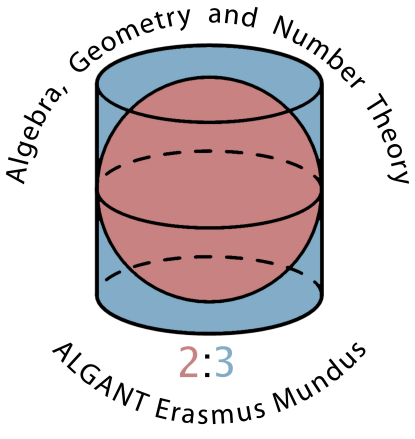
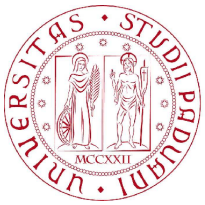


Gaussian primes after Fouvry and Iwaniec

Candidate
Dante Bonolis

Advisor
Etienne Fouvry



Academic year 2013/2014

Chapter 1

Introduction

From Fermat's work we know that any prime $p \equiv 1 \mod 4$, is of the form:

$$p = a^2 + b^2. \tag{1.1}$$

Fouvry and Iwaniec proved that there are infinitely many primes of the form $p = a^2 + (p')^2$ where p' is a prime number. Moreover they found an asymptotic distribution for prime numbers of this kind. In particular they proved the following theorem:

Theorem 1.0.1. *Let λ_l be complex numbers with $|\lambda_l| \leq 1$ then:*

$$\sum_{l^2+m^2 \leq x} \Lambda(l^2+m^2)\lambda_l = \sum_{l^2+m^2 \leq x} \psi(l)\lambda_l + O(x(\log x)^{-A}) \tag{1.2}$$

where Λ is the von Mangoldt function and:

$$\psi(l) = \prod_{(p,l)=1} \left(1 - \frac{\chi(p)}{p-1}\right) \tag{1.3}$$

where χ is the non trivial character modulus 4, A is any positive number and the implied constant depends only on A.

As an application of this theorem, if we choose $\lambda_l = \frac{\Lambda(l)}{\log x}$, we infer:

$$\sum_{l^2+m^2 \leq x} \Lambda(l^2+m^2)\Lambda(l) = H \sum_{l \leq \sqrt{x}} \gamma(l)\Lambda(l) + O(x(\log x)^{-A}) \tag{1.4}$$

where:

$$H = \prod_{p} \left(1 - \frac{\chi(p)}{p-1}\right),\ \gamma_l = \prod_{p|l} \left(1 - \frac{\chi(p)}{p-1}\right)^{-1} \sqrt{x - l^2}. \tag{1.5}$$

By summation by parts we deduce:

$$\sum_{l^2+m^2 \leq x} \Lambda(l^2+m^2)\Lambda(l) = \frac{\pi}{4} Hx + O(x(logx)^{-A}). \tag{1.6}$$

Chapter 2

Notation

We fix the notation that will be used from now on. Let $d > 1$ and k, l be integers. We define:

$$\rho_{k,l}(d) := \sum_{\nu^2+l^2 \equiv 0 \mod d} e\left(\frac{\nu k}{d}\right) \tag{2.1}$$

this function plays a fundamental role in the proof of the Main Theorem 1.0.1. so we present some properties. First of all notice that if $(d, l) = 1$ we can rewrite (2.1) as:

$$\rho_{k,l}(d) := \sum_{(\nu/l)^2+1 \equiv 0 \mod d} e\left(\frac{(\nu/l)kl}{d}\right) = \rho_{kl,1}(d) \tag{2.2}$$

If l and d are not coprime, let $(l, d) = ab^2$ with a square-free. We can write $d = ab^2d_1$ and $l = abl_1$ with $(al_1, d_1) = 1$. Suppose we have a solution to the equation $\nu^2+l^2 \equiv 0 \mod d$; if $m \equiv \nu \mod d$ then $ab \mid m$ which means that we can replace ν by $ab\nu_1$:

$$ab^2d_1 \mid (ab\nu_1)^2 + (abl_1)^2 \tag{2.3}$$

and so one obtains that $\nu_1^2 + l_1^2 \equiv 0 \mod d_1$, thus (2.1) becomes:

$$\rho_{k,l}(d) := \sum_{\substack{ab\nu_1 \mod ab^2d_1 \\ \nu_1^2+1_1^2 \equiv 0 \mod d_1}} e\left(\frac{ab\nu_1 k}{ab^2d_1}\right) = \sum_{\substack{\nu_1 \mod bd_1 \\ \nu_1^2+l_1^2 \equiv 0 \mod d_1}} e\left(\frac{\nu_1 k}{bd_1}\right). \tag{2.4}$$

Now we have to distinguish two cases:

i) $k \not\equiv 0 \mod b$. Then (2.4) vanishes by the orthogonality of the additive characters,

ii) $k = bk_1$, then (2.4) becomes:

$$\rho_{k,l}(d) = \sum_{\substack{\nu_1 \mod bd_1 \\ \nu_1^2+l_1^2 \equiv 0 \mod d_1}} e\left(\frac{\nu_1 k_1}{d_1}\right) = b\rho_{k_1,l_1}(d_1) = b\rho_{k_1l_1,1}(d_1) \tag{2.5}$$

We will denote $\rho_{0,l}(d) = \rho_l(d)$ and simply $\rho(d)$ when $k = 0$ and $l = 1$. Observe that by the previous discussion $\rho_l(d) = (r(d), l)\rho_l(d/(d, l^2))$ where $r(d)$ is the larger square which divides d. It is interesting to study the sums of the type:

$$\sum_{n \leq y} \frac{\mu(n)\rho_l(nc)}{n} \tag{2.6}$$

with $l, c \geq 1$. To semplify we will study only the case for $l = 1$, then by the previous remark we deduce from this simpler case the general one. The first step in order to obtain a bound for (2.6) is to study the sums:

$$\sum_{p \leq y} \frac{\rho(cp)\log(p)}{\rho(c)p} \tag{2.7}$$

Using the Siegel-Walfitz Theorem we get the following proposition:

Proposition 2.0.2. *For any $c \geq 1$ there exists a costant C_c such that:*

$$\sum_{p \leq y} \frac{\rho(cp)\log(p)}{\rho(c)p} = C_c + log(y) + O((\log y)^{-A}) \tag{2.8}$$

for any $A > 0$.

Proof. If $c = 1$ then we know that $\rho(p) = 2$ if $p \equiv 1 \mod 4$ and $\rho(p) = 0$ otherwise, and the result follows by the Siegel-Walfitz Theorem. If $c > 1$ it is enough to take:

$$C_c := C_1 + \sum_{p|c} \frac{\rho(pc)\log(p)}{\rho(c)p} - \frac{\rho(p)\log(p)}{p}. \tag{2.9}$$

In fact we deduce

$$\begin{aligned} C_c + log(y) + O((\log y)^{-A}) &= C_1 + \sum_{p|c} \frac{\rho(pc)\log(p)}{\rho(c)p} - \frac{\rho(p)\log(p)}{p} + \\ &+ log(y) + O((\log y)^{-A}) = \\ &= \sum_{p \leq y} \frac{\rho(p)\log(p)}{p} + O((\log y)^{-A}) + \\ &+ \sum_{\substack{p|c \\ p \leq y}} \frac{\rho(pc)\log(p)}{\rho(c)p} - \frac{\rho(p)\log(p)}{p} = \\ &= \sum_{p \leq y} \frac{\rho(cp)\log(p)}{\rho(c)p} + O((\log y)^{-A}) \end{aligned} \tag{2.10}$$

□

Let g be a multiplicative function such that:[1]

i) $0 \leq g(p) < 1$

ii) there exists a constant C such that for any $A > 0$, $\sum_{p \leq y} g(p)\log(p) = C + \log(y) + O((\log y)^{-A})$

then we have the bound $\sum_{n \leq y} \mu(n)g(b) \ll (\log y)^{-A}$. By this and the previous proposition we can deduce that:

$$\sum_{n \leq y} \frac{\mu(n)\rho(nc)}{n} \ll \rho(c)(\log y)^{-A} \tag{2.11}$$

We will introduce the classical notation of sieve theory: let $\{\lambda_l\}_{l \in \mathbb{N}}$ be any sequence of complex numbers and define:

$$a_n := \sum_{l^2+m^2=n} \lambda_l. \tag{2.12}$$

For $x > 1$ and for any $d \geq 1$ define:

$$A_d(x) := \sum_{\substack{n \leq x \\ n \equiv 0 \mod d}} a_n = \sum_{\substack{n \leq x \\ n \equiv 0 \mod d}} \sum_{l^2+m^2=n} \lambda_l = \sum_{l \leq \sqrt{x}} \lambda_l \sum_{\substack{l^2+m^2 \equiv 0 \mod d \\ l^2+m^2 \leq x}} 1. \tag{2.13}$$

It is reasonable to expect that the main term of $A_d(x)$ is:

$$M_d(x) = \frac{1}{d} \sum_{l^2+m^2 \leq x} \lambda_l \rho_l(d) \tag{2.14}$$

where $\rho_l(d)$ is like in (2.1). Then we define the error term $R_d(x) := A_d(x) - M_d(x)$. To finish one defines for every $D > 1$, $\quad R(x, D) := \sum_{d<D} |R_d(x)|$

[1] For a proof of this fact one can see [2] for the related argument

Chapter 3

The Vaughan's identity

We are interested in a sum of the type:

$$P(x) := \sum_{n \leq x} a_n \Lambda(n) \tag{3.1}$$

where a_n is any sequence of complex numbers. Using the following identity:

$$\Lambda(n) = -\sum_{d|n} \mu(d) \log d \tag{3.2}$$

we can rewrite $P(x)$ as :

$$P(x) = \sum_{d \leq x} \gamma_d A_d(x) \tag{3.3}$$

where $\gamma_d := -\mu(d)log(d)$. Suppose that we know a reasonable approximation of A_d, i.e. $A_d(x) = Xg(d)+r_d(x)$ where g is a multiplicative function such that $0 \leq g(p) < 1$ and $g(p) \ll 1/p$ for every prime p. Suppose moreover that

$$\sum_{p<y} g(p) = C + \log\log y + O((\log y)^{-A}) \tag{3.4}$$

for any $A > 0$. Then we can deduce that:

$$P(x) = XH(1 + O((\log x)^{-1})) + \sum_{d \leq x} \gamma_d r_d(x) \tag{3.5}$$

where $H = -\sum_d \mu(d)g(d)\log(d)$. In general, we cannot control the size of second term in (3.5) which could be larger than the main term. This problem is the well known parity problem. To solve this problem, Fouvry and Iwaniec used the Vaughan's identity. With this combinatorial device one exploits a bilinear form which can be well estimated thanks to the oscillation of the Moebius function. Thus fixing $y, z < x$, by Vaughan's identity we can write:

$$P(x) = P(z) + A(x, y; z) + B(x, y; z) \tag{3.6}$$

where[1]:

$$A(x,y;z) := \sum_{b\leq y}\Big\{A_b'(x) - A_b(x)\log b - \sum_{c\leq z}\Lambda(c)A_{bc}(x)\Big\} \tag{3.7}$$

and

$$B(x,y;z) := \sum_{bd<x, b>y}\mu(b)\Big(\sum_{c|d, c>z}\Lambda(c)\Big)a_{bd} \tag{3.8}$$

introducing the notation which we have presented before, with some technical passage we can infer[2]:

Proposition 3.0.3. *Let λ_l be complex numbers with $|\lambda_l| \leq 1$, suppose $0 < \epsilon < 1/3$ such that $y, z > x^\epsilon$ and $yz < x^{1-\epsilon}$ then:*

$$\sum_{l^2+m^2\leq x}\Lambda(l^2+m^2)\lambda_l = \sum_{l^2+m^2\leq x}\psi(l)\lambda_l + B(x,y,z) + R(x,y,z) + O_A(x(\log x)^{-A}) \tag{3.9}$$

with any $A \geq 2$ and where $B(x,y,z)$ is like in (3.8) and:

$$R(x,y,z) := \sum_{b\leq y}\mu(b)\Big\{R_b(x)\log\frac{x}{b} - \int_1^x R_b(t)\frac{dt}{t} - \sum_{c\leq z}\Lambda(c)R_{bc}(x)\Big\}. \tag{3.10}$$

[1] $A_b'(x) := \sum_{n\leq x, n\equiv 0 \mod b} a_n \log n$

[2] note that to prove this proposition we need the bound in 2.11 from the previous chapter

Chapter 4

The remainder term

The first step to conclude the proof of the Main Theorem 1.0.1. is to obtain a good estimate of $R(x,y,z)$ which appears in (3.10). One can easily obtain by the definition of $R(x,y,z)$ that:

$$|R(x,y,z)| \leq R(x,yz)\log(x) + \int_1^x R(t,y)\frac{dt}{t}. \tag{4.1}$$

Thus if we want a good bound for $R(x,y,z)$ we need to find a good bound for $R(x,D)$. Let's first see what happens if we try to estimate trivially $R(x,D)$. By definition we know that

$$R(x,D) := \sum_{d\leq D} |R_d(x)| \tag{4.2}$$

and we can trivially estimate $|R_d| \leq 4x^{\frac{1}{2}} \sum_l |\lambda_l| \frac{\rho_l(d)}{d}$, for $d \leq \sqrt{x}$. It follows that:

$$R(x,D) \ll x^{\frac{1}{2}+\epsilon}\|\lambda\|_1 \tag{4.3}$$

where $\|\lambda\|_1 := \sum_l |\lambda_l|$. However this bound is too large, in fact if we specialize (4.3) to our case we obtain that $R(x,y,z) \ll x^{1+\epsilon}$ which exceeds the main term.
The right way to approach this problem is to prove the following:

Lemma 4.0.4. *Let $\lambda_l \in \mathbb{C}$ be any complex numbers for $1 \leq l \leq \sqrt{x}$. Then for any $1 \leq D \leq x$ we have that:*

$$R(x,D) \ll x^{\frac{1}{2}+\epsilon} D^{\frac{1}{4}} \|\lambda\| \tag{4.4}$$

for any $\epsilon > 0$. Where $\|\lambda\|^2 := \sum_l |\lambda_l|^2$ and implied constant depends only on ϵ.

Remark that with Lemma 4.0.4 we can choose a very large distribution D in the setting of Proposition 3.0.3 In fact if $D = x^{1-\varepsilon}$ with $\varepsilon > 0$, an application of the bound (4.4) gives us:

$$R(x,D) \ll x^{1-\varepsilon'} \tag{4.5}$$

using (4.5) in (4.1) one obtains

$$R(x,y,z) \ll x^{1-\varepsilon'} \tag{4.6}$$

as we wanted.

Before proving Lemma 4.0.4 we need another lemma which gives an estimate for sums of the form:

$$A_d(f) = \sum_{n \equiv 0 \mod d} a_n f(n) \tag{4.7}$$

where $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is a smooth function which satisfies the following conditions:

i) $f(u) = 0$ if $u \geq x$

ii) $f^{(j)}(u) \ll \Delta^j$ if $1 < u < x$

with $x^{-1} \leq \Delta \leq 1$. Notice that in the summation (4.7) we can relax the condition $n \leq x$ because the size of n is controlled by the compact support of the function f. Using Poisson's summation formulas for (4.7) we get:

$$A_d(f) = \frac{1}{d} \sum_k \sum_l \lambda_l \rho_{k,l}(d) F_l\left(\frac{k}{d}\right) \tag{4.8}$$

where $\rho_{k,l}(d)$ is like in (2.1) and

$$F_l(z) = \int_{-\infty}^{\infty} f(l^2 + t^2) e(-zt) dt \tag{4.9}$$

As before we define a main term and an error term as follows:

$$M_d(f) := \frac{1}{d} \sum_l \lambda_l \rho_{0,l}(d) F_l(0) \tag{4.10}$$

with $R_d(f) := A_d(f) - M_d(f)$. With this notation we can state:

Lemma 4.0.5. *Let $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ smooth which satisfies (i) and (ii) for some $x^{-1} \leq \Delta \leq 1$; let λ_l any complex numbers for $1 \leq l \leq \sqrt{x}$. Then for any $1 \leq D \leq x$ we have that:*

$$\sum_{d \leq D} |R_d(f)| \ll x^{\frac{5}{4}+\epsilon} \Delta D^{\frac{1}{2}} \|\lambda\| \tag{4.11}$$

for any $\epsilon > 0$, the implied costant depends only on ϵ.

Proof. (sketch) Let's consider for any $d \leq D$ the remainder term:

$$R_d(f) = \frac{2}{d} \sum_{k=1}^{\infty} \sum_l \lambda_l \rho_{k,l}(d) F_l\left(\frac{k}{d}\right) \tag{4.12}$$

The basic idea of the proof is that the series in (4.12) converges rapidly and so we can truncate it. In fact by integrating by parts $F_l(z)$ and using the condition on the support and on the derivates of f one infers that:

$$F_l\left(\frac{k}{d}\right) \ll k^{-2} D^{-1}, \text{ if } k \geq K = D\Delta x^{\frac{1}{2}+\epsilon}. \tag{4.13}$$

Then (4.12) becomes:

$$R_d(f) = \frac{2}{d}\sum_{k=1}^{K}\sum_{l}\lambda_l\rho_{k,l}(d)F_l\left(\frac{k}{d}\right) + O\left(\frac{\|\lambda\|_1}{d}\right). \tag{4.14}$$

Using the fact that $e(t) = \cos(2\pi t) + i\sin(2\pi t)$, that $f(t^2+l^2)$ is a even function and that the integration domain in the definition of $F_l(z)$ is symmetric, we infer, after changing the variable $t = v\sqrt{(}x)/d$:

$$F_l\left(\frac{k}{d}\right) = 2\pi\sqrt{x}k^{-1}\int_0^\infty f(l^2+v^2xk^{-2})\cos\left(2\pi\frac{v\sqrt{x}}{d}\right)dv. \tag{4.15}$$

If we insert this in (4.12) we obtain:

$$\sum_{d\leq D} d|R_d(f)| \leq 4\sqrt{x}\sum_{d\leq D}\int_0^K \sum_{\substack{v<k<K\\ 0<l<\sqrt{x}}}\left|\lambda_l f(l^2+v^2xk^{-2})\rho_{k,l}(d)dv\right| + O(D\|\lambda\|_1) \tag{4.16}$$

note that we can truncate the integral in K because for $v > K$ we have that $vxk^{-2} \geq x$ and this implies that the inner sum vanishes by the assumption on f. Assume for now the following lemma which we will prove in the next chapter:

Lemma 4.0.6. *Let $\alpha_{k,l}$ any complex number then:*

$$\sum_{d\leq D}\left|\sum_{0<l\leq L}\sum_{0<k\leq K}\alpha_{k,l}\rho_{k,l}(d)\right| \leq 150(\log 3D)^3 D^{\frac{1}{2}}(D+KL)^{\frac{1}{2}}\|\tilde{\alpha}\| \tag{4.17}$$

where $\|\tilde{\alpha}\|^2 := \sum_{k,l}|\alpha_{k,l}|^2\tau(kl)$

Applying this lemma for $\alpha_{k,l} = \lambda_l f(l^2+vxk^{-2})$ one obtains:

$$\sum_{d\leq D} d|R_d(f)| \ll \|\tilde{\lambda}\|(DKx)^{\frac{1}{2}}(D+K\sqrt{x})^{\frac{1}{2}}(\log x)^4 + O(D\|\lambda\|_1). \tag{4.18}$$

The result follows if we replace $K = D\Delta x^{\frac{1}{2}+\epsilon}$ and if we observe that $D\|\lambda\|_1 \ll D\|\lambda\|x^{\frac{1}{4}}$.

□

We are finally ready to prove Lemma 4.0.4:

Proof. Fix $y < x$ and consider $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ which satisfies the following condition:

i) $f(u) = 1$ if $0 < u \leq x - y$

ii) $f(u) = 0$ if $u \geq x$

iii) $f^{(j)}(u) \ll y^{-j}$ if $x - y < u < x$

then by triangle inequality we infer:

$$|R_d(x)| \leq |A_d(f) - M_d(f)| + |A_d(x) - A_d(f)| + |M_d(x) - M_d(f)| \tag{4.19}$$

For the term $\sum_{d\leq D} |A_d(f) - M_d(f)|$ we can use Lemma 4.0.5, so now the problem is to give a bound for the other two terms. The idea now is to use the definition of f: up to y it is the characteristic function of the interval $(0, x-y]$ so the size of both terms is given by the terms for $n > x-y$: in fact one can write $A_d(x) = \sum_{n\equiv 0 \mod d} a_n \phi_{[0,x]}$[1] and this means that up to $x-y$ the terms in $A_d(x)$ and $A_d(x)$ must be equal. Thus we have

$$\sum_{d\leq D} |A_d(x) - A_d(f)| \ll \|\lambda\|(y^{\frac{1}{2}} + yx^{-\frac{1}{4}})x^{\epsilon} \tag{4.20}$$

and the same bound holds for $\sum_{d\leq D} |M_d(x) - M_d(f)|$. To finish the proof it is enough to put everything together and to choose $y = D^{\frac{1}{4}}x^{\frac{3}{4}}$. In this way the contribution of $\sum_{d\leq D} |A_d(x) - A_d(f)|$ and $\sum_{d\leq D} |M_d(x) - M_d(f)|$ become negligible and by the previous Lemma 4.0.5:

$$R(x, D) \ll x^{\frac{1}{2}+\epsilon} D^{\frac{1}{4}} \|\lambda\| \tag{4.21}$$

as we wanted. □

[1] $\phi_{[0,x]}$ is the characteristic function of the interval $(0, x]$

Chapter 5

Well-spaced points

During the proof of Lemma 4.0.5 we assumed:

Lemma 5.0.7. *Let $\alpha_{k,l}$ any complex number then:*

$$\sum_{d\leq D}\left|\sum_{0<l\leq L}\sum_{0<k\leq K}\alpha_{k,l}\rho_{k,l}(d)\right|\leq 150(\log 3D)^3 D^{\frac{1}{2}}(D+KL)^{\frac{1}{2}}\|\tilde{\alpha}\| \tag{5.1}$$

where $\|\tilde{\alpha}\|^2 := \sum_{k,l}|\alpha_{k,l}|^2\tau(kl)$

The strategy to prove this lemma is based on the large sieve applied on the arithmetic points of the form $\nu/d \mod 1$ where the ν runs over the roots of:

$$\nu^2+1=0 \mod d. \tag{5.2}$$

Fix $D>1$, if we consider ν_1/d_1 and ν_2/d_2 with $d_1,d_2<D$ and $\nu_1/d_1\neq\nu_2/d_2$ we trivially obtain:

$$\left\|\frac{\nu_1}{d_1}-\frac{\nu_2}{d_2}\right\|\geq\frac{1}{D^2}. \tag{5.3}$$

An apllication of the large sieve gives us:

Proposition 5.0.8. *For any complex number α_n we have:*

$$\sum_{d\leq D}\sum_{\nu^2+1\equiv 0 \mod d}\left|\sum_{n\leq N}\alpha_n e\left(\frac{\nu n}{d}\right)\right|^2\leq(D^2+N)\|\alpha\|^2 \tag{5.4}$$

where $\|\alpha\|^2 := \sum_n|\alpha_n|^2$

For our application this is not in enough, in fact the factor D^2 is too big to obtain the bound we need.
The key point to lower the exponent of D is to observe that we can associate to each point of this kind a primitive representation of d as sum of two square, i.e.:

$$d=s^2+r^2,\ (s,r)=1,\ -s<r<s \tag{5.5}$$

In fact if we have a primitive representation of d we can find a root of (5.2) by considering the solution of the equation $s\nu = r \mod d$. By this arithmetic property we can rewrite $\frac{\nu}{d} \mod 1$ in a different way:

$$\frac{\nu}{d} = \frac{r}{sd} - \frac{\overline{r}}{s} \mod 1, \text{ where } \overline{r}r = 1 \mod s. \tag{5.6}$$

In this way we replace ν/d with the sum of $\overline{r}/s$ (which has denominator much smaller then ν/d) and r/sd (which is negligible). Consider the points ν/d for which the corresponding r has fixed sign and $8D < d \leq 9D$, where D is fixed and greater than one, then by (5.6) and using the fact that $2\sqrt{2}\sqrt{D} < s \leq 3\sqrt{D}$, we infer:

$$\left\| \frac{\nu_1}{d_1} - \frac{\nu_2}{d_2} \right\| \geq \frac{1}{36D} \tag{5.7}$$

then thanks to the large sieve we obtain the following:

Lemma 5.0.9. *For any complex number α_n we have :*

$$\sum_{8D<d\leq 9D} \sum_{\nu^2+1\equiv 0 \mod d} \left| \sum_{n\leq N} \alpha_n e\left(\frac{\nu n}{d}\right) \right|^2 \leq 72(D+N)\|\alpha\|^2 \tag{5.8}$$

Note that if we apply Lemma 5.0.9 for $N = 1$ and $\alpha_1 = 1$ we get:

$$\sum_{8D<d\leq 9D} \rho(d) \leq 72(D+1) \tag{5.9}$$

From the Lemma 5.0.9 and this remark we can infer the following:

Corollary 5.0.10. *For any complex number α_n :*

$$\sum_{d\leq D} \sum_{\nu^2+1\equiv 0 \mod d} \left| \sum_{n\leq N} \alpha_n e\left(\frac{\nu n}{d}\right) \right| \leq 150 D^{\frac{1}{2}} (D+N)^{\frac{1}{2}} \|\alpha\| \tag{5.10}$$

Proof. First of all observe that we can split the interval $[1, D]$ in :

$$[1, D] = [D, \frac{8}{9}D) \cup [\frac{8}{9}D, \frac{8^2}{9}D) \cup [\frac{8^2}{9}D, \frac{8^3}{9}D)... \tag{5.11}$$

so if we use the Cauchy-Schwarz inequality for each interval, the Lemma 5.0.9, and the remark above to count the points ν/d inside a quarter of a disk then the result follows. □

If we define $a_n := \sum_{kl=n} \alpha_{k,l}$ then we obtain:

$$\sum_{d\leq D} \sum_{\nu^2+1\equiv 0 \mod d} \left| \sum_{0<l\leq L} \sum_{0<k\leq K} \alpha_{k,l} e\left(\frac{\nu kl}{d}\right) \right| \leq 150 D^{\frac{1}{2}} (D+KL)^{\frac{1}{2}} \|\tilde{\alpha}\|. \tag{5.12}$$

We are ready to prove Lemma 5.0.7:

Proof. The first step is to see what happens if one adds the condition $(d,l)=1$ in (5.12), and this can be done by the Moebius inversion formula. In fact one obtains:

$$\sum_{d\le D}\sum_{\nu^2+1\equiv 0 \mod d}\left|\sum_{\substack{0<l\le L\\(d,l)=1}}\sum_{0<k\le K}\alpha_{k,l}e\left(\frac{\nu kl}{d}\right)\right| \le \sum_{b\le D}\rho(b)\sum_{d\le Db^{-1}}\sum_{\nu^2+1\equiv 0 \mod d}\left|\sum_{0<l\le Lb^{-1}}\sum_{0<k\le K}\alpha_{k,l}e\left(\frac{\nu kl}{d}\right)\right| . \tag{5.13}$$

Using now (5.12) one can infer that

$$\sum_{d\le D}\sum_{\nu^2+1\equiv 0 \mod d}\left|\sum_{\substack{0<l\le L\\(d,l)=1}}\sum_{0<k\le K}\alpha_{k,l}e\left(\frac{\nu kl}{d}\right)\right| \le 150(\log 3D)D^{\frac{1}{2}}(D+KL)^{\frac{1}{2}}\|\tilde{\alpha}\| \tag{5.14}$$

where the factor $\log 3D$ comes out by the bound $\sum_{b\le D}\rho(b)b^{-1}$ which is obtained from (5.9) by partial summation. Thanks to the property of $\rho_{k,l}(d)$ which we have seen in the first chapter ((2.1),(2.4)) we can deduce that:

$$\sum_{d\le D}\left|\sum_{0<l\le L}\sum_{0<k\le K}\alpha_{k,l}\rho_{k,l}(d)\right| \le \sum_{ab^2d\le D} b \sum_{\nu^2+1\equiv 0 \mod d}\left|\sum_{\substack{0<l\le L(ab)^{-1}\\(l,d)=1}}\sum_{0<k\le Kb^{-1}}\alpha_{bk,abl}e\left(\frac{\nu kl}{d}\right)\right|. \tag{5.15}$$

Applying (5.14) we infer:

$$\sum_{ab^2d\le D} b \sum_{\nu^2+1\equiv 0 \mod d}\left|\sum_{\substack{0<l\le L(ab)^{-1}\\(l,d)=1}}\sum_{0<k\le Kb^{-1}}\alpha_{bk,abl}e\left(\frac{\nu kl}{d}\right)\right|. \le 150\left(\sum_{ab^2\le D}\frac{b}{ab^2}\right)(\log 3D)D^{\frac{1}{2}}(D+KL)^{\frac{1}{2}}\|\tilde{\alpha}\|. \tag{5.16}$$

To conclude the proof it's enough to observe that $\sum_{ab^2\le D}(ab)^{-1} < (\log 3D)^2$

□

Chapter 6

The bilinear form

The final step to finish the proof of the Main Theorem 1.0.1. is the estimation of the bilinear form:

$$B(x,y;z) := \sum_{bd<x, b>y} \mu(b)\Big(\sum_{c|d, c>z} \Lambda(c)\Big)a_{bd}. \tag{6.1}$$

We want to get:

$$B(x,y,z) \ll x\Delta(\log x)^5 \tag{6.2}$$

where $\Delta = (\log x)^{-A}$ for any $A > 5$. The crucial part of this argument is the oscillation of the Moebius function, which we will explain at the end of this chapter after having presented the strategy behind this last step.
Let's start by using the triangle inequality and the fact that $\sum_{c|d} \Lambda(c) = \log d$. We obtain:

$$|B(x,y;z)| \leq (\log x) \sum_{d>z} \Big| \sum_{y<b\leq x/d} \mu(b)a_{bd}\Big| \tag{6.3}$$

in order to separate the variables b and d, we break the sum into shorter sum

$$B(M,N) := \sum_{M<m\leq 2M} \Big| \sum_{N<n\leq N'} \mu(n)a_{mn}\Big| \tag{6.4}$$

where $N' := e^{\Delta}N$. Using these sums with $M = 2^j z$ and $N = e^{k\Delta}y$ we get:

$$B(x,y;z) \leq (\log x) \sum_{\substack{x\Delta<MN<x \\ M\geq y, N\geq z}} B(M,N) + O(x\Delta(\log x)^2) \tag{6.5}$$

where the error term comes out by the contribution the a_{bd}'s such that $bd \geq 2\Delta x$ and $e^{-2\Delta}x < bd \leq x$. Now observe that we have, by the condition on M and N, at most $2\Delta^{-1}(\log x)^2$ shorter sums $B(M,N)$, therefore to obtain (6.2) it is enough to prove that:

$$B(M,N) \ll \Delta^2 x(\log x)^2. \tag{6.6}$$

By a series of technical devices, which we skip, one can show that to find the bound (6.2) it is equivalent to the bound, for any $j > 0$:

$$C_{c,r}(M,N) \ll \frac{MN}{(\log N)^j} \tag{6.7}$$

for every $r < \Delta^{-2}$, $c < \Delta^{-4}$, $M \geq \Delta^4 z$, $N > \Delta^3 y$, and $x\Delta^5 < MN < x$, where[1]:

$$C_{c,r}(M,N) := \sideset{}{^*}\sum_{M<|w|^2\leq 2M} \Big| \sum_{N<|z|^2\leq N'} \mu(r|z|^2)\lambda(cz\cdot w)\Big|. \tag{6.8}$$

We present a general result for bilinear forms on gaussian integers. Let's first define the setting where we are going to work. Let $\alpha, \beta : \mathbb{Z} \longrightarrow \mathbb{C}$ two arithmetic functions on gaussian integers and suppose that:

i) the support of α is contained in $\{|z| \leq A\}$

ii) the support of β is contained in the anulus $\{B \leq w| \leq 2B\}$

with $1 < B \leq A$, then let $\lambda : \mathbb{Z} \longrightarrow \mathbb{C}$ and consider:

$$C(\alpha,\beta,\lambda) := \sideset{}{^*}\sum_{w}\sum_{z} \alpha(z)\beta(w)\lambda(z\cdot w) \tag{6.9}$$

By the assumption on α and β one can assume that the support of λ is contained in the interval $[-2AB, 2AB]$. In this context one can prove the following proposition

Proposition 6.0.11. *Let α, β and λ like before, let $\Lambda > 1$ such that:*

$$A^{\epsilon}\Lambda^7 < B < A\Lambda^{-1} \tag{6.10}$$

for some $\epsilon > 0$. Then we have

$$C(\alpha,\beta,\lambda) \ll \|\alpha\|\|\beta\|\|\lambda\|(AB)^{\frac{1}{2}}\Lambda^{-\frac{1}{2}} + \|\beta\|\|\lambda\|(AB)^{\frac{1}{2}}\Lambda^2\Big(\sum_{d<\Lambda^6} d^2 D_d(\alpha)\Big)^{\frac{1}{2}} \tag{6.11}$$

where

$$D_d(\alpha) := 2\pi A^{-2} \sum_{z_1\equiv z_2 \mod d} \alpha(z_1)\overline{\alpha}(z_2)\exp(-2\pi|z_1 - z_2|A^{-1}). \tag{6.12}$$

[1]The simbol $\sum^*$ means that we restrict the sum on the primitive elements

This result for bilinear form of the form (6.9) is not enough for our aim, in fact if one tries to trivially estimate $D_d(\alpha)$, one obtains by Cauchy-Schwarz:

$$\begin{aligned} D_d(\alpha) &\leq 2\pi A^{-2} \sum_{\delta \mod d} \Big(\sum_{z \equiv \delta \mod d} 1 \Big)^{\frac{1}{2}} \Big(\sum_{\substack{z_1 \equiv \delta \\ z_2 \equiv \delta}} |\alpha(z_1)|^2 |\alpha(z_2)|^2 \Big)^{\frac{1}{2}} \\ &\ll A^{-2} \sum_{\delta \mod d} \frac{A^2}{d^2} \Big(\sum_{z \equiv \delta} |\alpha(z)|^2 \Big) \\ &\ll \frac{\|\alpha\|^2}{d^2} \end{aligned} \tag{6.13}$$

But now if we insert this bound in (6.11) we get:

$$C(\alpha, \beta, \lambda) \ll \|\alpha\| \|\beta\| \|\lambda\| (AB)^{\frac{1}{2}} \Lambda^8 \tag{6.14}$$

Going back to our discussion, consider for example $C_{1,1}(M, N)$ to see why (6.14) it is too large. One can prove that if we choose $y = x^\theta$ and $z = x^\vartheta$ subject to $1/2 < \theta < 1$ and $0 < \vartheta < 1 - \theta$ the hypothesis of the Proposition 6.0.11 are satisfied for $\Lambda = (\log x)^j$ for any $j > 0$. But now if we apply (6.14) we have:

$$C_{1,1}(M, N) \ll x (\log x)^{8j} \tag{6.15}$$

which exceeds the main term in (3.10). But in our case $\alpha(z) = \mu(r|z|)$, and we have not used yet the oscillation of the Moebius function: this property will be crucial. In fact one can prove an equivalent version of the Theorem of Siegel-Walfitz for the Gaussian integers:

Theorem 6.0.12. *Let $\chi \mod q$ any Dirichlet's character on $\mathbb{Z}[i]$ with $q \in \mathbb{Z}$ then:*

$$\sum_{|z| \leq x} \mu(z) \chi(z) \ll_j x (\log(x))^{-j} \tag{6.16}$$

if one applies Theorem 6.0.12 adding the condition $(z, \overline{z}) \mid i - 1$ and $(z, r) = 1$ with $r < \Delta^{-2}$, one find that:

$$D_d(\mu(r|\cdot|)) \ll \frac{\|\mu(r|\cdot|)\|^2}{(\log N)^{-j}}. \tag{6.17}$$

Notice that (6.17) is not trivial if $d < (\log N)^j$. By this and the choice of $\Lambda = (\log N)^{j/23}$, the second term in (6.11) for $C(\mu(r|\cdot|, \beta, \lambda))^2$ becomes negligible and so one gets:

$$C_{c,r}(M, N) \leq \frac{MN}{(\log N)^j} \tag{6.18}$$

[2] β in this case the characteristic function of primitive Gaussian integers in the annulus $\{M < |w|^2 \leq 2M\}$ and $\lambda(z \cdot w) := \lambda_{z \cdot w}$

for $r < \Delta^{-2}$, $c < \Delta^{-4}$, $M \geq \Delta^4 z$, $N > \Delta^3 y$, and $x\Delta^5 < MN < x$ which is the bound we needed.
As we said before, in the argument above the oscillation of the Moebius function plays a crucial role. The technique used in [1] was the first prototype of a sieve which was developed after in [2]. We have already discussed at the beginning of chapter three the parity problem which apperas in the classical setting of sieve theory. In [2] the parity problem is solved for sequence $\{a_n\}$ of non-negative real numbers by adding in the classical setting the following assumption:

B) 1. *$\forall x > 1$, $\exists x^{2/3} < D < x$,$\exists 0 < \epsilon < \frac{1}{2}$, $\exists 2 < \delta(x), \Delta(x) < D^{1/2-\epsilon}$ such that $\forall \sqrt{D}\Delta^{-1} < N < \delta^{-1}\sqrt{x}$ and $\forall 1 \leq C \leq xD^{-1}$ we have:*

$$\sum_m \left| \sum_{\substack{N<n\leq 2N \\ mn\leq x}} \gamma(C,n)\mu(mn)a_{mn} \right| \ll X(\log x)^{-2^{22}}. \tag{6.19}$$

where:

$$\gamma(C,n) := \sum_{d|n, d\leq C} \mu(d) \tag{6.20}$$

This suggests us a reflection on the classical Selberg's example of the parity problem. Fix $C = 1$[3], let's define:

$$a_n := \frac{1}{2}(1 + \lambda(n)) \tag{6.21}$$

where here λ is Liouville's function. In practice a_n is the charateristc function of the numbers with a even number of prime factors. This means that the Moebius function which appears in (6.19) cannot vary so B) does not hold.

[3] this implies $\gamma(1,n) = 1$ for every n

Acknowledgements

First of all I want to thank my advisor, Etienne Fouvry, for giving me the opportunity to study this subject, for the help he gave me during these months and for his great availability.
I am also grateful to my family and my friends for their support in these years.

Bibliography

[1] E. Fouvry and H. Iwaniec, *Gaussian primes*, Acta Arithmetica (1997) 79 249-287

[2] J. Friedlander and H. Iwaniec, *Asymptotic sieve for primes*, Annales of Mathematics (1998) 148 1041-1065

[3] J. Friedlander and H. Iwaniec, *Opera De Cribro*, American Mathematical Soc. (2010)