Modular forms and equidistribution

Ramon M. Nunes

July 2, 2012

1 Introduction

The goal of these notes is to give a proof of the main theorem in [DFI], which is the following:

Theorem 1.1. We start with a polynomial with integer coefficients of degree 2, $P(x) = ax^2 + bx + c$ of negative discriminant ($b^2 - 4ac < 0$), more precisely $P(x) = x^2 + D$ where $D > 0$. For each p prime, we consider the equation $P(\nu) \equiv 0 \pmod{p}$. For each p and each solution $\nu_p \pmod{p}$, we take the well defined number $\frac{\nu_p}{p} \pmod 1$. Finally, we consider the set $E = \{\frac{\nu_p}{p}\} \subset \mathbb{R}/\mathbb{Z}$ where p ranges through all the prime numbers and ν_p over all the solutions $\pmod{p}$. We claim that the set E is equidistributed in $\mathbb{R}/\mathbb{Z}$. Alternatively, let's say $E_x = \{\frac{\nu_p}{p} \mid p \leq x\}$. Then, $\frac{|E_x \cap [a_1, a_2]|}{|E_x|}$ converges to $a_2 - a_1$ as $x \to \infty$ for every $0 \leq a_1 < a_2 \leq 1$.

We start with few lemmas:

Lemma 1.2. $|E_x| \sim \frac{x}{\log x}$

Proof. If we forget about the primes dividing D and $p = 2$ we have no problem because they are finitely many. Now we notice that we have solutions for $x^2 + D \equiv 0 \pmod{p}$ if and only if $(\frac{-D}{p}) = 1$ and in that case we have two distinct solutions, so, by the prime number theorem all we have to prove is that the number of $p \leq x$ satisfying this equation is roughly half of the primes $p \leq x$. This is the case because $p \mapsto (\frac{-D}{p})$ can be extended to a character $(\mathbb{Z}/4D\mathbb{Z})^* \to \mathbb{C}^*$ whose image is $\{1, -1\}$ so that the kernel is equal to the half of $|(\mathbb{Z}/4D\mathbb{Z})^*|$. In other words, roughly speaking half of the residues $\pmod{4D}$ have $(\frac{-D}{p}) = 1$ and the other half has $(\frac{-D}{p}) = -1$. Because of the generalized prime number theorem, the number of primes satisfying each of these equations is roughly $\frac{1}{2}(x/\log x)$. So we are done. □

The departure point of this proof is Weyl's criterion, which can be stated as:

Lemma 1.3 (Weyl)**.** Consider $\{x_n\}$ a sequence in $\mathbb{R}/\mathbb{Z}$. Then $\{x_n\}$ is equidistributed if and only if for every $h \in \mathbb{Z}$, $h \neq 0$, we have:

$$\sum_{n \leq x} e(hx_n) = o(x)$$

Proof. To say that a sequence is equidistributed is equivalent to say that if f is a real continuous function (or $f : \mathbb{R}/\mathbb{Z} \to \mathbb{C}$) then

$$\int_{[0,1]} f(x)dx = \lim_{x\to+\infty} \frac{1}{x} \sum_{n\leq x} f(x_n)$$

On the other hand, the condition from the lemma is equivalent to say that the fact above is true for the functions $f(x) = e(hx)$. Since the linear combinations of exponential functions form a dense set in $C(\mathbb{R}/\mathbb{Z})$, the result follows. □

Equivalently we should prove that:

$$\sum_{\frac{\nu}{p}\in E_x} e\left(\frac{h\nu}{p}\right) = o(|E_x|) = o\left(\frac{x}{\log x}\right)$$

or better:

$$\sum_{p\leq x} \rho_h(p) = \sum_{p\leq x} \sum_{P(\nu)\equiv 0 \,(\mathrm{mod}\ p)} e\left(\frac{h\nu}{p}\right) = o\left(\frac{x}{\log x}\right)$$

Remark 1.4. One important thing to say here is that we are going to consider $\rho_h(n)$ only for the n's relatively prime to D. Note that this does not change the equidistribution, since we are only giving up a finite number of primes. This will be important when we use the sieve methods because there we need some bounds for $\rho_h(n)$ which can not be obtained for the other elements. So that if $(n, D) \neq 1$, we consider $\rho_h(n) = 0$.

2 The oscillating sieve of Duke-Friedlander-Iwaniec

Let $\mathfrak{A} = (a_n)$ be a sequence of complex numbers. Here we are going to show the method of [DFI] for estimating $\sum_p a_p$ under conditions which are actually achievable in this case. One extra assumption we are making here is that $|a_n| \leq \tau(n)$ which will be the case when we apply the results of this section. We start with

Definition 2.1. We denote

$$S(\mathfrak{A}, z) = \sum_{\substack{n\leq x \\ (p,n)=1\ \forall p<z}} a_n$$

For example if we wanted to consider only primes in the sum, we could look at $S(\mathfrak{A}, x^{\frac{1}{2}})$. In real life, it is too much to ask for an estimative for $z = x^{\frac{1}{2}}$. So we try to avoid this taking z slightly smaller.

Lemma 2.1. Suppose we have $x^{\frac{1}{3}} < z \leq x^{\frac{1}{2}}$. We then have:

$$\sum_{p\leq x} a_p = S(\mathfrak{A}, z) + \mathcal{O}\left(\frac{x}{\log x} \log\left(\frac{\log x}{2\log z}\right)\right)$$

For $x \geq 2$, with an absolute implicit constant.

Proof. Note that in $S(\mathfrak{A}, z)$ we only have terms with at most two primes, because all of the primes have to be bigger than $x^{\frac{1}{3}}$. So that, by inclusion-exclusion, we have:

$$\sum_{p\leq x} a_p = S(\mathfrak{A}, z) - \sum_{\substack{pq\leq x \\ z\leq p\leq q}} a_{pq} + \sum_{p<z} a_p - a_1$$

Now we estimate the extra terms in the right side:

$$|\sum_{p<z} a_p| + |a_1| \leq 2\pi(z) + 1 \ll \frac{\sqrt{x}}{\log z}$$

by the prime number theorem and the fact that $a_n \leq \tau(n)$. In addition:

$$\begin{aligned}
|\sum_{pq\leq x}\sum_{z\leq p\leq q} a_{pq}| &\leq 4 \sum_{z\leq p\leq x^{\frac{1}{2}}} \sum 1 \\
&\leq 4 \sum_{z\leq p\leq x^{\frac{1}{2}}} \pi\left(\frac{x}{p}\right) \\
&< \frac{x}{\log x} \sum_{z\leq p\leq x^{\frac{1}{2}}} \frac{1}{p} \\
&\ll \frac{x}{\log x}(\log\log\sqrt{x} - \log\log z)
\end{aligned}$$

which concludes the proof. □

Let's say we have a sequence $\mathfrak{A} = (a_n)$, we define $\mathfrak{A}_p$ to be the sequence (b_n) where $b_n = a_{pn}$. With this new definition, we have:

Lemma 2.2. For every $w < z$, we have:

$$S(\mathfrak{A}, z) = S(\mathfrak{A}, w) - \sum_{w\leq p<z} S(\mathfrak{A}_p, w) + \sum_{w\leq q<p<z} S(\mathfrak{A}_{pq}, q)$$

Proof. Here we start with the Buchstab identity:

$$S(\mathfrak{A}, z) = S(\mathfrak{A}, w) - \sum_{w\leq p<z} S(\mathfrak{A}_p, p)$$

Which is a direct consequence of the fact that $S(\mathfrak{A}_p, p)$ is the sum of the a_n such that p is the smallest prime appearing. All we have to do now is apply the Buchstab's identity one more time for the term $\mathfrak{A}_p$ and obtain:

$$S(\mathfrak{A}, z) = S(\mathfrak{A}, w) - \sum_{w\leq p<z} S(\mathfrak{A}_p, w) + \sum_{w\leq q<p<z} S(\mathfrak{A}_{pq}, q)$$

□

We now take care of the first two summands of the equality in the lemma above.

Lemma 2.3. We have

$$S(\mathfrak{A}, w) - \sum_{w \le p < z} S(\mathfrak{A}_p, w) = \sum_{d|P(z)}^{!} \mu(d) A_d(x)$$

where $P(z) = \prod_{p<z} p$ and

$$A_d(x) = \sum_{\substack{n \le x \\ d|n}} a_n$$

And $\sum^{!}$ means we sum over d with at most one prime factor $\ge w$.

Proof. We can start from the equality below:

$$S(\mathfrak{A}, w) = \sum_{d|P(w)} \mu(d) A_d(x).$$

For proving it, we observe that:

$$\begin{aligned} RHS &= \sum_{d|P(w)} \sum_{\substack{n \le x \\ d|n}} \mu(d) a_n \\ &= \sum_{n \le x} \left(\sum_{d|(n,P(w))} \mu(d) \right) a_n \\ &= \sum_{\substack{n \le x \\ (n,P(w))=1}} a_n, \end{aligned}$$

where in the last line we used that $\sum_{d|n} \mu(d) = 1$ if $n = 1$ and 0 otherwise. Similarly, we have:

$$S(\mathfrak{A}_p, w) = \sum_{d|P(w)} \mu(d) A_{pd}(x)$$

Summing it for $w \le p < z$:

$$\sum_{w \le p < z} S(\mathfrak{A}_p, w) = \sum_{w \le p < z} \sum_{d|P(w)} -\mu(pd) A_{pd}(x)$$

Observe that the pd's above run exactly through those d dividing $P(z)$ with exactly one prime factor bigger than w, while $S(\mathfrak{A}, w)$ gives us the ones with none. □

We proceed now to estimate the sum with the ! sign. What we do is to cut in two parts according to whether d is large or small.

Lemma 2.4. Let $y \geq z$. We have

$$\sum_{\substack{d|P(z)\\ d>y}}^{!} \mu(d)A_d(x) \ll x(\log x)^8 \left(\frac{z}{y}\right)^{(\log w)^{-1}}$$

and

$$\sum_{\substack{d|P(z)\\ d\leq y}}^{!} \mu(d)A_d(x) \leq R(x,y) = \sum_{d\leq y} |A_d(x)|.$$

Proof. The second estimate is immediate. For the first one:

$$|\sum_{\substack{d|P(z)\\ d>y}}^{!} \mu(d)A_d(x)| \leq \sum_{\substack{d|P(z)\\ d>y}}^{!} \sum_{nd\leq x} \tau(nd).$$

Next we use the fact that $\tau(nd) \leq \tau(n)\tau(d)$ for every n and d. And the estimate

$$\sum_{n\leq x} \tau(n) \ll x \log x.$$

With that we get:

$$|\sum_{\substack{d|P(z)\\ d>y}}^{!} \mu(d)A_d(x)| \ll x \log x \sum_{\substack{d|P(z)\\ d>y}}^{!} \frac{\tau(d)}{d}.$$

The next step is a trick by Rankin in which we use in a very clever way the fact that $d > y$. Let ϵ be a parameter to be chosen later, then:

$$\sum_{\substack{d|P(z)\\ d>y}}^{!} \frac{\tau(d)}{d} \leq y^{-\epsilon} \sum_{d|P(z)}^{!} \tau(d)d^{-1+\epsilon}.$$

Using the multiplicativity of τ and the meaning of the ! symbol:

$$y^{-\epsilon} \sum_{d|P(z)}^{!} \tau(d)d^{-1+\epsilon} = y^{-\epsilon} \left\{1 + 2 \sum_{w\leq p<z} p^{-1+\epsilon}\right\} \prod_{p<w} (1 + 2p^{-1+\epsilon}).$$

First, we have

$$1 + 2 \sum_{w\leq p<z} p^{-1+\epsilon} \leq z^{\epsilon} \prod_{w\leq p<z} (1 + 2p^{-1}).$$

For the product, we choose $\epsilon = (\log w)^{-1}$ so that $p^{\epsilon} = \exp(\frac{\log p}{\log w}) < e$ for all $p < w$. So that:

$$\prod_{p<w} (1 + 2p^{-1+\epsilon}) \leq \prod_{p<w} (1 + 2ep^{-1}).$$

We now use that $1 + x < e^x$ for all $x > 0$ and the fact

$$\sum_{a\leq p<b} \frac{1}{p} < \log\left(\frac{\log b}{\log a}\right)$$

to obtain:

$$\sum_{\substack{d|P(z)\\ d>y}}{}^{!}\frac{\tau(d)}{d} \ll \left(\frac{z}{y}\right)^{\epsilon}\left(\frac{\log z}{\log w}\right)^{2}(\log w)^{7}.$$

Since $\left(\frac{\log z}{\log w}\right)^{2}(\log w)^{7} \leq (\log x)^{7}$, we have the result. □

We now pass to estimate the sum of the terms $S(\mathfrak{A}_{pq}, q)$. In this process we're gonna again introduce a new variable to separate the cases where q is big or small.

Lemma 2.5. Let v such that $x^{1/4} < v \leq x^{1/3}$. We have

$$\sum_{w\leq q<p<z} S(\mathfrak{A}_{pq}, q) = \sum_{\substack{mq\leq x\\ w\leq q\leq v\\ q<p_m<z}} \gamma_m a_{mq} + \mathcal{O}\left(\frac{x}{\log x}\log\left(\frac{\log x}{3\log v}\right)\right)$$

where p_m is the smallest prime divisor of m and

$$\gamma_m = \sum_{\substack{p|m\\ p<z}} 1$$

Proof. We write

$$\sum_{w\leq q<p<z} S(\mathfrak{A}_{pq}, q) = \sum_{v\leq q<p<z} S(\mathfrak{A}_{pq}, q) + \sum_{w\leq q<v} S(\mathfrak{A}_{pq}, q).$$

The second term here is the first one in the statement. We need then to focus on the first.

$$\sum_{v\leq q<p<z} S(\mathfrak{A}_{pq}, q) = \sum_{v\leq q<p<z} \sum_{\substack{pq|n\\ (n,P(q))=1}} a_n.$$

We then observe that the n's appearing are either of the form $n = pq$ or $n = pqr$ with r prime. Because all the primes dividing n have to be bigger than $v \geq x^{1/4}$ and $p > q$, which gives us the bound $a_n \leq \tau(n) \leq 8$. We now divide the sum in two parts according to whether $n = pq$ or $n = pqr$. We also observe that in this last case, since q is the smallest among the three primes dividing n, we get $q \leq x^{1/3}$. Finally

$$\begin{aligned}\sum_{w\leq q<p<z} S(\mathfrak{A}_{pq}, q) &< \sum_{v\leq q<p<z} 1 + \sum_{v\leq q<x^{1/3}} \sum_{q<p<z} \sum_{r\leq \frac{n}{pq}} 1\\ &< \left(\sum_{v\leq q<z} 1\right)^{2} + \sum_{v\leq q<x^{1/3}} \sum_{v<p<z} \frac{x}{pq\log x}\\ &\ll \left(\log\left(\frac{\log z}{\log v}\right)\right)^{2} + \frac{x}{\log x}\left(\log\left(\frac{\log z}{\log v}\right)\right) \sum_{v\leq q<x^{1/3}} \frac{1}{q}.\end{aligned}$$

And the result is proved observing that $1 < \frac{\log z}{\log v} < 2$. □

The following steps are a way of getting rid of the condition $q < p_m < z$ in $\sum \gamma_m a_{mq}$.

Lemma 2.6. Let $z \geq 1$. There exists a function $h : \mathbb{R} \to \mathbb{C}$, depending on z, such that for $1 \leq a, b \leq z$, the integral

$$\int_{\mathbb{R}} h(t) \left(\frac{a}{b}\right)^{it} dt$$

is equal to 1 if $a \leq b$ and 0 if not. In addition,

$$\int_{\mathbb{R}} |h(t)| dt \leq \log 2z.$$

Proof. Consider $g : \mathbb{R} \to [0, 1]$ given by:

$$g(u) = \begin{cases} uz & \text{if } 0 \leq u \leq 1/z \\ 1 & \text{if } 1/z \leq u \leq 1 \\ 1 + (1-u)z & \text{if } 1 \leq u \leq 1 + 1/z \\ 0 & \text{otherwise.} \end{cases}$$

If a, b are integers less than z, we have

$$g\left(\frac{a}{b}\right) = \begin{cases} 1 & \text{if } a \leq b \\ 0 & \text{otherwise.} \end{cases}$$

Now let $\hat{g}$ be the Mellin transform of g. Since $g(0) = 0$, we don't have a pole of $\hat{g}$ in $z = 0$ and $\hat{g}$ is holomorphic for $\Re(s) > -1$ and by Mellin inversion

$$g\left(\frac{a}{b}\right) = \frac{1}{2\pi i} \int_{(0)} \hat{g}(s) \left(\frac{a}{b}\right)^{-s} ds.$$

We now proceed to estimate $\hat{g}$ when $\Re(s) = 0$. First we have:

$$|\hat{g}(it)| \leq \int_0^{+\infty} \frac{g(u)}{u} du = \log\left(\frac{z+1}{z}\right)^{z+1}, \quad z \leq \log 4z.$$

Alternatively, for $s \neq 0$

$$\hat{g}(s) = \int_0^{\infty} g(u) u^{s-1} du = \frac{(z+1)^{s+1} - z^{s+1} - 1}{s(s+1)z^s},$$

which gives us

$$|\hat{g}(it)| \leq \frac{2z+2}{t+t^2} \leq \frac{4z}{t^2}.$$

Also, if we use the mean value inequality for $(z+1)^{s+1} - z^{s+1}$,

$$|\hat{g}(it)| \leq \max_{w \in [z, z+1]} \frac{|w^{it}|}{t} + \frac{1}{t+t^2} \leq \frac{2z}{t}.$$

We now combine aplly these three estimates for $|t| \leq (\log 4z)^{-1}$, $(\log 4z)^{-1} \leq |t| \leq 4z$ and $|t| > 4z$, respectively, to obtain

$$\int_{\mathbb{R}} |\hat{g}(it)|dt \leq 2 + \log 4z + \log\log 4z \ll \log 2z.$$

So, finally, we take $h(t) = (2\pi)^{-1}\hat{g}(it)$. The above assures that it satisfies all the necessary conditions. □

We then have the following immediate consequence:

Corollary 2.7. There exists $t \in \mathbb{R}$ such that for $x \geq 2$,

$$\sum_{\substack{mq \leq x \\ w \leq q < v \\ q < p_m < z}} \gamma_m a_{mq} \ll (\log x)^2 B(x)$$

$$B(x) = \sum_{(m,P(w))=1} \left| \sum_{w \leq q < v} a_{mq} q^{it} \right|.$$

for some $t \in \mathbb{R}$, the implicit constant being absolute.

Proof. Because of the lemma, we have

$$\sum_{\substack{mq \leq x \\ w \leq q < v \\ q < p_m < z}} \gamma_m a_{mq} = \int_{\mathbb{R}} B(x,t)h(t)dt.$$

Where

$$B(x,t) = \sum_{\substack{mq \leq x \\ w \leq q < v \\ q < p_m < z}} \gamma_m a_{mq} \left(\frac{p_m}{q}\right)^{-it}.$$

In such a way that

$$\begin{aligned}
\sum_{\substack{mq \leq x \\ w \leq q < v \\ q < p_m < z}} \gamma_m a_{mq} &\leq \int_{\mathbb{R}} |B(x,t)||h(t)|dt \\
&= |B(x,t_0)| \int_{\mathbb{R}} |h(t)|dt \quad \text{(Mean value inequality)} \\
&\leq (\log x)^2 \sum_{(m,P(w))=1} \left| \sum_{w \leq q < v} a_{mq} q^{it_0} \right|,
\end{aligned}$$

for some $t_0 \in \mathbb{R}$. Here we use both the bound from the lemma remembering that $\log 2z \leq \log x$ and the inequality $|\gamma_m| \ll \log x$.

To see why this last inequality holds, we notice that the extreme case is when we take $m = \prod_{p<u} p$ where u is the largest possible such that $m \leq x$. We then have

$$\gamma_m = u \sum_{p<u} \log p = \log m \leq \log x.$$

□

Finally, we can state (the proof is given by the previous lemmas) the main theorem of this section. Observe that we can write

$$B(x) = \sum_{mn \leq x} \beta_n \gamma_m a_{mn},$$

where $|\beta_n| \leq 1, |\gamma_n| \leq 1$, β_n is supported on the prime numbers and γ_m on the numbers without prime factor $< w$.

Theorem 2.8. Let $\mathfrak{A}$ be a sequence of complex numbers such that $|a_n| \leq \tau(n)$ for all $n \geq 1$. Let $x \geq 1$ and w, v, y, z parameters such that

$$w < x^{1/4} < v \leq x^{1/3} < z \leq y \leq x^{1/2} < x.$$

Then there exist complex numbers α_n, β_n, γ_n, such that $|\alpha_n| \leq 1$, $|\beta_n| \leq 1$, $|\gamma_n| \leq 1$, β_n is supported on prime numbers and γ_m on numbers without prime factor $< w$, and, furthermore

$$\sum_{p \leq x} a_p = R(x) + B(x)(\log x)^2 + \mathcal{O}(x(\log x)^8 (z/y)^{(\log w)^{-1}}) + \mathcal{O}\left(\frac{x}{\log x} \log\left(\frac{(\log x)^2}{6(\log v)(\log z)}\right)\right),$$

where

$$R(x) = \sum_{\substack{d \leq y \\ nd \leq x}} \alpha_d a_{nd}$$

$$B(x) = \sum_{\substack{mn \leq x \\ w \leq n < v}} \beta_n \gamma_m a_{mn}$$

and the implicit constants are absolute.

For a good application of the theorem, we need to have good bounds for $R(x)$, $B(x)$ and make good choices for the parameters. One (good) example is the following corollary.

Corollary 2.9. Let (a_n) like in the theorem above. Suppose that for every $\epsilon > 0$ we have

$$R(x) \ll_\epsilon \frac{x}{(\log x)^2},$$

$$B(x) \ll_\epsilon \frac{x}{(\log x)^4},$$

with the choice of parameters

$$w = \exp\left(\frac{\epsilon \log x}{10 \log\log x}\right), \quad v = x^{1/3-\epsilon}, \quad z = x^{1/2-2\epsilon}, \quad y = x^{1/2-\epsilon}.$$

Then, as $x \to +\infty$

$$\sum_{p \leq x} a_p = o\left(\frac{x}{\log x}\right).$$

Proof. Observe that

$$\frac{(\log x)^2}{\log v \log z} = -\log\{(1-3\epsilon)(1-4\epsilon)\} \ll \epsilon,$$

with an absolute constant if ϵ is small enough. The other substitutions follow similarly and finally we find out that it exists $C(\epsilon) > 0$ such that

$$\sum_{p \leq x} a_p \ll \frac{\epsilon x}{\log x} + \frac{C(\epsilon)x}{(\log x)^2},$$

with an absolute constant, giving us the result. □

3 Quadratic congruences and the upper half plane

We fix $D > 0$ and consider $P(x) = x^2 + D$ and like before,

$$\rho_h(n) = \sum_{\substack{\nu \in \mathbb{Z}/n\mathbb{Z} \\ P(\nu) \equiv 0 \,(\mathrm{mod}\ n)}} e\left(\frac{h\nu}{n}\right),$$

where $h \in \mathbb{Z}$, $h \neq 0$. For $d \geq 1$ our new object of interest will be

$$\mathcal{L}_d(x) = \sum_{d|n} \rho_h(n) \psi\left(\frac{n}{x}\right).$$

In this section, we interpret the sum above in terms of a certain Poincaré series and then, following the spectral theory of modular forms we estimate it in order to obtain the necessary bounds for the application of the sieve. We start with the tautological bijection between solutions of $b^2 + D \equiv 0 \pmod{n}$ and positive quadratic forms

$$q(x, y) = mx^2 + 2bxy + ny^2$$

of discriminant $D = mn - b^2$. Here the solutions are considered in $\mathbb{Z}$ (and not modulo n).

To parametrize this quadratic forms, we denote by $\mathcal{Q}$ the positive real quadratic forms of discriminant D and we make use of the natural action of $SL(2, \mathbb{Z})$ on $\mathcal{Q}$ given by

$$(\sigma.q)(x, y) = q(ax + \gamma y, \beta x + \delta y), \quad \sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2, \mathbb{R}).$$

Consider $\mathcal{Q}_{\mathbb{Z}} \subset \mathcal{Q}$ the subset of the positive integer quadratic forms with discriminant $D > 0$ and $SL(2, \mathbb{Z})\backslash\mathcal{Q}_{\mathbb{Z}}$ the set of orbits of the action. Furthermore, we pass to the complex upper half plane which can be used to parametrize the set $\mathcal{Q}$ by the map $\pi : \mathcal{Q} \to \mathbb{H}$ defined by

$$\pi(q) = \frac{b + i\sqrt{D}}{c}$$

if $q(x, y) = ax^2 + 2bxy + cy^2$. This the same as finding the root of $q(1, z) = 0$ with positive imaginary part. A good thing here is that the action of $SL(2, \mathbb{Z})$ in $\mathcal{Q}$ commutes with the natural action of $SL(2, \mathbb{Z})$ in $\mathbb{H}$ which is given by

$$\sigma.z = \frac{\alpha z + \beta}{\gamma z + \delta}$$

If we want to look for the equations (mod n), the condition on the last coefficient being equal to n can be translated in terms of the upper half plane as $\Im(z) = \frac{\sqrt{D}}{n}$. At this point we're going to have a closer look at the action of $SL(2,\mathbb{Z})$. More explicitly, we are going to consider a fundamental domain for this action and see what happens when we restrict ourselves to the points coming from $\mathcal{Q}_{\mathbb{Z}}$.

Proposition 3.1. (1) Let $F \subset \mathbb{H}$ be the open set

$$F = \{z = x + iy \mid |z| > 1, |x| < 1/2\}$$

and $\overline{F}$ its closure. Then $\forall z \in \mathbb{H}$, there is $w \in \overline{F}$ which is $SL(2,\mathbb{Z})$-equivalent to z. Moreover if $w \in F$, then w is unique and if $w \in \partial F$, it is unique except for multiplication by ± 1.

(2) The set $\Lambda = SL(2,\mathbb{Z})\backslash\mathcal{Q}_{\mathbb{Z}}$ is finite.

(3) For every $z \in \mathbb{H}$, the stabilizer $\Gamma_z = \{\sigma \in SL(2,\mathbb{R}) \mid \sigma z = z\}$ is finite with order at most 6.

This is a classical result which can be found in various sources, for example [K].

Up to now we're only considering solutions with $b \in \mathbb{Z}$. We have to somehow pass to consider them (mod n). Two solutions corresponding to the same one (mod n) differ by some multiple of n. Viewing them in the upper half plane, they look like $\frac{b+i\sqrt{D}}{n}$ and $\frac{b+ns+i\sqrt{D}}{n} = \frac{b+i\sqrt{D}}{n} + s$, i.e they differ only by an integer translation. With that in mind, we introduce the subgroup

$$B = \left\{ \pm \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \mid n \in \mathbb{Z} \right\},$$

which gives us exactly the elements of $SL(2,\mathbb{Z})$ which act by integer translations. We can now state the following proposition which will help us rewriting the sums $\mathcal{L}_d$ in terms of modular forms.

Proposition 3.2. Let $D > 0$ integer, $P(x) = x^2 + D$. For every $n \geq 1$, if $f(\nu)$ is a complex function defined on the solutions $\nu \in \mathbb{Z}/n\mathbb{Z}$ of $\nu^2 + D \equiv 0 \pmod{q}$, we have

$$\sum_{\nu \in \mathbb{Z}/n\mathbb{Z}} f(\nu) = \sum_{\substack{\sigma \in B\backslash SL(2,\mathbb{Z}) \\ \Im(\sigma z) = \sqrt{D}/n}} f(h\Re(\sigma z))$$

Observe that the $B\backslash SL(2,\mathbb{Z})$ appearing below the sum takes care of counting only once each solution (mod n).

The next step is incorporating the condition $d|n$. To do so, we now introduce Hecke congruence subgroups

$$\Gamma_0(d) = \left\{ \sigma = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in SL(2,\mathbb{Z}) \mid \gamma \equiv 0 \pmod{d} \right\}$$

for $d \geq 1$. Note that $\Gamma_0(1) = SL(2,\mathbb{Z})$ and $B \subset \Gamma_0(d)$ for all $d \geq 1$.

Let $\mathcal{Q}_d$ be the subset of $\mathcal{Q}_\mathbb{Z}$ given by the quadratic forms $q(x,y) = ax^2+bxy+cy^2$ such that $d|c$ or equivalently, $q(0,1) \equiv 0 \pmod{d}$. The point of introducing these congruence groups is that $\mathcal{Q}_d$ is invariant by the action of $\Gamma_0(d)$. Because

$$(\sigma q)(0,1) = q(\gamma,\delta) \equiv q(0,\delta) = \delta^2 q(0,1) \equiv 0 \pmod{d}$$

if $d \mid \gamma$, which is exactly the condition for $\Gamma_0(d)$. We next notice that $B\backslash SL(2,\mathbb{Z})$ can be written as $B\backslash\Gamma_0(d) \times \Gamma_0(d)\backslash SL(2,\mathbb{Z})$. So that we can write the identity above as

$$\sum_{\nu\in\mathbb{Z}/n\mathbb{Z}} f(\nu) = \sum_{z\in\Lambda} \frac{1}{|\Gamma_z|} \sum_{\substack{\tau\in\Gamma_0(d)\backslash\Gamma_0(1) \\ \Im(\sigma\tau z)=\sqrt{D}/n}} \sum_{\sigma\in B\backslash\Gamma_0(d)} f(h\Re(\sigma\tau z))$$

and by the discussion above, the condition $\Im(\sigma\tau z) = \sqrt{D}/n, d|n$ is equivalent to $\Im(\tau z) = \sqrt{D}/n, d|n$. Now we add up the respective identities as above for every n such that $d|n$ and we can then rewrite everything as

$$\sum_{d|n} \rho_h(n) = \sum_{z\in\Lambda} \frac{1}{|\Gamma_z|} \sum_{\tau\in\Gamma_0(d)\backslash\Gamma_0(1)}^{*} \sum_{\sigma\in B\backslash\Gamma_0(d)} e(h\Re(\sigma\tau z)).$$

where the symbol $*$ means that we only consider the elements (τ, z) such that $\Im(\tau z) = \frac{\sqrt{D}}{n}, d|n$.

Consider a test function of the type

$$F\left(\frac{2\pi h\sqrt{D}}{n}\right) = F(2\pi h\Im(\sigma\tau z))$$

with F of compact support (The particular way of writing is just for simplifying further formulas). Then we can slightly change the equation above to obtain

$$\sum_{d|n} \rho_h(n) F\left(\frac{2\pi h\sqrt{D}}{n}\right) = \sum_{z\in\Lambda} \frac{1}{|\Gamma_z|} \sum_{\tau\in\Gamma_0(d)\backslash\Gamma_0(1)}^{*} P_h(\tau z)$$

where

$$P_h(z) = \sum_{\sigma\in B\backslash\Gamma_0(d)} F(2\pi h\Im(\sigma z))e(h\Re(\sigma z)).$$

The function $P_h(z)$ is called a Poincaré series and will be the object which we will try to estimate in section 5.

We have a last lemma which estimates the number of τ's appearing in the Σ^* above.

Lemma 3.3. Let $D > 0$ and $d \geq 1$ integers without square factors. Let $z \in \Lambda$ and $q = ax^2 + bx + c$ the quadratic form associated to it. The number $m(d,z)$ of elements $\tau \in \Gamma_0(d)\backslash\Gamma_0(1)$, that appear in the sum above satisfies

$$m(d,z) \leq \tau(n)$$

Proof (just a mention). The proof uses the bijection between $\Gamma_0(d)\backslash\Gamma_0(1)$ and $\mathbb{P}^1(\mathbb{Z}/d\mathbb{Z})$ given by

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \mapsto (\gamma : \delta) \pmod{d}$$

The actual proof can be found in [K]. □

4 Automorphic forms and spectral theory

We start with $\Gamma < SL(2,\mathbb{Z})$ a discrete subgroup. We ask further that Γ has finite volume, i.e.

$$Vol(\Gamma\backslash\mathbb{H}) = \int_{F_\Gamma} \frac{dxdy}{y^2} < +\infty$$

Where F_Γ is a fundamental domain which will be considered fixed.

Let's take a look in such a fundamental domain. In general we cannot expect a fundamental domain to be compact, but it fails to be compact only for few points on the boundary where the domain accumulates, this points are what we call the cusps of Γ. They're not well defined but modulo Γ when we change the fundamental domain. To have a better view of the cusps we have:

Proposition 4.1. Let $\Gamma \in SL(2,\mathbb{R})$ such that $-\mathrm{Id}_{2\times 2} \in \Gamma$ and such that we have $\mathrm{vol}(\Gamma\backslash\mathbb{H}) < +\infty$. Then

(1) The cusps are in bijection with the points of $\mathbb{R}\cup\{\infty\}$ fixed by a parabolic element of Γ. There's only a finite number of cusps modulo Γ.

(2) For every cusp $\mathfrak{a}$, the stabilizer $\Gamma_\mathfrak{a} = \{\gamma \in \Gamma \mid \gamma\mathfrak{a} = \mathfrak{a}\}$ is isomorphic to $\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}$.

(3) Let $\mathfrak{a}$ be a cusp and $\tau_\mathfrak{a}$ a generator of $\Gamma_\mathfrak{a}$. There exists $\sigma_\mathfrak{a} \in SL(2,\mathbb{R})$ such that $\sigma_\mathfrak{a}\infty = \mathfrak{a}$ and

$$\sigma_\mathfrak{a}^{-1}\tau_\mathfrak{a}\sigma_\mathfrak{a} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = B.$$

Moreover, any other σ satisfying these two properties is of the form

$$\sigma = \pm\sigma_\mathfrak{a}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Sketch of proof. The point (1) is actually a very complicated result by Siegel [S. p.75-77]. For the other points, we just have to use the fact that Γ is discrete and that the parabolic elements fixing ∞ are of the form $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, $t \in \mathbb{Z}$ and at last we use the identity $\begin{pmatrix} a & u \\ 0 & a^{-1} \end{pmatrix}\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a^{-1} & -u \\ 0 & a \end{pmatrix} = \begin{pmatrix} 1 & a^2t \\ 0 & 1 \end{pmatrix}$. For the last assertion, we see that an element fixing ∞ commuting with $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ is always of that form. □

In our context an automorphic form is a function defined in the upper half plane satisfying some modularity over a certain subgroup of $SL(2,\mathbb{R})$ with some smooth properties. More precisely:

Definition 4.1. Let $\Gamma < SL(2,\mathbb{R})$ with finite volume.

(1) We let

$$\mathcal{A}(\Gamma) = \mathcal{A}(\Gamma\backslash\mathbb{H}) = \{f : \mathbb{H} \to \mathbb{C} \mid f(\gamma z) = f(z) \text{ for every } \gamma \in \Gamma\}$$

be the space of Γ-invariant functions.

(2) We let

$$\mathcal{A}^\infty(\Gamma) = \mathcal{A}^\infty(\Gamma\backslash\mathbb{H}) = \{f \in \mathcal{A}(\Gamma) \mid f \text{ is } C^\infty\}$$

be the space of smooth Γ-invariant functions.

(3) We let

$$\mathcal{A}^p(\Gamma) = \mathcal{A}^p(\Gamma\backslash\mathbb{H}) = \{f \in \mathcal{A}(\Gamma) \mid \exists N > 0 \text{ such that } f_{\mathfrak{a}} \ll y^N \text{ for every cusp } \mathfrak{a}, \text{ and every } Y \text{ sufficiently large}\}$$

be the space of automorphic forms (with moderated growth at every cusp).

(4) We let

$$\mathcal{D}(\Gamma) = \mathcal{D}(\Gamma\backslash\mathbb{H}) = \{f \in \mathcal{A}^\infty(\Gamma) \cap \mathcal{A}^p(\Gamma) \mid ||f||_2 < +\infty \text{ and } ||\Delta f||_\infty < +\infty\}$$

the domain of the automorphic laplacian.

Remark 4.2. Note that since $\text{vol}(\Gamma) < +\infty$, $L^\infty(\Gamma\backslash\mathbb{H}) \subset L^2(\Gamma\backslash\mathbb{H})$. In particular $\mathcal{D}(\Gamma) \subset L^2(\Gamma\backslash\mathbb{H})$.

The next proposition gives us some information about the Fourier expansion of a proper value of the laplacian operator.

Proposition 4.3. Recall that the hyperbolic laplacian is given by $\Delta f = -y^2(f_{xx} + f_{yy})$. Let $f \in \mathcal{A}^\infty(\Gamma) \cap \mathcal{A}^p(\Gamma)$ and $\mathfrak{a}$ a cusp of Γ. Suppose that $\Delta f = \lambda f, \lambda \in \mathbb{C}$. We then have

$$f_{\mathfrak{a}}(z) = b_{f,\mathfrak{a}}(0)y^s + c_{f,\mathfrak{a}}(0)y^{1-s} + \sqrt{y}\sum_{n\neq 0} b_{f,\mathfrak{a}}(n)K_{s-1/2}(2\pi|n|y)e(nx)$$

where $b_{f,\mathfrak{a}}(n), c_{f,\mathfrak{a}}(0) \in \mathbb{C}$ and $s(1-s) = \lambda$. The function K_ν is called the Bessel-MacDonald and given by

$$K_\nu(y) = \frac{1}{2}\int_0^{+\infty} \exp\left(-\frac{y}{2}\left(t + \frac{1}{t}\right)\right) t^{-\nu-1}dt$$

If $s = 1/2$, i.e $\lambda = 1/4$ we have to change y^{1-s} for $y^{1/2}\log y$.

Sketch of proof. Since $f_{\mathfrak{a}}(z) = f(\sigma_{\mathfrak{a}} z)$, we have $f_{\mathfrak{a}}(z+1) = f_{\mathfrak{a}}(z)$. So that we have a decomposition

$$f_{\mathfrak{a}}(z) = \sum_{n\in\mathbb{Z}} b_{f,\mathfrak{a}}(y,n)e(nx).$$

We next develop the equation $\Delta f = \lambda f$ and obtain a differential equation for each $b_{f,\mathfrak{a}}(y,n)$, precisely

$$g'' + \left(\frac{\lambda}{y^2} - 4\pi n^2\right) g = 0$$

For $n = 0$ we have the desired solutions. But for $n \neq 0$ we theoretically have two different solutions. But since as y approaches $+\infty$, the equations approaches $g'' - 4\pi n^2 g = 0$, we can choose two linearly independent solutions that each one is asymptotically equivalent to the solutions $e^{2\pi|n|y}$ and $e^{-2\pi|n|y}$. Since $b_{f,\mathfrak{a}}(y,n)$ is given by

$$\int_0^1 f_{\mathfrak{a}}(x+iy)e(-nx)dx$$

it should also satisfy the moderated growth, so that we only have a multiple of the second solution. Which is, as matter of fact, $\sqrt{(2\pi|n|y)}K_{s-1/2}(2\pi|n|y)$. □

In the sequel, we talk shortly about Eisenstein series.

Definition 4.2. Let $\mathfrak{a}$ be a cusp of Γ. The Eisenstein series defined by $\mathfrak{a}$ is the function

$$E_{\mathfrak{a}}(z,s) = \sum_{\gamma \in \Gamma_{\mathfrak{a}} \backslash \Gamma} \Im(\sigma_{\mathfrak{a}}^{-1}\gamma z)^s$$

This series can be shown to converge for $\Re(s) > 1$, but for the spectral theorem for automorphic forms, the functions that play a role are the Eisenstein functions for $\Re(s) = 1/2$. For this to make sense, we have:

Theorem 4.4. For every cusp $\mathfrak{a}$, the Eisenstein series $E_{\mathfrak{a}}(z,s)$ admits a meromophic continuation for $s \in \mathbb{C}$. It has a pole at $s = 1$ of constant residue

$$\mathrm{Res}_{s=1} E_{\mathfrak{a}}(z,s) = \frac{1}{\mathrm{vol}(\Gamma\backslash\mathbb{H})}$$

and apart from that, only a finite number of poles in the region $\Re(s) \geq 1/2$ which are real and for which the residue is a function of z which is a proper function of the laplacian with eigenvalue $s(1-s) \in [0, 1/4]$. Moreover if $\Gamma = \Gamma_0(d)$, these extra poles do not exist.

We are ready to state:

Theorem 4.5 (Spectral decomposition for Δ)**.** Let $q \geq 1$ be an integer, $f \in \mathcal{D}(\Gamma_0(q))$. We have

$$f(z) = \sum_{j\geq 0}(f|u_j)u_j(z) + \frac{1}{4\pi}\sum_{\mathfrak{a}}\int_{\mathbb{R}}(f|E_{\mathfrak{a}}(.,\tfrac{1}{2}+it))E_{\mathfrak{a}}(.,\tfrac{1}{2}+it)dt.$$

The series is absolutely and uniformly convergent in every compact set. Where the u_j are eigenfunctions in $L^2(\Gamma)$ for the laplacian with relative eigenvalues $\lambda_j \geq 0$ and

$$0 \leq \lambda_1 \leq \lambda_2 \leq \dots$$

5 Bounding Poincaré series

We now use the spectral decomposition for $P_h(z)$.

$$P_h(z) = \sum_{j\geq 0}(P_h|u_j)u_j(z) + \frac{1}{4\pi}\sum_{\mathfrak{a}}\int_{\mathbb{R}}(P_h|E_{\mathfrak{a}})E_{\mathfrak{a}}(z,\tfrac{1}{2}+it)dt.$$

We want to take a closer look to what these scalar products look like.

Lemma 5.1. Let $f \in \mathcal{A}^p(\Gamma_0(d))$ with Fourier expansion

$$f(z) = \sum_{n\in\mathbb{Z}} b_f(n,y)e(nx).$$

We then have

$$(P_h|f) = \int_0^{+\infty} F(2\pi hy)\overline{b_f(h,y)}\frac{dy}{y^2}.$$

Proof.

$$(P_h|f) = \int_{\Gamma_0(d)\backslash\mathbb{H}} \sum_{\sigma\in B\backslash\Gamma_0(d)} F(2\pi h\Im(\sigma z))e(h\Re(\sigma z))\overline{f(z)}\frac{dxdy}{y^2}.$$

Since f is modular over $\Gamma_0(d)$ and the hyperbolic measure is invariant by $SL(2,\mathbb{R})$ we obtain

$$\begin{aligned}(P_h|f) &= \int_{B\backslash\mathbb{H}} F(2\pi hy)e(hx)\overline{f(z)}\frac{dxdy}{y^2} \\ &= \int_0^{+\infty}\left(\int_0^1 F(2\pi hy)e(hx)\overline{f(z)}dx\right)\frac{dy}{y^2} = \int_0^{+\infty} F(2\pi hy)\overline{b_f(h,y)}\frac{dy}{y^2}\end{aligned}$$

□

As a corollary, if we write

$$\begin{aligned}u_0(z) &= \frac{1}{\sqrt{\mathrm{vol}(\Gamma_0(d)\backslash\mathbb{H})}} \\ u_j &= \sqrt{y}\sum_{n\in\mathbb{Z}} b_j(n)K_{s_j-1/2}(2\pi|n|y)e(nx), \quad j\geq 1 \\ E_{\mathfrak{a}}(z,\tfrac{1}{2}+it) &= b_{\mathfrak{a}}y^{\frac{1}{2}+it} + c_{\mathfrak{a}}y^{\frac{1}{2}-it} + \sqrt{y}\sum_{n\in\mathbb{Z}}\tau_{\mathfrak{a}}(n,\tfrac{1}{2}+it)K_{it}(2\pi|n|y)e(nx).\end{aligned}$$

Thus, if $s_j = \frac{1}{2}+it_j$,

$$(P_h|u_0) = 0; \quad (P_h|u_j) = (2\pi h)^{1/2}\tilde{F}(t_j)\overline{b_j(h)}$$

and for all cusps $\mathfrak{a}$ of $\Gamma_0(d)$ and every $t\in\mathbb{R}$,

$$(P_h|E_{\mathfrak{a}}(.,\tfrac{1}{2}+it)) = (2\pi h)^{1/2}\tilde{F}(t)\overline{\tau}_{\mathfrak{a}}(h,\tfrac{1}{2}+it)$$

where $\tilde{F}(t) = \int_0^{+\infty} F(y)K_{it}(y)y^{-\frac{3}{2}}dy$.

In the sequence we are going to need a definition of a flat function.

Definition 5.1. $F : [0,+\infty) \to \mathbb{C}$ be a function. We say that F is flat of parameter $Y \geq 1$ if

(1) F has compact support on $[Y^{-1}, 2Y^{-1}]$,

(2) $F \in C^\infty$ and for $0 \leq j \leq 4$ we have $||F^{(j)}||_\infty \leq Y^j$

The next lemma gives us a good way of finding new flat functions from old ones.

Lemma 5.2. Let f be a flat function of parameter Y.

(1) For every $\alpha > 0$, the function $g(x) = f(\alpha x)$ is flat of parameter αY.

(2) There exists a constant $c = c_4$ independent of f such that $g(x) = cf(1/x)$ is flat of parameter $(Y/2)^{-1}$.

Proof. (1) is direct since $g^{(n)}(x) = \alpha^n f^{(n)}(\alpha x)$. For (2), it's obvious that g has support in the interval $[Y/2, Y]$. For the bound, because of the chain rule, we know that $g^{(n)}(x)$ is a sum of terms $x^{-m} f^{(l)}(1/x)$ such that $m - l = n$ and $l \leq n$. Each of these terms is bounded by $(Y/2)^{-l} Y^m = 2^l Y^{-n} \leq 2^8 Y^{-n}$. So all we have to do is take a $c = d^{-1}$ with d bigger than 2^8 times the biggest number of terms as above appearing in the equation for $f^{(n)}$, $n \leq 4$. □

The following lemma takes care of estimating $\tilde{F}$ above.

Lemma 5.3. Let $t \in \mathbb{R}$, $\frac{1}{4} + t^2 > 0$ and F be a flat function of parameter Y. Then,

$$\tilde{F}(t) \ll (1 + |t|)^{-4} \sqrt{\frac{Y}{\cosh \pi t}} (Y^{it} + Y^{-it} + 2 \log(Y + 2))$$

Proof. This bound will be given by the juxtaposition of two bounds in different parts of the set $\{t^2 + 1/4 \geq 0\}$. Since F has compact support, we don't have problems with the existence of the integral. Moreover, we can consider the power series for K_{it} and the convergence can be taken to be uniform so that it converges with the integral. Well, we have

$$K_{it}(y) = i\pi \frac{I_{it}(y) - I_{-it}(y)}{\sinh \pi t}$$

where

$$I_{\pm it}(y) = \sum_{m \geq 0} \frac{(y/2)^{2m \pm it}}{m! \Gamma(m + 1 \pm it)}$$

We consider the part of $\tilde{F}$ coming from I_{it} and call it $\tilde{F}_+$. We're gonna estimate this. The other part being similar. We have

$$\tilde{F}_+ = \frac{i\pi 2^{-it}}{\sinh \pi t} \sum_{m \geq 0} \frac{\hat{F}(2m + it - 1/2)}{2^{2m} m! \Gamma(m + 1 + it)}$$

where $\hat{F}$ is the Mellin transform of F. Integrating by parts we obtain

$$\hat{F}(s) = \int_0^{+\infty} F(y) y^{s-1} dy = \frac{1}{s(s+1)(s+2)(s+3)} \int_0^{+\infty} F^{(4)}(y) y^{s+3} dy$$

We now use that F is flat of parameter Y to obtain

$$\hat{F}(s) \leq (1 + |s|)^{-4} Y^4 Y^{-1} (2Y^{-1})^{\sigma+3} \ll (1 + |t|)^{-4} (2Y^{-1})^{\sigma}$$

This allows us to estimate $\tilde{F}_+$

$$\tilde{F}_+(y) \ll (1 + |t|)^{-4} (\sin \pi t)^{-1} Y^{1/2} |Y^{it}| \sum_{m \geq 0} \frac{Y^{-2m}}{m! |\Gamma(m + 1 + it)|}$$

Because of the recurrence relation for Γ, we have

$$\Gamma(m+1+it) = (m+it)...(1+it)\Gamma(1+it) \gg m!|\Gamma(1+it)|$$

if $t \in \mathbb{R}$. That implies

$$\sum_{m\geq 0} \frac{Y^{-2m}}{m!|\Gamma(m+1+it)|} \ll \frac{1}{|\Gamma(1+it)|} \sum_{m\geq 0} \frac{Y^{-2m}}{m!} = e^{Y^{-2}} \frac{1}{|\Gamma(1+it)|}$$

Since also $|\Gamma(1+it)| = \sqrt{\frac{\pi t}{\sinh \pi t}}$, we have finally

$$\tilde{F}_+(y) \ll (1+|t|)^{-4} \sqrt{\frac{Y}{\sinh \pi t}} |Y^{it}|$$

From that we can get the result for $|t| \geq 1/4$ (for example). For $|t|$ close to zero, we use the bound

$$K_{it}(y) = -\log(y/2) + \mathcal{O}(1)$$

for $|t| \leq 1/4$. Which gives us the bound with factor $\log(Y+2)$. □

We're next going to use a Cauchy inequality to help estimating P_h. What we do is introduce a function $\iota(t)$ defined in the domain $(\frac{1}{4} + t^2 > 0)$ and obtain

$$|P_h(z)| \leq K_d(z)^{1/2} R_d(h)^{1/2}.$$

Where

$$K_d(z) = \sum \iota(t_j)|u_j(z)|^2 + \frac{1}{4\pi} \sum_{\mathfrak{a}} \int_{\mathbb{R}} \iota(t)|E_{\mathfrak{a}}(z, \tfrac{1}{2}+it)|^2 dt$$

and

$$R_d(h) = \sum \iota(t_j)^{-1}|(P_h|u_j)|^2 + \frac{1}{4\pi} \sum_{\mathfrak{a}} \int_{\mathbb{R}} \iota(t)^{-1}|(P_h|E_{\mathfrak{a}}(., \tfrac{1}{2}+it))|^2 dt$$

Now we have to estimate both K_d and R_d. First we consider K_d. For that we're gonna use the techniques of the Selberg trace formula. More precisely we have

Proposition 5.4. Let h be a holomorphic function in a strip $\{\Im(z) < \frac{1}{2} + \delta\}$ for some δ such that h is even and $h(t) \ll (1+|t|)^{-2-\delta}$. We consider further the following transforms

$$g(r) = \frac{1}{2\pi} \int_{\mathbb{R}} h(t) e^{irt} dt \quad \text{(Fourier Transform)}$$

$$q(v) = \frac{1}{2} g(\sqrt{v+1} - \sqrt{v}), v \geq 1$$

$$k(u) = -\frac{1}{\pi} \int_u^{+\infty} \frac{dq(v)}{\sqrt{v-u}}, u \geq 0$$

We then have

$$\sum_{\gamma \in \Gamma_0(d)} k(u(z, \gamma w)) = \sum_{j\geq 0} h(t_j) u_j(z) \overline{u_j(w)} + \frac{1}{4\pi} \sum_{\mathfrak{a}} \int_{\mathbb{R}} h(t) E_{\mathfrak{a}}(z, ...) \overline{E_{\mathfrak{a}}(w, ...)} dt$$

For every $z, w \in \mathbb{H}$. Where the series is absolutely and uniformly convergent in every compact.

$$u(z,w) = \frac{|z-w|^2}{4\Im(z)\Im(w)}.$$

This proposition becomes useful to us as soon as we choose $h(t) = \iota(t) = \frac{1}{1+t^2} - \frac{1}{4+t^2}$ which not only satisfies the conditions of the proposition but also gives us a positive transform k because we have $g(r) = e^{-r} - \frac{e^{-2r}}{2}$ which is positive and non increasing for $r \geq 0$. Finally, since $\Gamma_0(d) \subset \Gamma_0(1)$ it gives us $0 \leq \sum_{\gamma \in \Gamma_0(d)} \ldots \leq \sum_{\gamma \in \Gamma_0(1)} \ldots = K_1(z)$. In other words, we eliminate the dependence on d. Moreover K_1 is modular over $SL(2,\mathbb{Z})$ so that $K_d(\tau z) \leq K_1(\tau z) = K_1(z)$. What we have here is that $K_d(z) = \mathcal{O}(1)$.

Now, for R_d we have

$$R_d(h) \leq hY \left\{ \sum_{j \geq 1} \eta(t_j) \frac{|b_j(h)|^2}{\cosh \pi t_j} + \frac{1}{4\pi} \sum_{\mathfrak{a}} \int_{\mathbb{R}} \eta(t) \frac{|\tau_{\mathfrak{a}}|^2}{\cosh \pi t} dt \right\}.$$

where $\eta(t) = (Y^{2it} + Y^{-2it} + 4\log(Y+2)^2)\iota(t)$ and this inequality is true because of the bounding for $\tilde{F}$ above. To complete the estimate, we have

Theorem 5.5.

$$\sum_{j \geq 1} \eta(t_j) \frac{|b_j(h)|^2}{\cosh \pi t_j} + \frac{1}{4\pi} \sum_{\mathfrak{a}} \int_{\mathbb{R}} \eta(t) \frac{|\tau_{\mathfrak{a}}|^2}{\cosh \pi t} dt = f_0 + \sum_{d|c} c^{-1} S(h,h;c) f\left(\frac{4\pi h}{c}\right),$$

where

$$f_0 = \frac{1}{\pi} \int_{\mathbb{R}} \eta(t) \tanh(\pi t) t dt$$

$$f(x) = 2i \int_{\mathbb{R}} J_{2it}(x) \frac{\eta(t)}{\cosh(\pi t)} t dt$$

$$S(a,b;c) = \sum_{x \pmod{c}^*} e\left(\frac{ax + b\overline{x}}{c}\right)$$

and $J_\nu(z) = \sum_{m \geq 0} \frac{(-1)^m (z/2)^{\nu+2m}}{m!\Gamma(m+1+\nu)}$

Sketch of proof. The idea is considering

$$U_m(z,w) = \sum_{\gamma \in B \backslash \Gamma_0(d)} (4\pi m \Im(\gamma z))^w e(m\Re(\gamma z))$$

and its spectral decomposition. □

Before we proceed, we shall state two more estimates which are going to be needed in the next theorem:

- $J_\nu(y) \ll e^{\pi|\Im(\nu)|} y^{\Re(\nu)}$ for all $y > 0$ and $0 \leq \Re(\nu) \leq 2$.
- $|S(a,b;c)| \leq (a,b,c)^{1/2} \tau(c) c^{1/2}$

The first one comes from the Stirling's formula for the Γ function. The second one is the Weil's bound for Kloosterman's sums.

With all those things in mind, we can finally state the main theorem of this section.

Theorem 5.6. (1) Let $D > 0$ be fixed. We have

$$R_d(h) \leq (hY)(\log(Y+2))^2\{1 + (hY)^{1/2}d^{-1}\tau(d)\tau(h)^2(h,d)^{1/2}\}$$

for every $d \geq 1, h \geq 1$, the implicit constant being absolute.

(2) Let $D > 0$ and $d \geq 1$ integers without square factors. We have

$$\sum_{n \equiv 0 \,(\mathrm{mod}\ n)} \rho_h(n) F\left(\frac{2\pi h\sqrt{D}}{n}\right) \ll_D (hY)^{1/2}\{1+(hY)^{1/4}d^{-1}(h,d)^{1/4}\}(\log(Y+2))^2\tau(d)^2\tau(h)$$

For every F flat function of parameter Y and every $h \geq 1$, the implicit constant depending only on D.

Proof. (2) is a consequence of what we've done in this chapter, and the final considerations of section 3. For (1), by the previous theorem, we first notice that by the formula for f_0, we have

$$f_0 \ll \log(Y+2)^2 \ll \log(Y+2)^4.$$

We then have to estimate $f(x)$. We write $s = 1/2 + it$ and then

$$f(x) = \int_{(1/2)} J_{2(s-1/2)}(x)\frac{\eta(i(s-1/2))}{\cos\pi(s-1/2)}(s-1/2)ds$$

We majorate the terms in the integrand in the strip $\{1/2 \leq \sigma < 1\}$ with the aid of the bound for J_ν mentioned in the remark above. We thus obtain that the integrand in $f(x)$, $I(x,t)$, satisfies $I(x,t) \ll x^{2\sigma-1}(Y^{2\sigma-1})(1+|s|)^{-3}$. Thus, $I(x,t)$ has a polynomial growth, which tells us we can actually change the line of integration, giving us

$$f(x) \ll x^{2\sigma-1}(Y^{2\sigma-1} + Y^{1-2\sigma} + 4\log(Y+2)^2)$$

But since $Y \geq 1$ and $4(\log(Y+2))^2 > 1$, we have, actually, $f(x) \ll (xY)^{2\sigma-1}\log(Y+2)^2$. Together with the Weil's bound:

$$\begin{aligned}\sum_{d|c} c^{-1}S(h,h;c)f\left(\frac{4\pi h}{c}\right) &\ll (hY)^{2\sigma-1}(\log(Y+2)^2)\sum_{l\geq 1}(ld)^{1/2-2\sigma}(h,ld)^{1/2}\tau(ld) \\ &< (hY)^{2\sigma-1}(\log(Y+2)^2)d^{1/2-2\sigma}\tau(d)(h,d)\sum_{l\geq 1}(h,l)\tau(l)l^{1/2-2\sigma}\end{aligned}$$

To evaluate the $l-$sum, we take $\sigma > 3/4$ so that $(2\sigma - 1/2 > 1)$ and simply use the formula

$(h,l) = \sum_{\delta|(h,l)} \phi(\delta) = \sum_{\delta|h} \chi_{\delta|l}\phi(\delta)$, which gives us

$$\begin{aligned}\sum_{l\geq 1}(h,l)\tau(l)l^{1/2-2\sigma} &= \sum_{\delta|h}\sum_{l\geq 1}\phi(\delta)\chi_{\delta|l}\tau(l)^{1/2-2\sigma}\\ &= \sum_{\delta|h}\phi(\delta)\sum_{m\geq 1}\tau(\delta m)(\delta m)^{1/2-2\sigma}\\ &\leq \sum_{\delta|h}\phi(\delta)\delta^{1/2-2\sigma}\zeta(2\sigma-1/2)^2\\ &\leq \zeta(2\sigma-1/2)^2\sum_{\delta|h}\tau(\delta) \leq \zeta(2\sigma-1/2)^2\tau(h)^2.\end{aligned}$$

Since, $\zeta(1+h) \ll \frac{1}{h}$ in a bounded interval $0 < h \leq A < \infty$, then, for $3/4 < \sigma < 1$,

$$\sum_{d|c}\frac{1}{c}S(h,h;c)f\left(\frac{4\pi h}{c}\right) \ll \frac{(hY)^{2\sigma-1}(\log(Y+2))^2}{(2\sigma-1/2)^2}\tau(h)^2 d^{1/2-2\sigma}\tau(d)(h,d)^{1/2}.$$

We conclude by taking $\sigma = \frac{3}{4} + \frac{C}{\log hY}$ with C small enough. □

6 Conclusion

Here, we can finally apply the estimatives of last chapter and the sieve methods from section 2 to deduce the main theorem. Let $D > 0$ be a fixed integer and $h \neq 0$. We'd like to estimate the Weyl's sums

$$S_{h,D}(x) = \sum_{p\leq x}\rho_h(p) = \sum_{p\leq x}\sum_{\nu^2+D\equiv 0 \pmod p} e\left(\frac{h\nu}{p}\right)$$

We can reduce ourselves to the case where D has no square factors because of the following inequality

$$S_{h,m^2D}(x) = S_{mh,D}(x) + \mathcal{O}(1)$$

for $m > 0$, where the $\mathcal{O}(1)$ part comes from the p's dividing m. Which are, in fact, finite. Since what we've got with the Poincaré series is not yet the above sum above but with some test functions involved. It remains yet to do some technical work. For the use of the sieve methods, we're going to need some bounds. For that we have

Lemma 6.1. $\rho_h(n) \leq \tau(n)$.

Proof. We have the trivial bound

$$\rho_{h,D}(n) \leq U(n) = \{x \pmod n \mid x^2 + D \equiv 0 \pmod n\}$$

Since the function in the right is multiplicative, as well as the τ function, we can prove only for $n = p^k$. Suppose you have $x^2 \equiv y^2 \equiv -D \pmod{p^k} \Rightarrow p \mid (x+y)(x-y)$. We have two cases: If $p \neq 2$, p dividing both would imply $d \mid x$, but we know $p \nmid D$. So we need p^k dividing either $(x-y)$ or $(x+y)$ which implies $x \equiv \pm y \pmod{p^k}$ which means

only two solutions. Now suppose $p = 2$ and $k \geq 3$ (for 2 and 4 the result is immediate). We then have again $2^k|(x+y)(x-y)$. Since the difference is $2y$ which is not a multiple of 4, one of them can have at most one factor 2 dividing it. So that we should have 2^k dividing either $(x+y)$ or $(x-y)$ which implies $y = x$, $x + 2^k$, x or $2^k - x$ (4 solutions). Since $\tau(2^k) = 2k + 1 > 4$ if $k \geq 3$, we're done. □

Proposition 6.2. Let $h \neq 0$ be a fixed integer, $x \geq 2$ and $g : (0, +\infty) \to [0, 1]$ a C^∞ function such that $\text{supp}(g) \subset [1/2, 1]$ and for $0 \leq j \leq 4$ and $x \in \mathbb{R}$ we have

$$|g^{(j)}(x)| \leq 1.$$

Then we have

$$\sum_p \rho_h(p) g\left(\frac{p}{x}\right)$$

uniformly in terms of g, which is to say that, for every $\epsilon > 0$, there is x_0 depending only on h, D and ϵ such that for $x \geq x_0$,

$$\left|\sum_{p \leq x} \rho_h(p) g\left(\frac{p}{x}\right)\right| \leq \frac{\epsilon x}{\log x}.$$

Proof. To prove this proposition, we will use the final corollary from Section 2 to the sequence

$$a_n = \rho_h(n) g\left(\frac{n}{x}\right)$$

where g is a function as in the proposition. We certainly have $|a_n| \leq \tau(n)$ so that what is left is to prove the estimates for linear forms $R(x)$ and bilinear forms $B(x)$. The uniformity in g comes from the way the corollary was conceived.

Linear Forms

We write

$$R(x) = \sum_{\substack{d \leq y \\ md \leq x}} \alpha_d \sum_m \rho_d(md) g\left(\frac{md}{x}\right).$$

Clearly, we have

$$|R(x)| \leq \sum_{d \leq y} \left|\sum_{d|m} \rho_d(m) g(m/x)\right|,$$

since $|\alpha_d| \leq 1$. The sum inside the absolute value can be seen as a sum which can be estimated using the final theorem of the previous section with F given by

$$F(y) = g\left(\frac{2\pi h\sqrt{D}}{xy}\right).$$

The good thing here is that $g(x)$ is flat of parameter 2 (because $1 \leq 2^i$) and because of the theorem about flat functions, there is a constant c independent of g such that $c \cdot F$ is also flat of parameter $Y = \frac{x}{2\pi h\sqrt{D}}$. Consequently, by the final theorem from the previous section,

$$\sum_{m \equiv 0 \,(\mathrm{mod}\ d)} \rho_h(m) g(m/x) \ll_{D,h} (x^{1/2} + x^{3/4} d^{1/2})(\log x)^2 \tau(d)^2,$$

the constant depending only on D and h. Adding up these estimates for $d \leq y$, one obtains

$$R(x) \ll_{D,h} x^{1/2}(\log x)^2 y (\log y)^3 + y^{1/2} x^{3/4} (\log x)^2 y^{1/2} (\log y)^3.$$

The above bounds for the sums with the divisor functions can be obtained by technique of shifting the integration line. Now if we take $y = x^{1/2-\epsilon}$,

$$R(x) \ll_{D,h} x^{1-\epsilon/2} (\log x)^5$$

which is already a satisfactory bound.

Bilinear forms

We start from

$$B(x) = \sum_{\substack{mn \leq x \\ w \leq n \leq v}} \beta_n \gamma_m a_{mn}.$$

Remember that β_n has support on the prime numbers. This will be of extreme importance in the following steps.

Lemma 6.3. Let $w = \exp\left(\frac{\epsilon \log x}{10 \log\log x}\right)$. Then

$$|B(x)| \leq B^*(x) + \mathcal{O}\left(\frac{x}{(\log x)^A}\right)$$

for every $A > 0$, the constant depending on A, where

$$B^* = \sum_m \left| \sum_{w \leq n < v} \beta_n \rho_n(mn) g(mn/x) \right|$$

This lemma is telling us that we can 'forget' about the case $(m, n) \neq 1$ (i.e. $n|m$, because $\beta_n = 0$ otherwise).

Proof. As we observed, the remainder is

$$\begin{aligned}
\sum_m \left| \sum_{w \leq n < v} \beta_n a_{mn} \right| &\leq \sum_{w \leq n < v} \sum_{m \leq x n^{-2}} \tau(mn^2) \\
&\leq 3 \sum_{w \leq n < v} \sum_{m \leq x n^{-2}} \tau(m) \\
&\leq x(\log x) \sum_{w \leq n < v} \frac{1}{n^2} \\
&\leq x(\log x) w^{-1}.
\end{aligned}$$

where once again we use the bound for sums of function obtained the same way. Since $w \gg_A (\log x)^A$ for every $A > 0$ the proof is finished. □

For $B^*(x)$, because of the condition $(m,n) = 1$, a root of $P(\nu) \equiv 0 \pmod{mn}$ is equivalent, by the Chinese Remainder Theorem to a pair (δ, ν) where $P(\delta) \equiv 0 \pmod{m}$ and $P(\nu) \equiv 0 \pmod{n}$. Now, with this in mind, we observe that

$$B^*(x) \leq \sum_m \left| \sum_{P(\delta)\equiv 0 \pmod m} \sum_{\substack{w\leq n<v \\ (m,n)=1}} \beta_n g\left(\frac{mn}{x}\right) \sum_{\substack{P(\nu)\equiv 0 \pmod{mn} \\ \nu\equiv\delta \pmod{mn}}} e\left(\frac{h\nu}{mn}\right) \right|$$

We next use the Cauchy-Schwarz inequality twice but before we add a term $\frac{1}{\sqrt{m}} \cdot \sqrt{m}$, which gives us

$$\begin{aligned} B^*(x)^2 &\leq \left(\sum_m \frac{1}{m}\right) \sum_m m \left| \sum_{P(\delta)\equiv 0 \pmod m} \sum_{\substack{w\leq n<v \\ (m,n)=1}} \beta_n g\left(\frac{mn}{x}\right) \sum_{\substack{P(\nu)\equiv 0 \pmod{mn} \\ \nu\equiv\delta \pmod{mn}}} e\left(\frac{h\nu}{mn}\right) \right|^2 \\ &\leq (\log x) \sum_m \tau(m) m \sum_{P(\delta)\equiv 0 \pmod m} \left| \sum_{\substack{w\leq n<v \\ (m,n)=1}} \beta_n g\left(\frac{mn}{x}\right) \sum_{\substack{P(\nu)\equiv 0 \pmod{mn} \\ \nu\equiv\delta \pmod{mn}}} e\left(\frac{h\nu}{mn}\right) \right|^2 \\ &< (\log x) \sum_{n_1,n_2} \beta_{n_1}\overline{\beta_{n_2}} \sum_{(m,n_1n_2)} \tau(m) m g\left(\frac{mn_1}{x}\right) g\left(\frac{mn_2}{x}\right) \\ &\qquad \times \sum_{P(\delta)\equiv 0 \pmod m} \sum_{\substack{P(\nu_i)\equiv 0 \pmod{mn_i} \\ \nu_i=\delta \pmod{n_i}}} e\left(\frac{h\nu_1}{mn_1} - \frac{h\nu_2}{mn_2}\right). \end{aligned}$$

Lemma 6.4. The contribution Δ of the diagonal terms $(n_1 = n_2)$ satisifies

$$\Delta \ll x^2 (\log x)^4 w^{-1}$$

Proof. Take $n = n_1 = n_2$. The sum over δ and ν_i is bounded by $\tau(m)\tau(n)^2$ and

$$\begin{aligned} \Delta &\leq (\log x) \sum_{w\leq n<v} |\beta_n|^2 \sum_{(m,n)} m\tau(m) g\left|\left(\frac{mn}{x}\right)\right|^2 \tau(m)\tau(n)^2 \\ &\ll \sum_{w\leq n<v} \frac{x}{n} \sum_{m\leq x/n} \tau(m)^2 \\ &< (\log x) \sum_{w\leq n<v} n^{-2} (\log x)^3 \ll x^2 (\log x)^4 w^{-1}. \end{aligned}$$

For the sum of $\tau(m)^2$ we use again the technique of shifting the integration line. □

Now we consider the other case $n_1 \neq n_2$. Since n_1, n_2 are primes, we have $(n_1, n_2) = 1$. Let's denote the corresponding sum by $C^*(n_1, n_2)$. Since $(n_1, n_2) = 1$ the triplet (δ, ν_1, ν_2)

corresponds to a root $\nu \pmod{mn_1n_2}$. More precisely

$$\sum_{P(\delta)\equiv 0 \pmod m} \sum_{\substack{P(\nu_i)\equiv 0 \pmod{mn_i} \\ \nu_i \equiv \delta \pmod m}} e\left(\frac{h\nu_1}{mn_1} - \frac{h\nu_2}{mn_2}\right) = \sum_{P(\nu)\equiv 0 \pmod{mn_1n_2}} e\left(\frac{h(n_2-n_1)\nu}{mn_1n_2}\right)$$
$$= \rho_{h(n_2-n_1)}(mn_1n_2),$$

so that

$$C^*(n_1,n_2) = (\log x) \sum_{(m,n_1n_2)} \tau(m) m g\left(\frac{mn_1}{x}\right) g\left(\frac{mn_2}{x}\right) \rho_{h(n_2-n_1)}(mn_1n_2)$$
$$= (\log x) \sum_{(m,n_1n_2)} \tau(m) j\left(\frac{mn_1n_2}{x}\right) \rho_{h(n_2-n_1)}(mn_1n_2)$$

where

$$j(y) = \frac{xy}{n_1n_2} g\left(\frac{y}{n_2}\right) g\left(\frac{y}{n_1}\right).$$

Right now, we almost have a Poincaré series, we just have to add again the terms where m is not coprime with the n_i's.

Lemma 6.5. Suppose $n_1 < n_2$ we have

$$C^*(n_1,n_2) = C(n_1,n_2) + \mathcal{O}(x^2(\log x)^4 n_1^{-3})$$

where

$$C(n_1,n_2) = (\log x) \sum_{n_1n_2|m} \tau\left(\frac{m}{n_1n_2}\right) j\left(\frac{m}{x}\right) \rho_{h(n_2-n_1)}(m).$$

Proof. The remainder is bounded by

$$(\log x)\left(\sum_{m\equiv 0 \pmod{n_1}} \tau(m) j\left(\frac{mn_1n_2}{x}\right) \rho_{h(n_2-n_1)}(m) + \sum_{m\equiv 0 \pmod{n_2}} \tau(m) j\left(\frac{mn_1n_2}{x}\right) \rho_{h(n_2-n_1)}(m)\right).$$

We can estimate both terms the same way:

$$\left|\sum_{m\equiv 0 \pmod n} \tau(m) j\left(\frac{mn_1n_2}{x}\right) \rho_{h(n_2-n_1)}(m)\right| \le \sum_m \tau(m) mn\ g\left(\frac{mnn_1}{x}\right) g\left(\frac{mnn_2}{x}\right) \tau(mnn_1n_2)$$
$$\ll n \sum_{m\le x(nn_1)^{-1}} m\tau(m)^2$$
$$\ll x^2(\log x)^3 (nn_1^2)^{-1},$$

where $n = n_1$ or n_2. Since $n_1 < n_2$, we are done. □

To calculate the bound for $C(n_1, n_2)$, the first thing we do is to estimate the term $\tau(m)\log x$ by x^{ϵ_1}. So that we have

$$x^{\epsilon_1} \sum_{n_1 n_2 | m} j\left(\frac{m}{x}\right) \rho_{h(n_2-n_1)}(m).$$

The next step is to prove that the function $j(y)$ is flat of some parameter, or at least, a multiple of a flat function. Suppose for the moment that $n_1 \leq n_2$ (the other case is analogous). Notice that $j(y)$ has support in the interval $[\frac{n_2}{2}, n_1] \subset [\frac{\sqrt{n_1 n_2}}{2}, \sqrt{n_1 n_2}]$ if $n_1 \leq n_2 \leq 2n_1$ and vanishes otherwise. In the first case,

- The function $f(y) = \frac{y}{\sqrt{n_1 n_2}}$ satisfies $|f^{(i)}(y)| \leq (n_1 n_2)^{-i/2}$ for every $y \in [\frac{\sqrt{n_1 n_2}}{2}, \sqrt{n_1 n_2}]$ because of a direct computation.
- The function $g_1(y) = g(y/n_1)$ satisfies $g_1^{(i)}(y) \leq n_1^{-i}$ for every $y \in [\frac{\sqrt{n_1 n_2}}{2}, \sqrt{n_1 n_2}]$, $i \leq 4$ because of the bounds for g.
- The function $g_2(y) = g(y/n_2)$ satisfies $g_2^{(i)}(y) \leq n_2^{-i}$ for every $y \in [\frac{\sqrt{n_1 n_2}}{2}, \sqrt{n_1 n_2}]$, $i \leq 4$ because of the bounds for g.

By the Leibniz rule we obtain that if $\tilde{f}(y) = f(y)g_1(y)g_2(y)$, then

$$\begin{aligned}
\tilde{f}^{(i)}(y) &\leq (n_1^{-1} + n_2^{-1} + (n_1 n_2)^{-1/2})^i \\
&= (n_1 n_2)^{-i/2} \left(\sqrt{\frac{n_2}{n_1}} + \sqrt{\frac{n_1}{n_2}} + 1\right)^i \\
&\leq \left(\frac{2+\sqrt{2}}{2}\right)^4 \left(\frac{2}{\sqrt{n_1 n_2}}\right)
\end{aligned}$$

for $i \leq 4$. Now, we write

$$\tilde{F}(u) = j\left(\frac{2\pi\tilde{h}\sqrt{D}}{xu}\right)$$

where $\tilde{h} = h(n_2 - n_1)$. By the lemma about flat functions, we find out that $\tilde{F}$ is of the form

$$c\frac{x}{\sqrt{n_1 n_2}} k(u) \quad \text{where}$$

where c is an absolute constant and k is flat of parameter $\frac{x\sqrt{n_1 n_2}}{4\pi\tilde{h}\sqrt{D}}$. The estimates of the final theorem from the last section gives us

$$\begin{aligned}
C(n_1, n_2) &\ll_D \frac{x^{1+\epsilon_1}}{\sqrt{n_1 n_2}} \{x^{1/2}(n_1 n_2)^{1/4} + (\tilde{h}, \tilde{d})^{1/4} x^{3/4} (n_1 n_2)^{-1/8}\} (\log x)^2 \tau(h)\tau(n_2 - n_1)\tau(n_1 n_2)^2 \\
&\ll_D \tau(h) x^{2\epsilon_1} (x^{3/2}(n_1 n_2)^{-1/4} + h^{1/4} x^{7/4} (n_1 n_2)^{-5/8})
\end{aligned}$$

Here used the estimates $\tau(n_2 - n_1)\tau(n_1 n_2)^2 (\log x)^2 \ll x^{\epsilon_1}$ because $n_2 - n_1$, $n_1 n_2 \leq x^2$ and $(\tilde{h}, n_1 n_2) = (h, n_1 n_2) \leq h$ because n_1 and n_2 are distinct primes.

Now we sum over all $n_1 \neq n_2$ and we get that the first summand is bounded by

$$x^{3/2+2\epsilon_1} \left| \sum_{n \leq v} n^{-1/4} \right|^2 \ll x^{3/2+2\epsilon_1} v^{3/2} \leq x^{2+2\epsilon_1 - 3\epsilon/2}$$

and the second by

$$x^{7/4+2\epsilon_1}\left|\sum_{n\leq v} n^{-5/8}\right|^2 \ll x^{7/4+2\epsilon_1}v^{3/4} \leq x^{2+2\epsilon_1-3\epsilon/4},$$

where the implicit constant depends on D, ϵ_1 and h. If we take $2\epsilon_1 < \epsilon/2$ we get that both bounds above can be taken as $x^{2-\epsilon/4}$.

It still remains to compute the terms in the passage from $C^*(n_1, n_2)$ to $C(n_1, n_2)$. Once again we remember that these terms are only non zero if $\frac{1}{2} \leq \frac{n_1}{n_2} \leq 2$, so that the term $\mathcal{O}(x^2(\log x)^2 n^{-3})$ appears at most $4n$ times, leading to a term

$$\mathcal{O}\left(x^2(\log x)^2 \sum_{w\leq n} 4n^{-2}\right) \ll x^2(\log x)^2 w^{-1}.$$

Putting all the pieces together:

$$|B^*(x)| \ll x^2(\log x)^3 w^{-1} + x^{2-\epsilon/4} + x^2(\log x)^2 w^{-1}$$

Since $w \gg (\log x)^A$ for any $A > 0$, we can actually get

$$|B^*(x)| \ll \frac{x}{(\log x)^A},$$

also for any $A > 0$. Our early estimate shows that we can get the same bound for $|B(x)|$ instead of $|B^*(x)|$, which finishes our estimation of the bilinear forms and enables us to prove the last proposition. □

Proof (of the proposition). Here we use the final corollary of Section 2. The bounds that we obtained are even better than what is needed. □

At last, we can give a proof of the main theorem.

Proof (of the main theorem). As we already know, by the Weyl's criterion, we just have to prove

$$\sum_{p\leq x} \rho_h(p) = o\left(\frac{x}{\log x}\right).$$

We fix $0 < \delta < 1$ and take ψ to be a smooth majorant of the indicator function for the interval $[1, x]$ with support on $[0, (1+\delta)x]$. We have, for $x \geq 2$,

$$\sum_{p\leq x} \rho_h(p) = \sum_{p\leq x} \rho_h(p)\psi(p) = \mathcal{O}(\pi((1+\delta)x) - \pi(x)) = \sum_{p} \rho_h(p)\psi(p) + \mathcal{O}\left(\frac{\delta x}{\log x}\right)$$

with an absolute implicit constant, by the Prime Number Theorem. Now we make a partition of unity, obtaining a sequence of functions ψ_N, $0 \leq \psi_N \leq 1$ such that each ψ_N is flat with compact support in $[\frac{(1+\delta)x}{2^{N+1}}, \frac{(1+\delta)x}{2^N}]$ and

$$\psi = \sum_N \psi_N.$$

The number of values of N is $\ll \log x$ because after that we get

$$\frac{(1+\delta)x}{2^N} < 1.$$

The contribution from the N such that $2^{-N}(1+\delta)x < \sqrt{x}$ is clearly $< \sqrt{x}$. For all the remaining ones, we get that the function g given by

$$g(y) = \psi_N\left(\frac{(1+\delta)x}{2^N y}\right)$$

satisfies (up to multiplying by an absolute constant c) the condition in the last proposition. We then get

$$\left|\sum_{p\leq x} \rho_h(p)\psi_N(p)\right| \leq \frac{c\epsilon(1+\delta)x}{2^N \log x}$$

for $x \geq x_0^2$. Summing over N, we obtain

$$\left|\sum_{p\leq x} \rho_h(p)\psi_N(p)\right| \leq \frac{4c\epsilon x}{\log x}.$$

Thus getting

$$\left|\sum_{p\leq x} \rho_h(p)\right| \leq \frac{4c\epsilon x}{\log x} + \mathcal{O}\left(\frac{\delta x}{\log x}\right) + \sqrt{x}.$$

Since δ and ϵ can be taken arbitrarily small, we have the result. □

References

[DFI] W. Duke, J. B. Friedlander and H. Iwaniec, *Equidistribution of roots of a quadratic congruence to prime moduli*, Annals of Mathematics, 141 (1995), 423-441.

[I] H. Iwaniec, *Spectral Methods of Automorphic Forms* (2nd ed.), American Mathematical Society. ISBN 978-0-8218-3160-1.

[K] E. Kowalski, *Un cours de théorie analytique des nombres*, SMF, Cours spécialisés, ISBN: 2-85629-161-9.

[S] C. L. Siegel, *Some remarks on discontinuous groups*, Annals of Mathematics, 46 (1945), p. 708-718.