invertible residue classes.

Proof. We simply write (with an obvious notation)

$$\sum_{a \bmod^* q} S(a/q) = \sum_{n\leqslant N} u_n c_q(n) = \mu(q)S(0)$$

and consequently, when q is square-free, we have

$$|S(0)| \leqslant \sum_{a \bmod^* q} |S(a/q)|. \tag{1.13}$$

Cauchy's inequality ends the proof. □

We will again see (1.13) in Chapter 9 in Eq. (9.4).

2 Local estimates, the hard way through

We work in this chapter on the relations between functions over $\mathbb{Z}/q\mathbb{Z}$, with a support maybe restricted to some subset $\mathcal{K}_q$, and functions over $\mathbb{Z}/d\mathbb{Z}$ when d divides q. We develop a "local" L^2-formalism that shares several similarities with the Fourier formalism, either over the additive group $(\mathbb{Z}/q\mathbb{Z},+)$ or over the multiplicative group $((\mathbb{Z}/q\mathbb{Z})^*,\times)$. Some geometric notions are required to handle a more general situation which we introduce in this chapter. Of this setting, we shall use Theorem 1.1 intensively, and some of the vocabulary.

A general theory of "characters" and "conductors" is developed in [35]. It also appears in [41].

2.1. Introduction

Let $\mathscr{F}(\mathcal{K}_q)$ be the vector space of functions from some subset $\mathcal{K}_q \subset \mathbb{Z}/q\mathbb{Z}$ to $\mathbb{C}$. We shall regularly consider functions of $\mathscr{F}(\mathcal{K}_q)$ as functions from $\mathscr{F}(\mathbb{Z}/q\mathbb{Z})$ that vanish outside $\mathcal{K}_q$.

When $f \in \mathscr{F}(\mathbb{Z}/q\mathbb{Z})$, we may decompose it according to additive characters via the Fourier transform:

$$\begin{aligned} f(n) &= \sum_{a \bmod q} \Big(\frac{1}{q} \sum_{c \bmod q} f(c)e(-ac/q)\Big)e(na/q) \\ &= \sum_{d|q} \sum_{a \bmod^* d} \Big(\frac{1}{q} \sum_{c \bmod q} f(c)e(-ac/d)\Big)e(na/d) = \sum_{d|q} f_d(n) \end{aligned} \tag{2.1}$$

say, where f_d is a function that depends only on n modulo d. We may do the same when the support of f lies in $(\mathbb{Z}/q\mathbb{Z})^*$ with multiplicative Dirichlet characters. When n is prime to q, we have:

$$\begin{aligned} f(n) &= \sum_{\chi \bmod q} \Big(\frac{1}{\varphi(q)} \sum_{c \bmod^* q} f(c)\overline{\chi}(c)\Big)\chi(n) \\ &= \sum_{d|q} \sum_{\chi \bmod^* d} \Big(\frac{1}{\varphi(q)} \sum_{c \bmod q} f(c)\overline{\chi(c)}\Big)\chi(n) = \sum_{d|q} f_d^*(n) \end{aligned} \tag{2.2}$$

say, where again the f_d^*'s are functions that depend only on n modulo d. We used above $a \bmod^* d$ to say that a is running through every invertible classes modulo d, and later $\chi \bmod^* d$ to say that χ is running through every character of conductor d. We may confuse the character, say χ_1, modulo q induced by a character χ, and χ itself because c and n can be restricted to being prime to q, in which case we have $\chi_1(c) = \chi(c)$ and $\chi_1(n) = \chi(n)$.

[35] O. Ramaré, 2007, "An explicit result of the sum of seven cubes".
[41] O. Ramaré, 2022, "The number of rationals determined by large sets of sifted integers".

2.2. Some geometry modulo q

We consider a collection $(\mathcal{K}_q)_{q\in\mathcal{Q}}$ that is such that:

- We have $\mathcal{K}_q \subset \mathbb{Z}/q\mathbb{Z}$. This is an obvious requirement, but is put to clarify the setting for readers that would skip too many paragraphs!
- $\mathcal{Q}$ is some divisor-closed set* of moduli, maybe $\mathbb{N}\backslash\{0\}$, or $\{q \leqslant Q\}$ or the set of square-free integers. The choice of this set is usually of no consequence, it is introduced only to ease usage, so no effort should be made at this level.
- The sequence should be *consistent*, i.e. when $d|q$, we have $\mathcal{K}_d = \mathcal{K}_q/d\mathbb{Z}$. This could be stated with a lot of symbols, introducing the canonical projection $\sigma_{q\to d}$ from $\mathbb{Z}/q\mathbb{Z}$ to $\mathbb{Z}/d\mathbb{Z}$ and saying that $\sigma_{q\to d}(\mathcal{K}_q) = \mathcal{K}_d$.

We call such a collection a *compact set*. This terminology has some history: if we are to consider the profinite completion $\mathbb{Z}$ along $\mathbb{Q}$, our sequence indeed lifts as a topological compact.

This being set, we need two regularity notions:

- First, the geometric equivalent of multiplicativity. A compact set is said to be *multiplicative*† whenever $\mathcal{K}_{q_1} \times \mathcal{K}_{q_2} \simeq \mathcal{K}_{q_1q_2}$ as soon as q_1 and q_2 are coprime. This is to say that the arrow
$$\begin{array}{rcl} \mathcal{K}_{q_1q_2} & \to & \mathcal{K}_{q_1} \times \mathcal{K}_{q_2} \\ x & \mapsto & (x \bmod q_1, x \bmod q_2) \end{array}$$
is one-to-one. This is a very natural notion which is easily checked. All our examples will be multiplicative.
- We recall the *Johnsen-Gallagher condition* introduced in the previous chapter. We say that this condition holds whenever
$$\forall d|q, \forall y \in \mathcal{K}_d, \#\{x \in \mathcal{K}_q : x \equiv y[d\} = |\mathcal{K}_q|/|\mathcal{K}_d|. \tag{1.4}$$
This is equivalent to saying that the number of preimages in $\mathcal{K}_q$ of any point y of $\mathcal{K}_d$ does not depend on y. When the set of moduli $\mathcal{Q}$ contains only square-free moduli, which is the case in most sieving situations, this condition automatically holds.

We finally endow $\mathscr{F}(\mathcal{K}_q)$ with a hermitian product given by (1.5). When $\mathcal{K}_q = \mathbb{Z}/q\mathbb{Z}$, we simply write $[f,g]_q$. Notice that if f and g are in $\mathscr{F}(\mathcal{K}_q)$ and are considered as elements of $\mathscr{F}(\mathbb{Z}/q\mathbb{Z})$ that vanish outside $\mathcal{K}_q$, then $[f,g]_{K_q} = (q/|\mathcal{K}_q|)[f,g]_q$. In particular orthogonality is preserved.

*I.e. when $d|q$ and $q \in \mathcal{Q}$, we should have $d \in \mathcal{Q}$. In particular, 1 always belongs to $\mathcal{Q}$.

†In earlier work, I used *multiplivatively split* instead of the simpler *multiplicative*.

2.3. Local couplings

Our next task is to link together the arithmetic modulo distinct moduli. Let us assume that the compact set $(\mathcal{K}_q)_{q\in\mathcal{Q}}$ satisfies the Johnsen-Gallagher condition (1.4). To do so, we consider the usual lift when $d|q$:

$$\begin{array}{rccl} L_{\tilde{q}}^{\tilde{d}}: & \mathscr{F}(\mathcal{K}_d) & \to & \mathscr{F}(\mathcal{K}_q) \\ & f & \mapsto & f\circ\sigma_{q\to d}: \begin{array}{ccc} \mathcal{K}_q & \to & \mathbb{C} \\ x & \mapsto & f(x \bmod d) \end{array} \end{array} \tag{2.3}$$

This function is a natural one. The reader may wonder why we chose $\tilde{q}$ instead of q; it will avoid troubles later on. In order to further compare the Hermitian structures, we consider the operator $J_{\tilde{d}}^{\tilde{q}}$ from $\mathscr{F}(\mathcal{K}_q)$ to $\mathscr{F}(\mathcal{K}_d)$ which associates to $f\in\mathscr{F}(\mathcal{K}_q)$ the function

$$J_{\tilde{d}}^{\tilde{q}}(f):\mathcal{K}_d\to\mathbb{C},\qquad x\mapsto\frac{|\mathcal{K}_d|}{|\mathcal{K}_q|}\sum_{\substack{n\in\mathcal{K}_q\\ n\equiv x[d]}} f(n). \tag{2.4}$$

This operator satisfies the fundamental:

$$[L_{\tilde{q}}^{\tilde{d}}(f)|g]_q=[f|J_{\tilde{d}}^{\tilde{q}}(g)]_d. \tag{2.5}$$

Proof. We simply check directly that

$$\begin{aligned}[L_{\tilde{q}}^{\tilde{d}}(f),g]_q&=\frac{1}{|\mathcal{K}_q|}\sum_{x\in\mathcal{K}_d}f(x)\sum_{\substack{n\in\mathcal{K}_q\\ n\equiv x[d]}}\overline{g(n)}\\ &=\frac{1}{|\mathcal{K}_d|}\sum_{x\in\mathcal{K}_d}f(x)\overline{\left(\frac{|\mathcal{K}_d|}{|\mathcal{K}_q|}\sum_{\substack{n\in\mathcal{K}_q\\ n\equiv x[d]}}g(n)\right)}\end{aligned}$$

as required. □

Thus the maps $L_{\tilde{q}}^{\tilde{d}}$ and $J_{\tilde{d}}^{\tilde{q}}$ are adjoint to each other, even if the reader may be unfamiliar with the concept when applied to linear functions that are *not* homomorphisms. Let us define

$$U_{\tilde{q}\to\tilde{d}}=L_{\tilde{q}}^{\tilde{d}}J_{\tilde{d}}^{\tilde{q}}. \tag{2.6}$$

The next section is devoted to understanding these operators. Note that they depend on $\mathcal{K}$, so whenever required, we shall add this parameter as in $U_{\tilde{q}\to\tilde{d}}(f;\mathcal{K})$.

2.4. The Fourier structure

Let us assume that the compact set $(\mathcal{K}_q)_{q\in\mathcal{Q}}$ satisfies the Johnsen-Gallagher condition (1.4). We start with the following fundamental property.

Lemma 2.1
The operator $U_{\tilde{q}\to\tilde{d}}$ is Hermitian. Furthermore, $U_{\tilde{q}\to\tilde{d_1}}$ and $U_{\tilde{q}\to\tilde{d_2}}$ commute with each other and we have

$$U_{\tilde{q}\to\tilde{d_1}}U_{\tilde{q}\to\tilde{d_2}} = U_{\tilde{q}\to\widetilde{(d_1,d_2)}}. \tag{2.7}$$

Proof. The Hermitian property is quickly proved:

$$\begin{aligned}[U_{\tilde{q}\to\tilde{d}}(f)|g]_q &= [L^{\tilde{d}}_{\tilde{q}}J^{\tilde{q}}_{\tilde{d}}(f)|g]_q = [J^{\tilde{q}}_{\tilde{d}}(f)|J^{\tilde{q}}_{\tilde{d}}(g)]_q \\ &= \overline{[J^{\tilde{q}}_{\tilde{d}}(g)|J^{\tilde{q}}_{\tilde{d}}(f)]_q} = \overline{[L^{\tilde{d}}_{\tilde{q}}J^{\tilde{q}}_{\tilde{d}}(g)|f]_q} = [f|L^{\tilde{d}}_{\tilde{q}}J^{\tilde{q}}_{\tilde{d}}(g)]_q,\end{aligned}$$

where, in fact, we have not used any property of $\mathcal{K}$. The commuting property requires more hypothesis. By using the definition of $U_{\tilde{q}\to\tilde{d_1}}$, we find that

$$U_{\tilde{q}\to\tilde{d_1}}U_{\tilde{q}\to\tilde{d_2}}(f)(x) = \frac{|\mathcal{K}_{d_1}|}{|\mathcal{K}_q|}\sum_{\substack{n\in\mathcal{K}_q\\ n\equiv x[d_1]}} U_{\tilde{q}\to\tilde{d_2}}(f)(n)$$

into which we substitute the definition of $U_{\tilde{q}\to\tilde{d_2}}$ to obtain

$$\begin{aligned}U_{\tilde{q}\to\tilde{d_1}}U_{\tilde{q}\to\tilde{d_2}}(f)(x) &= \frac{|\mathcal{K}_{d_1}|}{|\mathcal{K}_q|}\sum_{\substack{n\in\mathcal{K}_q\\ n\equiv x[d_1]}}\frac{|\mathcal{K}_{d_2}|}{|\mathcal{K}_q|}\sum_{\substack{m\in\mathcal{K}_q\\ m\equiv n[d_2]}} f(m) \\ &= \frac{|\mathcal{K}_{d_1}||\mathcal{K}_{d_2}|}{|\mathcal{K}_q|}\sum_{m\in\mathcal{K}_q} W(m;x)f(m),\end{aligned}$$

say, where we have written $W(m;x)$ to denote

$$W(m;x) = \sum_{\substack{n\in\mathcal{K}_q\\ n\equiv x[d_1], n\equiv m[d_2]}} 1. \tag{2.8}$$

The reader will check that $W(m;x) = 0$ when m and x are not congruent modulo (d_1,d_2) and that $W(m;x) = |\mathcal{K}_{d_1}||\mathcal{K}_{d_2}|/|\mathcal{K}_{(d_1,d_2)}|$ when they are. This proves (2.7), and consequently the fact that the operators $U_{\tilde{q}\to\tilde{d}}$ commute with each other. Note that this argument depends crucially on the split multiplicativity of $\mathcal{K}$. □

A consequence of the above lemma is that $U_{\tilde{q}\to\tilde{d}}$ is a Hermitian projection. Let us further define

$$U_{\tilde{q}\to d} = \sum_{\ell|d} \mu(d/\ell) U_{\tilde{q}\to\tilde{\ell}}. \tag{2.9}$$

The main structure theorem is the following:

Theorem 2.2

The operators $(U_{\tilde{q}\to d})_{d|q}$ are two by two orthogonal Hermitian projections. For each divisor r of q, we further have

$$U_{\tilde{q}\to\tilde{r}} = \sum_{d|r} U_{\tilde{q}\to d}.$$

Note that $U_{\tilde{q}\to\tilde{q}}$ is the identity.

Proof. On applying the preceding lemma, we get

$$\begin{aligned} U_{\tilde{q}\to d_1} U_{\tilde{q}\to d_2} &= \sum_{\ell_1|d_1,\ell_2|d_2} \mu(d_1/\ell_1)\mu(d_2/\ell_2) U_{\tilde{q}\to\widetilde{(\ell_1,\ell_2)}} \\ &= \sum_{t|(d_1,d_2)} \Bigg(\sum_{\substack{\ell_1|d_1,\ell_2|d_2 \\ (\ell_1,\ell_2)=t}} \mu(d/\ell_1)\mu(d/\ell_2) \Bigg) U_{\tilde{q}\to\tilde{t}}. \end{aligned}$$

The inner coefficient is multiplicative and is readily seen to vanish when $d_1 \neq d_2$ and to equal $\mu(d_1/t)$ otherwise, thus establishing that $U_{\tilde{q}\to d}$ is indeed a projection and that $U_{\tilde{q}\to d_1}$ and $U_{\tilde{q}\to d_2}$ are orthogonal when $d_1 \neq d_2$. The remaining statements follow. □

Theorem 2.2 is the main basis of what now follows. Let us set

$$\mathfrak{M}(\tilde{q}\to d) = U_{\tilde{q}\to d}\mathscr{F}(\mathcal{K}_q) \tag{2.10}$$

which we endow with the scalar product of $\mathscr{F}(\mathcal{K}_q)$. This set depends on q: it is made of functions over $\mathcal{K}_q$ but this dependence proved to be immaterial in the next lemma.

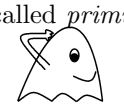
Lemma 2.3

When $d|q$, the arrows in

$$\mathfrak{M}(\tilde{q}\to d) \quad \underset{L^{\tilde{d}}_{\tilde{q}}}{\overset{J^{\tilde{q}}_{\tilde{d}}}{\rightleftarrows}} \quad \mathfrak{M}(\tilde{d}\to d)$$

are isometries, inverses of one another.

This lemma legitimates a special name for $\mathfrak{M}(\tilde{d}\to d)$, which we simply call $\mathfrak{M}(d)$.Its elements will sometimes be called *primitive*.



Proof. We first note that $L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d} J_{\tilde{d}}^{\tilde{q}}(F) = U_{\tilde{d}\to d}(F)$, which in passing proves that $L_{\tilde{q}}^{\tilde{d}}\mathfrak{M}(\tilde{d}\to d) = \mathfrak{M}(\tilde{q}\to d)$. Next, given any two elements $U_{\tilde{d}\to d}(f)$ and $U_{\tilde{d}\to d}(g)$ of $\mathfrak{M}(\tilde{d}\to d)$, we have

$$[L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d}(f)|L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d}(g)]_d = [L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d} J_{\tilde{d}}^{\tilde{q}}(F)|L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d} J_{\tilde{d}}^{\tilde{q}}(G)]_d$$

if we write $f = J_{\tilde{d}}^{\tilde{q}}(F)$ and $g = J_{\tilde{d}}^{\tilde{q}}(G)$. We continue simply:

$$\begin{aligned}[L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d} J_{\tilde{d}}^{\tilde{q}}(F)|L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d} J_{\tilde{d}}^{\tilde{q}}(G)]_d &= [U_{\tilde{q}\to d}(F)|U_{\tilde{q}\to d}(G)]_d = [U_{\tilde{q}\to d}(F)|G]_d \\ &= [L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d} J_{\tilde{d}}^{\tilde{q}}(F)|G]_d = [U_{\tilde{d}\to d} J_{\tilde{d}}^{\tilde{q}}(F)|J_{\tilde{d}}^{\tilde{q}}(G)]_d \\ &= [U_{\tilde{d}\to d}(f)|g]_d = [U_{\tilde{d}\to d}(f)|U_{\tilde{d}\to d}(g)]_d\end{aligned}$$

indeed establishing that the restriction of $L_{\tilde{q}}^{\tilde{d}}$ to $\mathfrak{M}(\tilde{q}\to d)$ is an isometry. To show that both operators are inverses of each other, we note that

$$L_{\tilde{q}}^{\tilde{d}}\big(J_{\tilde{d}}^{\tilde{q}} L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d}(f)\big) = U_{\tilde{q}\to d} U_{\tilde{q}\to d}(F) = U_{\tilde{q}\to d}(F) = L_{\tilde{q}}^{\tilde{d}}\left(U_{\tilde{d}\to d}(f)\right)$$

and since L is an injection, this indeed implies that $J_{\tilde{d}}^{\tilde{q}} L_{\tilde{q}}^{\tilde{d}} U_{\tilde{d}\to d}(f) = U_{\tilde{d}\to d}(f)$. The reverse equation

$$L_{\tilde{q}}^{\tilde{d}} J_{\tilde{d}}^{\tilde{q}} U_{\tilde{q}\to d}(f') = U_{\tilde{q}\to d}(f')$$

is quickly proved. □

Thus in the relation

$$\mathscr{F}(\mathcal{K}_q) = \bigoplus_{r|q} \mathfrak{M}(\tilde{q}\to r) \tag{2.11}$$

we may regroup $\bigoplus_{r|d}\mathfrak{M}(\tilde{q}\to r)$ for some divisor d of q and identify it with $\mathscr{F}(\mathcal{K}_d)$ via L or J, and this identification respects each summand. We may then identify $\mathscr{F}(\mathcal{K}_d)$ with the set of functions of $\mathscr{F}(\mathcal{K}_q)$ that depend only on the class of the variable modulo d, and $\mathfrak{M}(\tilde{q}\to r)$ as being the functions that depend only on the class of the variable modulo r, where r is minimal subject to this condition. Naturally, r is some kind of *conductor*.

We may split f according to (2.11), which we term *decomposing f in Fourier components*, and this is done via

$$f = \sum_{r|q} U_{\tilde{q}\to r}(f). \tag{2.12}$$

Note finally that

$$\big\|U_{\tilde{q}\to q}(f)\big\|_q^2 = \frac{1}{|\mathcal{K}_q|}\sum_{n\in\mathcal{K}_q}\left|\sum_{d|q}\mu(q/d)\frac{|\mathcal{K}_d|}{|\mathcal{K}_q|}\sum_{m\equiv n[d]} f(m)\right|^2. \tag{2.13}$$

2.5. Reduction to local properties

Let us assume that the compact set $(\mathcal{K}_q)_{q\in\mathcal{Q}}$ satisfies the Johnsen-Gallagher condition (1.4). Given a sequence $(u_n)_{n\geqslant 1}$ carried by $\mathcal{K}$ up to level D^*, we consider

$$\Delta_d(u)(n) = |\mathcal{K}_d| \sum_{m\equiv n[d]} u_m \tag{2.14}$$

which is a function of $\mathscr{F}(\mathcal{K}_d)$ provided $d \leqslant D$, which we assume. We have chosen this normalisation because it yields

$$J_{\tilde{d}}^{\tilde{q}}\Delta_q = \Delta_d, \tag{2.15}$$

allowing us to use either notion. In particular, it implies that

$$\left\|U_{\tilde{q}\to d}(\Delta_q(u))\right\|_q^2 = \left\|U_{\tilde{d}\to d}(\Delta_d(u))\right\|_d^2 \tag{2.16}$$

whenever $d|q$.

Though we will not use it, expressing explicitely the adjoint of Δ_q is easy. We consider

$$\begin{array}{rcccl} \nabla_q & : & \mathscr{F}(\mathbb{Z}/q\mathbb{Z}) & \longrightarrow & \mathscr{F}([1,N]) \\ & & h & \mapsto & \nabla_q(h) : \ [1,N] \longrightarrow \mathbb{C} \\ & & & & \qquad\qquad x \mapsto h(x \bmod q) \end{array} \tag{2.17}$$

which satisfies

$$[\Delta_q(h_1)|h_2]_q = [h_1|\nabla_q(h_2)], \tag{2.18}$$

(where $[h,g] = \sum_{n\leqslant N} h(n)\overline{g(n)}$) justifying again our scaling in the definition of Δ_q. Note further that

$$\forall d|q, \quad \nabla_q(L_{\tilde{q}}^{\tilde{d}}(h)) = \nabla_d(h), \tag{2.19}$$

both properties stated with obvious notations.

2.6. Some explicit expression in the reference case

When the compact set $\mathcal{K}$ is $(\mathbb{Z}/q\mathbb{Z})_q$, also denoted $\hat{\mathbb{Z}}$, we have at our disposal the usual Fourier decomposition

$$f(n) = \sum_{d|q} \sum_{a \bmod^* d} \hat{f}(q, a/d) e(na/d) \tag{2.20}$$

where we have set

$$\hat{f}(q, a/d) = \frac{1}{q} \sum_{n \bmod q} f(n) e(-na/d). \tag{2.21}$$

*I.e. such that $\forall n$ for which $u_n \neq 0$, we have $n \in \mathcal{K}_d$ for every $d \leqslant D$.

This decomposition is the one given by (2.12), for we immediately check that

$$U_{\tilde{q}\to d}(f;\hat{\mathbb{Z}})(n) = \sum_{a \bmod^* d} \hat{f}(q, a/d)e(na/d). \tag{2.22}$$

Let us give a formal proof of this last equality.

Proof. Using (2.9), we infer that

$$U_{\tilde{q}\to d}(f;\hat{\mathbb{Z}})(n) = \sum_{r|d} \mu(d/r)\frac{r}{q} \sum_{m\equiv n[r]} f(m) = \frac{1}{q}\sum_{m \bmod q} f(m) \sum_{\substack{r|d \\ r|m-n}} r\mu(d/r)$$

where we identify the summation over r as the Ramanujan sum $c_d(n-m)$. On using the expression of this Ramanujan sum as a sum of exponentials, we get

$$\begin{aligned} U_{\tilde{q}\to d}(f;\hat{\mathbb{Z}})(n) &= \frac{1}{q}\sum_{m \bmod q} f(m)c_d(n-m) \\ &= \frac{1}{q}\sum_{a \bmod d}\sum_{m \bmod q} f(m)e(-ma/d)e(na/d) \end{aligned}$$

where the reader will swiftly recognize (2.22). □

The above formula applies when f is a function modulo q, while in practice, we deal with functions on the integers. So let us see what happens when $f = \Delta_q(u)$. Here is the main result in this situation that we for instance use for the large sieve inequality.

$$\left\|U_{\tilde{q}\to q}(\Delta_q(u);\hat{\mathbb{Z}})\right\|_q^2 = \sum_{a \bmod^* q}\left|\sum_m u_m e(-am/q)\right|^2. \tag{2.23}$$

Proof. We follow the definitions one by one to reach

$$\begin{aligned} U_{\tilde{q}\to q}(\Delta_q(u);\hat{\mathbb{Z}})(n) &= \sum_{a \bmod^* q}\left(\frac{1}{q}\sum_{b \bmod q} q \sum_{m\equiv b[q]} u_m e(-ab/q)\right)e(na/q) \\ &= \sum_{a \bmod^* q}\sum_m u_m e(-am/q)\, e(na/q). \end{aligned}$$

We are now ready to compute $\|U_{\tilde{q}\to q}(\Delta_q(u);\hat{\mathbb{Z}})\|_q^2$. We find that

$$\left\|U_{\tilde{q}\to q}(\Delta_q(u);\hat{\mathbb{Z}})\right\|_q^2 = \frac{1}{q}\sum_{n \bmod q}\left|\sum_{a \bmod^* q}\sum_m u_m e(-am/q)\, e(na/q)\right|^2$$

i.e.

$$\left\|U_{\tilde{q}\to q}(\Delta_q(u);\hat{\mathbb{Z}})\right\|_q^2 = \sum_{a \bmod^* q}\left|\sum_m u_m e(-am/q)\right|^2$$

as we have claimed. □

2.7. A local lower bound

Let us assume that the compact set $(\mathcal{K}_q)_{q\in\mathcal{Q}}$ satisfies the Johnsen-Gallagher condition (1.4). In this section, we use $U_{\tilde{q}\to d}(\cdot;\hat{\mathbb{Z}})$ for the sequence of operators U associated to $\mathcal{K}=(\mathbb{Z}/d\mathbb{Z})_d$, denote by $\|\cdot\|_q$ the corresponding norm (this norm depends on the ambient compact set), and also $\Delta(\cdot;\hat{\mathbb{Z}})$ for the projection operator. We reserve $U_{\tilde{q}\to d}$ for the operators associated with $\mathcal{K}$, and exceptionally here $\|\cdot\|_{\mathcal{K}_q}$ for the relevant norm. Note that if f is in $\mathscr{F}(\mathcal{K}_d)$, then $\|f\|_q^2=|\mathcal{K}_q|\|f\|_{\mathcal{K}_q}^2/q$ and $\Delta_d(\cdot;\hat{\mathbb{Z}})=d\Delta_d/|\mathcal{K}_d|$. When we know that our sequence is carried by a smaller compact set $\mathcal{K}$, we may introduce this information via the following transformation (see (2.9)):

$$\begin{aligned}\big\|U_{\tilde{q}\to q}(\Delta_q(u;\hat{\mathbb{Z}});\hat{\mathbb{Z}})\big\|_q^2&=\sum_{d|q}\mu(q/d)\big\|\Delta_d(u;\hat{\mathbb{Z}})\big\|_d^2=\sum_{d|q}\mu(q/d)\frac{d}{|\mathcal{K}_d|}\big\|\Delta_d(u)\big\|_{\mathcal{K}_d}^2\\&=\sum_{d|q}\mu(q/d)\frac{d}{|\mathcal{K}_d|}\sum_{r|d}\big\|U_{\tilde{r}\to r}(\Delta_r(u))\big\|_{\mathcal{K}_r}^2.\end{aligned}$$

Let us continue this line of thoughts, though only locally, i.e. modulo q. On recalling (2.23), the above equation gives us the formula:

$$\begin{aligned}&\sum_{a\bmod^* q}\Big|\sum_n u_n e(na/q)\Big|^2\\&\qquad=\sum_{r|q}\prod_{\substack{p^\alpha\|q\\p^\alpha\|r}}\frac{p^\alpha}{|\mathcal{K}_{p^\alpha}|}\prod_{\substack{p^\alpha\|q\\p^\alpha\nmid r}}\Big(\frac{p^\alpha}{|\mathcal{K}_{p^\alpha}|}-\frac{p^{\alpha-1}}{|\mathcal{K}_{p^{\alpha-1}}|}\Big)\big\|U_{\tilde{r}\to r}(\Delta_r(u))\big\|_{\mathcal{K}_r}^2.\end{aligned}\tag{2.24}$$

We have therefore reached a decomposition of the quadratic form on the LHS on the space of sequences carried by $\mathcal{K}_q$ as a sum of the norm of orthogonal Hermitian projections. When q is square-free, Huxley in Section 6 of [20] proves and uses such a decomposition. Here is another proof of Theorem 1.1 based on these ideas.

Second proof of Theorem 1.1. The proof is easier to follow when q is square-free. When $\sum_n u_n=0$, the contribution of $r=1$ in (2.24) vanishes, leading to:

$$\sum_{a\bmod^* q}\Big|\sum_n u_n e(na/q)\Big|^2\geqslant\prod_{p|q}\frac{p-|\mathcal{K}_p|}{|\mathcal{K}_p|}\min_{1<r|q}\frac{r}{\prod_{p|r}(p-|\mathcal{K}_p|)}\sum_{r|q}\big\|U_{\tilde{r}\to r}\Delta_r(u)\big\|_{\mathcal{K}_r}^2.$$

The reader will swiftly complete the proof of this case by using

$$\sum_{r|q}\big\|U_{\tilde{r}\to r}\Delta_r(u)\big\|_{\mathcal{K}_r}^2=|\mathcal{K}_q|\sum_{b\in\mathcal{K}_q}\Big|\sum_{n\equiv b[q]}u_n\Big|^2,$$

and adapt the proof to the case of non-square-free modulus q. □

[20] M. Huxley, 1972, "Irregularity in sifted sequences".

3 Local estimates, size condition

Up to now, we did not investigate precisely what happens at the place at infinity. The approach we follow here is due to Selberg to prove the large sieve inequality; in particular he built the function $C_{[-M;M],\delta}$ given below. It turned out that Beurling had already achieved such a construction in the late 1930's without publishing. This explains why this function is now referred to as the Beurling-Selberg function.

The next exposition is based on the paper [54] of Vaaler (see also the joint work [14] by Graham & Vaaler). We explicitate several formulas by simply reading carefully this paper.

3.1. Some special functions

Let us follow the classical paper [54] by J. Vaaler. We first introduce some notation and define*

$$\hat{J}(t) = \begin{cases} 1 & \text{if } t = 0, \\ \pi t(1-|t|)\cot \pi t + |t| & \text{if } 0 < |t| < 1, \\ 0 & \text{if } 1 \leqslant |t|. \end{cases} \tag{3.1}$$

As per [54, Theorem 6], the function $\hat{J}$ is even, non-negative, continuously differentiable, and strictly decreasing on $[0,1]$. We also have $J(x) = H'(x)/2$ where H is defined in (3.2) below. We continue with the definitions from the same paper:

$$K(z) = \left(\frac{\sin \pi z}{\pi z}\right)^2, H(z) = \left(\frac{\sin \pi z}{\pi}\right)^2 \left(\sum_{m\in\mathbb{Z}} \frac{\operatorname{sgn}(m)}{(z-m)^2} + \frac{2}{z}\right) \tag{3.2}$$

(with $\operatorname{sgn} 0 = 1$), and the Beurling-Selberg function $B(x) = K(x) + H(x)$. We finally set

$$2C_{[a,b],\delta}(x) = B(\delta(b-x)) + B(\delta(x-a)) \geqslant 2 \cdot 1_{x\in[a,b]}. \tag{3.3}$$

Theorem 3.1

Let $\delta > 0$ and $M \geqslant 0$ be two parameters. The function $C_{[-M,M],\delta}$ is an upper bound for the characteristic function of $[-M, M]$. When $|t| \leqslant \delta$, we

[54] J. Vaaler, 1985, "Some Extremal Functions in Fourier Analysis".
[14] S. Graham and J. Vaaler, 1981, "A class of extremal functions for the Fourier transform".
*The Fourier transform $\hat{f}$ is defined by $\hat{f}(t) = \int_{-\infty}^{\infty} f(u)e(ut)du$.

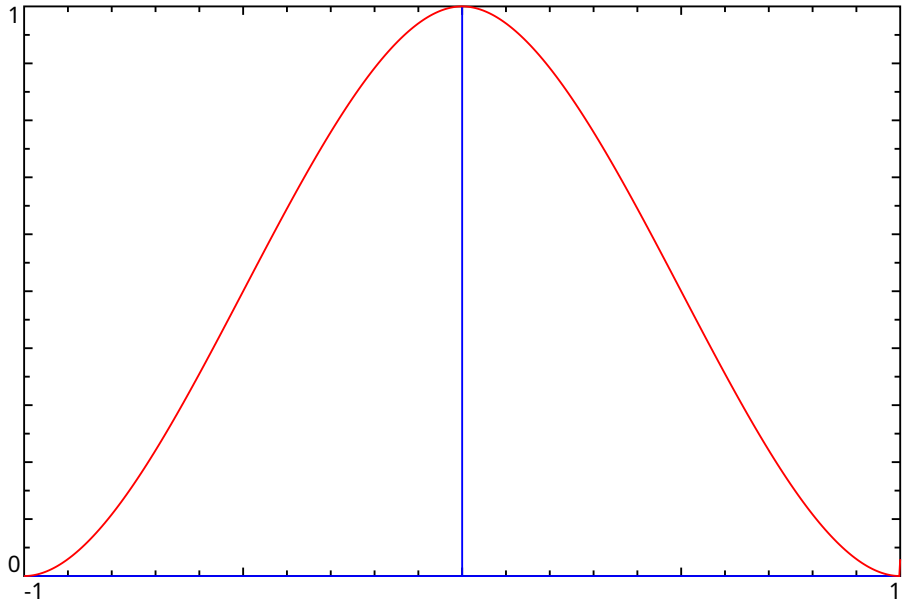


Figure 3.1: $\hat{J}(t)$

have

$$\hat{C}_{[-M,M],\delta}(t) = \delta^{-1}(1-|\delta^{-1}t|)\cos 2\pi Mt + \frac{\hat{J}(\delta^{-1}t)}{\pi t}\sin 2\pi Mt.$$

When $|t| \geqslant \delta$, we have $\hat{C}_{[-M,M],\delta}(t) = 0$. We have $|\hat{C}_{[-M,M],\delta}(t)| \leqslant \hat{C}_{[-M,M],\delta}(0) = 2M + \delta^{-1}$.

Notice that $\hat{C}_{[-M,M],\delta}(t) = \delta^{-1}\hat{C}_{[-\delta M,\delta M],1}(\delta^{-1}t)$.

Proof. Let us start with somewhat more generality and write

$$2C_{[a,b],\delta}(x) = 21_{x\in[a,b]} + B(\delta(b-x)) - \operatorname{sgn}\delta(b-x) + B(\delta(x-a)) - \operatorname{sgn}\delta(x-a).$$

Set $F(x) = B(x) - \operatorname{sgn} x$ and $F_2(x) = F(\delta(b-x))$. By [54, Corollary 7], we have

$$\hat{F}(t) = (1-|t|)1_{|t|\leqslant 1} + \frac{1}{i\pi t}\big(\hat{J}(t)-1\big)$$

and classically $\hat{F}_2(t) = e(bt)\delta^{-1}\hat{F}(-t/\delta)$. This leads to

$$2\hat{C}_{[a,b],\delta}(t) = \frac{e(bt)-e(at)}{i\pi t} + \delta^{-1}e(bt)\hat{F}(-\delta^{-1}t) + \delta^{-1}e(at)\hat{F}(\delta^{-1}t).$$

[54] J. Vaaler, 1985, "Some Extremal Functions in Fourier Analysis".

Let us proceed by specializing $a = -b = -M$, we get

$$2\delta \hat{C}_{[-M,M],\delta}(t) = (e(Mt) + e(-Mt))(1 - |\delta^{-1}t|) + \frac{e(Mt) - e(-Mt)}{i\pi\delta^{-1}t}\hat{J}(\delta^{-1}t).$$

The lemma follows readily from this expression. □

Lemma 3.2

Let $\lambda \in [0, 1]$ be a positive parameter. When $|u| \leqslant 1$, we have

$$\hat{C}_{[-\lambda,\lambda],1}(u) = (1 - |u|)\cos 2\pi\lambda u + \frac{\hat{J}(u)}{\pi u}\sin 2\pi\lambda u.$$

When $|u| \geqslant 1$, we have $\hat{C}_{[-\lambda,\lambda],1}(u) = 0$.
When $|y| \leqslant \lambda$, we have $C_{[-\lambda,\lambda],1}(y) \geqslant 1$.

Proof. By construction, we have $C_{[-\lambda,\lambda],1}(y) \geqslant 1$ when $|y| \leqslant \lambda$. When $\lambda \leqslant 1/4$, the function $(1 - |u|)\cos 2\pi\lambda u$ is decreasing over $[0, 1]$. The same holds true for the functions $\hat{J}(u)$ and $\sin(2\pi\lambda u)/u$, hence for $\hat{C}_{[-\lambda,\lambda],1}(u)$. We readily find that

$$C_{[-\lambda,\lambda],1}(y) = \int_{-1}^{1} \hat{C}_{[-\lambda,\lambda],1}(u)e(-uy)du = 2\int_{0}^{1} \hat{C}_{[-\lambda,\lambda],1}(u)\cos(2\pi uy)du.$$

The lemma follows readily. □

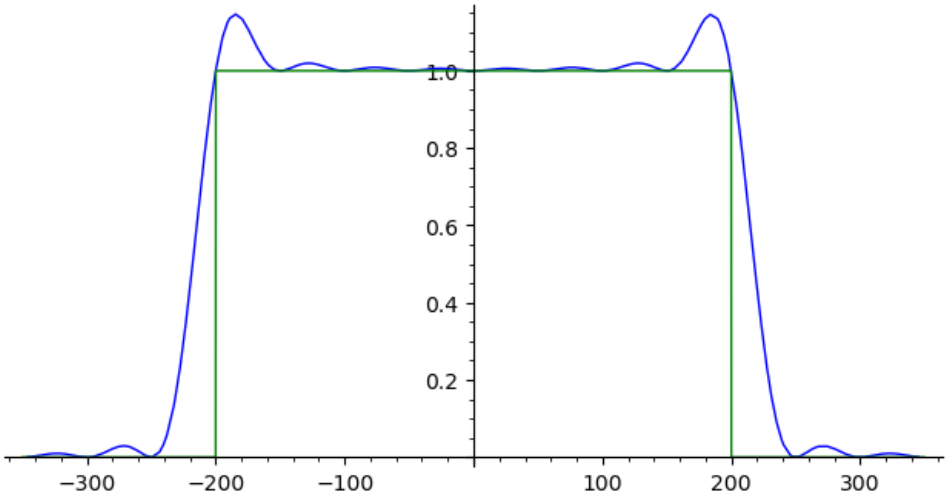


Figure 3.2: $C_{[-250,250],1/50}(u)$

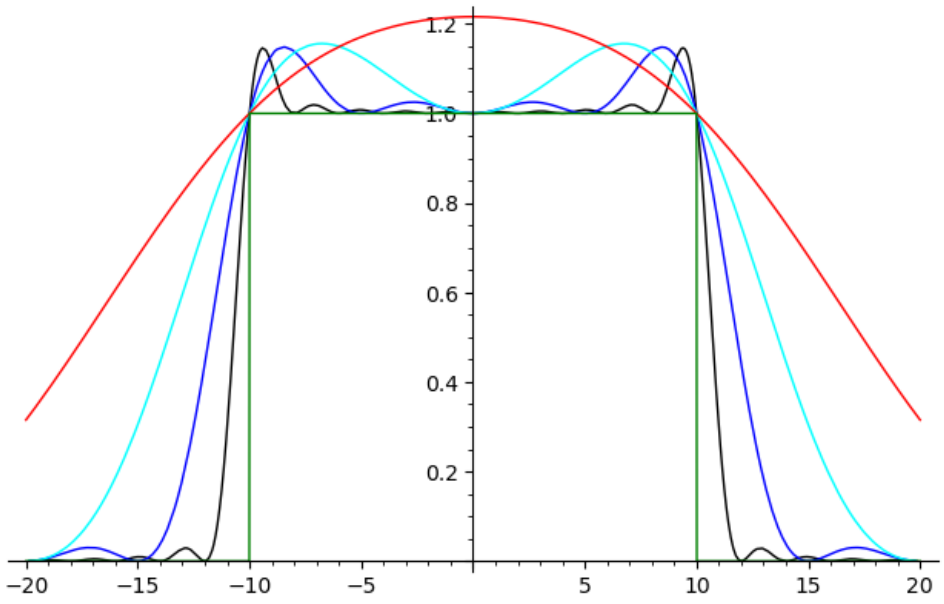


Figure 3.3: $C_{[-10,10],1/2}(u)$ in black, $C_{[-10,10],1/5}(u)$ in blue, $C_{[-10,10],1/10}(u)$ in cyan and $C_{[-10,10],1/20}(u)$ in red

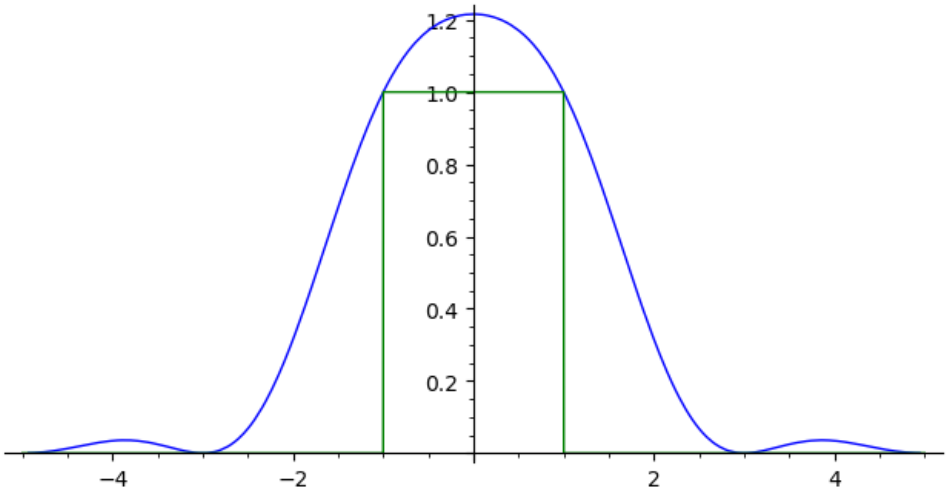


Figure 3.4: $C_{[-1,1],1/2}(u)$

3.2. A quadratic form

This section is somehow off-topic, but prepares to a problem we will find later on. We consider the quadratic form

$$Q((u_n)) = \int_{-\delta}^{\delta} |S(\beta)|^2 d\beta = \sum_{m,n} u_m \overline{u_n} \frac{\sin(2\pi(m-n)\delta)}{\pi(m-n)}. \tag{3.4}$$

To be in accordance with other works, we assume that N is an integer.*

Eigenvectors and eigenvalues

The eigenvectors of this quadratic form are called the *discrete prolate spheroidal sequences*, DPSS in short. Following D. Slepian in [50], we write

$$\sum_m \frac{\sin(2\pi(n-m)\delta)}{\pi(n-m)} v_{m-1}^{(k)}(N,\delta) = \lambda_k(N,\delta) v_{n-1}^{(k)}(N,\delta). \tag{3.5}$$

These eigenvalues are normalized so that

$$\sum_n v_{n-1}^{(k)}(N,\delta)^2 = 1, \tag{3.6}$$

and

$$\sum_n v_{n-1}^{(k)}(N,\delta) \geqslant 0, \quad \sum_n (N+1-2n) v_{n-1}^{(k)}(N,\delta) \geqslant 0. \tag{3.7}$$

They are furthermore ordered in $k \in \{0, \cdots, N-1\}$ with decreasing values of $\lambda_k(N,\delta)$. The reader will find numerous papers and numerics on these sequences. (Voir [17])

A minimization

As it turns out, the problem will shall meet is to find the best constant $C(N,\delta)$ such that

$$\int_{-\delta}^{\delta} |S(\beta)|^2 d\beta \geqslant C(N,\delta) |S(0)|^2. \tag{3.8}$$

And we would even be able to restrict this question to sequences (u_n) that are non-negative, obtaining a constant $C_+(N,\delta)$. A consequence of the later study (see Section 6.8) will show that

$$C_+(N,\delta) \geqslant \left(1 - \frac{2}{\sqrt{\delta N}}\right)/N. \tag{3.9}$$

*Except that D. Slepian considers sequences $(u_n)_{0\leqslant n\leqslant N-1}$ while we have $(u_n)_{1\leqslant n\leqslant N}$. This explains several of the -1 the reader will see in the indices.

[50] D. Slepian, 1978, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case".

[17] F. A. Gruenbaum, 1981, "Eigenvectors of a Toeplitz matrix: Discrete version of the prolate spheroidal wave functions".

It is easy to show that (when $1/2 \geqslant \delta$)

$$1/N \geqslant C(N,\delta) \geqslant C_+(N,\delta). \tag{3.10}$$

This is simply proved by noticing that (when

$$N = \int_{-1/2}^{1/2} \Big|\sum_n e(n\beta)\Big|^2 d\beta \geqslant \int_{-\delta}^{\delta} \Big|\sum_n e(n\beta)\Big|^2 d\beta \geqslant C(N,\delta)\Big|\sum_n 1\Big|^2.$$

Numerical optimization with positivity conditions

We may consider this problem with Kuhn-Tucker conditions, i.e. examine the function

$$H_{N,\delta}((u_n),\lambda) = 2\delta \sum_{m,n} u_m u_n \operatorname{sinc}(2\pi(m-n)\delta) - \sum_n \lambda_n u_n + \mu\Big(\sum_n u_n - 1\Big). \tag{3.11}$$

For every local minima of $2\delta\sum_{m,n} u_m u_n \operatorname{sinc}(2\pi(m-n)\delta)$ under the conditions $-u_n \leqslant 0$ for every n and $\sum_n u_n = 1$, there exists* for every n some $\lambda_n \geqslant 0$ and a real μ such that

$$\begin{cases} \forall m, \quad 2\delta\displaystyle\sum_n u_n \operatorname{sinc}(2\pi(m-n)\delta) = \lambda_m - \mu, \\ \forall m, \quad u_m \geqslant 0 \quad \text{and} \quad \lambda_m u_m = 0, \\ \displaystyle\sum_n u_n = 1. \end{cases} \tag{3.12}$$

In which case the minimum is readily checked to be $-\mu$. This means that there exists a subset $I \subset \{1,\cdots,N\}$ of indices that are such that $u_m = 0$ and that $\lambda_m = 0$ when $m \notin I$. Since in each case, we get a linear system of $N - |I| + 1$ unknowns (the $+1$ is for the μ-variable), with $N-|I|+1$ equations. Such a system is likely to be non-degenerate (its determinant is a real analytic function of δ, so it cannot vanish on any non-trivial interval without vanishing everywhere; this implies that we need only move δ by a tiny amount to get a non-zero determinant).

We get the next table for $NC_+(N,\delta)$:

N	$\delta = 0.2$	$\delta = 0.1$
10	0.8488...	0.7105...
11	0.8593...	0.7333...
12	0.8567...	0.7474...
13	0.8608...	0.7471...
14	0.8794...	0.7458...
15	0.8882...	0.7426...
16	0.8896...	0.7563...
17	0.8852...	0.7773...
18	0.8973...	0.7943...
19	0.9063...	0.8063...
20	0.9098...	0.8143...

*In our present problem, this necessary condition can furthermore be proved to be sufficient.

And a table for $NC_+(N,\delta)/(1-2/\sqrt{\delta N})$:

N	19	20	21	22
$\delta = 0.3$	5.8499...	5.1885...21	4.6818...	4.3113...

Sadly enough, this last data is barely significant as $N\delta$ is not large enough. A major slowing factor is that we have to investigate every possible subset of indices. Notice however that, when an admissible optimum (i.e. a minimum with non-negative coordinates) has been found for a given subset, it is not required to investigate the subsets of this subset. This remark does not speed the process, as it turns out that checking that a given subset does not belong to the list of cleared ones takes too much time.

Here is the Pari-GP script used to compute the above tables.

```
{getmin(listindices, delta) =
   my(locN = length(listindices), mymat = matrix(locN + 1),
      vecimage = vector(locN + 1), res, isok = 1, firstnonzeron = 1);

   \\ Initialization:
   for(m = 1, locN,
      for(n = m, locN,
         mymat[m, n] = sinc(2*Pi*(listindices[m]-listindices[n])*delta)*2*delta;
         mymat[n, m] = mymat[m, n]));
   for(m = 1, locN, mymat[m, locN + 1] = 1);
   for(n = 1, locN, mymat[locN + 1, n] = 1);

   for(n = 1, locN, vecimage[n] = 0);
   vecimage[locN + 1] = 1;

   \\ Get possible minimum:
   res = matsolve(mymat, vecimage~);

   \\ Analyze the result:
   while(res[firstnonzeron] == 0, firstnonzeron++);
   if(res[firstnonzeron] < 0, res = -res);
   for(n = firstnonzeron + 1, locN,
      if(res[n] < 0, isok = 0; break,));
   return([isok, - res[locN + 1]]);}

{work(N, delta, DoTell = 1) =
   my(locres, res = 1);
   forsubset(N, listindices,
      if(length(listindices) == 0,,
         locres = getmin(listindices, delta);
         if(locres[1] == 1,
            if(DoTell == 1,
              print("Minimum at ", listindices, " = ", locres[2]),);
            res = min(locres[2], res),)));
   return(res);}
```

3.3. Reproducing the plots

The figures in this chapter have been drawn by Sage, see [45]. Here is the script "`Curves.sage`" we have used which we load with the command `load("Curves.sage")`.

```
import math

begInterval = -10 # We bound above the characteristic function of
endInterval = 10  # [begInterval, endInterval]

begDraw = -20     # The plot axis is [begDraw, endDraw]
endDraw = 20

mydelta = 1/10
myapectratio = 20 # This should be around (endDraw - begDraw)/2.

def Bz(z, Nmax=100000): # An approximation of
    if z.is_integer():  # the Beurling-Selberg function.
        return(1)
    else:
        aux = 2/z +1/z^2
        for n in [1..Nmax]:
            aux += 1/(z-n)^2-1/(z+n)^2
        return(aux*(sin(float(pi)*z)/float(pi))^2)

def Majo(z, a, b, delta, Nmax=100000):
    return((Bz(delta*(z-a), Nmax)+Bz(delta*(b-z), Nmax))/2)

def CharFun(z, a, b): # A generic version of SpeFun
    if a <= z <= b:
        return(1)
    else:
        return(0)

def SpeFun(z):  # The characteristic function to bound above
    return(CharFun(z, begInterval,endInterval))

plot1a = plot(Majo(x, begInterval, endInterval, 1/2, 6000),
              (x, begDraw, endDraw),
              color = "black", aspect_ratio = myapectratio)
plot1b = plot(Majo(x, begInterval, endInterval, 1/5, 6000),
              (x, begDraw, endDraw),
              color = "blue", aspect_ratio = myapectratio)
```

[45] W. Stein et al., 2024, *Sage Mathematics Software (Version 9.5)*.

```
plot1c = plot(Majo(x, begInterval, endInterval, 1/10, 6000),
              (x, begDraw, endDraw),
              color = "cyan", aspect_ratio = myapectratio)
plot1d = plot(Majo(x, begInterval, endInterval, 1/20, 6000),
              (x, begDraw, endDraw),
              color = "red", aspect_ratio = myapectratio)

plot3 = plot(SpeFun, (x, begDraw, endDraw),
             color = "green", aspect_ratio = myapectratio)

(plot1a + plot1b + plot1c + plot1d + plot3).show()
```

4 The large sieve inequality

The *large sieve* terminology appears for the first time in the paper [23] by Y. Linnik. This concept has evolved largely since then. An historical account may be found in the book [28] by H. Montgomery. See also the paper [27] by the same author.

4.1. The large sieve inequality

Let us start with a complex vector space $\mathcal{H}$ endowed with a hermitian form $[f|g]$, left linear and right sesquilinear. To be consistent with later notations, the norm of φ is denoted by $\|\varphi\|_2$. Our first lemma reads as follows

Lemma 4.1
For any finite family (φ_i^*) of points of $\mathcal{H}$, and with $M_i = \sum_j |[\varphi_i^*|\varphi_j^*]|$, we have

$$\sum_i M_i^{-1} |[f|\varphi_i^*]|^2 \leqslant \|f\|_2^2.$$

So that if $[\varphi_i^*|\varphi_j^*]$ is small for $i \neq j$ then M_i is close to $\|\varphi_i^*\|_2^2$. This lemma is due to Selberg, as mentioned in Section 2 of [4] and in [3].

Proof. For the proof, we simply write

$$\Big\| f - \sum_i \xi_i \varphi_i^* \Big\|_2^2 \geqslant 0$$

and expand the square. We take care of $\| \sum_i \xi_i \varphi_i^* \|_2^2$ by using

$$\Big\| \sum_i \xi_i \varphi_i^* \Big\|_2^2 = \sum_{i,j} \xi_i \overline{\xi_j} [\varphi_i^* | \varphi_j^*]$$

to which we simply apply $2|\xi_i \overline{\xi_j}| \leqslant |\xi_i|^2 + |\xi_j|^2$. We have reached

$$\|f\|_2^2 - 2\Re \sum_i \overline{\xi_i} [f|\varphi_i^*] + \sum_i M_i |\xi_i|^2 \geqslant 0.$$

We finally select $\xi_i = [f|\varphi_i^*]/M_i$. The lemma readily follows. □

[23] Y. Linnik, 1941, "The large sieve".
[28] H. Montgomery, 1971, "Topics in Multiplicative Number Theory".
[27] H. Montgomery, 1978, "The analytic principle of the large sieve".
[4] E. Bombieri, 1987/1974, "Le grand crible dans la théorie analytique des nombres".
[3] E. Bombieri, 1971, "A note on the large sieve".

Theorem 4.2

Let $\mathcal{X}$ be a finite set of points of $\mathbb{R}/\mathbb{Z}$. Set $\delta = \min\{\|x - x'\|, x \neq x' \in \mathcal{X}\}$. For any sequence of complex numbers $(u_n)_{1\leqslant n\leqslant N}$[a], we have

$$\sum_{x\in\mathcal{X}}\Big|\sum_{n\leqslant N} u_n e(nx)\Big|^2 \leqslant \sum_n |u_n|^2(N-1+\delta^{-1}).$$

[a] The parameter N is *not* restricted to integer values.

The theorem in this version is due to Selberg. The same year and by a different method, a marginally weaker version (without the -1 on the right) was proved by Montgomery & Vaughan in [29].

Proof. Instead of localising $n \in [1, N]$, let us set $M = [(N-1)/2]$ and use $n \in [-M, N-M-1] \cap \mathbb{Z} \subset [-M, M]$. For each $x \in \mathcal{X}$, we define

$$\forall n \in \mathbb{Z}, \quad \varphi_x(n) = e(nx)\sqrt{C_{[-M,M],\delta}(n)}. \tag{4.1}$$

We also define $f(n) = u_n$ for $n \in [-M, M]$ and appeal to Selberg's Lemma 4.1. We readily find that

$$\sum_{n\in\mathbb{Z}} \varphi_x(n)\overline{\varphi_{x'}(n)} = \sum_{n\in\mathbb{Z}} e(n(x-x'))C_{[-M,M],\delta}(n) = \hat{C}_{[-M,M],\delta}(x-x')$$

by Poisson's formula. When x and x' are in $\mathcal{X}$, this quantity has value $2M + \delta^{-1}$ when $x = x'$ and vanishes otherwise. This proves our inequality. □

4.2. A global inequality

Here is the common consequence of Theorem 4.2 we shall use.

Theorem 4.3

For any sequence of complex numbers $(u_n)_{1\leqslant n\leqslant N}$, we have

$$\sum_{q\leqslant Q}\sum_{a \bmod^* q}\Big|\sum_{n\leqslant N} u_n e(na/q)\Big|^2 \leqslant \sum_n |u_n|^2(N-1+Q^2).$$

[29] H. Montgomery and R. Vaughan, 1973, "The large sieve".

4.3. Montgomery's sieve and the Brun-Titchmarsh inequality

Joining the global inequality given in Theorem 4.3 together with the local ones given in Theorem 1.1 gives Montomery's sieve. The reader should go through Section 2.2.

Theorem 4.4

Let $(\mathcal{K}_q)_{q\leqslant Q}$ be a multiplicative compact set, i.e. a consistent sequence of multiplicative subsets of $\mathbb{Z}/q\mathbb{Z}$ satisfying the Gallagher-Johnsen condition. Let Z be the number of integers n from $[M+1, M+N]$ that are such that $n \in \mathcal{K}_q$ for every $q \leqslant Q$. We have

$$Z \leqslant \frac{N-1+Q^2}{G^*(Q)} \quad \text{where} \quad G^*(Q) = \sum_{q\leqslant Q} \prod_{p^\alpha \| q} \Big(\frac{p^\alpha}{|\mathcal{K}_{p^\alpha}|} - \frac{p^{\alpha-1}}{|\mathcal{K}_{p^{\alpha-1}}|} \Big).$$

Theorem $\mathscr{A}$ is a consequence of the above, in a strong form. With the above Theorem 4.4, one gets $2+o(1)$ instead of 2. In such a strong form, it is Theorem 2 in [29] by H. Montgomery & R. C. Vaughan.

4.4. Reminder on the G-function for the primes

In the context of the Selberg sieve, there appears a family of G-functions. The reader may find the details in the book [18] by H. Halberstam & H.-E. Richert, or in [34, 36]. For the primes, the main player is

$$G(z) = \sum_{q\leqslant z} \frac{\mu^2(q)}{\varphi(q)}. \tag{4.2}$$

It is evaluated in several places. Here is a simple lemma.

Lemma 4.5. We have $G(z) \geqslant \log z$.

Proof. We find that

$$\frac{\mu^2(q)}{\varphi(q)} = \prod_{p|q} \frac{1}{1-\frac{1}{p}} = \sum_{\substack{n\geqslant 1 \\ p|n \Longrightarrow p|q}} \frac{1}{n}$$

[18] H. Halberstam and H.-E. Richert, 1974, *Sieve methods.*
[34] D. S. Ramana and O. Ramaré, 2025, *Arithmetical aspects of the large sieve inequality – II.*
[36] O. Ramaré, 2009, *Arithmetical aspects of the large sieve inequality.*

and therefore, on setting $k(n) = \prod_{p|n} p$, we find that

$$\frac{\mu^2(q)}{\varphi(q)} = \sum_{\ell : k(\ell)=q} \frac{1}{\ell}.$$

We may check this identity by noticing that we indeed get a sum of $1/\ell$, with $k(\ell) = q$, that every such integer indeed appears, and that it appears only once. Consequently, we get

$$G(z) = \sum_{q \leqslant z} \mu^2(q) \sum_{\substack{n \geqslant 1 \\ p|n \Longrightarrow p|q}} \frac{1}{n} = \sum_{\substack{n \geqslant 1 \\ k(n) \leqslant z}} \frac{1}{n} \geqslant \sum_{n \leqslant z} \frac{1}{n} \geqslant \log z.$$

This ends the proof. □

The function G belongs to the more general family

$$G_k(z) = \sum_{\substack{q \leqslant z \\ (q,k)=1}} \frac{\mu^2(q)}{\varphi(q)}. \tag{4.3}$$

We shall use next simple inequality.

Lemma 4.6. We have $\frac{\varphi(k)}{k} G(zk) \geqslant G_k(z) \geqslant \frac{\varphi(k)}{k} \log z$.

This can be found, for instance, page 23 in the proof Brun-Titchmarsh inequality in [4] by E. Bombieri and has its origin in the paper [24] by J.E. van Lint and H.E. Richert (around Eq. (1.1)–(1.3) therein).

Proof. We have

$$\begin{aligned} G(z) &= \sum_{d|k} \sum_{\substack{n \leqslant z \\ (n,k)=d}} \frac{\mu^2(n)}{\varphi(n)} = \sum_{d|k} \frac{\mu^2(d)}{\varphi(d)} \sum_{\substack{n \leqslant z/d \\ (n,k)=1}} \frac{\mu^2(n)}{\varphi(n)} \\ &= \sum_{d|k} \frac{\mu^2(d)}{\varphi(d)} G_k(z/d) \begin{cases} \leqslant \dfrac{k}{\varphi(k)} G_k(z), \\ \geqslant \dfrac{k}{\varphi(k)} G_k(z/k). \end{cases} \end{aligned}$$

This inequality together with Lemma 4.5 leads to the announced inequalities. □

In Chapter 8, we shall also require an upper bound. The litterature has a large amount of work on very precise estimates (see for instance Theorem 3.1 in [39]). Let us prove a quick estimate as an exercise.

[4] E. Bombieri, 1987/1974, "Le grand crible dans la théorie analytique des nombres".
[24] J. van Lint and H. Richert, 1965, "On primes in arithmetic progressions".
[39] O. Ramaré, 2019, "Explicit average orders: News and Problems".

Lemma 4.7. We have $G(z) \leqslant \log(20z)$.

Proof. We write

$$G(z) = \sum_{q\leqslant z} \frac{\mu^2(q)}{q} \sum_{d|q} \frac{1}{\varphi(d)} = \sum_{d\leqslant z} \frac{\mu^2(d)}{\varphi(d)} \sum_{\substack{q\leqslant z\\ d|q}} \frac{\mu^2(q)}{q}$$
$$= \sum_{d\leqslant z} \frac{\mu^2(d)}{d\varphi(d)} \sum_{\substack{q\leqslant z/d\\ (q,d)=1}} \frac{\mu^2(q)}{q}.$$

By mimicking the proof of Lemma 4.6, we readily find that

$$\sum_{\substack{q\leqslant z/d\\ (q,d)=1}} \frac{\mu^2(q)}{q} \leqslant \prod_{p|d} \frac{p}{p+1} \sum_{q\leqslant z} \frac{\mu^2(q)}{q}. \tag{4.4}$$

This gives us

$$G(z) \leqslant \sum_{d\leqslant z} \frac{\mu^2(d)}{\prod_{p|d}(p^2-1)} \sum_{q\leqslant z} \frac{\mu^2(q)}{q} \leqslant \prod_{p\geqslant 1} \left(1 + \frac{1}{p^2-1}\right) \sum_{q\leqslant z} \frac{\mu^2(q)}{q}$$
$$\leqslant \frac{\pi^2}{6} \sum_{q\leqslant z} \frac{\mu^2(q)}{q}.$$

We now find that

$$\sum_{q\leqslant Q} \mu^2(q) = \sum_{q\leqslant Q} \sum_{d^2|q} \mu(d) = \sum_{d\leqslant \sqrt{Q}} \mu(d) \left(\frac{Q}{d^2} + \mathcal{O}^*(1)\right)$$
$$\leqslant \left(\frac{6}{\pi^2} + \frac{1}{\sqrt{Q}-1}\right) Q + \sqrt{Q} \leqslant \frac{6}{\pi^2} Q + 3\sqrt{Q}.$$

This inequality is readily proved for $Q \geqslant 10$ and extended by direct checking. Therefore

$$\sum_{q\leqslant z} \frac{\mu^2(q)}{q} = \sum_{q\leqslant z} \mu^2(q) \int_q^z \frac{dt}{t^2} + \frac{1}{z} \sum_{q\leqslant z} \mu^2(q) = \int_1^z \left(\sum_{q\leqslant t} \mu^2(q)\right) \frac{dt}{t^2} + \frac{1}{z} \sum_{q\leqslant z} \mu^2(q)$$
$$\leqslant \int_1^z \left(\frac{6}{\pi^2} t + 3\sqrt{t}\right) \frac{dt}{t^2} + \frac{1}{z} \left(\frac{6}{\pi^2} z + 3\sqrt{z}\right)$$
$$\leqslant \frac{6}{\pi^2} \log z + 6 + \frac{6}{\pi^2} + 3 \leqslant \frac{6}{\pi^2} \log(20z).$$

The proof of complete. □

4.5. Proof of a weak form of Theorem $\mathscr{A}$

We want to use Theorem 4.4. We consider the interval $I = [(M-\ell)/k, (M+N-\ell)/k]$ and, for every prime $p \leqslant \sqrt{N/k}$, we set

$$\begin{cases} \mathcal{K}_p = \mathbb{Z}/p\mathbb{Z}^* - \ell k^{-1} & \text{when } (p,k)=1, \\ \mathcal{K}_p = \mathbb{Z}/p\mathbb{Z} & \text{when } p|k. \end{cases} \tag{4.5}$$

Then let $p' \in [M+1, M+N]$ be a prime $> \sqrt{N/k}$ that is $\equiv \ell[k]$ and let us define $n = (p'-\ell)/k$. For any $p \leqslant \sqrt{N/k}$, we have $p' \in \mathbb{Z}/p\mathbb{Z}^*$, and therefore $n \in \mathcal{K}_p$. We define $\mathcal{K}_q$ by multiplicativity*.

After some thoughts, we see that Theorem 4.4 tells us that, for any $Q \leqslant \sqrt{N/k}$, we have

$$\sum_{\substack{M<p\leqslant M+N \\ p\equiv \ell[k]}} 1 \leqslant \sqrt{N/k} + \frac{\frac{N}{k} - 1 + Q^2}{G_k(Q)}.$$

The reader will readily conclude from there, with the choice $Q^2 = N/k$.

4.6. A stronger form of Theorem $\mathscr{A}$

We may use Theorem 1.3 rather than Theorem 1.2. This leads to the next result that has not yet appeared in print.

Theorem 4.8

When ℓ is prime to k, $M \geqslant 0$ and $N > k$ are real numbers, we have

$$\sum_{\substack{q\leqslant Q \\ q\in\mathscr{S} \\ (q,k)=1}} \frac{\varphi(q)}{q} \log\frac{Q}{q} \sum_{a \bmod^* q} \Big| \sum_{\substack{M<p\leqslant M+N \\ p\equiv\ell[k]}} u_p e(pa/q) \Big|^2 \leqslant \frac{N+kQ^2}{\varphi(k)} \sum_{\substack{M<p\leqslant M+N \\ p\equiv\ell[k]}} |u_p|^2$$

where $\mathscr{S}$ is the set of *square-full* integers.

Let us recall that an integer in said to be *square-full* when $p|n \implies p^2|n$. This sequence is referenced on the Online Encyclopedy of Integer Sequences [31] as A013929. The readers will check that there are $\zeta(3/2)(1+o(1))\sqrt{Q}$ square-full integers below Q.

Theorem 4.8 is seen to indeed be an extension (up to the $1+o(1)$) of Theorem $\mathscr{A}$ by selecting $Q = \sqrt{N/k}/\log(N/k)$, $u_p = 1$ and $q = 1$. A less

*Though not needed, we may define $\mathcal{K}_{p^2}$ as the reverse image of $\mathcal{K}_p$ by the canonical projection $\sigma_{p^2\to p}$ recalled in Section 2.2, and similarly for $\mathcal{K}_{p^\nu}$.

[31] OEIS Foundation Inc., 2019, *The On-Line Encyclopedia of Integer Sequence.*

common way to recover this result is to select again $Q = \sqrt{N/k}/\log(N/k)$, but $u_p = e(p/4)$ and $q = 4$. And though we take another path, we again loose the now famous factor 2.

5 From the Brun-Titchmarsh Theorem to Siegel zeros

5.1. Siegel zero and Brun-Titchmarsh theorem

We follow here an idea of Motohashi in [30], though by other means. The fact that beating the constant 2 in the Brun-Titchmarsh theorem could "remove" a possible Siegel zero seems to have been first conjectured by Rodosskiĭ as noted by Klimov in [21]. See also the paper [49] by Siebert.

Theorem $\mathscr{B}$

There exist two effective constants c_3 and c_4, such that for $q \geqslant c_4$, the following two conditions are equivalent.

(a) There exist a constant $\xi > 0$ such that for any ℓ prime to q, we have, with $X = q^{c_3}$:

$$\sum_{\substack{X<p\leqslant 2X \\ p\equiv \ell[q]}} 1 \leqslant \frac{2-\xi}{\phi(q)} \sum_{X<p\leqslant 2X} 1. \tag{5.1}$$

(b) For any real character modulo q, we have $L(1,\chi) \gg 1/\log q$.

Such a statement is always tricky. For instance, we indeed use characters and not only primitive characters. We will not prove the reverse implication in these notes.

Proof that (a) implies (b). We follow Ramachandra, Sankaranarayanan & Srinivas in [33]. By summing our upper bound over all ℓ such that $\chi(\ell) = -1$, we discover that the number of primes in $(X, 2X]$ with $\chi(p) = 1$ is at least

$$\xi \sum_{X<p\leqslant 2X} 1/2.$$

Consider next $G(s) = \zeta(s)L(s,\chi) = \sum_{n\geqslant 1} g(n)n^{-s}$ where $g(n) = (1 \star \chi)(n)$. Note that $g(n)$ is non-negative. Note further that $g(p) = 2$ when $\chi(p) = 1$, from which we infer that

$$\sum_{X<n\leqslant 2X} g(n) \geqslant \sum_{X<p\leqslant 2X} g(p) \gg \xi X/\log X.$$

[30] Y. Motohashi, 1979, "A note on Siegel's zeros".
[21] N. I. Klimov, 1961, "Almost prime numbers".
[49] H. Siebert, 1983, "Sieve methods and Siegel's zeros".
[33] K. Ramachandra, A. Sankaranarayanan, and K. Srinivas, 1996, "Ramanujan's lattice point problem, prime number theory and other remarks".

This readily yields

$$\frac{1}{2i\pi}\int_{c-i\infty}^{c+i\infty} G(s+1)\Gamma(s)\big((2X)^s - X^s\big)ds$$
$$= \sum_{n\geqslant 1}\frac{g(n)}{n}\big(e^{-n/(2X)} - e^{-n/X}\big) \gg \sum_{X<n\leqslant 2X}\frac{g(n)}{X} \gg \xi/\log X.$$

Next, shifting the path of integration in the above integral to $\Re s = -1/4$, we see that it is

$$L(1,\chi)\log 2 + \mathcal{O}\left(X^{-1/4}\int_{c-i\infty}^{c+i\infty}|G(s+1)\Gamma(s)|ds\right).$$

The exponential decay of $\Gamma(s)$ in vertical strips (a consequence of the Stirling formula) as well as the bound $|G(3/4+it)| \ll q^{1/4}(1+|t|)$ ensures us that this last error term is at most $\mathcal{O}((q/X)^{1/4})$, which in turn is $\mathcal{O}(q^{-1})$ since $c_3 \geqslant 5$. In conclusion, we find that $L(1,\chi) \gg 1/\log X \gg 1/\log q$ as required. □

Thus, improving on the constant 2 in the Brun-Titchmarsh theorem when X is a power of q would remove any Siegel zero. Note that we use only the Brun-Titchmarsh theorem for the initial range.

Drawing on similar ideas in [1], Basquin established a theorem linking an effective lower bound for $L(1,\chi)$ of the shape $1/q^c$ for some $c \in]0,1/2]$ with the improvement on the constant 2 in the Brun-Titchmarsh theorem, but in a different range for X.

Theorem 5.1

Let $c > 0$ be a parameter. The following three problems are equivalent:

1. For every $\varepsilon > 0$, and every real character χ, prove in an effective way that $L(1,\chi) \gg q^{-c-\varepsilon}$.
2. For every $\varepsilon > 0$, prove (5.1) for every $q \leqslant (\log X)^{(1/c)-\varepsilon}$.
3. For every $\varepsilon > 0$, prove in an effective way that $\psi(X;q,\ell) \sim X/\phi(q)$ for every $q \leqslant (\log X)^{(1/c)-\varepsilon}$.

This statement also tells us that, if we are able to beat the factor 2 in the upper bound, then a much stronger conclusion follows, namely an equivalent for $\psi(X;q,\ell)$. This situation is comparable to what happens with the elementary proof of the prime number theorem, a proof this time heavily linked to the *parity principle*. See the papers [48] and [47] by Selberg and the paper [9] from Erdős.

[1] J. Basquin, 2006, "Mémoire de DEA, Lille".
[48] A. Selberg, 1949, "An elementary proof of the prime-number theorem".
[47] A. Selberg, 1949, "An elementary proof of Dirichlet's theorem about primes in an arithmetic progression".
[9] P. Erdös, 1949, "On a new method in elementary number theory which leads to an elementary proof of the prime number theorem".

6 Montgomery's sieve from Parseval

6.1. Introduction

In this chapter, we propose (yet) another proof of the arithmetical form of the large sieve inequality recalled below. Its advantage is that it remains very close to the Parseval equality (in particular we do not use any duality) if we are to believe in sensible behaviour of our quantities but still misses a factor 2 at the end. This factor 2 is the one occuring in the Brun-Titchmarsh inequality; improving on it would have major consequences, for instance on the class number problem. The defect that transpires has already been observed by D. Goldston in the case of the initial segment of the primes in [13]. Let us indicate to the readers that P. Gallagher has already given in [12] (see Eq. (3') therein) a proof of the large sieve inequality by employing the Parseval identity as we do, but his proof looses a factor π with respect, a loss that we avoid.

We are mainly concerned with the case when $u_n \neq 0$ only when $\ell + kn$ is a prime number for some integer $\ell \in \{1, \cdots, k\}$ such that $(\ell, k) = 1$. The modulus k is only assumed to be non-negative and indeed, the case $k = 1$ has been decisive in designing the proofs. Then, for any prime p that is prime to k, the class of n modulo p avoids the class ℓk^{-1} mod p or else we have $\ell + kn = p$. We say that (u_n) is (k, Q)-*prime-like* if there exists ℓ as above such that $\ell + kn$ is prime to every prime $\leqslant Q$. This is a non-standard definition but it avoids to much generality. The characteristic function of the primes between $Q + 1$ and N is $(1, Q)$-prime-like. Given $k \geqslant 1$ and $c \in \{1, \cdots, k\}$ prime to k, the characteristic function of the primes congruent to ℓ and lying between Q and kN is (k, Q)-prime-like, as is any subsequence of it.

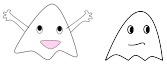
Theorem 6.1

Let (u_n) be a non-negative (k, Q)-prime like sequence. We have

$$\frac{G_k(Q)}{1 + Q^2/V} \int_{-\infty}^{\infty} \Big| \sum_{|n-t| \leqslant V} u_n \Big|^2 \frac{dt}{4V^2} \leqslant \sum_n |u_n|^2.$$

The readers may be interested in the expression of the LHS provided by (6.14). On invoking Lemma 6.4, we may deduce Theorem $\mathscr{C}$ from this result.

[13] D. A. Goldston, 2000, "The major arcs approximation of an exponential sum over primes".
[12] P. Gallagher, 1967, "The large sieve".



Corollary 6.2

Let (u_n) be a non-negative (k, Q)-prime like sequence. We have

$$G_k(Q)\Big(1-\frac{2\sqrt{2}Q}{\sqrt{N}}\Big)\Big|\sum_n u_n\Big|^2 \leqslant N\sum_n |u_n|^2.$$

From which the Brun-Titchmarsh inequality, i.e. Theorem $\mathscr{A}$ with $2+o(1)$, follows readily. We use the next notation in this whole chapter.

$$S(\alpha)=\sum_{n\leqslant N} u(n)e(n\alpha), \quad (e(y)=\exp(2i\pi y)). \tag{6.1}$$

This is more general than (2).

6.2. A large sieve inequality alternative

Let us start with a consequence of the Parseval equality that has a similar flavour.

Lemma 6.3

For any $Q \geqslant 1$, we have

$$\sum_{q\leqslant Q}\sum_{a \bmod^* q}\int_{-\frac{1}{2Q^2}}^{\frac{1}{2Q^2}}\Big|S\Big(\frac{a}{q}+\beta\Big)\Big|^2 d\beta \leqslant \sum_{n\leqslant N}|u_n|^2.$$

This inequality is immediate for anyone used to the Circle Method. Notice the absence of the factor $N+Q^2$ that appears in Theorem 4.3. The price to pay is to have an integral around a/q. As a comparaison, let us mention the inequality

$$\sum_{q\leqslant Q}\sum_{a \bmod^* q}\sum_{\substack{\ell\in\mathbb{Z}\\ 2q|\ell|\leqslant Q}}\Big|S\Big(\frac{a}{q}+\frac{\ell}{Q^2}\Big)\Big|^2 \leqslant \sum_{n\leqslant N}|u_n|^2(N+Q^2). \tag{6.2}$$

proved in Theorem 1.8 of [34].

Proof. We use the Parseval identity together with the subsets

$$I_Q(a/q)=\frac{a}{q}+\Big[\frac{-1}{q(q+Q)},\frac{1}{q(q+Q)}\Big], q\leqslant Q.$$

[34] D. S. Ramana and O. Ramaré, 2025, *Arithmetical aspects of the large sieve inequality – II.*

These subsets are disjoint as, with an obvious notation,

$$\left|\frac{a}{q}+\beta-\frac{\tilde{a}}{\tilde{q}}-\tilde{\beta}\right| > \frac{|a\tilde{q}-\tilde{a}q|}{q\tilde{q}} - \frac{1}{q(q+Q)} - \frac{1}{\tilde{q}(\tilde{q}+Q)}$$
$$> \frac{1}{q\tilde{q}} - \frac{1}{q(q+\tilde{q})} - \frac{1}{\tilde{q}(\tilde{q}+q)} = 0.$$

The Kloosterman arcs decomposition would also leads to a proof of our lemma. □

6.3. Splitting the range

We prove in this preliminary section a lemma that will be used later on.

Lemma 6.4

When the function a is Riemann-integrable and supported on $[-V, V]$, and $(u_m)_{m\leqslant N}$ are complex numbers, we have

$$\int_{-\infty}^{\infty}\Big|\sum_m u_m a(t-m)\Big|^2 dt = \frac{|\sigma|^2}{N}\left(1-\frac{2V}{N}\right)\Big|\sum_m u_m\Big|^2 + \int_{-V}^{N+V}\left|\sum_m u_m a(t-m) - \frac{\sigma\sum_m u_m}{N}\right|^2 dt.$$

where $\sigma = \int_{-V}^{V} a(t)dt$.

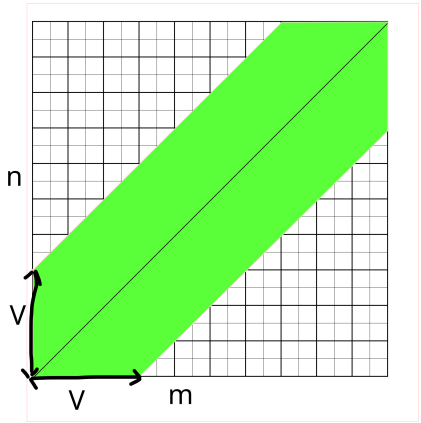
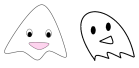


Figure 6.1: Domain $|m-n| \leqslant V$



Proof. Let us set $X = \sum_m u_m$. We first find that

$$\int_{-V}^{N+V} \sum_m u_m a(t-m)dt = \sum_m u_m \int_{-V}^{N+V} a(t-m)dt = \sigma X.$$

Consequently, we find that

$$\begin{aligned}
&\int_{-V}^{N+V} \Big|\sum_m u_m a(t-m) - \frac{\sigma X}{N}\Big|^2 dt \\
&= \int_{-V}^{N+V} \Big|\sum_{m\leqslant V} u_m a(t-m)\Big|^2 dt - 2\Re\frac{\overline{\sigma X}}{N}\int_{-V}^{N+V}\sum_m u_m a(t-m)dt + \frac{|\sigma|^2(N+2V)}{N^2}|X|^2 \\
&= \int_{-V}^{N+V} \Big|\sum_m u_m a(t-m)\Big|^2 dt - \frac{|\sigma|^2}{N}|X|^2 + \frac{2V|\sigma|^2|X|^2}{N^2}.
\end{aligned}$$

The lemma follows readily. □

6.4. Using the error term

Lemma 6.5

Let $(u_m)_{m\leqslant N}$ be a sequence of non-negative numbers. We have

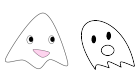
$$\begin{aligned}
&\frac{N}{4V^2(1-2V/N)}\int_{-\infty}^{\infty}\Big|\sum_{|m-t|\leqslant V} u_m\Big|^2 dt \\
&\qquad\geqslant |S(0)|^2 + \Big|S(\beta) - \frac{S(0)}{N}\int_{-V}^{N+V} e(\beta t)dt + \mathcal{O}(\beta V S(0))\Big|^2.
\end{aligned}$$

Proof. Let us notice that $S(0) = \sum_m u_m = \sum_m |u_m|$. We set

$$\Delta(N,V) = \int_{-V}^{N+V}\Big|\sum_{|m-t|\leqslant V} u_m - \frac{2VS(0)}{N}\Big|^2 dt. \tag{6.3}$$

We swiftly obtain:

$$\begin{aligned}
\Delta(N,V) &= \int_{-V}^{N+V}\Big|\sum_{|m-t|\leqslant V} u_m\big(e((m-t)\beta) + \mathcal{O}^*(2\pi\beta V)\big) - \frac{2VS(0)}{N}\Big|^2 dt \\
&= \int_{-V}^{N+V}\Big|\sum_{|m-t|\leqslant V} u_m\big(e(m\beta) + \mathcal{O}^*(2\pi\beta V)\big) - \frac{2VS(0)e(\beta t)}{N}\Big|^2 dt
\end{aligned}$$



We may then use Cauchy's inequality (in reverse) to infer that

$$\begin{aligned}(N+2V)\Delta(N,V) &\geqslant \left|\int_{-V}^{N+V}\Big(\sum_{|m-t|\leqslant V} u_m\big(e(m\beta)+\mathcal{O}^*(2\pi\beta V)\big)-\frac{2VS(0)e(\beta t)}{N}\Big)dt\right|^2\\ &\geqslant \left|2V\sum_m u_m\big(e(m\beta)+\mathcal{O}^*(2\pi\beta V)\big)-\frac{2VS(0)}{N}\int_{-V}^{N+V}e(\beta t)dt\right|^2\\ &\geqslant \left|2VS(\beta)+\mathcal{O}^*(2\pi\beta V^2S(0))-\frac{2VS(0)}{N}\int_{-V}^{N+V}e(\beta t)dt\right|^2\\ &\geqslant 4V^2\left|S(\beta)-\frac{S(0)}{N}\Big(\int_{-V}^{N+V}e(\beta t)dt+\mathcal{O}^*(2\pi\beta VN)\Big)\right|^2.\end{aligned}$$

We finally only have to plus this inequality in Lemma 6.4, and simplify the expression to get the present lemma. □

6.5. A family of Fourier transforms

Let us start with an easy lemma.

Lemma 6.6

Let us set $a_0(t)=1_{|t|\leqslant V}$. When $y\in[0,2V]$, we find that

$$\int_{-\infty}^{\infty}a_0(t)a_0(t-y)dt=\max(0,2V-|y|)=\int_{|y|}^{2V}1dt.$$

We introduce a family of functions in (6.4) to clarify our process. The main case will be $a=a_0$ as defined above. As it turns out, it is also the one that governs the proof of Lemma 6.7.

Lemma 6.7

Let a be a non-negative even Rieman-integrable function in $L^2(\mathbb{R})\cap L^1(\mathbb{R})$ that is non-increasing on $[0,\infty)$. The function

$$(a\star a)(y)=\int_{-\infty}^{+\infty}a(t)a(y-t)dt$$

is even and non-increasing on $[0,\infty)$.

Proof. Let us start with the next stream of identities.

$$(a \star a)(y) = \int_{-\infty}^{+\infty} a(t)a(t-y)dt = \int_{-\infty}^{+\infty} a(t+y)a(t)dt = \int_{-\infty}^{+\infty} a(-t-y)a(-t)dt$$
$$= \int_{-\infty}^{+\infty} a((-y)-t)a(t)dt = (a \star a)(-y).$$

This shows that the function $(a \star a)$ is indeed even. When $a(t) = 1_{|t| \leqslant T}$, we have $(a \star a)(y) = (T - |y|)^+$ which is indeed non-increasing when $y \geqslant 0$. This property extends to linear combinations of similar functions, provided the coefficients are non-negative. The proof is complete. □

Lemma 6.8

Let a be a function as in Lemma 6.7 and let us define the non-negative function b by $\int_{-\infty}^{+\infty} a(t)a(t-y)dt = \int_{|y|}^{\infty} b(t)dt$. For $\delta > 0$, we consider

$$\hat{D}_{a,\delta}(u) = \int_0^{\infty} \hat{C}_{[-\lambda,\lambda],\delta}(u)b(\lambda)d\lambda \tag{6.4}$$

where $C_{[-\lambda,\lambda],\delta}$ is defined in Eq. (3.3). We have $\hat{D}_{a,\delta}(u) = 0$ when $|u| \geqslant \delta$, then $\hat{D}_{a,\delta}(u) \leqslant \delta^{-1}A(a) + B(a)$ where $A(a) = \|a\|_2^2$, $B(a) = \|a\|_1^2$ and

$$D_{a,\delta}(y) \geqslant \int_{|y|}^{\infty} b(\lambda)d\lambda = \int_{-\infty}^{+\infty} a(t)a(t-y)dt. \tag{6.5}$$

Proof. Most of this lemma is a rather trivial consequence of Lemma 3.1 and the properties of the different players. $\hat{D}_{a,\delta}(u) \leqslant \int_0^{\infty}(\delta^{-1} + 2\lambda)b(\lambda)d\lambda = \delta^{-1}A(a) + B(a)$. We find at first that $B(a) = 2\int_0^{\infty} \lambda b(\lambda)d\lambda$. Let us express this quantity in terms of a. Let us set

$$\gamma(y) = \int_y^{\infty} b(t)dt. \tag{6.6}$$

We have

$$B(a) = 2\int_0^{\infty} \lambda b(\lambda)d\lambda = -2\int_0^{\infty} \lambda\gamma'(\lambda)d\lambda = 2\int_0^{\infty} \gamma(\lambda)d\lambda$$
$$= \int_0^{\infty}\int_{-\infty}^{\infty} a(t)a(t-\lambda)dt\,d\lambda + \int_0^{\infty}\int_{-\infty}^{\infty} a(t)a(t+\lambda)dt\,d\lambda = \left(\int_{-\infty}^{\infty} a(t)dt\right)^2.$$

The lemma follows swiftly. □

6.6. Explicit expression for $\hat{D}_{a_0,\delta}(u)$

When $a = a_0$, we may select $b_0(t) = 1_{t \leqslant 2V}$. We find that

$$A(a_0) = 2V, \quad B(a_0) = 4V^2. \tag{6.7}$$

Lemma 6.9
Let $\delta > 0$ and $V \geqslant 0$ be two parameters. When $|t| \leqslant \delta$, we define

$$\hat{D}_{a_0,\delta}(t) = \delta^{-1}(1 - |\delta^{-1}t|)\frac{\sin(4\pi Vt)}{2\pi t} + \hat{J}(\delta^{-1}t)\frac{\sin^2(2\pi Vt)}{(\pi t)^2}.$$

and extend this definition by $\hat{D}_{a_0,\delta}(t) = 0$ when $|t| \geqslant \delta$. We have

$$\forall t \in \mathbb{R}, \quad \hat{D}_{a_0,\delta}(t) \leqslant (2\delta^{-1} + 4V)V.$$

Proof. By (6.4), we find that

$$\begin{aligned}\hat{D}_{a_0,\delta}(t) &= \delta^{-1}(1 - |\delta^{-1}t|)\int_0^{2V} \cos(2\pi\lambda t)d\lambda + \frac{\hat{J}(\delta^{-1}t)}{\pi t}\int_0^{2V} \sin(2\pi\lambda t)d\lambda \\ &= \delta^{-1}(1 - |\delta^{-1}t|)\frac{\sin(4\pi Vt)}{2\pi t} + \frac{\hat{J}(\delta^{-1}t)}{\pi t}\frac{1 - \cos(4\pi Vt)}{2\pi t}. \\ &= \delta^{-1}(1 - |\delta^{-1}t|)\frac{\sin(4\pi Vt)}{2\pi t} + \hat{J}(\delta^{-1}t)\frac{\sin^2(2\pi Vt)}{(\pi t)^2}.\end{aligned}$$

We readily deduce from this expression that

$$\hat{D}_{a_0[V],\delta}(t) = \delta^{-2}\hat{D}_{a_0[\delta V],1}(\delta^{-1}t) \tag{6.8}$$

where $a_0[W]$ denotes the function a_0 with the parameter V being equal to W. Here is the expression for $\hat{D}_{a_0[W],1}(t)$, when $0 \leqslant t \leqslant 1$:

$$\begin{aligned}\hat{D}_{a_0[W],1}(t) &= (1-t)\frac{\sin(4\pi Wt)}{2\pi t} + \hat{J}(t)\frac{\sin^2(2\pi Wt)}{(\pi t)^2} \\ &= (1-t)\frac{\sin(4\pi Wt)}{2\pi t} + (1-t)\cot(\pi t)\frac{\sin^2(2\pi Wt)}{\pi t} + \frac{\sin^2(2\pi Wt)}{\pi^2 t}.\end{aligned}$$

The plot below may give the feeling that this function in non-negative but we show that this is not the case in Figure 6.6. The proof is complete. □

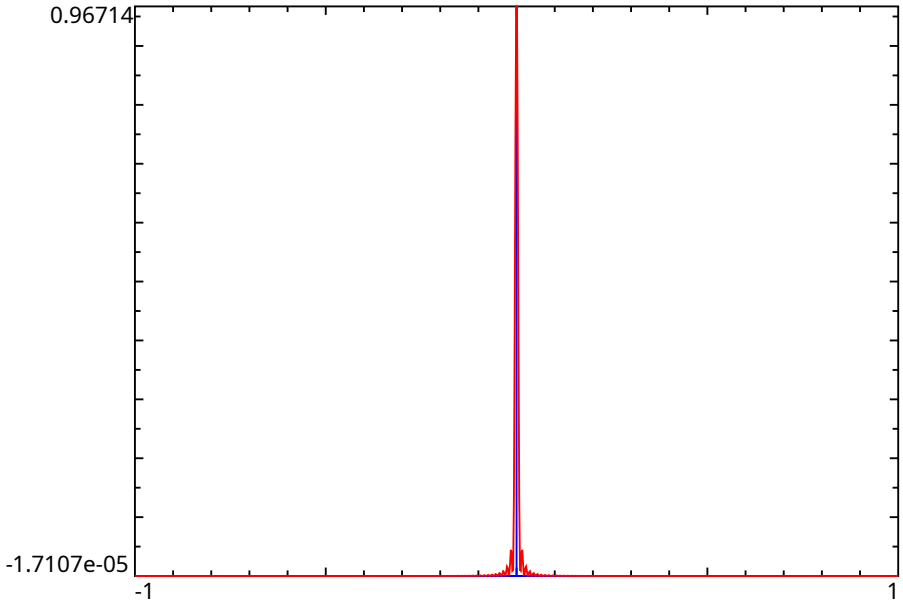


Figure 6.2: $\hat{D}_{a_0[\delta V],1}(u)/(2\delta V + 4(\delta V)^2)$ with $\delta V = 50$

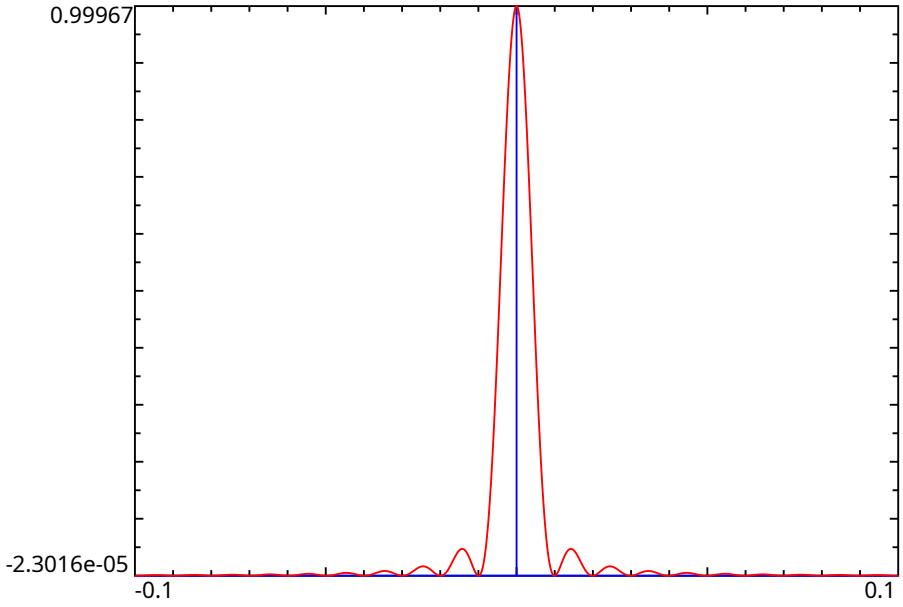


Figure 6.3: $\hat{D}_{a_0[\delta V],1}(u)/(2\delta V + 4(\delta V)^2)$ with $\delta V = 50$

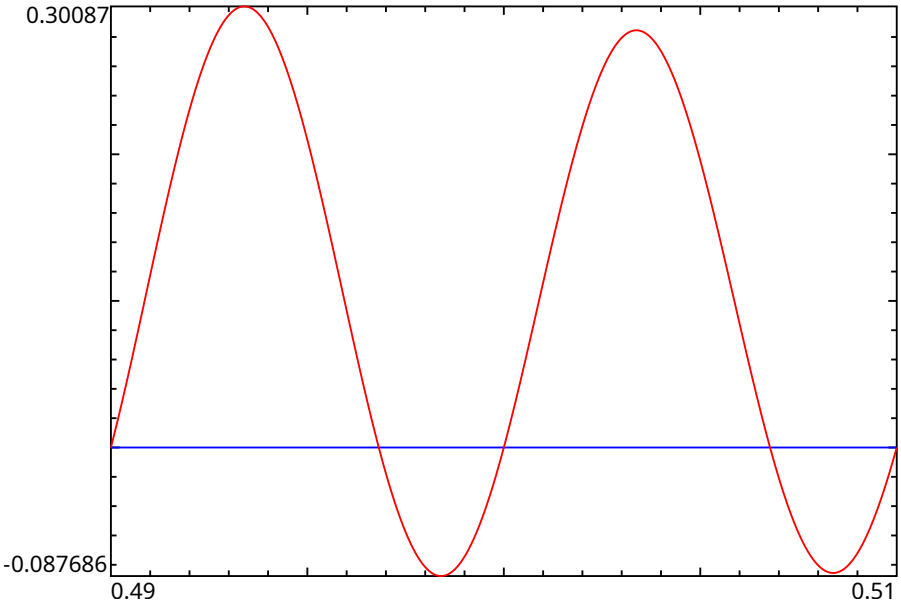


Figure 6.4: $\hat{D}_{a_0[\delta V],1}(u)$ with $\delta V = 50$

6.7. Base Camp

Proof of Theorem 6.1. Let us first notice, by using Theorem 1.2, that

$$\sum_{a \bmod^* q} \int_{-\delta}^{\delta} \left| S\left(\frac{a}{q} + \beta\right) \right|^2 d\beta \geqslant \frac{\mu^2(q)}{\varphi(q)} \int_{-\delta}^{\delta} |S(\beta)|^2 d\beta.$$

Lemma 6.8 with $\delta = 1/(2Q^2)$ has the inequality:

$$1_{[-\delta,\delta]}(u) \geqslant \frac{\hat{D}_{a,\delta}(u)}{\delta^{-1}A(a) + B(a)}.$$

This gives us

$$\int_{-\delta}^{\delta} |S(\beta)|^2 d\beta \geqslant \int_{-\delta}^{\delta} |S(\beta)|^2 \frac{\hat{D}_{a,\delta}(\beta)}{\delta^{-1}A(a) + B(a)} d\beta. \tag{6.9}$$

The integral may be extended to $\beta \in \mathbb{R}$. We furthermore find that

$$\begin{aligned} \sum_{a \bmod^* q} \int_{-\delta}^{\delta} \left| S\left(\frac{a}{q} + \beta\right) \right|^2 d\beta &\geqslant \frac{\mu^2(q)}{\varphi(q)} \int_{-\infty}^{\infty} |S(\beta)|^2 \frac{\hat{D}_{a,\delta}(\beta)}{\delta^{-1}A(a) + B(a)} d\beta \\ &\geqslant \frac{\mu^2(q)}{\varphi(q)(\delta^{-1}A(a) + B(a))} \sum_{m,n} u_m \overline{u_n} D_{a,\delta}(m-n). \end{aligned}$$

As (u_n) is assumed to be non-negative, we may bound below $D_{a,\delta}(y)$ by $\int_{-\infty}^{+\infty} a(t)a(t-y)dt$, getting

$$\sum_{a \bmod^* q} \int_{-\delta}^{\delta} \left| S\left(\frac{a}{q}+\beta\right)\right|^2 d\beta \geqslant \frac{\mu^2(q)/\varphi(q)}{\delta^{-1}A(a)+B(a)} \int_{-\infty}^{\infty} \left|\sum_m u_m a(t-m)\right|^2 dt. \quad (6.10)$$

We conclude by appealing to Lemma 6.4 and by noticing that $\sigma^2 = B(a)$. Our theorem readily follows by specializing a to a_0. □

Now that this proof is over, the readers may go back to Section 3.2 and see its relevance.

6.8. Proof of Corollary 6.2 and of (3.9)

We use Theorem 6.1 together with Lemma 6.4 to infer that

$$\frac{G_k(Q)}{1+Q^2/V}\left(1-\frac{2V}{N}\right)\left|\sum_n u_n\right|^2 \leqslant N\sum_n |u_n|^2.$$

The parameter V needs to be optimized. On setting $b = 2Q^2/N$ and $w = N/(2V)$, we readily check that the optimal value of w is given by

$$w = 1+\sqrt{1+b^{-1}}. \quad (6.11)$$

With this value, we find that

$$\frac{1}{1+Q^2/V}\left(1-\frac{2V}{N}\right) = \frac{1}{1+bw}\left(1-\frac{1}{w}\right) = \frac{\sqrt{1+b}}{(2+b)\sqrt{b}+\sqrt{1+b}(1+2b)}.$$

We check that, when $b \geqslant 0$, the next inequality holds true:

$$\frac{\sqrt{1+b}}{\sqrt{1+b}+(2+b)\sqrt{b}+2b\sqrt{1+b}} \geqslant 1-2\sqrt{b}.$$

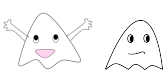
Our corollary follows swiftly from there, as well as the inequality (3.9).

6.9. Following Vaaler

The function $D_{a_0,\delta}(y)$ is an upper bound for $f(y) = (2V-|y|)^+$. Rather than the construction we followed, we could rely on Corollary 12 of [54]. Let us start by noticing that the total variation $V_f(y)$ of f on $(-\infty, y]$ is given by

$$V_f(y) = \max(4V, y+2V, 0) \quad (6.12)$$

[54] J. Vaaler, 1985, "Some Extremal Functions in Fourier Analysis".



so that $dV_f(y) = 1_{|y|\leqslant 2V} = b_0(|y|)$. The majorant proposed by Vaaler is thus (recall the notation $F_\delta(x) = \delta F(\delta x)$ given at the beginning of Section 4 in [54], so that $\hat{F}_\delta(t) = \hat{F}(t/\delta)$):

$$M(y) = \int_{-\infty}^{\infty} f(t)J_\delta(y-t)dt + (2\delta)^{-1}\int_{-\infty}^{\infty} dV_f(t)K_\delta(y-t)dt.$$

We directly get

$$\begin{aligned}\hat{M}(u) &= \hat{f}(u)\hat{J}(u/\delta) + (1/2)\widehat{dV_f}(u)\hat{K}(u/\delta)\\ &= \left(\frac{\sin 2\pi V u}{\pi u}\right)^2 \hat{J}(u/\delta) + \frac{\sin 4\pi V u}{2\pi u}\delta^{-1}(1-|u/\delta|)^+\end{aligned}$$

which is precisely $\hat{D}_{a_0,\delta}(u)$, as given in Lemma 6.9. The advantage of the construction proposed by Vaaler is that we can dispense with our monotonicity hypothesis on a, provided we introduce the total variation as above.

6.10. A technological remark

We readily find that

$$\widehat{a * a}(\beta) = |\hat{a}(\beta)|^2 \tag{6.13}$$

and

$$(a*a)(y) = \int_{-\infty}^{\infty} |\hat{a}(\beta)|^2 e(-y\beta)d\beta.$$

We may therefore write

$$\int_{-\infty}^{\infty}\Big|\sum_m u_m a(t-m)\Big|^2 dt = \int_{-\infty}^{\infty}\Big|S(\beta)\hat{a}(\beta)|^2 d\beta.$$

The differences with (6.9) is that the kernel $|\hat{a}(\beta)|^2$ is non-negative while it is not the case of $\hat{D}_{a,\delta}(\beta)$, but the interval of integration is now infinite. When $a = a_0$, we get

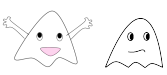
$$\int_{-\infty}^{\infty}\Big|\sum_{|n-t|\leqslant V} u_n\Big|^2 \frac{dt}{4V^2} = \int_{-\infty}^{\infty}\left(\frac{\sin 2\pi V\beta}{2\pi V\beta}\right)^2 |S(\beta)|^2 d\beta. \tag{6.14}$$

We may change a.

Lemma 6.10

With $a = a_0^{(*k)} = a_{k-1}$, the k-th convolution product of a_0, we have

$$B(a_k) \leqslant (2V)^{2(k+1)} \quad \text{and} \quad A(a_k) \leqslant (2V)^{2k+1}.$$



Proof. We have directly $B(a_k) = \|a_k\|_1^2 \leqslant \|a_0\|_1^{2(k+1)}$. For the L_2-norm, we use the inequality $\|f * g\|_2^2 \leqslant \|f\|_2^2 \|g\|_1^2$ repeatedly to reach our result. □

On using Lemma 6.10, we obtain

$$\sum_{a \bmod^* q} \int_{-\delta}^{\delta} \left| S\left(\frac{a}{q} + \beta\right) \right|^2 d\beta \geqslant \frac{\mu^2(q)/\varphi(q)}{1 + \delta^{-1}/(2V)} \int_{-\infty}^{\infty} \left(\frac{\sin 2\pi V\beta}{2\pi V\beta}\right)^{2(k+1)} |S(\beta)|^2 d\beta.$$

These inequalities are in fact weaker as k increases as

$$\left(\frac{\sin 2\pi V\beta}{2\pi V\beta}\right)^{2(k+1)} \leqslant \left(\frac{\sin 2\pi V\beta}{2\pi V\beta}\right)^{2k}.$$

6.11. Computing $C_{[-\lambda,\lambda],\delta}$, $\hat{C}_{[-\lambda,\lambda],\delta}$ and $\hat{D}_{a_0,\delta}$

Here are some functions in Pari-GP that the readers may use to compute the players occuring in this chapter.

```
\\ C:
{B(y) =
   my(res = 0.0, maxm = 2000);
   res = sum(m = 0, maxm, (sinc(Pi*(y-m)))^2);
   res += sinc(Pi*y)^2*(1+2*Pi*y);
   res += -sum(m = 1, maxm, (sinc(Pi*(y+m)))^2);
   return(res);}

{myfun(lambda, x) = return(B(lambda-x)+B(x+lambda));}

\\ Jhat, Chat and Dhat:
{Jhat(u) =
   if(abs(u) > 1,
      return(0),
      if(abs(u) > 0.01,
         return(Pi*u*(1-abs(u))*cotan(Pi*u)+abs(u)),
         return((1-abs(u))*cos(Pi*u)/sinc(Pi*u)+abs(u))));}

{Chat(lambda, u) =
   if(abs(u) <= 0.00000000000001,
      return(1+2*lambda),
      return((1-abs(u))*cos(2*Pi*lambda*u)
             + sin(2*Pi*lambda*u)/u/Pi*Jhat(u)));}

{Dhat(deltaV, t) =
```

```
   my(res);
   if(abs(t) > 1,
      return(0),
      res = (1-abs(t))*2*deltaV*sinc(4*Pi*deltaV*t);
      res += Jhat(t)*4*deltaV^2*sinc(2*Pi*deltaV*t)^2;
      return(res));}

{normalizedDhat(deltaV, t) =
   return(Dhat(deltaV, t)/deltaV/(2+4*deltaV));}

\\ s = plothexport("svg", t=-0.1,0.1, normalizedDhat(50, t));
\\ write("Dhat-50-0dot1.svg", s)
\\ This gives a plot in svg format. To convert it in pdf-format:
\\ inkscape --export-type=pdf Dhat-50-0dot1.svg
```

6.12. Addendum: two conditional estimates

Here is a typical regularity estimate one can prove if we assume that the factor 2 is indeed required.

Theorem 6.11
If

$$S(0) = \frac{2N}{\log N} + \mathcal{O}\Big(\frac{\log\log N}{\log N}\Big)$$

for some infinite family $\mathscr{N}$ of N's, then, for any $A \geqslant 1$, we have

$$S(\beta) = \frac{2}{\log N}\int_0^N e(\beta t)dt + \mathcal{O}_A\Big(S(0)\sqrt{\frac{\log\log N}{\log N}}\Big)$$

for $|\beta| \leqslant (\log N)^A/N$ and $N \in \mathscr{N}$.

The approximation is rather weak, but the range in β is very large. This approximation is immediate when $|\beta| \leqslant o(1)/N$.

Proof. We select $V = N/(\log N)^{A+2}$ and $Q^2 = V/\log N$ in Theorem 6.1. We then use Lemma 6.5. The result follows readily. □

Let us end this section with a result that states that irregularity of disbribution of the u_m's in small intervals implies that the factor 2 can be diminished.

Theorem 6.12

Let $A \geqslant 1$ be given. Let us set $U = \sum_m u_m$ and, for any $\epsilon_1, \epsilon_2 > 0$ and $V \in [N/\log N)^A, N]$:

$$\mathscr{T}(V, \epsilon_1) = \left\{ t \in [-V, N+V] : \left| \sum_{|m-t| \leqslant V} u_m - \frac{2VU}{N} \right| \geqslant \epsilon_1 2VU/N \right\}. \quad (6.15)$$

If for every $N \in \mathscr{N}$, we can find a parameter $V \in [N/(\log N)^A, N]$ such that $\mathrm{meas}(\mathscr{T}(V, \epsilon_1)) > \epsilon_2 N$, then for those N's, we have

$$\sum_m u_m \leqslant \frac{(2 + o_A(1) - \epsilon_1^2 \epsilon_2)N}{\log N}.$$

Though these two precise theorems are novel, they are in line with the Deuring-Heilbronn phenomenom (though this phenomenom concerns only a very special situation) and a result of a similar flavour can be readily infered from Theorem 15.2 and 15.3 of [36].

[36] O. Ramaré, 2009, *Arithmetical aspects of the large sieve inequality.*

7 An Enveloping sieve

We fix two real parameters $z_0 \leqslant z$ and consider the sole case of prime numbers. Let us set

$$P(z_0) = \prod_{p<z_0} p. \tag{7.1}$$

It is easy to reproduce the analysis of [37, Section 3] as far as exact formulae are concerned, but one gets easily sidetracked towards slightly different formulae. The reader may for instance compare [40, Lemma 4.2] and [37, (4.1.14)]. Similar material is also the topic of [36, Chapter 12]. So we present a complete analysis in our special case. Here is the main end-product we shall use.

Theorem 7.1

Let $z_0 \leqslant z$ be two parameters. There exists an upper bound $\beta_{z_0,z}$ of the characteristic function of those integers that do not have any prime factor in the interval $[z_0, z)$. The function $\beta_{z_0,z}$ admits the expansion:

$$\beta_{z_0,z}(n) = \sum_{\substack{q\leqslant z^2,\\ q|P(z)/P(z_0)}} w_q(z;z_0)c_q(n)$$

where $c_q(n)$ is the Ramanujan sum and where

$$w_q(z;z_0) = \frac{\mu(q)}{\varphi(q)G(z;z_0)}\frac{G_{[q]}(z;z_0)}{G(z;z_0)},$$

with the definition

$$G_{[q]}(z;z_0) = \sum_{\substack{\ell\leqslant z/\sqrt{q},\\ (\ell,qP(z_0))=1}} \frac{\mu^2(\ell)}{\varphi(\ell)}\xi_q(z/\ell) \tag{7.2}$$

and

$$\xi_q(y) = \sum_{\substack{q_1q_2q_3=q,\\ q_1q_3\leqslant y,\\ q_2q_3\leqslant y}} \frac{\mu(q_3)\varphi_2(q_3)}{\varphi(q_3)} \quad \text{where} \quad \varphi_2(q_3) = \prod_{p|q}(p-2).$$

Notice that $\xi_q(y) = q/\varphi(q)$ when $y \geqslant q$ and that $|\xi_q(y)| \leqslant 3^{\omega(q)}$ always.

Remark 1 *The factor $G_{[q]}(z;z_0)/G(z;z_0)$ should be looked upon as a mild perturbation. It can be shown to be equivalent to 1 as q goes to infinity, and, in*

[37] O. Ramaré and I. Ruzsa, 2001, "Additive properties of dense subsets of sifted sequences".
[40] O. Ramaré, 1995, "On Snirel'man's constant".

general, it only introduces technicalities.

Proof of Theorem 7.1. We split the proof in three steps. We follow closely Section 3 of [37]. See also [36, Chapter 11].

Building the upper bound

The initial idea of the Selberg sieve is to consider the family

$$\beta_{z_0,z}(n) = \Bigg(\sum_{\substack{d:d|n,\\ (d,P(z_0))=1,\\ d\leqslant z}} \lambda_d \Bigg)^2 \tag{7.3}$$

for arbitrary real coefficients λ_d that are only constrained by the condition $\lambda_1 = 1$. Indeed, for any such set of of coefficients, the resulting function is non-negative and takes the value 1 at integers n that have no prime factors dividing $P(z)/P(z_0)$. After an optimization step that we skip, one reaches the choice

$$\lambda_d = 1_{(d,P(z_0))=1} \frac{\mu(d) d G_d(z/d; z_0)}{\varphi(d) G(z; z_0)} \tag{7.4}$$

where G_d is given by (4.3) (notice that, indeed, $\lambda_1 = 1$). From now onward, we reserve the notation λ_d for this special choice. Though we shall not use it, notice that Lemma 4.6 implies the bound $|\lambda_d| \leqslant 1$.

We develop the square above and get

$$\begin{aligned} \beta_{z_0,z}(n) &= \sum_{\substack{d_1,d_2,\\ [d_1,d_2]|n}} \lambda_{d_1}\lambda_{d_2} = \sum_{d_1,d_2} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]} \sum_{q|[d_1,d_2]} \sum_{a \bmod^* q} e(na/q) \\ &= \sum_{\substack{q\leqslant z^2,\\ (q,P(z_0))=1}} w_q(z; z_0) c_q(n) \end{aligned}$$

where

$$w_q(z; z_0) = \sum_{q|[d_1,d_2]} \frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]}. \tag{7.5}$$

Note that $w_q(z; z_0) = 0$ when q does not divide $P(z)/P(z_0)$, and in particular when it is not squarefree. Let us assume now that $q|P(z)/P(z_0)$.

Expliciting $w_q(z; z_0)$

We introduce the definition (7.4) of the λ_d's and obtain

$$G(z; z_0)^2 w_q(z; z_0) = \sum_{\substack{\ell_1,\ell_2\leqslant z,\\ (\ell_1\ell_2,P(z_0))=1}} \frac{\mu^2(\ell_1)}{\varphi(\ell_1)} \frac{\mu^2(\ell_2)}{\varphi(\ell_2)} \sum_{\substack{q|[d_1,d_2],\\ d_1|\ell_1,d_2|\ell_2}} \frac{d_1\mu(d_1)d_2\mu(d_2)}{[d_1,d_2]}.$$

[37] O. Ramaré and I. Ruzsa, 2001, "Additive properties of dense subsets of sifted sequences".
[36] O. Ramaré, 2009, *Arithmetical aspects of the large sieve inequality.*

The inner sum vanishes is ℓ_1 has a prime factor prime to $q\ell_2$, and similarly for ℓ_2. Furthermore, we need to have $q|[\ell_1, \ell_2]$ for the inner sum not to be empty. Whence we may write $\ell_1 = q_1q_3\ell$ and $\ell_2 = q_2q_3\ell$ where $(\ell, q) = 1$ and $q = q_1q_2q_3$. The part of the inner sum corresponding to ℓ has value $\prod_{p|\ell}(p - 2 + 1) = \varphi(\ell)$. We have reached

$$G(z; z_0)^2 w_q(z; z_0) = \sum_{\substack{\ell \leqslant z,\\ (\ell, qP(z_0))=1}} \frac{\mu^2(\ell)}{\varphi(\ell)} \sum_{\substack{q_1q_2q_3=q,\\ q_1q_3\ell \leqslant z,\\ q_2q_3\ell \leqslant z}} \frac{1}{\varphi(q)\varphi(q_3)} \sum_{\substack{q|[d_1,d_2],\\ d_1|q_1q_3,\\ d_2|q_2q_3}} \frac{d_1\mu(d_1)d_2\mu(d_2)}{[d_1, d_2]}.$$

In this last inner sum, we have necessarily $d_1 = q_1d_1'$ and $d_2 = q_2d_2'$, so $q_3 = [d_1', d_2']$. We may thus write

$$G(z; z_0)^2 w_q(z; z_0) = \sum_{\substack{\ell \leqslant z,\\ (\ell, qP(z_0))=1}} \frac{\mu^2(\ell)}{\varphi(\ell)} \sum_{\substack{q_1q_2q_3=q,\\ q_1q_3\ell \leqslant z,\\ q_2q_3\ell \leqslant z}} \frac{\mu(q)\mu(q_3)}{\varphi(q)\varphi(q_3)} \sum_{\substack{d_1', d_2':\\ q_3=[d_1',d_2']}} \frac{d_1'\mu(d_1')d_2'\mu(d_2')}{[d_1', d_2']}.$$

This last inner sum has value $\varphi_2(q_3)$, whence

$$G(z; z_0)^2 w_q(z; z_0) = \frac{\mu(q)}{\varphi(q)} \sum_{\substack{\ell \leqslant z,\\ (\ell, qP(z_0))=1}} \frac{\mu^2(\ell)}{\varphi(\ell)} \sum_{\substack{q_1q_2q_3=q,\\ q_1q_3\ell \leqslant z,\\ q_2q_3\ell \leqslant z}} \frac{\mu(q_3)\varphi_2(q_3)}{\varphi(q_3)}$$

as announced. The size conditions are readily seen to imply that $\ell \leqslant z/\sqrt{q}$. □

Lemma 7.2. When $4 \leqslant z_0 \leqslant z$, we have $|w_q(z; z_0)| \leqslant 6 \dfrac{\log z_0}{\sqrt{q} \log z}$.

Proof. We deduce from the definition the estimate $|\xi_q(y)| \leqslant 3^{\omega(q)}$, and thus

$$|G(z; z_0)w_q(z; z_0)| \leqslant 3^{\omega(q)}/\varphi(q). \tag{7.6}$$

As $z_0 > 3$, we may assume that q is prime to 6, since otherwise $w_q(z; z_0) = 0$. We use Lemma 7.5 to get

$$\begin{aligned} |w_q(z; z_0)| &\leqslant \prod_{p \geqslant 5} \max\left(\frac{3\sqrt{p}}{p-1}, 1\right) \frac{1}{G(z; z_0)\sqrt{q}} \leqslant \frac{2.23 \times e^\gamma \log 2z_0}{\sqrt{q} \log z} \\ &\leqslant \frac{2.23 \times e^\gamma \log z_0}{\sqrt{q} \log z} \left(1 + \frac{\log 2}{\log 4}\right) \leqslant 6 \frac{\log z_0}{\sqrt{q} \log z} \end{aligned}$$

It has been enough, in the Euler product, to consider the primes $p = 5$ and $p = 7$. The lemma follows swiftly. □

7.1. Handling the G-functions

We define

$$G_d(y;z_0)=\sum_{\substack{\ell\leqslant y,\\(\ell,dP(z_0))=1}}\frac{\mu^2(\ell)}{\varphi(\ell)},\quad G(y;z_0)=G_1(y;z_0). \tag{7.7}$$

When $z_0=2$, these functions are classical in sieve theory (see for instance Equation (1.3) of [19, Chapter 3] by H. Halberstam and H.E. Richert)and we shall in fact reduce the analysis to this case. So we use the specific notation:

$$G(y)=G(y;2)=\sum_{\ell\leqslant y}\frac{\mu^2(\ell)}{\varphi(\ell)}. \tag{7.8}$$

Here are three lemmas we will combine for their evaluations.

Lemma 7.3. We have $\displaystyle\prod_{p<z_0}\frac{p}{p-1}G(z;z_0)\geqslant G(z)$.

Lemma 7.4. When $z_0\geqslant 2$, we have $\displaystyle\prod_{p<z_0}\frac{p-1}{p}\geqslant\frac{e^{-\gamma}}{\log(2z_0)}$.

Lemma 7.5. When $z_0\geqslant 2$, we have $G(z;z_0)\geqslant e^{-\gamma}\dfrac{\log z}{\log(2z_0)}$.

Proof of Lemma 7.3. Let us change the notation in Lemma 4.6 and use z_1 rather than z_0. Then Lemma 7.3 follows from Lemma 4.6 with $z_1=2$ and $d=P(z_0)$, and on recalling definition (7.8). □

Proof of Lemma 7.4. In [44, Theorem 8] by J.B. Rosser & L. Schoenfeld, we find the estimate $\displaystyle\prod_{p<z_0}\frac{p}{p-1}<e^{\gamma}\log z_0\left(1+\frac{1}{2\log^2 z_0}\right)$ which is valid when $z_0>286$. Hence, when $z_0>286$ we find that

$$e^{\gamma}\log(2z_0)\prod_{p<z_0}\frac{p-1}{p}\geqslant\left(1+\frac{\log 2}{\log z_0}\right)\left(1+\frac{1}{2\log^2 z_0}\right)^{-1}\geqslant 1$$

[19] H. Halberstam and H. Richert, 1981, "Almost-primes in short intervals".
[44] J. Rosser and L. Schoenfeld, 1962, "Approximate formulas for some functions of prime numbers".

as $\log 2 \geqslant 1/2$. A direct inspection using Pari-GP [32] establishes the stated inequality for the remaining values of z_0. □

Proof of Lemma 7.5. Though the proof that follows does not require it, let us notice that the lemma is obvious when $\log z_0 \geqslant e^{-\gamma} \log z$, as $G(z; z_0) \geqslant 1$ (consider the contribution of the summand $\ell = 1$ in (7.7)). For the proof, simply combine Lemma 7.3 together with Lemma 7.4. □

[32], 2014, *PARI/GP, version* `2.7.0`.

8 Large Sieve for Primes

Theorem 8.1

Let $\mathcal{X}$ be a δ-well spaced subset of $\mathbb{R}/\mathbb{Z}$ and $N \geqslant 1000$. Let $(u_p)_{p\leqslant N}$ be a sequence of complex numbers. We have

$$\sum_{x\in\mathcal{X}}\Big|\sum_{p\leqslant N} u_p e(xp)\Big|^2 \leqslant 280\frac{N+\delta^{-1}}{\log N}\log(2|\mathcal{X}|)\sum_{p\leqslant N}|u_p|^2.$$

Let us recall that a set $\mathcal{X}\subset\mathbb{R}/\mathbb{Z}$ is said to be δ-well spaced when $\min_{x\neq x'\in\mathcal{X}}|x-x'|_{\mathbb{Z}}\geqslant\delta$, where $|y|_{\mathbb{Z}}=\min_{k\in\mathbb{Z}}|y-k|$ denotes in a rather unusual manner the distance to the nearest integer. In most applications, δ^{-1} is smaller than N.

B.J. Green & T. Tao's result in [15] relates to a similar inequality though with a larger dependence in $|\mathcal{X}|$ that the $\log(2|\mathcal{X}|)$ we have here. We shall prove this inequality in dual form in Theorem 8.2.

Though Theorem 8.1 is an L^2-estimate, a fundamental *maximal* character is hidden in the fact that the set $\mathcal{X}$ may be chosen freely.

8.1. The fundamental estimate

By a classical duality argument, the proof of Theorem 8.1 will be a straightforward consequence of the next inequality (see Theorem 6 of Chapter 7 in [26] by H. L. Montgomery).

Theorem 8.2

Let $N\geqslant 1000$. Let $\mathcal{B}$ be a δ-well spaced subset of $\mathbb{R}/\mathbb{Z}$. For any function f on $\mathcal{B}$, we have

$$\sum_{p\leqslant N}\Big|\sum_{b\in\mathcal{B}} f(b)e(bp)\Big|^2 \leqslant 280(N+\delta^{-1})\|f\|_2^2\frac{\log(2\|f\|_1^2/\|f\|_2^2)}{\log N}.$$

where $\|f\|_q^q=\sum_{b\in\mathcal{B}}|f(b)|^q$ for any positive q.

Proof. Let us first notice that $\|f\|_1^2\geqslant\|f\|_2^2$. Let $z=N^{1/4}$ and

$$z_0=\left(2\frac{\|f\|_1^2}{\|f\|_2^2}\right)^2\geqslant 4.$$

[15] B. Green and T. Tao, 2006, "Restriction theory of the Selberg sieve, with applications".
[26] H. Montgomery, 1994, *Ten lectures on the interface between analytic number theory and harmonic analysis.*

We have $z_0 \leqslant z$ when $\|f\|_1^2/\|f\|_2^2 \leqslant N^{1/8}/2$. When this condition is not met, we use the dual of the usual large sieve inequality (see [27] by H.L. Montgomery) to infer that

$$\sum_{p\leqslant N}\Big|\sum_{b\in\mathcal{B}} f(b)e(bp)\Big|^2 \leqslant (N+\delta^{-1})\|f\|_2^2$$
$$\leqslant (N+\delta^{-1})\|f\|_2^2\frac{\log(2\|f\|_1^2/\|f\|_2^2)}{\log(N^{1/8})}.$$

This establishes our inequality in this case. Henceforth, we assume that $z_0 \leqslant z$. We discard the small primes trivially:

$$\sum_{p\leqslant z}\Big|\sum_{b\in\mathcal{B}} f(b)e(bp)\Big|^2 \leqslant z\|f\|_1^2 \leqslant N^{3/8}\|f\|_2^2/\sqrt{8}$$
$$\leqslant N\frac{\|f\|_2^2\log(2\|f\|_1^2/\|f\|_2^2)}{\log N}\frac{\log N}{\sqrt{8}N^{5/8}\log 2}$$
$$\leqslant \frac{N}{2880}\frac{\|f\|_2^2\log(2\|f\|_1^2/\|f\|_2^2)}{\log N}.$$

Let us now define

$$W = \sum_{z<p\leqslant N}\Big|\sum_{b\in\mathcal{B}} f(b)e(bp)\Big|^2. \tag{8.1}$$

We bound above the characteristic function of those primes by our enveloping sieve and further majorize the characteristic function of the interval $[1,N]$ by a function $\psi(u) = C_{[-N/2,N/2],\delta_1}(u-(N/2))$ (see Theorem 3.1) of Fourier transform supported by $[-\delta_1,\delta_1]$ where $\delta_1 = \min(\delta, 1/(2z^4))$, and which is such that $\hat{\psi}(0) = N+\delta_1^{-1}$. This leads to

$$W \leqslant \sum_{\substack{q\leqslant z^2,\\(q,P(z_0))=1}} w_q(z;z_0)\sum_{a \bmod^* q}\sum_{b_1,b_2} f(b_1)\overline{f(b_2)}\sum_{n\in\mathbb{Z}} e((b_1-b_2)n)e(an/q)\psi(n).$$

We split this quantity according to whether $q < z_0$ or not:

$$W = W(q<z_0) + W(q\geqslant z_0).$$

When $q \geqslant z_0$, Poisson summation formula tells us that the inner sum is also $\sum_{m\in\mathbb{Z}}\hat{\psi}(b_1-b_2-(a/q)+m)$. The sum over b_1, b_2 and n is thus

$$\leqslant (N+\delta_1^{-1})\sum_{b_1,b_2} f(b_1)\overline{f(b_2)}\#\big\{(a/q)/\|b_1-b_2+a/q\|<\delta_1\big\}.$$

[27] H. Montgomery, 1978, "The analytic principle of the large sieve".

Given (b_1, b_2), at most one a/q may work, since $1/z^4 > 2\delta_1$. By bounding above $w_q(z; z_0)$ by Lemma 7.2, we see that

$$W(q \geqslant z_0) \leqslant 6(N + \delta_1^{-1}) \frac{\|f\|_1^2 \log z_0}{\sqrt{z_0} \log z}$$
$$\leqslant \frac{6}{\sqrt{2}} (N + \delta_1^{-1}) \frac{\|f\|_2^2 \log z_0}{\log z}. \tag{8.2}$$

When $w_q(z; z_0) \neq 0$, we have $q | P(z)/P(z_0)$; on adding the condition $q < z_0$, only $q = 1$ remains. Since $\mathcal{B}$ is δ-well-spaced and $w_1(z; z_0) = 1/G(z; z_0)$, Lemma 7.5 leads to

$$W(q < z_0) \leqslant (N + \delta_1^{-1}) \frac{e^\gamma \|f\|_2^2 \log 2z_0}{\log z}.$$

We check that $(N + \delta_1^{-1}) \leqslant \frac{N+4N}{N}(N + \delta^{-1})$. We finally get

$$\sum_{p \leqslant N} \Big| \sum_{b \in \mathcal{B}} f(b) e(bn) \Big|^2 \leqslant \left(\frac{1}{2880} + 5 \times 2 \times 4 \times \left(\frac{6}{\sqrt{2}} + e^\gamma \left(1 + \frac{\log 2}{\log z_0} \right) \right) \right)$$
$$\times (N + \delta^{-1}) \|f\|_2^2 \frac{\log(2\|f\|_1^2/\|f\|_2^2)}{\log N}.$$

The proof of the theorem follows readily. □

8.2. Proof of Theorem 8.1

Proof of Theorem 8.1. This is a classical argument of duality. We write

$$V = \sum_{x \in \mathcal{X}} \Big| \sum_{p \leqslant N} u_p e(xp) \Big|^2 = \sum_{x \in \mathcal{X}} \sum_{p \leqslant N} u_p \overline{S(x)} e(xp)$$

where $S(x) = \sum_{1 \leqslant p \leqslant N} u_p e(xp)$. On using the Cauchy-Schwarz inequality, we get

$$V^2 \leqslant \sum_{p \leqslant N} |u_p|^2 \sum_{p \leqslant N} \Big| \sum_{x \in \mathcal{X}} \overline{S(x)} e(xp) \Big|^2.$$

We invoke Theorem 8.2 and notice to control $\|S\|_1^2/\|S\|_2^2$ that

$$\Big(\sum_{x \in \mathcal{X}} |\overline{S(x)}| \Big)^2 \leqslant |\mathcal{X}| \sum_{x \in \mathcal{X}} |\overline{S(x)}|^2.$$

This leads to

$$V^2 \leqslant 280 \, \frac{N + \delta^{-1}}{\log N} \sum_{p \leqslant N} |u_p|^2 \sum_{x \in \mathcal{X}} |S(x)|^2 \log(2|\mathcal{X}|).$$

On simplifying by $\sum_{x \in \mathcal{X}} |S(x)|^2$ (after discussing whether it vanishes or not), we get our estimate. □

8.3. Another proof for Farey fractions

Theorem 8.1 handles arbitrary subsets $\mathcal{X}$, but an inequality valid only for subsets of the Farey sequence is enough for Theorem 9.4. As it is simpler, let us record it here. It relies on the next lemma.

Lemma 8.3

Let $(u_p)_{\sqrt{N}<p\leqslant N}$ be a sequence of complex numbers. For any $Q \leqslant \sqrt{N}$, we have

$$\sum_{q\leqslant Q}\sum_{a \bmod^* q}\left|\sum_p u_p e(pa/q)\right|^2 = \sum_{q\leqslant Q}\frac{q}{\varphi(q)}G_q(Q/q)\sum_{\chi \bmod^* q}\left|\sum_p u_p\chi(p)\right|^2$$

where G_q is defined in (4.3) and "$\chi \bmod^* q$" means that χ runs through primitives characters modulo q.

This identity is somehow contained in the paper [5] by E. Bombieri & H. Davenport. A general version of it, valid for any sieve situation can be found in Theorem 4 of [37].

Proof. Let us define

$$S(c;q) = \sum_{p\equiv c[q]} u_p.$$

We readily find that (see (2.2)), when $q < \sqrt{N}$:

$$\begin{aligned} S(q;b) &= \sum_{d|q}\sum_{\chi \bmod^* d}\left(\frac{1}{\varphi(q)}\sum_{c \bmod q} S(q;c)\overline{\chi(c)}\right)\chi(b) \\ &= \sum_{d|q}\sum_{\chi \bmod^* d}\left(\frac{1}{\varphi(q)}\sum_{c \bmod^* q} S(q;c)\overline{\chi(c)}\right)\chi(b), \end{aligned}$$

the last line being valid because our sequence is supported on integers prime to q. Since $\chi(c)$ depends only on c modulo d, we may modify the above in the form

$$\begin{aligned} S(q;b) &= \frac{1}{\varphi(q)}\sum_{d|q}\sum_{\chi \bmod^* d}\left(\sum_{c \bmod^* q} S(d;c)\overline{\chi(c)}\right)\chi(b), \\ &= \frac{1}{\varphi(q)}\sum_{d|q}\sum_{\chi \bmod^* d} S(\overline{\chi})\chi(b), \end{aligned}$$

say. Consequently, we find that

$$\varphi(q)^2\sum_{b \bmod^* q}|S(q,b)|^2 = \varphi(q)\sum_{d|q}\sum_{\chi \bmod^* d}|S(\overline{\chi})|^2.$$

[5] E. Bombieri and H. Davenport, 1968, "On the large sieve method".
[37] O. Ramaré and I. Ruzsa, 2001, "Additive properties of dense subsets of sifted sequences".

Similarly, and setting $S(a/q) = \sum_p u_p e(pa/q)$, we find that

$$q^2 \sum_{b \bmod^* q} |S(q,b)|^2 = q \sum_{d|q} \sum_{a \bmod^* d} |S(a/d)|^2.$$

A comparison of the two yields the identity

$$\sum_{d|q} \sum_{a \bmod^* d} |S(a/d)|^2 = \frac{q}{\varphi(q)} \sum_{d|q} \sum_{\chi \bmod^* d} |S(\overline{\chi})|^2.$$

On using the Moebius inversion formula, this implies that

$$\begin{aligned}\sum_{a \bmod^* \ell} |S(a/\ell)|^2 &= \sum_{q|\ell} \mu(\ell/q) \frac{q}{\varphi(q)} \sum_{d|q} \sum_{\chi \bmod^* d} |S(\overline{\chi})|^2 \\ &= \sum_{d|\ell} \sum_{d|q|\ell} \mu(\ell/q) \frac{q}{\varphi(q)} \sum_{\chi \bmod^* d} |S(\overline{\chi})|^2.\end{aligned}$$

The arithmetical coefficient may be computed by multiplicativity. We swiftly find that, with the power of the prime p dividing ℓ being α while the one dividing d is $\beta \leqslant \alpha$:

$$\sum_{p^\beta | q | p^\alpha} \mu(p^\alpha/q) \frac{q}{\varphi(q)} = \begin{cases} p^\alpha/\varphi(p^\alpha) & \text{when } \beta = \alpha, \\ 0 & \text{when } \beta < \alpha \text{ and } \alpha \geqslant 2, \\ 1/(p-1) & \text{when } \beta < \alpha \text{ and } \alpha = 1. \end{cases}$$

The last case occurs only when $\beta = 0$. When this contribution does not vanish, we may therefore write $\ell = dm$ where $(m,d) = 1$ and $\mu^2(m) = 1$ and the value is $d/(\varphi(d)\varphi(m))$. Let us now sum over ℓ. We find that

$$\sum_{\ell \leqslant Q} \sum_{a \bmod^* \ell} |S(a/\ell)|^2 = \sum_{d \leqslant Q} \sum_{\substack{m \leqslant Q/d \\ (m,d)=1}} \frac{d}{\varphi(d)} \frac{\mu^2(m)}{\varphi(m)} \sum_{\chi \bmod^* d} |S(\overline{\chi})|^2.$$

This is exactly what we claimed. □

Please notice that the above proof does not use the value of the Gauss sums, contrarily to the one of Bombieri & Davenport, which is one of the reasons it may be extended to a larger setting.

We are now ready to prove our inequality.

Theorem 8.4

Let $(u_p)_{\sqrt{N} < p \leqslant N}$ be a sequence of complex numbers. For any $Q \leqslant \sqrt{N}$, we have

$$\sum_{\ell \leqslant Q} \sum_{a \bmod^* \ell} |S(a/\ell)|^2 \leqslant 8 \frac{N \log(20Q)}{\log N} \sum_p |u_p|^2.$$

Proof. When $Q \geqslant N^{1/4}$, we have $(4\log Q)/\log N \geqslant 1$ and the large sieve inequality gives us

$$\sum_{\ell\leqslant Q}\sum_{a \bmod^* \ell} |S(a/\ell)|^2 \leqslant (N+Q^2)\sum_p |u_p|^2 \leqslant 8\frac{N\log Q}{\log N}\sum_p |u_p|^2$$

as required. When $Q \leqslant N^{1/4}$, we employ Lemma 8.3 to write:

$$\begin{aligned}
\sum_{\ell\leqslant Q}\sum_{a \bmod^* \ell} |S(a/\ell)|^2 &= \sum_{q\leqslant Q}\frac{q}{\varphi(q)}G_q(Q/q)\sum_{\chi \bmod^* q}\left|\sum_p u_p\chi(p)\right|^2 \\
&= \sum_{q\leqslant Q}\frac{q}{\varphi(q)}\frac{G_q(Q/q)}{G_q(\sqrt{N}/q)}G_q(\sqrt{N}/q)\sum_{\chi \bmod^* q}\left|\sum_p u_p\chi(p)\right|^2 \\
&\leqslant \max_{q\leqslant Q}\frac{G_q(Q/q)}{G_q(\sqrt{N}/q)}\sum_{q\leqslant\sqrt{N}}\frac{q}{\varphi(q)}G_q(\sqrt{N}/q)\sum_{\chi \bmod^* q}\left|\sum_p u_p\chi(p)\right|^2 \\
&\leqslant \max_{q\leqslant Q}\frac{G_q(Q/q)}{G_q(\sqrt{N}/q)}\sum_{\ell\leqslant\sqrt{N}}\sum_{a \bmod^* \ell}|S(a/\ell)|^2.
\end{aligned}$$

We readily find, by using Lemma 4.6, that

$$\max_{q\leqslant Q}\frac{G_q(Q/q)}{G_q(\sqrt{N}/q)} \leqslant \frac{G(Q)}{\log(\sqrt{N}/Q)} \leqslant \frac{\log(20Q)}{(1/4)\log N}$$

by Lemmas 4.5 and 4.7. □

9 Primes and Cusps

During their investigations on the primes, B. Green in [16] and B. Green & T. Tao in [15] were led to consider the large values of the trigonometric polynomial built on some dense subset of the primes. Let us start by a definition.

Definition 9.1

Let $\mathcal{P}^* \subset [1, N]$ be a subset of the primes, and let $A \geqslant 1$ be given. We define the *A-spectum* of $\mathcal{P}^*$ by

$$\mathscr{C}(\mathcal{P}^*, A) = \left\{ \alpha \in \mathbb{R}/\mathbb{Z} : \left| \sum_{p \in \mathcal{P}^*} e(\alpha p) \right| \geqslant |\mathcal{P}^*|/A \right\} \tag{9.1}$$

where $|\mathcal{P}^*|$ denotes the cardinality of $\mathcal{P}^*$. Each point in this A-spectrum is called an *A-cusp*.

As the involved trigonometric polynomial is continuous, the set $\mathscr{C}(\mathcal{P}^*, A)$ is closed, hence compact, and is more precisely a finite union of arcs. The word *spectrum* is in accordance with several works, see e.g. Section 3.4 of [46] by T. Sanders or Section 4.6 of the book [52] by T. Tao & V. Vu. The reader should notice a common variation: the right-hand side is often N/A rather than $|\mathcal{P}^*|/A$ as above.

On the number of cusps

Theorem $\mathscr{D}$

There exist positive constants C_1 and C_2 such that the following holds. Define $D(\mathcal{P}^*, A)$ to be the maximal cardinality of a $1/N$-well spaced subset of $\mathscr{C}(\mathcal{P}^*, A)$ and $K = N/(|\mathcal{P}^*| \log N)$. We have, for $A \leqslant \sqrt{N}$,

$$\frac{C_1 A^2}{K \log(2A)} \leqslant D(\mathcal{P}^*, A) \leqslant C_2 A^2 K \log(2A).$$

Furthermore, we have

$$C_1 A \leqslant \int_1^A \frac{D(\mathcal{P}^*, a)}{a^2} da, \quad \int_1^A \frac{D(\mathcal{P}^*, a)}{a^3} da \leqslant C_2 K \log(2A).$$

[16] B. Green, 2005, "Roth's theorem in the primes".
[15] B. Green and T. Tao, 2006, "Restriction theory of the Selberg sieve, with applications".
[46] T. Sanders, 2011, "On Roth's theorem on progressions".
[52] T. Tao and V. Vu, 2008, "John-type theorems for generalized arithmetic progressions and iterated sumsets".

The last two inequalities tell us that $D(\mathcal{P}^*, A)$ can neither remain of order $A^2/\log(2A)$ nor of order $A^2\log(2A)$.

9.1. Cusps are scarce

Given a subset $\mathcal{P}^* \subset [\sqrt{N}, N]$ of the primes, of cardinality $N/(K\log N)$, B. Green & T. Tao proved in [15] that the cardinality of a $1/N$-well spaced subset of $\mathscr{C}(\mathcal{P}^*, A)$ is $\mathcal{O}_\varepsilon((A^2K)^{1+\varepsilon})$, for every positive ε. We are somewhat more precise in the next result.

Theorem 9.2
Let $\mathcal{X}$ be a $1/N$-well spaced subset such that

$$\mathcal{X} \subset \mathscr{C}(\mathcal{P}^*, A) = \Big\{\alpha \in \mathbb{R}/\mathbb{Z} : \Big|\sum_{p\in\mathcal{P}^*} e(\alpha p)\Big| \geqslant |\mathcal{P}^*|/A\Big\}.$$

With $K = N/(|\mathcal{P}^*|\log N)$, the set $\mathcal{X}$ satisfies

$$|\mathcal{X}| \leqslant (4e^\gamma + o(1))A^2K\log(2A).$$

When $N \geqslant 10^4$, such a set contains at most $19A^2K\log(2A)$ points.

In particular, the measure of $\mathscr{C}(\mathcal{P}^*, A)$ is $\mathcal{O}(A^2(\log A)/N)$, as N grows.

Lemma 9.3
Let $N \geqslant 10^4$ and $\mathcal{X}$ be a δ-well spaced subset of $\mathbb{R}/\mathbb{Z}$. Let $(u_p)_{p\leqslant N}$ be a sequence of complex numbers. We have, for $A \geqslant 1$,

$$\#\Big\{x \in \mathcal{X} : \Big|\sum_{p\leqslant N} u_p e(xp)\Big| \geqslant V/A\Big\} \leqslant 19A^2\log(2A),$$

where V is defined by

$$V = \Big(\frac{N+\delta^{-1}}{\log N}\sum_{p\leqslant N}|u_p|^2\Big)^{1/2}. \tag{9.2}$$

The constant 19 may be replaced by $4e^\gamma + o(1) = 7.12\cdots$ when $N \to \infty$ and $A \to \infty$.

We used the notation $\#S$ to denote the cardinality of the set S.

[15] B. Green and T. Tao, 2006, "Restriction theory of the Selberg sieve, with applications".

Proof of Lemma 9.3. This is a trivial application of Markov's inequality. □

Proof of Theorem 9.2. After the change of variable $A \mapsto A\sqrt{K(1+(N\delta)^{-1})}$, Lemma 9.3 tells us that

$$\#\Big\{x \in \mathcal{X} : \Big|\sum_{\substack{p\leqslant N\\ p\in\mathcal{P}^*}} e(xp)\Big| \geqslant \frac{N}{AK\log N}\Big\} \leqslant 19A^2K(1+(N\delta)^{-1}). \tag{9.3}$$

Theorem 9.2 follows swiftly from there. □

9.2. Getting many cusps

Theorem 9.4

Notation being as in Definition 9.1. Let $N \geqslant 10^4$, $A \in [2, \sqrt{N}]$, $B \in [1, A]$ and $\xi \in \mathscr{C}(\mathcal{P}^*, B)$ be given. Define $t^*(\alpha) = |T^*(\alpha)|/T^*(0)$. The set

$$\mathcal{F} = \big\{\xi + (a/q) : a \bmod^* q, q \leqslant A/B\big\} \cap \mathscr{C}(\mathcal{P}^*, A)$$

satisfies

$$|\mathcal{F}| \geqslant A^2/(6800B^4Z^2K\log A),$$

where $Z \in [1/B, 1]$ is the maximum value of $t^*(\xi + (a/q))$ for $\xi + (a/q) \in \mathcal{F}$.

It is simpler to grasp the relative scope of this result by first examining the next corollary.

Corollary 9.5

Notation being as in Definition 9.1. Let $N \geqslant 10^4$, and $A \in [2, \sqrt{N}]$ be given. Define $t^*(\alpha) = |T^*(\alpha)|/T^*(0)$. The set

$$\mathcal{F} = \big\{(a/q) : a \bmod^* q, q \leqslant A\big\} \cap \mathscr{C}(\mathcal{P}^*, A)$$

satisfies

$$|\mathcal{F}| \geqslant A^2/(6800K\log A).$$

Lemma 9.6

Notation being as in Definition 9.1, and assuming $N \geqslant 10$, the following holds:

- When α lies in $\mathscr{C}(\mathcal{P}^*, A)$, so do $-\alpha$ and $\frac{1}{2}+\alpha$.
- We have $\{\frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\} \subset \mathscr{C}(\mathcal{P}^*, 2)$
- For every $\xi \in \mathbb{R}/\mathbb{Z}$ with $T^*(\xi) \neq 0$ and every square-free positive integer $q < \sqrt{N}$ such that $\varphi(q) \leqslant AT^*(0)/|T^*(\xi)|$, there exists a, coprime with q, such that $\xi + (a/q)$ lies in $\mathscr{C}(\mathcal{P}^*, A)$.

Proof. Let us recall that $\mathcal{P}^* \in [\sqrt{N}, N]$, so that the elements of $\mathcal{P}^*$ are prime to the moduli 2, 3 and q appearing in the three claimed properties. The first item is a consequence of the facts that (1) the characteristic function $\mathbf{1}_{\mathcal{P}^*}$ of $\mathcal{P}^*$ is real valued and (2) that every member of $\mathcal{P}^*$ is odd. Concerning the third item, it is enough to notice the inequality

$$\sum_{a \bmod^* q} \Big| \sum_{p \in \mathcal{P}^*} e(p(\xi + (a/q)) \Big| \geqslant \Big| \sum_{a \bmod^* q} \sum_{p \in \mathcal{P}^*} e(p(\xi + (a/q)) \Big| = \mu^2(q)|T^*(\xi)| \quad (9.4)$$

with $\xi = 0$, since $c_q(p) = \mu(q)$ where $c_q(n)$ denotes the value of the Ramanujan sum at n. The second item follows from this same inequality applied to $q = 3$, on noticing that the absolute values of the involved trigonometric polynomial at $1/3$ and $2/3$ are the same. □

The proof of Theorem 9.4 relies on extracting information from the stream of inequalities (9.4), for varying q's. We concentrate on the major case $\xi = 0$.

Proof of Corollary 9.5. Let us set $Q = A/2 \geqslant 1$ and

$$\mathcal{Q} = \big\{ q : q \leqslant Q, \mu^2(q) = 1 \big\}. \quad (9.5)$$

The non-negative variables $t^*(a/q)$ are bounded above by 1 and satisfy by (9.4):

$$\forall q \in \mathcal{Q}, \quad \sum_{a \bmod^* q} t^*(a/q) \geqslant 1$$

as well as, by Theorem 9.2 and since the points in $\mathcal{F}$ are $1/N$-well spaced,

$$\forall C \geqslant 1, \quad \#\big\{ a/q : q \in \mathcal{Q}, t^*(a/q) \geqslant 1/C \big\} \leqslant 19KC^2 \log 2C.$$

Let us further introduce the variables

$$x(q, C) = \#\big\{ a \leqslant q : (a, q) = 1 \quad \& \quad 1/C > t^*(a/q) \geqslant 1/(C+1) \big\}.$$

We thus get

$$\forall q \in \mathcal{Q}, \quad \sum_{C \geqslant 1} \frac{1}{C} x(q, C) \geqslant 1.$$

Let us assume momentarily that A is a positive integer and degrade the above inequality into

$$\sum_{1 \leqslant C \leqslant A} \frac{1}{C} x(q, C) + \Big(\varphi(q) - \sum_{1 \leqslant C \leqslant A} x(q, C) \Big) \frac{1}{A} \geqslant 1,$$

i.e., when $q \leqslant A/2$,

$$\sum_{1\leqslant C\leqslant A}\left(\frac{1}{C}-\frac{1}{A}\right)x(q,C) \geqslant 1-\frac{\varphi(q)}{A} \geqslant 1/2.$$

We now sum over $q \in \mathcal{Q}$, getting, by Lemma 9.7 found in the final section, that

$$\sum_{1\leqslant C\leqslant A}\left(\frac{1}{C}-\frac{1}{A}\right)\sum_{q\in\mathcal{Q}} x(q,C) \geqslant A/4.$$

Let us set

$$X(C)=\sum_{D\leqslant C}\sum_{q\in\mathcal{Q}} x(q,D), \quad (X(0)=0)$$

which therefore satisfies

$$\sum_{1\leqslant C\leqslant A}\left(\frac{1}{C}-\frac{1}{A}\right)(X(C)-X(C-1)) \geqslant \frac{A}{4}.$$

On reshuffling the left-hand side, we obtain:

$$\sum_{1\leqslant C\leqslant A-1} X(C)\left(\frac{1}{C}-\frac{1}{C+1}\right) \geqslant \frac{A}{4}.$$

The non-decreasing function $C \mapsto X(C)$ is therefore constrainted by the two inequalities:

$$\sum_{1\leqslant C\leqslant A-1}\frac{X(C)}{C(C+1)} \geqslant \frac{A}{4}, \quad X(C) \leqslant 19K(C+1)^2\log(2C+2). \tag{9.6}$$

Let us split the first sum at $C_0 = \lceil \theta A\rceil \leqslant A$. We deduce from the above that

$$19K\log(2A+2)\sum_{1\leqslant C\leqslant \theta A}\frac{C+1}{C} + X(A-1)\left(\frac{1}{C_0}-\frac{1}{A}\right) \geqslant \frac{A}{4}$$

and thus

$$38K\theta A\log(2A+2)+\frac{X(A)}{\theta A} \geqslant \frac{A}{4}.$$

This calls for

$$\theta^{-1}=\sqrt{\frac{A^2}{X(A)}38K\log(2A+2)} \geqslant 1 \tag{9.7}$$

and this choice leads to

$$2\sqrt{X(A)38K\log(2A+2)} \geqslant A/4$$

This finally amounts to $X(A) \gg \frac{A^2}{K\log(A+1)}$. This concludes the proof. □

Proof of Theorem $\mathscr{D}$. The point-wise upper bound is proved in Theorem 9.2, while the lower one comes from Corollary 9.5. A closer inspection of both proofs discloses the integral inequalities that follow. The first one comes from (9.6) while the second one comes from Theorem 8.1 through

$$\sum_{1\leqslant a\leqslant A} \frac{D(\mathcal{P}^*, a) - D(\mathcal{P}^*, a-1)}{a^2} \ll \log(2A)$$

on setting $D(\mathcal{P}^*, 0) = 0$. We leave the details to the readers. □

9.3. Auxiliaries

Lemma 9.7. For any real number $Q \geqslant 1$, we have $\sum_{q\leqslant Q} \mu^2(q) \geqslant Q/2$.

Be cautious: the lower bound of K. Rogers given in [43] is only valid for *integer* values of Q, though this is not specified.

Proof. When $Q \geqslant 1664$, this is a consequence of [6, Théorème 3] by H. Cohen & F. Dress which asserts that $\sum_{q\leqslant Q} \mu^2(q) = 6\pi^{-2}Q + \mathcal{O}^*(0.1333\sqrt{Q})$. A straightforward numerical check concludes the proof. □

[43] K. Rogers, 1964, "The Schnirelmann density of the squarefree integers".
[6] H. Cohen and F. Dress, 1988, "Estimations numériques du reste de la fonction sommatoire relative aux entiers sans facteur carré".

10 Two examples

10.1. Results

The case of the full sequence of primes

We may approximate the trigonometric polynomial T on the primes via a local model, i.e. write

$$T(\alpha) = \sum_{p \leqslant N} e(p\alpha) = \frac{1}{V(z_0)\log N} \sum_{\substack{n \leqslant N \\ (n, P(z_0))=1}} e(n\alpha) + \mathcal{O}\Big(\frac{N}{z_0 \log N}\Big) \qquad (10.1)$$

where $z_0 \leqslant \log\log N$ is a parameter at our disposal and

$$P(z_0) = \prod_{p<z_0} p, \quad V(z_0) = \prod_{p<z_0} \Big(1 - \frac{1}{p}\Big). \qquad (10.2)$$

Eq. (10.1) is proved in next section. As a consequence and when limiting our study to A small, say $A = o(\log\log N)$, we find that the set of A-cusps of T is a union of arcs around points from $\{a/q : (a,q) = 1, q|P(z_0)\}$, for some z_0 (chosen for the error term in (10.1) to become $< N/(A\log N)$). Around a/q and when $q|P(z_0)$, say in $\alpha = \frac{a}{q} + \beta$, we readily find in Lemma 10.1 that

$$|T(\alpha)| \leqslant \min\Big(\frac{\mu^2(q)N}{\varphi(q)\log N}, \frac{P(z_0)}{|P(z_0)\beta|_{\mathbb{Z}}}\Big) + \mathcal{O}\Big(\frac{N}{z_0\log N}\Big).$$

Furthermore, on adapting the proof of P. Erdős from [10]*, we readily find that

$$\sum_{\varphi(q)\leqslant A} \mu^2(q)\varphi(q) \sim \frac{A^2}{2}. \qquad (10.3)$$

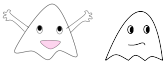
Gathering our results, we find that $\mathscr{C}(\mathcal{P} \cap [1,N], A)$ is a union of about $(A^2/2)$ arcs of length $\mathcal{O}(1/N)$.

$\diamond$ 1 $\diamond$ This settles the question when $A \leqslant \log\log N$. But what about larger values of A? Figure 10.1 is built as follows: for each angle α, we plot a segment between the center 0 and the point on the circle $e(\alpha)$, starting in 0 and of length

$$\big(|T(\alpha)|/T(0)\big)^{1/4}.$$

[10] P. Erdös, 1945, "Some remarks on Euler's ϕ-function and some related problems".

*The result of Erdős does not seem to include the square-free condition on q. But it may be there already, or in the later papers [8] by R. Dressler or [2] by P. Bateman.



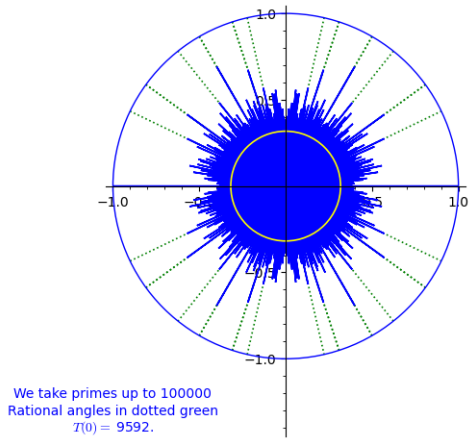


Figure 10.1: Plot of $(|T(\alpha)|T(0)^{-1})^{1/4}$. The dotted lines from the origin correspond to rational angles with a denominator $\leqslant 10$.

The smaller circle in yellow has radius $1/T(0)^{1/8}$. The reason for this doctoring is that the yellow circle would have been too small without taking this fourth power. Figure 10.1 exhibits the presence of very sharp cusps. We produced it by using Sage, cf [45].

A subsequence exhibiting non-rational cusps

A part of the Farey sequence, namely the points with square-free denominators, appears as cusps for the full sequence of primes. We found the presence of this sequence in many numerical examples*. Let us give an example where non-rational cusps appear. We select

$$\mathcal{P}_0^* = \{p : \{p\sqrt{2}\} \leqslant 1/2\}$$

where $\{x\}$ denotes the fractional part of the real number x. By detecting the condition $\{p\sqrt{2}\} \leqslant 1/2$ through Fourier analysis, we readily find that

$$T_0^*(\alpha) = \sum_{\substack{p\leqslant N\\ \{p\sqrt{2}\}\leqslant 1/2}} e(p\alpha) = \frac{1}{V(z_0)\log N} \sum_{\substack{n\leqslant N\\ \{n\sqrt{2}\}\leqslant 1/2\\ (n,P(z_0))=1}} e(n\alpha) + \mathcal{O}\Big(\frac{N}{z_0\log N}\Big) \qquad (10.4)$$

where $z_0 \leqslant \log\log N$ and $P(z_0) = \prod_{p<z_0} p$ as above. This is proved in next section. Figure 10.2 also exhibits the presence of very sharp cusps.

[45] W. Stein et al., 2024, *Sage Mathematics Software (Version 9.5)*.

*We tried random subsequences of primes of relative density 1/2, then Ramanujan primes and, finally we selected successively one prime out of two.

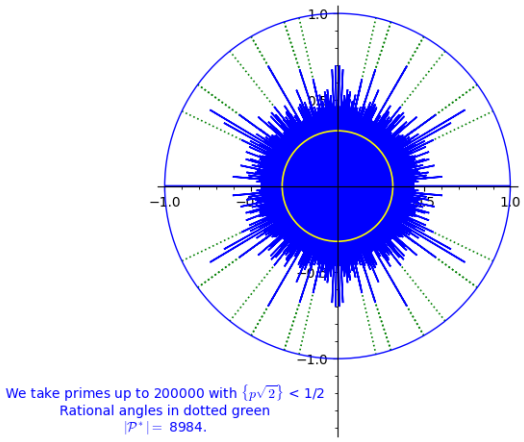


Figure 10.2: Plot of $(|T_0^*(\alpha)|/T_0^*(0))^{1/4}$. The dotted segments from the origin correspond to rational angles with a denominator $\leqslant 10$: some new cusps (not aligned with a dotted segment) arise.

10.2. Proofs

Let us start with a classical estimate.

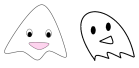
Lemma 10.1

When $N \geqslant 2$, $\alpha = (a/q) + \beta$ with a prime to q and $|q\beta| \leqslant N/(\log N)^A$, we have

$$T(\alpha) = \sum_{p \leqslant N} e(p\alpha) = \frac{\mu(q)}{\varphi(q) \log N} \sum_{n \leqslant N} e(n\beta) + \mathcal{O}\left(\frac{N}{(\log N)^2}\right).$$

Proof. A fast track for this proof is to combine [55, Theorem 3.1] of the book of R.C. Vaughan together with an obvious adaptation of [55, Lemma 3.1] in the same book. Since we do not weigh the primes by $\log p$, we have to limit our saving to $\mathcal{O}(N/(\log N)^2)$. A more detailled proof may be found in the book [42] by M. Rassias, combining Theorem 2.3.3 and Theorem 2.3.12 therein. Let us also mention that the book [11] by T. Estermann contains also the required proof, by combining Theorem 58 therein together with Eq. (152), also therein. □

[55] R. Vaughan, 1981, *The Hardy-Littlewood method.*
[42] M. T. Rassias, 2017, *Goldbach's problem.*
[11] T. Estermann, 1952, *Introduction to modern Prime Number Theory.*



Lemma 10.2

When $\alpha = (a/q) + \beta$ with a prime to q and $|q\beta| \leqslant 1/(2P(z_0))$, we have

$$\sum_{\substack{n\leqslant N\\(n,P(z_0))=1}} e(n\alpha) = \frac{\mu(q)V(z_0)}{\varphi(q)} \sum_{n\leqslant N} e(n\beta) + \mathcal{O}\Big(qP(z_0) + \frac{N|\beta|P(z_0)^2}{\varphi(q)}\Big).$$

Proof. Let us use the shortcut $M_0 = P(z_0)$. By detecting the coprimality condition through the Moebius function, we readily find that

$$\begin{aligned} L(\alpha) = \sum_{\substack{n\leqslant N\\(n,P(z_0))=1}} e(n\alpha) &= \sum_{d|M_0} \mu(d) \sum_{m\leqslant N/d} e(dm\alpha) \\ &= \sum_{\substack{d|M_0\\q|d}} \mu(d) \sum_{m\leqslant N/d} e(dm\alpha) + \mathcal{O}\Big(\sum_{\substack{d|M_0\\q\nmid d}} 1/|d\alpha|_{\mathbb{Z}}\Big). \end{aligned}$$

In the error term, we have $1/|d\alpha|_{\mathbb{Z}} \ll q$. In the main term (which may be non-zero only when $q|M_0$), we have $e(dm\alpha) = e(dm\beta)$ for each summand, so we may replace α by β in the main term. This quantity is therefore independant on a on which we may average, getting

$$L(\alpha) = \frac{\mu(q)}{\varphi(q)} L(\beta) + \mathcal{O}(qM_0).$$

To proceed, we use the above with $q = 1$ and notice that $e(dm\beta) - e((dm+k)\beta) \ll |k\beta|$. Consequently, we obtain

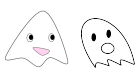
$$d\, e(dm\beta) = \sum_{(d-1)m<n\leqslant dm} e(n\beta) + \mathcal{O}(d^2|\beta|).$$

We therefore find that

$$\begin{aligned} d \sum_{\substack{n\leqslant N\\d|n}} e(n\beta) &= \sum_{n\leqslant d[N/d]} e(n\beta) + \mathcal{O}(dN|\beta|) \\ &= \sum_{n\leqslant N} e(n\beta) - \sum_{d[N/d]<n\leqslant N} e(n\beta) + \mathcal{O}(dN|\beta|) \\ &= \sum_{n\leqslant N} e(n\beta) + \mathcal{O}(d + dN|\beta|) \end{aligned}$$

from which we infer that $L(\beta) = V(z_0) \sum_{n\leqslant N} e(n\beta) + \mathcal{O}(M_0 q^{-1} + N|\beta|M_0^2)$, ending the proof. □

Proof of Eq. (10.1). We compare the expressions of Lemmas 10.1 and 10.2 to get the result. □



Lemma 10.3

Let $I \subset \mathbb{R}/\mathbb{Z}$ be an interval and χ_I be its indicator function. For each positive integer H there exist coefficients $a_H(h)$ and C_h for $-H \leqslant h \leqslant H$ with $|a_H(h)| \leqslant \min(|I|, 1/(|h|\pi))$ and $|C_h| \leqslant 1$ such that the trigonometric polynomial

$$\chi^*_{I,H}(t) = |I| + \sum_{0<|h|\leqslant H} a_H(h) e(ht) \tag{10.5}$$

satisfies, for every $t \in \mathbb{R}/\mathbb{Z}$,

$$|\chi_I(t) - \chi^*_{I,H}(t)| \leqslant \frac{1}{H+1} \sum_{|h|\leqslant H} C_h \Big(1 - \frac{|h|}{H+1}\Big) e(ht). \tag{10.6}$$

We also have $|a_H(h)| \leqslant \min(|I|, 1/(|h|+1))$.

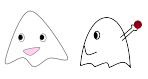
This is, up to a trivial change of notation, a specialisation of [54, Theorem 19] by J. Vaaler, reproduced in [25, Lemma 6.2] by M. Madritsch & R. Tichy.

Proof of Eq. (10.4). We detect the condition $\{p\sqrt{2}\} \leqslant 1/2$ in $T_0^*(\alpha)$ by using Lemma 10.3. We then employ the local model expansion (10.1) of $T(\alpha)$ to infer the expression

$$\begin{aligned} T_0^*(\alpha) = \sum_{\substack{p\leqslant N \\ \{p\sqrt{2}\}\leqslant 1/2}} e(p\alpha) &= \frac{(1/2)}{V(z_0)\log N} \sum_{\substack{2\leqslant n\leqslant N \\ (n,P(z_0))=1}} e(n\alpha) \\ &+ \sum_{\substack{|h|\leqslant H \\ h \text{ odd}}} \frac{1}{i\pi h} \frac{1}{V(z_0)\log N} \sum_{\substack{2\leqslant n\leqslant N \\ (n,P(z_0))=1}} e\big(n(h\sqrt{2}+\alpha)\big) + \mathcal{O}\bigg(\Big(\frac{1}{H} + \frac{1}{z_0}\Big)\frac{N}{\log N}\bigg). \end{aligned} \tag{10.7}$$

We may then exchange the summations over h and n and use Lemma 10.3 in a reverse manner, then select $H = z_0$. We reach in this manner the claimed formula. □

[54] J. Vaaler, 1985, "Some Extremal Functions in Fourier Analysis".
[25] M. G. Madritsch and R. F. Tichy, 2019, "Multidimensional van der Corput sets and small fractional parts of polynomials".



Notation

The notation used throughout these notes is standard ... in one way or the other! Here is a guideline:

— $e(y) = \exp(2i\pi y)$.

— The use of the letter p for a variable always implies this variable is a prime number.

— $[d, d']$ stands for the lcm and (d, d') for the gcd of d and d'.

— $|\mathcal{A}|$ stands for the cardinality of the set $\mathcal{A}$ while $\mathbf{1}_{\mathcal{A}}$ stands for its characteristic function.

— $\mathbf{1}$ denotes a characteristic function in one way or another. For instance, $\mathbf{1}_{\mathcal{K}_d}$ is 1 if $n \in \mathcal{K}_d$ and 0 otherwise, but we could also write it as $\mathbf{1}_{n\in\mathcal{K}_d}$, closer to what is often called the Dirac δ-symbol. We also use $\mathbf{1}_{(n,d)=1}$ and $\mathbf{1}_{q=q'}$.

— $\mathcal{P}$ is the set of prime numbers.

— $q\|d$ means that q divides d in such a way that q and d/q are coprime. In words we shall say that q *divides* d *exactly*.

— The squarefree kernel of the integer $d = \prod_i p_i^{\alpha_i}$ is $\prod_i p_i$, the product of all prime factors of d.

— $\omega(d)$ is the number of prime factors of d, counted without multiplicity.

— $\varphi(d)$ is the Euler totient, i.e. the cardinality of the multiplicative group of $\mathbb{Z}/d\mathbb{Z}$.

— $\tau(d)$ is the number of positive divisors of d.

— $\tau_k(d)$ is the number of k-tuples of (positive) integers $(d_1, \cdots, d_k)$ such that $d_1 \cdots d_k = d$, so that $\tau_2 = \tau$.

— $\mu(d)$ is the Moebius function, that is 0 when d is divisible by a square > 1 and otherwise $(-1)^r$ otherwise, where r is the number of prime factors of d.

— $c_q(n)$ is the Ramanujan sum. It is the sum of $e(an/q)$ over all a modulo q that are prime to q.

— $\Lambda(n)$ is van Mangoldt function: which is $\log p$ is n is a power of the prime p and 0 otherwise.

— The notation $f = \mathcal{O}_A(g)$ means that there exists a constant B such that $|f| \leqslant Bg$ but that this constant may depend on A. When we put in several parameters as subscripts, it simply means the implied constant depends on all of them.

— The notation $f = \mathcal{O}^*(g)$ means that $|f| \leqslant g$, that is a $\mathcal{O}$-like notation, but with an implied constant equal to 1.

— The notation $f \star g$ denotes the arithmetic convolution of f and g, that is to say the function h on positive integers such that $h(d) = \sum_{q|d} f(q)g(d/q)$. exists for every real number x.

— $\mathcal{U}$ is the compact set $(\mathcal{U}_d)_d$ where, for each d, $\mathcal{U}_d$ is the set of invertible elements modulo d.

— The letter ψ is used in two different context: either to denote the summatory function of the van Mangoldt function, that is to say $\psi(x) = \sum_{n \leqslant x} \Lambda(n)$, with the variation $\psi(x, \chi) = \sum_{n \leqslant x} \chi(n)\Lambda(n)$.

— We used the Chebyshev functions ϑ and ψ as well as their variations $\vartheta(x; \chi)$, $\vartheta(x; q, a)$, $\psi(x, \chi)$ and $\psi(x; q, a)$.

Bibliography

[1] J. Basquin. "Mémoire de DEA, Lille". In: (2006), pp. 1–37 (cit. on p. 44).

[2] P. T. Bateman. "The distribution of values of the Euler function". In: *Acta Arith.* 21 (1972), pp. 329–345. DOI: 10.4064/aa-21-1-329-345 (cit. on p. 77).

[3] E. Bombieri. "A note on the large sieve". In: *Acta Arith.* 18 (1971), pp. 401–404 (cit. on p. 35).

[4] E. Bombieri. "Le grand crible dans la théorie analytique des nombres". In: *Astérisque* 18 (1987/1974), 103pp (cit. on pp. 35, 38).

[5] E. Bombieri and H. Davenport. "On the large sieve method". In: *Abh. aus Zahlentheorie und Analysis zur Erinnerung an Edmund Landau* Deut. Verlag Wiss., Berlin (1968), pp. 11–22 (cit. on p. 68).

[6] H. Cohen and F. Dress. "Estimations numériques du reste de la fonction sommatoire relative aux entiers sans facteur carré". In: *Prépublications mathématiques d'Orsay : Colloque de théorie analytique des nombres, Marseille* (1988), pp. 73–76 (cit. on p. 76).

[7] I. V. Čulanovskiĭ. "Certain estimates connected with a new method of Selberg in elementary number theory". In: *Doklady Akad. Nauk SSSR (N.S.)* 63 (1948), pp. 491–494 (cit. on p. 4).

[8] R. E. Dressler. "A density which counts multiplicity". In: *Pacific J. Math.* 34 (1970), pp. 371–378. ISSN: 0030-8730,1945-5844. URL: http://projecteuclid.org.ezproxy.math.cnrs.fr/euclid.pjm/1102976431 (cit. on p. 77).

[9] P. Erdös. "On a new method in elementary number theory which leads to an elementary proof of the prime number theorem". In: *Proc. Natl. Acad. Sci. USA* 35 (1949), pp. 374–384 (cit. on p. 44).

[10] P. Erdös. "Some remarks on Euler's ϕ-function and some related problems". In: *Bull. Amer. Math. Soc.* 51 (1945), pp. 540–544. ISSN: 0002-9904. DOI: 10.1090/S0002-9904-1945-08390-6 (cit. on p. 77).

[11] T. Estermann. *Introduction to modern Prime Number Theory.* Cambridge Univ. Press, 1952 (cit. on p. 79).

[12] P. Gallagher. "The large sieve". In: *Mathematika* 14 (1967), pp. 14–20 (cit. on p. 45).

[13] D. A. Goldston. "The major arcs approximation of an exponential sum over primes". In: *Acta Arith.* 92.2 (2000), pp. 169–179 (cit. on p. 45).

[14] S. Graham and J. Vaaler. "A class of extremal functions for the Fourier transform". In: *Trans. Amer. Math. Soc.* 265.1 (1981), pp. 283–302 (cit. on p. 25).

[15] B. Green and T. Tao. "Restriction theory of the Selberg sieve, with applications". In: *J. Théor. Nombres Bordx* 18.1 (2006), pp. 147–182 (cit. on pp. 65, 71, 72).

[16] B. Green. "Roth's theorem in the primes". In: *Ann. of Math.* 3.161 (2005), pp. 1609–1636 (cit. on p. 71).

[17] F. A. Gruenbaum. "Eigenvectors of a Toeplitz matrix: Discrete version of the prolate spheroidal wave functions". English. In: *SIAM J. Algebraic Discrete Methods* 2 (1981), pp. 136–141. ISSN: 0196-5212. DOI: 10.1137/0602017 (cit. on p. 29).

[18] H. Halberstam and H.-E. Richert. *Sieve methods*. London Mathematical Society Monographs, No. 4. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], London-New York, 1974, xiv+364 pp. (loose errata) (cit. on p. 37).

[19] H. Halberstam and H. Richert. "Almost-primes in short intervals". In: *[A] Recent progress in analytic number theory, Symp. Durham 1979* 1 (1981), pp. 69–101 (cit. on p. 62).

[20] M. Huxley. "Irregularity in sifted sequences". In: *J. Number Theory* 4 (1972), pp. 437–454 (cit. on pp. 10, 23).

[21] N. I. Klimov. "Almost prime numbers". In: *Uspehi Mat. Nauk* 16.3(99) (1961), pp. 181–188 (cit. on p. 43).

[22] Y. Linnik. "The dispersion method in binary additive problems". In: *Leningrad* (1961), 208pp (cit. on p. 4).

[23] Y. Linnik. "The large sieve". In: *Doklady Akad. Nauk SSSR* 30 (1941), pp. 292–294 (cit. on p. 35).

[24] J. van Lint and H. Richert. "On primes in arithmetic progressions". In: *Acta Arith.* 11 (1965), pp. 209–216 (cit. on p. 38).

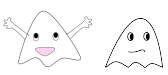
[25] M. G. Madritsch and R. F. Tichy. "Multidimensional van der Corput sets and small fractional parts of polynomials". In: *Mathematika* 65.2 (2019), pp. 400–435. DOI: 10.1112/s0025579318000529 (cit. on p. 81).

[26] H. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*. Vol. 84. CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1994, pp. xiv+220. ISBN: 0-8218-0737-4 (cit. on p. 65).

[27] H. Montgomery. "The analytic principle of the large sieve". In: *Bull. Amer. Math. Soc.* 84.4 (1978), pp. 547–567 (cit. on pp. 35, 66).

[28] H. Montgomery. "Topics in Multiplicative Number Theory". In: *Lecture Notes in Mathematics (Berlin)* 227 (1971), 178pp (cit. on p. 35).

[29] H. Montgomery and R. Vaughan. "The large sieve". In: *Mathematika* 20.2 (1973), pp. 119–133 (cit. on pp. 4, 36, 37).



[30] Y. Motohashi. "A note on Siegel's zeros". In: *Proc. Jap. Acad., Ser. A* 55 (1979), pp. 190–192 (cit. on pp. 5, 43).

[31] OEIS Foundation Inc. *The On-Line Encyclopedia of Integer Sequence*. http://oeis.org/. 2019 (cit. on p. 40).

[32] *PARI/GP, version* 2.7.0. http://pari.math.u-bordeaux.fr/. The PARI Group. Bordeaux, 2014 (cit. on p. 63).

[33] K. Ramachandra, A. Sankaranarayanan, and K. Srinivas. "Ramanujan's lattice point problem, prime number theory and other remarks". In: *Hardy and Ramanujan journal* 19 (1996) (cit. on p. 43).

[34] D. S. Ramana and O. Ramaré. *Arithmetical aspects of the large sieve inequality – II*. Texts and Readings in Mathematics. Hindustan Book Agency and Springer, 2025, xiv + 240pp (cit. on pp. 37, 46).

[35] O. Ramaré. "An explicit result of the sum of seven cubes". In: *Manuscripta Math.* 124.1 (2007), pp. 59–75 (cit. on p. 15).

[36] O. Ramaré. *Arithmetical aspects of the large sieve inequality*. Vol. 1. Harish-Chandra Research Institute Lecture Notes. With the collaboration of D. S. Ramana. New Delhi: Hindustan Book Agency, 2009, pp. x+201. ISBN: 978-81-85931-90-6 (cit. on pp. 37, 58–60).

[37] O. Ramaré and I. Ruzsa. "Additive properties of dense subsets of sifted sequences". In: *J. Théorie N. Bordeaux* 13 (2001), pp. 559–581 (cit. on pp. 59, 60, 68).

[38] O. Ramaré and J.-C. Schlage-Puchta. "Improving on the Brun-Titchmarsh theorem". In: *Acta Arith.* 131.4 (2008), pp. 351–366 (cit. on p. 4).

[39] O. Ramaré. "Explicit average orders: News and Problems". In: *Number theory week 2017*. Vol. 118. Banach Center Publ. Polish Acad. Sci. Inst. Math., Warsaw, 2019, pp. 153–176 (cit. on p. 38).

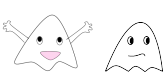
[40] O. Ramaré. "On Snirel'man's constant". In: *Ann. Scu. Norm. Pisa* 22 (1995), pp. 645–706 (cit. on p. 59).

[41] O. Ramaré. "The number of rationals determined by large sets of sifted integers". In: *Proc. Indian Acad. Sci. Math. Sci.* 132.2 (2022), Paper No. 62, 13. DOI: 10.1007/s12044-022-00698-z (cit. on p. 15).

[42] M. T. Rassias. *Goldbach's problem*. Selected topics, With a foreword by Jörg Brüdern and Preda Mihăilescu. Springer, Cham, 2017, pp. xv+122. ISBN: 978-3-319-57912-2; 978-3-319-57914-6. DOI: 10.1007/978-3-319-57914-6 (cit. on p. 79).

[43] K. Rogers. "The Schnirelmann density of the squarefree integers". In: *Proc. Amer. Math. Soc.* 15 (1964), pp. 515–516. DOI: 10.2307/2034736 (cit. on p. 76).

[44] J. Rosser and L. Schoenfeld. "Approximate formulas for some functions of prime numbers". In: *Illinois J. Math.* 6 (1962), pp. 64–94 (cit. on p. 62).



[45] W. Stein et al. *Sage Mathematics Software (Version 9.5)*. `http://www.sagemath.org`. The Sage Development Team. 2024 (cit. on pp. 32, 78).

[46] T. Sanders. "On Roth's theorem on progressions". In: *Ann. of Math. (2)* 174.1 (2011), pp. 619–636. DOI: 10.4007/annals.2011.174.1.20 (cit. on pp. 6, 71).

[47] A. Selberg. "An elementary proof of Dirichlet's theorem about primes in an arithmetic progression". In: *Ann. Math.* 50.2 (1949), pp. 297–304 (cit. on p. 44).

[48] A. Selberg. "An elementary proof of the prime-number theorem". In: *Ann. Math.* 50.2 (1949), pp. 305–313 (cit. on p. 44).

[49] H. Siebert. "Sieve methods and Siegel's zeros". In: *Studies in pure mathematics*. Birkhäuser, Basel, 1983, pp. 659–668. ISBN: 3-7643-1288-2 (cit. on p. 43).

[50] D. Slepian. "Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case". English. In: *Bell Syst. Tech. J.* 57 (1978), pp. 1371–1430. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1978.tb02104.x (cit. on p. 29).

[51] J. M. Steele. *The Cauchy-Schwarz master class*. AMS/MAA Problem Books Series. An introduction to the art of mathematical inequalities. Mathematical Association of America, Washington, DC; Cambridge University Press, Cambridge, 2004, pp. x+306. ISBN: 0-521-83775-8; 0-521-54677-X. DOI: 10.1017/CBO9780511817106 (cit. on p. 11).

[52] T. Tao and V. Vu. "John-type theorems for generalized arithmetic progressions and iterated sumsets". In: *Adv. Math.* 219.2 (2008), pp. 428–449. DOI: 10.1016/j.aim.2008.05.002 (cit. on p. 71).

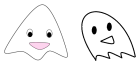
[53] E. Titchmarsh. "A divisor problem." English. In: *Rendiconti Palermo* 54 (1930), pp. 414–429 (cit. on p. 4).

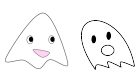
[54] J. Vaaler. "Some Extremal Functions in Fourier Analysis". In: *Bull. A. M. S.* 12 (1985), pp. 183–216 (cit. on pp. 25, 26, 54, 55, 81).

[55] R. Vaughan. *The Hardy-Littlewood method*. Vol. 80. Cambridge Tracts in Mathematics. Cambridge: Cambridge University Press, 1981, pp. xi+172. ISBN: 0-521-23439-5 (cit. on p. 79).

[56] T. Yamada. *Explicit improvements of the Brun-Titchmarsh theorem for arbitrary intervals*. 2023. arXiv: 2312.16090 [math.NT]. URL: https://arxiv.org/abs/2312.16090 (cit. on p. 4).

Index





We present in this series of lectures how one may sieve from the large sieve inequality. We shall rapidly establish this inequality and derive Montgomery's bound, and in particular the Brun-Titchmarsh Theorem. The factor 2 that seems to be a loss in this inequality will be shown to be linked with possible Siegel zeros. We will proceed by proving (a variant of) this theorem by starting directly from the Parseval identity on $\mathbb{R}/\mathbb{Z}$ and follow a path that seems potentially optimal but that will still lead to the same loss of a factor 2.

Thefore large values of the Fourier polynomial on the primes, say in some interval, probably do not behave as expected and other phases may intervene. In order to investigate this possibility, we prove a sharp large sieve inequality for this trigonometric polynomial when evaluated on a small subset by using an enveloping sieve. On calling loosely a *cusp* a point where our Fourier polynomial takes a large value, a consequence of our inequality is that many rational points are indeed cusps and that any other cusp is accompanied by a large stream of rational translates that are also cusps.