

Algebra & Number Theory

Volume 12
2018
No. 3

Mean square in the prime geodesic theorem

Giacomo Cherubini and João Guerreiro



Mean square in the prime geodesic theorem

Giacomo Cherubini and João Guerreiro

We prove upper bounds for the mean square of the remainder in the prime geodesic theorem, for every cofinite Fuchsian group, which improve on average on the best known pointwise bounds. The proof relies on the Selberg trace formula. For the modular group we prove a refined upper bound by using the Kuznetsov trace formula.

1. Introduction

There is a striking similarity between the distribution of lengths of primitive closed geodesics on the modular surface and prime numbers. If we consider the asymptotic of the associated counting function, the problem of finding optimal upper bounds on the remainder is intriguing, especially because the relevant zeta function is known to satisfy the corresponding Riemann hypothesis.

We start with reviewing briefly the framework of the problem (for a more detailed introduction we refer to [Sarnak 1980; Iwaniec 1995, §10.9]), and then state our results. Since the definitions extend to every finite volume Riemann surface, we work in this generality, and only at a later point do we specialize to the modular surface.

Every cofinite Fuchsian group Γ acts on the hyperbolic plane \mathbb{H} by linear fractional transformations, and every hyperbolic element $g \in \Gamma$ is conjugated over $\mathrm{SL}_2(\mathbb{R})$ to a matrix of the form

$$\begin{pmatrix} \lambda^{1/2} & 0 \\ 0 & \lambda^{-1/2} \end{pmatrix}, \quad \lambda \in \mathbb{R}, \lambda > 1.$$

The trace of g is therefore $\mathrm{Tr}(g) = \lambda^{1/2} + \lambda^{-1/2}$, and we define its norm to be $N(g) = \lambda$. Since the trace and the norm are constant on the conjugacy class of g , if we set $P = \{\gamma g \gamma^{-1}, \gamma \in \Gamma\}$, we can define the trace and the norm of P by $\mathrm{Tr}(P) = \mathrm{Tr}(g)$ and $NP = N(g)$. Moreover, we say that g (and P) is primitive if g cannot be expressed as a positive power (greater than one) of another element $g_0 \in \Gamma$. Every class P can then be written as $P = P_0^\nu$ for some primitive conjugacy class P_0 and some positive integer $\nu \geq 1$.

For $X > 0$ define the counting function

$$\pi_\Gamma(X) = \#\{P_0 : NP_0 \leq X\}. \tag{1-1}$$

This work was supported by the Max Planck Institut für Mathematik. We would like to record our thanks to the Institute for the excellent working conditions.

MSC2010: primary 11F72; secondary 11L05, 11M36.

Keywords: prime geodesic theorem, Selberg trace formula, Kuznetsov trace formula, Kloosterman sums.

A hyperbolic conjugacy class P determines a closed geodesic on the Riemann surface $\Gamma \backslash \mathbb{H}$, and its length is given exactly by $\log NP$. For this reason we say that $\pi_\Gamma(X)$ in (1-1) counts primitive hyperbolic conjugacy classes in Γ , lengths of primitive geodesics on $\Gamma \backslash \mathbb{H}$, or, by abusing language, “prime” geodesics.

It is a result from the middle of the last century (see e.g., Huber [1961], although the result was already known to Selberg) that as $X \rightarrow \infty$ we have the asymptotic formula

$$\pi_\Gamma(X) \sim \text{li}(X), \quad (1-2)$$

where $\text{li}(X)$ is the logarithmic integral. The asymptotic coincides with that of the prime counting function, since we have

$$\pi(X) = \#\{p \leq X : p \text{ prime}\} \sim \text{li}(X) \quad \text{as } X \rightarrow \infty. \quad (1-3)$$

The structure of the two problems is also similar. The study of primes is strictly related to the study of the zeros of the Riemann zeta function $\zeta(s)$; on the other hand, for primitive geodesics we need to study the zeros of the Selberg zeta function $Z_\Gamma(s)$, which is defined for $\Re(s) > 1$ by

$$Z_\Gamma(s) = \prod_{P_0} \prod_{\nu=0}^{\infty} (1 - (NP_0)^{-\nu-s}),$$

the outer product ranging over primitive hyperbolic classes in Γ , and extends to a meromorphic function on the complex plane.

It is an interesting question to determine the correct rate of the approximation in (1-2) and (1-3), namely to prove optimal upper bounds on the differences

$$\pi(X) - \text{li}(X) \quad \text{and} \quad \pi_\Gamma(X) - \text{li}(X). \quad (1-4)$$

In the case of primes, the conjectural estimate $|\pi(X) - \text{li}(X)| \ll X^{1/2} \log X$ is equivalent to the Riemann hypothesis (see e.g., [Ivić 1985, Theorem 12.3]).

To understand what type of control we should expect on the difference on the right in (1-4), we introduce the weighted counting function

$$\psi_\Gamma(X) = \sum_{NP \leq X} \Lambda_\Gamma(NP),$$

where $\Lambda_\Gamma(NP) = \log(NP_0)$ if P is a power of the primitive conjugacy class P_0 and $\Lambda_\Gamma(x) = 0$ otherwise. The function $\psi_\Gamma(X)$ is the analogous of the classical summatory von Mangoldt function $\psi(X)$ in the theory of primes, and studying $\pi_\Gamma(X)$ is equivalent to study $\psi_\Gamma(X)$, but it is easier to work with the latter.

The asymptotic (1-2) for $\pi_\Gamma(X)$ translates into $\psi_\Gamma(X) \sim X$ as X tends to infinity. The spectral theory of automorphic forms provides a finite number of additional secondary terms and we define the complete

main term $M_\Gamma(X)$ to be the function

$$M_\Gamma(X) = \sum_{\frac{1}{2} < s_j \leq 1} \frac{X^{s_j}}{s_j}, \tag{1-5}$$

where s_j is defined by $\lambda_j = s_j(1 - s_j)$, λ_j are the eigenvalues of the Laplace operator Δ acting on $L^2(\Gamma \backslash \mathbb{H})$, and the sum is restricted to the *small eigenvalues*, that is, those satisfying $0 \leq \lambda_j < \frac{1}{4}$. Since the eigenvalues of Δ form a discrete set with no accumulation points, the sum in (1-5) is finite. The remainder $P_\Gamma(X)$ is then defined as

$$P_\Gamma(X) = \psi_\Gamma(X) - M_\Gamma(X).$$

Various types of pointwise upper bounds have been proved in the past years on $P_\Gamma(X)$, and we review them in Remark 1.5 below. In a different direction, it is possible to study the moments of $P_\Gamma(X)$, and this has been neglected so far.

In the case of the prime number theorem, it is a classical result (assuming the Riemann hypothesis) that the normalized remainder

$$\frac{\psi(e^y) - e^y}{e^{y/2}}$$

admits moments of every order and a limiting distribution with exponentially small tails (see [Wintner 1935, p. 242] and [Rubinstein and Sarnak 1994, Theorem 1.2]). The aim of this paper is to prove the following estimate on the second moment of $P_\Gamma(X)$.

Theorem 1.1. *Let Γ be a cofinite Fuchsian group. For $A \gg 1$ we have*

$$\frac{1}{A} \int_A^{2A} |P_\Gamma(X)|^2 dX \ll A^{4/3}.$$

Remark 1.2. By construction the zeros of the Selberg zeta function correspond to the eigenvalues of Δ . The zeros corresponding to small eigenvalues give rise to the secondary terms in the definition of $M_\Gamma(X)$, and there are no other zeros in the half plane $\Re(s) > \frac{1}{2}$. In this sense $Z_\Gamma(s)$ satisfies an analogous of the Riemann hypothesis, and the fact suggests that we might have

$$P_\Gamma(X) \ll X^{1/2+\varepsilon} \quad \forall \varepsilon > 0, \tag{1-6}$$

since this is the case for the primes under RH. However, the same method of proof fails, due to the abundance of zeros of $Z_\Gamma(s)$ (which in turn is related to the fact that $Z_\Gamma(s)$ is a function of order two, while $\zeta(s)$ is of order one), and only gives

$$P_\Gamma(X) \ll X^{3/4}. \tag{1-7}$$

For a proof we refer to [Iwaniec 1995, Theorem 10.5]. On the other hand, it is possible to prove that (1-6), if true, is optimal, since we have the Omega result by Hejhal [1983, Theorem 3.8, p. 477, and note 18, p. 503]

$$P_\Gamma(X) = \Omega_\pm(X^{1/2-\delta}) \quad \forall \delta > 0,$$

which can be strengthened to $\Omega_{\pm}(X^{1/2}(\log \log X)^{1/2})$ for cocompact groups and congruence groups.

Remark 1.3. Recently Koyama [2016] and Avdispahić [2017; 2018] have tried an approach to find upper bounds on $P_{\Gamma}(X)$ using a lemma of Gallagher [1970, Lemma 1]. Their results are of the type

$$P_{\Gamma}(X) \ll X^{\eta+\varepsilon} \quad \forall \varepsilon > 0,$$

where $\eta = \frac{7}{10}$ or $\eta = \frac{3}{4}$ for Γ cocompact or ordinary cofinite, respectively, and with X outside a set $B \subseteq [1, \infty)$ of finite logarithmic measure, that is, such that $\int_B X^{-1} dX < \infty$. Theorem 1.1 improves on this to the extent that η can be taken to be any number greater than $\frac{2}{3}$. Indeed, a direct consequence of Theorem 1.1 is that for every cofinite Fuchsian group and for every $\varepsilon > 0$, the set

$$B = \{X \geq 1 : |P_{\Gamma}(X)| \geq X^{2/3}(\log X)^{1+\varepsilon}\}$$

has finite logarithmic measure. Iwaniec [1984a, p. 187] claims without proof that “it is easy to prove that $P_{\Gamma}(X) \ll X^{2/3}$ for almost all X ”. Theorem 1.1 proves his claim (up to ε) for X outside a set of finite logarithmic measure.

In the case of the modular group we can prove a stronger bound than for the general cofinite case. In this case we can exploit the Kuznetsov trace formula, obtaining the following.

Theorem 1.4. *Let $G = \mathrm{PSL}(2, \mathbb{Z})$. For $A \gg 1$ and every $\varepsilon > 0$,*

$$\frac{1}{A} \int_A^{2A} |P_G(X)|^2 dX \ll A^{5/4+\varepsilon}.$$

Remark 1.5. The estimate in (1-7) is currently the best known pointwise upper bound in the case of general cofinite Fuchsian groups. For the modular group and congruence groups it is possible to prove sharper estimates. Iwaniec [1984b, Theorem 2], Luo and Sarnak [1995, Theorem 1.4], Cai [2002, p. 62], and Soundararajan and Young [2013, Theorem 1.1] have worked on reducing the exponent for the modular group $G = \mathrm{PSL}(2, \mathbb{Z})$. The currently best known result is

$$P_G(X) \ll X^{\eta+\varepsilon}, \quad \text{for } \eta = \frac{25}{36} \text{ and every } \varepsilon > 0,$$

due to Soundararajan and Young. The exponent $\frac{7}{10}$, proved by Luo and Sarnak, holds also for congruence groups, see [Luo et al. 1995, Corollary 1.2], and for cocompact groups coming from quaternion algebras, see [Koyama 1998].

Observe that not only is the exponent $\eta = \frac{1}{2}$ out of reach, but it seems to be a hard problem reaching and, afterwards, breaking the barrier $\eta = \frac{2}{3}$. Iwaniec [1984a, p. 188; 1984b, (12)] suggested that this follows from the assumption of the Lindelöf hypothesis for Rankin L-functions. In Theorem 1.1 we prove the bound with the exponent $\frac{2}{3}$ on average, unconditionally, for every cofinite Fuchsian group. Theorem 1.4 is saying that $P_G(X) \ll X^{5/8+\varepsilon}$ on average. This goes halfway between the trivial bound (1-7) and the conjectural estimate (1-6).

In the proof of Theorem 1.4 we use a truncated formula for $\psi_\Gamma(X)$, proved by Iwaniec [1984b, Lemma 1], condensing the analysis of the Selberg trace formula, and relating $P_\Gamma(X)$ to the spectral exponential sum

$$R(X, T) = \sum_{t_j \leq T} X^{it_j}.$$

The trivial bound for this sum, in view of Weyl’s law (see e.g., [Iwaniec 1995, Corollary 11.2] or [Venkov 1990, Theorem 7.3]), is $|R(X, T)| \ll T^2$. Reducing the exponent of T is possible at the cost of some power of X (see [Iwaniec 1984b; Luo and Sarnak 1995]), and Petridis and Risager conjectured [2017, Conjecture 2.2] that we should have square root cancellation, that is,

$$R(X, T) \ll T^{1+\varepsilon} X^\varepsilon$$

for every $\varepsilon > 0$. This would give the conjectural bound (1-6) for the modular group. In the appendix to [Petridis and Risager 2017] Laaksonen provides numerics that support this conjecture. In the proof of Theorem 1.4 (see Proposition 4.5) we prove that we have

$$\frac{1}{A} \int_A^{2A} |R(X, T)|^2 dX \ll T^{2+\varepsilon} A^{1/4+\varepsilon},$$

and hence $|R(X, T)| \ll T^{1+\varepsilon} X^{1/8+\varepsilon}$ on average.

The mean square estimate in Theorem 1.4 reduces to proving a similar estimate for certain weighted sums of Kloosterman sums. We obtain the desired bounds in a relatively clean way by applying the Hardy–Littlewood–Pólya inequality [Hardy et al. 1934, Theorem 381, p. 288] in a special case: for $0 < \lambda < 1$, $\lambda = 2(1 - p^{-1})$, and $\{a_r\}$ a sequence of nonnegative numbers, then we have

$$\sum_{\substack{r,s \\ r \neq s}} \frac{a_r a_s}{|r - s|^\lambda} \ll_\lambda \left(\sum_r a_r^p \right)^{2/p}. \tag{1-8}$$

With more work it is possible to sharpen Theorem 1.4 by replacing A^ε by some power of $\log A$. To do this, one can use a version of (1-8) with explicit implied constant proved by Carneiro and Vaaler [2010, Corollary 7.2, (7.20)], or the extremal case of (1-8) with $(\lambda, p) = (1, 2)$, and a logarithmic correction, proved by Li and Villavert [2011].

We also note that we simply use the Weil bound when estimating the weighted sums of Kloosterman sums and we do not exploit any cancellation amongst the Kloosterman sums. Exploring this phenomenon could lead to a power-saving improvement of Theorem 1.4.

For the proof of Theorem 1.1 we use the Selberg trace formula with a suitably chosen test function that fulfills our needs. In the general case of Γ cofinite, unlike in the modular group case, Iwaniec’s truncated formula for $\psi_\Gamma(X)$ is not available. It is probably possible to prove an analogous result, but we preferred to work directly with the trace formula. In fact, in order to prove his formula, Iwaniec uses some arithmetic information on the structure of the lengths of closed geodesics on the modular surface

[Iwaniec 1984b, Lemma 4], namely their connection with primitive binary quadratic forms (proved by Sarnak [1982, §1]), and it is unclear whether a similar argument works for the general cofinite case.

Remark 1.6. We do not discuss the first moment of $P_\Gamma(X)$. Following the argument of [Phillips and Rudnick 1994, Theorems 1.1 and 1.4] it should be possible to prove that for every cofinite Fuchsian group Γ there exists a constant L_Γ such that we have

$$\lim_{A \rightarrow \infty} \frac{1}{\log A} \int_1^A \left(\frac{P_\Gamma(X)}{X^{1/2}} \right) \frac{dX}{X} = L_\Gamma.$$

2. Proof of Theorem 1.1

In this section we explain the structure of the proof of Theorem 1.1. We decided to keep this section to a more colloquial tone, and to relegate the technical computations to the next section. Hence the proof of the theorem is completed in Section 3.

The starting point of our analysis is the Selberg trace formula, that we recall from the book of Iwaniec [1995, Theorem 10.2]. An even function g is said to be an admissible test function in the trace formula if its Fourier transform $h(t)$ (see (2-4) for the convention used to define the Fourier transform) satisfies the conditions [Iwaniec 1995, (1.63)]

$$\begin{aligned} h(t) &\text{ is even,} \\ h(t) &\text{ is holomorphic in the strip } |\Im(t)| \leq \frac{1}{2} + \varepsilon, \\ h(t) &\ll (1 + |t|)^{-2-\varepsilon} \text{ in the strip.} \end{aligned} \tag{2-1}$$

For an admissible test function g , the Selberg trace formula is the identity

$$\sum_P \frac{g(\log NP)}{2 \sinh((\log NP)/2)} \Lambda_\Gamma(NP) = IE + EE + PE + DS + CS + AL, \tag{2-2}$$

where the sum on the left runs over hyperbolic conjugacy classes in Γ , and the terms appearing on the right are explained as follows.

The term IE denotes a contribution coming from the identity element, which forms a conjugacy class on its own. We have

$$IE = -\frac{\text{vol}(\Gamma \backslash \mathbb{H})}{4\pi} \int_{-\infty}^{+\infty} t h(t) \tanh(\pi t) dt.$$

The term EE denotes a contribution from the elliptic conjugacy classes in Γ (there are only finitely many such classes). Denote by \mathcal{R} a primitive elliptic conjugacy class, and let $m = m_{\mathcal{R}} > 1$ be the order of \mathcal{R} . We have

$$EE = -\sum_{\mathcal{R}} \sum_{1 \leq \ell < m} \left(2m \sin\left(\frac{\pi \ell}{m}\right) \right)^{-1} \int_{-\infty}^{+\infty} h(t) \frac{\cosh\left(\pi\left(1 - \frac{2\ell}{m}\right)t\right)}{\cosh(\pi t)} dt.$$

The term PE denotes a contribution from the parabolic conjugacy classes, which are associated to the cusps of Γ . Let \mathfrak{C} denote the total number of inequivalent cusps. Then we have

$$PE = \frac{\mathfrak{C}}{2\pi} \int_{-\infty}^{+\infty} h(t)\psi(1+it) dt + \mathfrak{C}g(0) \log 2.$$

Here $\psi(z)$ is the digamma function, i.e., the logarithmic derivative of the gamma function, $\psi(z) = (\Gamma'/\Gamma)(z)$. The term DS corresponds to the discrete spectrum, and is given by

$$DS = \sum_{t_j} h(t_j),$$

where the sum runs over the spectral parameter t_j , associated to the eigenvalue λ_j of the Laplace operator via the identity $\lambda_j = \frac{1}{4} + t_j^2$ (here and in the following we assume that $t_j \in [0, \frac{1}{2}]$ for $\lambda_j \in [0, \frac{1}{4}]$, and $t_j > 0$ for $\lambda_j > \frac{1}{4}$). The term CS comes from the continuous spectrum, and is given by

$$CS = \frac{1}{4\pi} \int_{-\infty}^{+\infty} h(t) \frac{-\varphi'}{\varphi}(\frac{1}{2} + it) dt,$$

where $\varphi(s)$ is the scattering determinant for the group Γ [Iwaniec 1995, p. 140]. Finally, the term AL is a single term that comes from a combination of the spectral part and the geometric part [Iwaniec 1995, (10.11) and (10.17)], and it is defined by $AL = \frac{1}{4}h(0) \text{Tr}(\Phi(\frac{1}{2}) - \mathbf{I})$, where $\Phi(s)$ is the scattering matrix associated to Γ .

A first naive choice of test function in the Selberg trace formula (2-2) is

$$g_\rho(x) = 2 \sinh(\frac{1}{2}|x|) \mathbf{1}_{[0,\rho]}(|x|), \tag{2-3}$$

where $\rho = \log X$. If we choose $g(x)$ in this way, then the left hand side of (2-2) reduces exactly to the function $\psi_\Gamma(X)$. Unfortunately the function $g_\rho(x)$ in (2-3) is not an admissible test function, and we need therefore to take a suitable approximation of it. An analysis of the right hand side in (2-2) then leads to the desired results. To see that g_ρ is not admissible in the trace formula, consider its Fourier transform $h_\rho(t)$. We have

$$h_\rho(t) = \int_{\mathbb{R}} g_\rho(x) e^{-itx} dx = 2 \int_0^\rho \sinh(\frac{1}{2}x) (e^{itx} + e^{-itx}) dx. \tag{2-4}$$

The integral can be computed directly, which gives, for $t = \pm \frac{i}{2}$,

$$h_\rho(\pm \frac{i}{2}) = 4 \sinh^2(\frac{1}{2}\rho), \tag{2-5}$$

and, for $t \neq \pm \frac{i}{2}$,

$$h_\rho(t) = \frac{2}{\frac{1}{2} + it} \cosh(\rho(\frac{1}{2} + it)) + \frac{2}{\frac{1}{2} - it} \cosh(\rho(\frac{1}{2} - it)) - \frac{2}{\frac{1}{4} + t^2}. \tag{2-6}$$

The last integral in (2-4) shows that $h_\rho(t)$ is even and entire in t , but from (2-6) we see that we only have $h_\rho(t) \ll t^{-1}$ as t tends to infinity, and so we do not have sufficient decay as required in (2-1).

We construct two functions $g_{\pm}(x)$ that approximate from above and below the function g_{ρ} , and are admissible in the Selberg trace formula. Let $q(x)$ be an even, smooth, nonnegative function on \mathbb{R} with compact support contained in $[-1, 1]$ and unit mass (i.e., $\|q\|_1 = 1$). Let $0 < \delta < \frac{1}{4}$ and define

$$q_{\delta}(x) = \frac{1}{\delta}q\left(\frac{x}{\delta}\right).$$

For $\rho > 1$ define g_{\pm} to be the convolution product of the shifted function $g_{\rho \pm \delta}$ with the function q_{δ} , namely

$$g_{\pm}(x) = (g_{\rho \pm \delta} * q_{\delta})(x) = \int_{\mathbb{R}} g_{\rho \pm \delta}(x - y)q_{\delta}(y) dy. \tag{2-7}$$

Taking convolution products has the advantage that the Fourier transform of the convolution is the pointwise product of the Fourier transforms of the two factors. Hence if we denote by $\hat{q}_{\delta}(t)$ the Fourier transform of q_{δ} , that is,

$$\hat{q}_{\delta}(t) = \int_{\mathbb{R}} q_{\delta}(x)e^{-itx} dx, \tag{2-8}$$

and by h_{\pm} the Fourier transform of g_{\pm} , then we obtain

$$h_{\pm}(t) = h_{\rho \pm \delta}(t)\hat{q}_{\delta}(t). \tag{2-9}$$

Since the function $\hat{q}_{\delta}(t)$ is entire and satisfies, for $|\Im(t)| \leq M < \infty$,

$$\hat{q}_{\delta}(t) \ll \frac{1}{1 + \delta^k |t|^k} \quad \forall k \geq 0$$

(see Lemma 3.1), we conclude that $h_{\pm}(t)$ is an entire function and satisfies $h_{\pm}(t) \ll (1 + |t|)^{-2-\varepsilon}$ for some $\varepsilon > 0$ in the strip $|\Im(t)| \leq \frac{1}{2} + \varepsilon$, as required in (2-1). This shows that the function g_{\pm} is an admissible test function in the trace formula. By construction, the function $g_{\pm}(x)$ is supported on $|x| \in [0, \rho + \delta \pm \delta]$. Moreover we have the inequalities (see Lemma 3.2), for $x \geq 0$,

$$g_{-}(x) + O(\delta e^{x/2} \mathbf{1}_{[0, \rho]}(x)) \leq g_{\rho}(x) \leq g_{+}(x) + O(\delta e^{x/2} \mathbf{1}_{[0, \rho]}(x)).$$

If we set

$$\psi_{\pm}(X) = \sum_P \frac{g_{\pm}(\log NP)}{2 \sinh((\log NP)/2)} \Lambda_{\Gamma}(NP),$$

then using the asymptotic $\psi_{\Gamma}(X) \sim X$ we conclude that we have the inequalities

$$\psi_{-}(X) + O(\delta X) \leq \psi_{\Gamma}(X) \leq \psi_{+}(X) + O(\delta X).$$

From this we deduce that we have

$$\frac{1}{A} \int_A^{2A} |P(X)|^2 dX \ll \frac{1}{A} \int_A^{2A} |\psi_{\pm}(X) - M_{\Gamma}(X)|^2 dX + O(\delta^2 A^2), \tag{2-10}$$

and in order to prove Theorem 1.1 we give bounds on the right hand side in (2-10). We found convenient to pass to the logarithmic variable $\rho = \log X$, since this simplifies slightly the computations. Moreover,

we insert a weight function in the integral to pass from the sharp mean square to a smooth one. Consider a smooth, nonnegative real function $w(\rho)$, compactly supported in $[-\varepsilon, 1 + \varepsilon]$, for $0 < \varepsilon < \frac{1}{4}$. In addition, assume that $0 \leq w(\rho) \leq 1$, and that $w(\rho) = 1$ for $\rho \in [0, 1]$. Let

$$w_R(\rho) = w(\rho - R). \tag{2-11}$$

In view of the inequality

$$\frac{1}{A} \int_A^{2A} |\psi_{\pm}(X) - M_{\Gamma}(X)|^2 dX \ll \int_{\mathbb{R}} |\psi_{\pm}(e^{\rho}) - M_{\Gamma}(e^{\rho})|^2 w_R(\rho) d\rho, \tag{2-12}$$

where $R = \log A$, we see that in order to estimate the last integral in (2-10) it suffices to upper bound the integral on the right in (2-12). In the following, abusing notation, we write $\psi_{\pm}(\rho)$ in place of $\psi_{\pm}(e^{\rho})$, and $M_{\Gamma}(\rho)$ in place of $M_{\Gamma}(e^{\rho})$.

At this point we can exploit the Selberg trace formula to analyze ψ_{\pm} . The terms of the discrete spectrum DS associated to the small eigenvalues $\lambda_j \in [0, \frac{1}{4}]$ (corresponding to the spectral parameter t_j in the interval $[0, \frac{i}{2}]$, and for technical convenience we include the eigenvalue $\lambda_j = \frac{1}{4}$) need particular care. We define $M_{\pm}(\rho)$ to be the sum of such terms, namely

$$M_{\pm}(\rho) := \sum_{t_j \in [0, \frac{i}{2}]} h_{\pm}(t_j).$$

In view of the definition of the complete main term in (1-5) it is easy to prove (see Lemma 3.3) that we have

$$M_{\pm}(\rho) = M_{\Gamma}(\rho) + O(\delta e^{\rho} + e^{\rho/2}).$$

Hence we have

$$\begin{aligned} \psi_{\pm}(\rho) - M_{\Gamma}(\rho) &= \psi_{\pm}(\rho) - M_{\pm}(\rho) + O(\delta e^{\rho} + e^{\rho/2}) \\ &= IE_{\pm} + EE_{\pm} + PE_{\pm} + DS'_{\pm} + CS_{\pm} + AL_{\pm} + O(\delta e^{\rho} + e^{\rho/2}), \end{aligned} \tag{2-13}$$

where the term DS'_{\pm} denotes now the contribution from the discrete spectrum restricted to the eigenvalues $\lambda_j > \frac{1}{4}$. The first three terms and the term AL_{\pm} in (2-13) can be bounded pointwise (see Lemma 3.5–3.7) by

$$|IE_{\pm}| + |EE_{\pm}| + |PE_{\pm}| + |AL_{\pm}| \ll e^{\rho/2} + \log(\delta^{-1}).$$

Squaring and integrating in (2-13) we obtain therefore

$$\begin{aligned} &\int_{\mathbb{R}} |\psi_{\pm}(\rho) - M_{\Gamma}(\rho)|^2 w_R(\rho) d\rho \\ &\ll \int_{\mathbb{R}} |DS'_{\pm}|^2 w_R(\rho) d\rho + \int_{\mathbb{R}} |CS_{\pm}|^2 w_R(\rho) d\rho + O(e^R + \delta^2 e^{2R} + \log^2(\delta^{-1})). \end{aligned} \tag{2-14}$$

In Propositions 3.10 and 3.11 we prove that the following estimate holds:

$$\int_{\mathbb{R}} |DS'_{\pm}|^2 w_R(\rho) d\rho + \int_{\mathbb{R}} |CS_{\pm}|^2 w_R(\rho) d\rho \ll \frac{e^R}{\delta} + e^{R/2} \log^2(\delta^{-1}). \tag{2-15}$$

Combining (2-14) and (2-15) we obtain

$$\int_{\mathbb{R}} |\psi_{\pm}(\rho) - M_{\Gamma}(\rho)|^2 w_R(\rho) d\rho \ll \frac{e^R}{\delta} + e^{R/2} \log^2(\delta^{-1}) + \delta^2 e^{2R},$$

and choosing $\delta = e^{-R/3}$ to optimize the first and the last term, we arrive at the bound

$$\int_{\mathbb{R}} |\psi_{\pm}(s) - M_{\Gamma}(s)|^2 w_R(\rho) d\rho \ll e^{4R/3}.$$

Recalling (2-10) and (2-12), and setting $R = \log A$, we conclude that we have

$$\frac{1}{A} \int_A^{2A} |P(X)|^2 dX \ll \int_{\mathbb{R}} |\psi_{\pm}(\rho) - M_{\Gamma}(\rho)|^2 w_R(\rho) d\rho + O(\delta^2 A^2) \ll A^{4/3}.$$

This proves Theorem 1.1.

3. Technical lemmata

In this section we prove the auxiliary results needed in Section 2 to prove Theorem 1.1. We start with a simple computation to bound the function $\hat{q}_{\delta}(t)$ and its derivatives.

Lemma 3.1. *Let $0 < \delta < \frac{1}{4}$, $\hat{q}_{\delta}(t)$ as in (2-8), and let $t \in \mathbb{R}$. Let $j, k \in \mathbb{N}$, $j, k \geq 0$. We have*

$$\left| \frac{d^j \hat{q}_{\delta}}{dt^j}(t) \right| \ll \frac{\delta^j}{1 + |\delta t|^k},$$

and the implied constant depends on j and k .

Proof. From the definition of $\hat{q}_{\delta}(t)$ we have

$$\frac{d^j \hat{q}_{\delta}}{dt^j}(t) = \frac{d^j}{dt^j} \int_{\mathbb{R}} q(x) e^{-it\delta x} dx = (-i\delta)^j \int_{\mathbb{R}} x^j q(x) e^{-it\delta x} dx. \tag{3-1}$$

Bounding in absolute value we get

$$\frac{d^j \hat{q}_{\delta}}{dt^j}(t) \ll \delta^j. \tag{3-2}$$

Integrating by parts in the last integral of (3-1), and using that $q(x)$ is smooth and has compact support, we obtain instead

$$\frac{d^j \hat{q}_{\delta}}{dt^j}(t) = (-1)^k \frac{(-i\delta)^j}{(-it\delta)^k} \int_{\mathbb{R}} e^{-it\delta x} \frac{d^k}{dx^k} (x^j q(x)) dx \ll \frac{\delta^j}{|\delta t|^k}. \tag{3-3}$$

Combining (3-2) and (3-3) we obtain

$$\frac{d^j \hat{q}_{\delta}}{dt^j}(t) \ll \delta^j \min(1, |\delta t|^{-k}) \ll \frac{\delta^j}{1 + |\delta t|^k}.$$

This proves the lemma. □

The function $\hat{q}_\delta(t)$ is used to construct the convolution product $g_\pm = g_{\rho \pm \delta} * q_\delta$ that approximate the function g_ρ . In the next lemma we show that g_ρ is bounded from above and below by g_+ and g_- respectively, up to a small error.

Lemma 3.2. *Let $\rho > 1$ and $0 < \delta < \frac{1}{4}$. Let $g_\rho(x)$ and $g_\pm(x)$ as in (2-3) and (2-7) respectively. For $x \geq 0$ we have*

$$g_-(x) + O(\delta e^{x/2} \mathbf{1}_{[0, \rho]}(x)) \leq g_\rho(x) \leq g_+(x) + O(\delta e^{x/2} \mathbf{1}_{[0, \rho]}(x)). \tag{3-4}$$

Proof. Observe that $g_\pm = g_\rho * q_\delta$, and since both factors in the convolution product are nonnegative, we conclude that $g_\pm(x) \geq 0$ for every $x \in \mathbb{R}$. Observe also that g_\pm is supported on the compact set $\{|x| \leq \rho + \delta \pm \delta\}$. By definition of g_\pm we can therefore write, for $x \geq 0$,

$$g_\pm(x) = 2 \int_{\mathbb{R}} \mathbf{1}_{[0, \rho \pm \delta]}(|x - y|) \sinh\left(\frac{1}{2}|x - y|\right) q_\delta(y) dy = 2 \cdot \mathbf{1}_{[0, \rho + \delta \pm \delta]}(x) \int_Q \sinh\left(\frac{1}{2}|x - y|\right) q_\delta(y) dy,$$

where $Q = [\max\{-\delta, x - (\rho \pm \delta)\}, \delta]$. For $0 \leq x \leq \delta$ we have

$$g_\rho(x) = O(\delta) \quad \text{and} \quad g_\pm(x) = O(\delta),$$

so that (3-4) holds trivially.

For $x > \delta$ we have $x - y > 0$ for every $y \in Q$, and using the addition formula for the hyperbolic sine we can write

$$\begin{aligned} g_\pm(x) &= 2 \cdot \mathbf{1}_{[0, \rho + \delta \pm \delta]}(x) \int_Q \left[\cosh\left(\frac{1}{2}y\right) \sinh\left(\frac{1}{2}x\right) - \cosh\left(\frac{1}{2}x\right) \sinh\left(\frac{1}{2}y\right) \right] q_\delta(y) dy \\ &= 2 \cdot \mathbf{1}_{[0, \rho + \delta \pm \delta]}(x) \sinh\left(\frac{1}{2}x\right) \int_Q q_\delta(y) dy + O(\delta e^{x/2} \mathbf{1}_{[0, \rho + \delta \pm \delta]}(x)). \end{aligned} \tag{3-5}$$

Now for $\delta < x < \rho \pm \delta - \delta$ we have $Q = [-\delta, \delta]$, so that (3-5) reduces to

$$g_\pm(x) = g_\rho(x) + O(\delta e^{x/2} \mathbf{1}_{[0, \rho + \delta \pm \delta]}(x)).$$

For $\rho - 2\delta \leq x \leq \rho$ we can bound by positivity in (3-5)

$$g_-(x) \leq g_\rho(x) + O(\delta e^{x/2} \mathbf{1}_{[0, \rho]}(x)),$$

so that the first inequality in (3-4) holds in this case. Finally for $\rho \leq x \leq \rho + 2\delta$ we observe that $g_\rho(x) = 0$ and $g_+(x) \geq 0$, and so we conclude that the second inequality in (3-4) holds in this case. This proves the lemma. \square

The function $h_\pm(t)$ associated to $g_\pm(x)$ gives, for t corresponding to small eigenvalues, the terms appearing in the definition of the main term $M_\Gamma(X)$ in (1-5), up to small error. We prove this in the next lemma.

Lemma 3.3. *Let $\rho > 1$, $0 < \delta < \frac{1}{4}$, and let $h_\pm(t)$ as in (2-9). Let $t_j \in (0, \frac{i}{2}]$ be the spectral parameter associated to the eigenvalue $\lambda_j \in [0, \frac{1}{4})$. There exists $0 < \varepsilon_\Gamma < \frac{1}{4}$ such that*

$$h_\pm(t_j) = \frac{e^{\rho(1/2 + |t_j|)}}{\frac{1}{2} + |t_j|} + O(\delta e^\rho + e^{\rho(1/2 - \varepsilon_\Gamma)}). \tag{3-6}$$

Proof. Recall that there are only finitely many eigenvalues λ_j in $(0, \frac{1}{4})$, and therefore there exists $0 < \varepsilon_\Gamma < \frac{1}{4}$ such that $|t_j| \geq \varepsilon_\Gamma$ for every $\lambda_j \in (0, \frac{1}{4})$. Observe that for $|t| \leq \frac{1}{2}$ we have

$$\hat{q}_\delta(t) = \int_{\mathbb{R}} q(x)e^{-it\delta x} dx = 1 + O(\delta|t|).$$

The claim then follows from (2-5) and (2-6), and from the fact that $h_\pm(t) = h_{\rho \pm \delta}(t)\hat{q}_\delta(t)$. □

Remark 3.4. Equation (3-6) also holds for $\lambda_j = \frac{1}{4}$, except that we have to multiply the first term by a factor of 2.

The next three lemmata show that in the problem of estimating $P_\Gamma(X)$ using the Selberg trace formula we can neglect the terms coming from the identity class, the elliptic classes, and the parabolic classes, as they contribute a small quantity.

Lemma 3.5. *Let $\rho > 1$, $0 < \delta < \frac{1}{4}$, and let g_\pm and h_\pm as in (2-7) and (2-9) respectively. Then we have*

$$\int_{-\infty}^{+\infty} th_\pm(t) \tanh(\pi t) dt \ll \frac{e^{\rho/2}}{\rho^2} + \log(\delta^{-1}).$$

Proof. Recall that we have $h_\pm(t) = h_\rho(t)\hat{q}_\delta(t)$, and that by Lemma 3.1 we have $\hat{q}_\delta(t) \ll (1 + |\delta t|^k)^{-1}$ for every $k \geq 0$. Using (2-6) and the definition of the hyperbolic cosine we can write

$$h_{\rho \pm \delta}(t) = \frac{2}{1 + 2it} (e^{(\rho \pm \delta)(1/2+it)} + e^{-(\rho \pm \delta)(1/2+it)}) + \frac{2}{1 - 2it} (e^{(\rho \pm \delta)(1/2-it)} + e^{-(\rho \pm \delta)(1/2-it)}) - \frac{2}{\frac{1}{4} + t^2}. \tag{3-7}$$

Bounding in absolute value the integrand associated to the last term in (3-7), we obtain

$$\int_{\mathbb{R}} \left| \frac{t \hat{q}_\delta(t) \tanh(\pi t)}{\frac{1}{4} + t^2} \right| dt \ll \int_{\mathbb{R}} \frac{dt}{(1 + |t|)(1 + |\delta t|)} \ll \log(\delta^{-1}) + 1. \tag{3-8}$$

Now consider the integrand associated to the first term in (3-7) We integrate by parts twice and obtain

$$\begin{aligned} I &:= 2 \int_{\mathbb{R}} \frac{t \hat{q}_\delta(t) \tanh(\pi t) e^{(\rho \pm \delta)(1/2+it)}}{1 + 2it} dt \\ &= 2 \sum_{j=1}^2 (-1)^{j-1} \frac{e^{(\rho \pm \delta)(1/2+it)}}{(i(\rho \pm \delta))^j} \frac{d^{j-1}}{dt^{j-1}} \left(\frac{t \hat{q}_\delta(t) \tanh(\pi t)}{1 + 2it} \right) \Big|_{t=-\infty}^{+\infty} \\ &\quad + 2 \int_{\mathbb{R}} \frac{e^{(\rho \pm \delta)(1/2+it)}}{(i(\rho \pm \delta))^2} \frac{d^2}{dt^2} \left(\frac{t \hat{q}_\delta(t) \tanh(\pi t)}{1 + 2it} \right) dt. \end{aligned}$$

By Lemma 3.1 we see that the boundary terms vanish, and we can bound

$$\frac{d^2}{dt^2} \left(\frac{t \hat{q}_\delta(t) \tanh(\pi t)}{1 + 2it} \right) \ll \frac{1}{(1 + |\delta t|^k)} \left(\delta^2 + \frac{1}{1 + |t|^2} \right),$$

for every $k \geq 0$, with implied constant depending on k . Fixing $k > 1$ we obtain

$$I \ll \frac{e^{\rho/2}}{\rho^2} \int_{\mathbb{R}} \frac{1}{(1 + |\delta t|^k)} \left(\delta^2 + \frac{1}{1 + |t|^2} \right) dt \ll \frac{e^{\rho/2}}{\rho^2}. \tag{3-9}$$

The other terms in (3-7) are treated similarly, and are bounded by the same quantity. Adding (3-8) and (3-9) we obtain the desired estimate. \square

Lemma 3.6. *Let $\rho > 1, 0 < \delta < \frac{1}{4}$, and h_{\pm} as in (2-9). Let $m, \ell \in \mathbb{N}$, with $m \geq 2$ and $1 \leq \ell < m$. Then*

$$\int_{\mathbb{R}} h_{\pm}(t) \frac{\cosh(\pi(1 - 2\ell/m)t)}{\cosh(\pi t)} dt \ll \frac{e^{\rho/2}}{\rho^2},$$

with implied constant that depends on m and ℓ .

Proof. Recall that $h_{\pm}(t) = h_{\rho \pm \delta}(t) \hat{q}_{\delta}(t)$, and use (3-7) to express $h_{\rho \pm \delta}(t)$. The integrand associated to the last term in (3-7) is bounded by

$$\int_{\mathbb{R}} \frac{\cosh(\pi(1 - 2\ell/m)t)}{(1 + t^2) \cosh(\pi t)} dt \ll \int_{\mathbb{R}} \frac{dt}{1 + t^2} \ll 1.$$

Now consider the term $\exp((\rho \pm \delta)(\frac{1}{2} + it))$ in (3-7). The corresponding integral contributes

$$\begin{aligned} J &:= 2 \int_{\mathbb{R}} \frac{e^{(\rho \pm \delta)(1/2 + it)} \hat{q}_{\delta}(t) \cosh(\pi(1 - 2\ell/m)t)}{(1 + 2it) \cosh(\pi t)} dt \\ &= 2 \sum_{j=1}^2 (-1)^{j-1} \frac{e^{(\rho \pm \delta)(1/2 + it)}}{(i(\rho \pm \delta))^j} \frac{d^{j-1}}{dt^{j-1}} \left(\frac{\hat{q}_{\delta}(t) \cosh(\pi t(1 - 2\ell/m))}{(1 + 2it) \cosh(\pi t)} \right) \Big|_{t=-\infty}^{+\infty} \\ &\quad + 2 \int_{\mathbb{R}} \frac{e^{(\rho \pm \delta)(1/2 + it)}}{(i(\rho \pm \delta))^2} \frac{d^2}{dt^2} \left(\frac{\hat{q}_{\delta}(t) \cosh(\pi t(1 - 2\ell/m))}{(1 + 2it) \cosh(\pi t)} \right) dt. \end{aligned}$$

The boundary terms vanish, and by Lemma 3.1 we can bound

$$\frac{d^2}{dt^2} \left(\frac{\hat{q}_{\delta}(t) \cosh(\pi t(1 - 2\ell/m))}{(1 + 2it) \cosh(\pi t)} \right) \ll \frac{1}{(1 + |t|)(1 + |\delta t|^k)} \left(\frac{1}{1 + |t|^2} + \delta^2 \right)$$

where the implied constant depends on m and ℓ . Hence we get

$$J \ll \frac{e^{\rho/2}}{\rho^2} \int_{\mathbb{R}} \frac{1}{(1 + |t|)(1 + |\delta t|^k)} \left(\frac{1}{1 + |t|^2} + \delta^2 \right) dt \ll \frac{e^{\rho/2}}{\rho^2}.$$

The terms associated to the other exponentials in (3-7) are treated similarly, and are bounded by the same quantity. \square

Lemma 3.7. *Let $\rho > 1, 0 < \delta < \frac{1}{4}$, and let g_{\pm} and h_{\pm} as in (2-7) and (2-9). Then*

$$g_{\pm}(0) \log 2 + \frac{1}{2\pi} \int_{-\infty}^{+\infty} h_{\pm}(t) \psi(1 + it) dt \ll \delta e^{\rho/2} + \log(\delta^{-1}).$$

Proof. We use here the formula given in [Iwaniec 1995, (10.17)] to get back to an integral involving g_{\pm} . We have

$$g_{\pm}(0) \log 2 + \frac{1}{2\pi} \int_{-\infty}^{+\infty} h_{\pm}(t) \psi(1 + it) dt = \frac{h_{\pm}(0)}{4} - \gamma g_{\pm}(0) + \int_0^{\infty} \log(\sinh(\frac{1}{2}x)) dg_{\pm}(x). \quad (3-10)$$

By (3-7) we know that

$$\frac{h_{\pm}(0)}{4} = e^{\rho/2} + O(\delta e^{\rho/2} + 1), \tag{3-11}$$

and by definition of $g_{\pm}(x)$ we have $g_{\pm}(x) = O(\delta)$ for $|x| \leq \delta$. The last integral in (3-10) is analyzed as follows. For $x \in [0, \delta]$ we bound

$$\frac{d}{dx}g_{\pm}(x) = (g_{\rho \pm \delta} * q'_{\delta})(x) \ll \int_{-\delta}^{\delta} \sinh(\frac{1}{2}|x - y|)|q'_{\delta}(y)| dy \ll \frac{1}{\delta} \sinh(\delta)\|q'\|_1 \ll 1,$$

so that we have, uniformly in δ ,

$$\int_0^{\delta} \log(\sinh(\frac{1}{2}x)) dg_{\pm}(x) \ll \int_0^{\delta} |\log(\sinh(\frac{1}{2}x))| dx \ll 1.$$

For $x > \delta$ we integrate by parts obtaining

$$\int_{\delta}^{\infty} \log(\sinh(\frac{1}{2}x)) dg_{\pm}(x) = -g_{\pm}(\delta) \log(\sinh(\frac{1}{2}\delta)) - \frac{1}{2} \int_{\delta}^{\infty} \frac{g_{\pm}(x)}{\sinh(\frac{1}{2}x)} \cosh(\frac{1}{2}x) dx.$$

Since $g_{\pm}(\delta) \ll \delta$, the boundary term is bounded by $O(\delta \log(\delta^{-1}))$. If we write $\cosh(\frac{x}{2}) = \sinh(\frac{x}{2}) + e^{-x/2}$, the integral associated to $e^{-x/2}$ can be bounded by

$$\int_{\delta}^{\infty} \frac{g_{\pm}(x)e^{-x/2}}{\sinh(\frac{1}{2}x)} dx \ll \int_{\delta}^1 \frac{dx}{x} + \int_1^{\infty} e^{-x/2} dx \ll \log(\delta^{-1}) + 1.$$

Finally we have

$$-\frac{1}{2} \int_{\delta}^{\infty} g_{\pm}(x) dx = -2 \cosh(\frac{1}{2}(\rho \pm \delta)) + O(1) = -e^{\rho/2} + O(\delta e^{\rho/2} + 1).$$

The exponential cancels with the first term in (3-11), and combining the other estimates we obtain the claim. □

We turn now our attention to finding upper bounds for the mean square of the spectral side in the Selberg trace formula. We start with an estimate for the integral of $h_{\pm}(t_1)h_{\pm}(t_2)$.

Lemma 3.8. *Let $0 < \delta < \frac{1}{4}$, let $h_{\rho \pm \delta}(t)$ be as in (2-9), and let $w_R(\rho)$ be as in (2-11). Let $R > 1$, and $t_1, t_2 \in \mathbb{R}$. Then we have*

$$\int_{\mathbb{R}} h_{\rho \pm \delta}(t_1)h_{\rho \pm \delta}(t_2)w_R(\rho) d\rho \ll \frac{e^R v(t_1)v(t_2)}{1 + |t_1 - t_2|^2} + \frac{e^R v(t_1)v(t_2)}{1 + |t_1 + t_2|^2} + e^{R/2} v(t_1^2)v(t_2^2),$$

where $v(t) = (1 + |t|)^{-1}$, and the implied constant does not depend on δ .

Proof. In order to express $h_{\rho \pm \delta}$ we consider again (3-7).

Multiplying $h_{\rho \pm \delta}(t_1)$ with $h_{\rho \pm \delta}(t_2)$, the product of the first exponential in (3-7) (from the factor $h_{\rho \pm \delta}(t_1)$) and the last term from the factor $h_{\rho \pm \delta}(t_2)$ contributes

$$\int_{\mathbb{R}} \frac{4e^{(\rho \pm \delta)(1/2 + it_1)}}{(1 + 2it_1)(\frac{1}{4} + t_2^2)} w_R(\rho) d\rho = \int_{\mathbb{R}} \frac{2e^{(\rho \pm \delta)(1/2 + it_1)} w'_R(\rho)}{(\frac{1}{2} + it_1)^2 (\frac{1}{4} + t_2^2)} d\rho \ll e^{R/2} v(t_1^2) v(t_2^2) \|w'\|_1 \ll e^{R/2} v(t_1^2) v(t_2^2). \quad (3-12)$$

Now consider the product of the first exponential in (3-7) for $h_{\rho \pm \delta}(t_1)$ and the same term for $h_{\rho \pm \delta}(t_2)$. This contributes

$$\int_{\mathbb{R}} \frac{4e^{(\rho \pm \delta)(1 + i(t_1 + t_2))}}{(1 + 2it_1)(1 + 2it_2)} w_R(\rho) d\rho \ll e^R v(t_1) v(t_2) \min(1, |t_1 + t_2|^{-2} \|w''\|_1) \ll \frac{e^R v(t_1) v(t_2)}{1 + |t_1 + t_2|^2}. \quad (3-13)$$

The other terms in the product $h_{\rho \pm \delta}(t_1)h_{\rho \pm \delta}(t_2)$ are bounded similarly by (3-12) and (3-13), except that we need to replace $|t_1 + t_2|$ by $|t_1 - t_2|$ when we integrate the product $\exp((\rho \pm \delta)(\pm 1 \pm i(t_1 - t_2)))$. This concludes the proof. \square

In order to exploit at best the bound proved in the previous lemma, we estimate the size of the spectrum on unit intervals.

Lemma 3.9. *Let Γ be a cofinite Fuchsian group, let $\varphi(s)$ be the scattering determinant associated to Γ , and let $T > 1$. We have*

$$\#\{T \leq t_j \leq T + 1\} + \int_{T \leq |t| \leq T+1} \left| \frac{-\varphi'}{\varphi} \left(\frac{1}{2} + it \right) \right| dt \ll T.$$

Proof. Recall Weyl’s law in its strong form (see [Venkov 1990, Theorem 7.3])

$$\#\{t_j \leq T\} + \frac{1}{4\pi} \int_{-T}^T \frac{-\varphi'}{\varphi} \left(\frac{1}{2} + it \right) dt = \frac{\text{vol}(\Gamma \backslash \mathbb{H})}{4\pi} T^2 + \frac{\mathfrak{C}}{\pi} T \log T + O(T),$$

where we recall that \mathfrak{C} is the number of inequivalent cusps of Γ . Consider the equation above at the point $T + 1$, and subtract from it the same quantity for T . In order to shorten notation we write $f(t) = -(\varphi'/\varphi)(\frac{1}{2} + it)$. We get

$$\#\{T \leq t_j \leq T + 1\} + \frac{1}{4\pi} \int_{-T-1}^{-T} f(t) dt + \frac{1}{4\pi} \int_T^{T+1} f(t) dt = O(T). \quad (3-14)$$

The function $f(t)$ is bounded from below by a constant k that depends on the group (this follows from the Maass–Selberg relations, see [Iwaniec 1995, (10.9)]). Hence we have

$$\int_{-T-1}^{-T} f(t) dt \geq k.$$

Since the number $\#\{T \leq t_j \leq T + 1\}$ is nonnegative, we can write

$$\int_T^{T+1} f(t) dt \leq \int_T^{T+1} f(t) dt + \int_{-T-1}^{-T} f(t) dt - k + 4\pi \cdot \#\{T \leq t_j \leq T + 1\} = O(T) - k = O(T).$$

Again from the fact that $f(t)$ is bounded from below by a constant, we infer that in fact we have

$$\int_T^{T+1} |f(t)| dt \ll T.$$

Similarly we find

$$\int_{-T-1}^{-T} |f(t)| dt \ll T,$$

and from (3-14) we conclude that we also have $\#\{T \leq t_j \leq T + 1\} \ll T$. This proves the lemma. \square

At this point we can prove bounds on the mean square of the discrete and continuous spectrum in the Selberg trace formula. We discuss first the discrete spectrum.

Proposition 3.10. *Let $0 < \delta < \frac{1}{4}$, let h_{\pm} as in (2-9), and let $R > 1$. We have*

$$\int_{\mathbb{R}} \left| \sum_{t_j > 0} h_{\pm}(t_j) \right|^2 w_R(\rho) d\rho \ll \frac{e^R}{\delta} + e^{R/2} \log^2(\delta^{-1}).$$

Proof. Recall that $h_{\pm}(t) = h_{\rho \pm \delta}(t) \hat{q}_{\delta}(t)$, and that $\hat{q}_{\delta}(t) \ll (1 + |\delta t|^k)^{-1}$ for every $k \geq 0$. Using Lemma 3.9 this implies that the series

$$\sum_{t_j > 0} h_{\pm}(t_j)$$

is absolutely convergent, and we can write

$$\int_{\mathbb{R}} \left| \sum_{t_j > 0} h_{\pm}(t_j) \right|^2 w_R(\rho) d\rho = \sum_{t_j > 0} \sum_{t_{\ell} > 0} \hat{q}_{\delta}(t_j) \hat{q}_{\delta}(t_{\ell}) \int_{\mathbb{R}} h_{\rho \pm \delta}(t_j) h_{\rho \pm \delta}(t_{\ell}) w_R(\rho) d\rho.$$

By Lemma 3.8 we can estimate the integral and bound the double sum by

$$e^R \sum_{t_j, t_{\ell} > 0} \frac{|\hat{q}_{\delta}(t_j) \hat{q}_{\delta}(t_{\ell})| v(t_1) v(t_2)}{1 + |t_j - t_{\ell}|^2} + e^{R/2} \sum_{t_j, t_{\ell} > 0} |\hat{q}_{\delta}(t_j) \hat{q}_{\delta}(t_{\ell})| v(t_1^2) v(t_2^2), \tag{3-15}$$

where $v(t) = (1 + |t|)^{-1}$. Using Lemma 3.1 to bound $\hat{q}_{\delta}(t) \ll (1 + |\delta t|^k)^{-1}$ for every $k \geq 0$, we can estimate the second sum in (3-15) by

$$e^{R/2} \left(\sum_{t_j \leq \delta^{-1}} \frac{1}{t_j^2} + \sum_{t_j > \delta^{-1}} \frac{1}{\delta t_j^3} \right)^2 \ll e^{R/2} (\log^2(\delta^{-1}) + 1).$$

Now consider the first sum in (3-15). By symmetry and positivity, we can consider only the sum over $t_{\ell} \geq t_j$. Moreover we split the sum in order to optimize the bounds available. Consider a unit neighborhood of the diagonal $t_{\ell} = t_j$. Using Lemma 3.9 we can estimate

$$e^R \sum_{t_j > 0} \sum_{t_j \leq t_{\ell} \leq t_j + 1} \frac{|\hat{q}_{\delta}(t_j) \hat{q}_{\delta}(t_{\ell})|}{(1 + |t_j|)(1 + |t_{\ell}|)} \ll e^R \sum_{t_j \leq \delta^{-1}} \frac{1}{t_j} \sum_{t_j \leq t_{\ell} \leq t_j + 1} \frac{1}{t_{\ell}} + \frac{e^R}{\delta^2} \sum_{t_j > \delta^{-1}} \frac{1}{t_j^2} \sum_{t_j \leq t_{\ell} \leq t_j + 1} \frac{1}{t_{\ell}^2} \ll \frac{e^R}{\delta}.$$

The tail of the double sum, that is, the range $t_j > \delta^{-1}$ and $t_\ell > t_j + 1$, can be analyzed (we follow here the same method as in Cramér [1922]) by using a unit interval decomposition for the sum over t_ℓ , together with Lemma 3.9, to get

$$\frac{e^R}{\delta^2} \sum_{t_j > \delta^{-1}} \frac{1}{t_j^2} \sum_{t_\ell > t_j + 1} \frac{1}{t_\ell^2 |t_\ell - t_j|^2} \ll \frac{e^R}{\delta^2} \sum_{t_j > \delta^{-1}} \frac{1}{t_j^2} \sum_{k=1}^{\infty} \frac{1}{(t_j + k)k^2} \ll \frac{e^R}{\delta^2} \sum_{t_j > \delta^{-1}} \frac{1}{t_j^3} \ll \frac{e^R}{\delta}.$$

Finally, the range $t_j \leq \delta^{-1}$ and $t_\ell > t_j + 1$ is bounded by

$$\ll e^R \sum_{t_j \leq \delta^{-1}} \frac{1}{t_j} \sum_{t_\ell > t_j + 1} \frac{1}{t_\ell |t_\ell - t_j|^2} \ll e^R \sum_{t_j \leq \delta^{-1}} \frac{1}{t_j} \sum_{k=1}^{\infty} \frac{1}{k^2} \ll \frac{e^R}{\delta}.$$

We conclude that we have

$$\int_{\mathbb{R}} \left| \sum_{t_j > 0} h_{\pm}(t_j) \right|^2 w_R(\rho) d\rho \ll \frac{e^R}{\delta} + e^{R/2} \log^2(\delta^{-1}),$$

as claimed. □

The analysis of the continuous spectrum is similar, and we obtain the same bounds. With the proposition below we conclude the list of auxiliary results needed to prove Theorem 1.1.

Proposition 3.11. *Let $0 < \delta < \frac{1}{4}$ and h_{\pm} be as in (2-9). Let $R \geq 1$. Then*

$$\int_{\mathbb{R}} \left| \int_{\mathbb{R}} h_{\pm}(t) \frac{-\varphi'}{\varphi} \left(\frac{1}{2} + it \right) \right|^2 w_R(\rho) d\rho \ll \frac{e^R}{\delta} + e^{R/2} \log^2(\delta^{-1}). \tag{3-16}$$

Proof. Recall that we have $h_{\pm}(t) = h_{\rho \pm \delta}(t) \hat{q}_{\delta}(t)$, and that $\hat{q}_{\delta}(t) \ll (1 + |\delta t|^k)^{-1}$ for every $k \geq 0$. For simplicity we write

$$f(t) = \frac{-\varphi'}{\varphi} \left(\frac{1}{2} + it \right).$$

Let J denote the integral in (3-16). Due to the decay properties of \hat{q}_{δ} and to Lemma 3.9, J is absolutely convergent, and so we can write

$$J = \int_{\mathbb{R}} \int_{\mathbb{R}} f(t_1) \overline{f(t_2)} \hat{q}_{\delta}(t_1) \hat{q}_{\delta}(t_2) \int_{\mathbb{R}} h_{\rho \pm \delta}(t_1) h_{\rho \pm \delta}(t_2) w_R(\rho) d\rho dt_1 dt_2.$$

The innermost integral is bounded using Lemma 3.8. This gives

$$J \ll e^R \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|f(t_1) f(t_2) \hat{q}_{\delta}(t_2) \hat{q}_{\delta}(t_2)|}{(1 + |t_1|)(1 + |t_2|)(1 + |t_1 - t_2|^2)} dt_1 dt_2 + e^{R/2} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{|f(t_1) f(t_2) \hat{q}_{\delta}(t_2) \hat{q}_{\delta}(t_2)|}{(1 + t_1^2)(1 + t_2^2)} dt_1 dt_2. \tag{3-17}$$

The second integral in (3-17) is bounded by

$$\left(\int_0^{\delta^{-1}} \frac{|f(t)|}{1 + t^2} dt + \int_{\delta^{-1}}^{\infty} \frac{|f(t)|}{\delta t^3} dt \right)^2 \ll (\log(\delta^{-1}) + 1)^2 \ll \log^2(\delta^{-1}),$$

where in the first inequality we have used a unit interval decomposition of the domain of integration, and Lemma 3.9 to bound the integral of $|f(t)|$ in unit intervals. Now consider the first integral in (3-17). By symmetry and positivity we can consider only the integral over $t_2 \geq t_1 \geq 0$. A unit neighborhood of the diagonal $t_1 = t_2$ gives

$$\int_0^{\delta^{-1}} \frac{|f(t_1)|}{1+t_1} \int_{t_1}^{t_1+1} \frac{|f(t_2)|}{1+t_2} dt_2 dt_1 + \int_{\delta^{-1}}^{\infty} \frac{|f(t_1)|}{\delta^2 t_1^3} \int_{t_1}^{t_1+1} \frac{|f(t_2)|}{\delta^2 t_2^3} dt_2 dt_1 \ll \int_0^{\delta^{-1}} \frac{|f(t_1)|}{1+t_1} dt_1 + \int_{\delta^{-1}}^{\infty} \frac{|f(t_1)|}{\delta^4 t_1^4} dt_1 \ll \frac{1}{\delta}.$$

The tail of the double integral, that is, the range $t_1 \geq \delta^{-1}$ and $t_2 \geq t_1 + 1$, can be bounded as follows:

$$\int_{\delta^{-1}}^{\infty} \frac{|f(t_1)|}{\delta^2 t_1^3} \int_{t_1+1}^{\infty} \frac{|f(t_2)|}{\delta^2 t_2^3 |t_2 - t_1|^2} dt_2 dt_1 \ll \int_{\delta^{-1}}^{\infty} \frac{|f(t_1)|}{\delta^4 t_1^5} dt_1 \ll 1.$$

The range $t_1 \leq \delta^{-1}$ and $t_2 \geq t_1 + 1$ contributes

$$\int_0^{\delta^{-1}} \frac{|f(t_1)|}{1+t_1} \int_{t_1+1}^{\infty} \frac{|f(t_2)|}{(1+t_2)(1+\delta^2 t_2^2) |t_2 - t_1|^2} dt_2 dt_1 \ll \int_0^{\delta^{-1}} \frac{|f(t_1)|}{1+t_1} dt_1 \ll \delta^{-1}.$$

Summarizing, we have showed that we have the bound

$$J \ll \frac{e^R}{\delta} + e^{R/2} \log^2(\delta^{-1}),$$

which is what we wanted. This proves the proposition. □

4. Modular group

This section is devoted to the proof of Theorem 1.4, concerning the case $G = \text{PSL}_2(\mathbb{Z})$. Our approach starts with a lemma of Iwaniec [1984b, Lemma 1], which gives

$$\psi_G(X) = X + 2\Re\left(\sum_{t_j \leq T} \frac{X^{1/2+it_j}}{\frac{1}{2} + it_j}\right) + O\left(\frac{X}{T} \log^2 X\right),$$

where $1 \leq T \leq X^{1/2}(\log X)^{-2}$ (and it is understood that the sum runs over $t_j > 0$). Note that in this case the only small eigenvalue of Δ is $\lambda = 0$, so that $M_G(X) = X$ and $P_G(X) = \psi_G(X) - X$. From the equation above we deduce that

$$\frac{1}{A} \int_A^{2A} |P_G(X)|^2 dX \ll \frac{1}{A} \int_A^{2A} \left| \sum_{t_j \leq T} \frac{X^{1/2+it_j}}{\frac{1}{2} + it_j} \right|^2 dX + \frac{A^2}{T^2} \log^2 A. \tag{4-1}$$

We now describe the outline of the proof of Theorem 1.4. The main idea is to use the Kuznetsov trace formula to estimate the mean square of

$$\sum_{t_j \leq T} \frac{X^{1/2+it_j}}{\frac{1}{2} + it_j}$$

or, equivalently by partial summation, the mean square of

$$R(X, T) = \sum_{t_j \leq T} X^{it_j}. \tag{4-2}$$

Therefore, we start by setting up the Kuznetsov trace formula with a test function that gives us a smooth version of the sum in (4-2), and we estimate the mean square of the weighted sum of Kloosterman sums that show up in its geometric side (Lemma 4.1). This allows us to bound the spectral side of the trace formula (Lemma 4.2). In order to control the behavior of the Fourier coefficients in the spectral sums of the trace formula we use a smoothed average of these sums. This way we obtain (Lemma 4.4) a mean square estimate of a smoothed version of $R(X, T)$, from which we extract a mean square estimate for the sharp sum (Proposition 4.5).

We now start by setting up the Kuznetsov trace formula. Let $\phi(x)$ be a smooth function on $[0, \infty]$ such that

$$\begin{aligned} |\phi(x)| &\ll x, & x \rightarrow 0, \\ |\phi^{(l)}(x)| &\ll x^{-3}, & x \rightarrow \infty, \end{aligned}$$

for $l = 0, 1, 2, 3$. Define

$$\begin{aligned} \phi_0 &= \frac{1}{2\pi} \int_0^\infty J_0(y)\phi(y) dy, \\ \phi_B(x) &= \int_0^1 \int_0^\infty \xi x J_0(\xi x) J_0(\xi y)\phi(y) dy d\xi, \\ \phi_H(x) &= \int_1^\infty \int_0^\infty \xi x J_0(\xi x) J_0(\xi y)\phi(y) dy d\xi, \\ \hat{\phi}(t) &= \frac{\pi}{2i \sinh \pi t} \int_0^\infty (J_{2it}(x) - J_{-2it}(x))\phi(x) \frac{dx}{x}, \end{aligned}$$

where J_ν is the Bessel function of the first kind and order ν . By the properties of the Hankel transform we have

$$\phi(x) = \phi_B(x) + \phi_H(x).$$

We now choose ϕ as in [Luo and Sarnak 1995]. For $X, T > 1$ we set

$$\begin{aligned} \phi_{X,T}(x) &= \frac{-\sinh \beta}{\pi} x \exp(ix \cosh \beta), \\ 2\beta &= \log X + \frac{i}{T}, \end{aligned}$$

and apply the Kuznetsov trace formula [1980, Theorem 1] with $\phi_{X,T}$ as the test function. Let $\{f_i\}_{i=1}^\infty$ be an orthonormal basis of Maass cusp forms for $SL_2(\mathbb{Z})$, with eigenvalues $\lambda_j = \frac{1}{4} + t_j^2$. These cusp forms have Fourier expansions [Kuznetsov 1980, (2.10)]

$$f_j(z) = \sqrt{y} \sum_{n=1}^\infty \rho_j(n) K_{it_j}(2\pi ny) \cos(2\pi nx),$$

where $K_\nu(x)$ is the K -Bessel function. For $l_1, l_2 \geq 1$ the trace formula reads

$$\sum_{t_j} \hat{\phi}(t_j) v_j(l_1) \overline{v_j(l_2)} + \frac{2}{\pi} \int_0^\infty \frac{\hat{\phi}(t)}{|\zeta(1+2it)|^2} d_{it}(l_1) d_{it}(l_2) dt = \delta_{l_1, l_2} \phi_0 + \sum_{c=1}^\infty \frac{S(l_1, l_2, c)}{c} \phi_H\left(\frac{4\pi\sqrt{l_1 l_2}}{c}\right),$$

where $\rho_j(l) = v_j(l) \cosh(\pi t_j)^{1/2}$, $d_{it}(l) = \sum_{d_1 d_2 = l} (d_1/d_2)^{it}$, and $S(l_1, l_2, c)$ is the classical Kloostermann sum. By [Luo and Sarnak 1995, p. 234] we have

$$\hat{\phi}_{X,T}(t_j) = X^{it_j} e^{-t_j/T} + O(e^{-\pi t_j}), \tag{4-3}$$

$$(\phi_{X,T})_0 \ll X^{-1/2}, \tag{4-4}$$

$$\frac{2}{\pi} \int_0^\infty \frac{\hat{\phi}_{X,T}(t)}{|\zeta(1+2it)|^2} (d_{it}(n))^2 dt \ll T \log^2 T d^2(n), \tag{4-5}$$

$$S_n((\phi_{X,T})_B) \ll n^{1/2} X^{-1/2} \log^2 n, \tag{4-6}$$

where

$$S_n(\psi) = \sum_{c=1}^\infty \frac{S(n, n, c)}{c} \psi\left(\frac{4\pi n}{c}\right). \tag{4-7}$$

Analyzing the right hand side of the Kuznetsov trace formula, we prove a bound for the mean square of $S_n(\phi_{X,T})$ as follows.

Lemma 4.1. *Let $A, T > 2$ and let n be a positive integer. Then, for any $\varepsilon > 0$,*

$$\frac{1}{A} \int_A^{2A} |S_n(\phi_{X,T})|^2 dX \ll_\varepsilon (nA^{1/2} + T^2)(An)^\varepsilon.$$

We postpone the proof of Lemma 4.1 to the end of the section, and we show here how to recover a similar bound on the spectral side of the Kuznetsov trace formula.

Lemma 4.2. *Let $A, T > 2$ and let n be a positive integer. Then, for any $\varepsilon > 0$,*

$$\frac{1}{A} \int_A^{2A} \left| \sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right|^2 dX \ll_\varepsilon (nA^{1/2} + T^2)(AnT)^\varepsilon.$$

Proof. By the trace formula and the bounds in equations (4-4)–(4-6) we deduce

$$\frac{1}{A} \int_A^{2A} \left| \sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right|^2 dX \ll \frac{1}{A} \int_A^{2A} |S_n(\phi_{X,T})|^2 dX + T^2 \log^4 T d^4(n) + \frac{n \log A \log^4 n}{A}.$$

Observing that $d^4(n) \ll n^\varepsilon$, the claim follows from Lemma 4.1. □

Once we have bounds for a fixed n , we average over $n \in [N, 2N]$. Let $h(\xi)$ be a smooth function supported in $[N, 2N]$, whose derivatives satisfy

$$|h^{(p)}(\xi)| \ll N^{-p} \quad \text{for } p = 0, 1, 2, \dots$$

and such that

$$\int_{-\infty}^{+\infty} h(\xi) d\xi = N.$$

The next lemma is a result of Luo and Sarnak [1995].

Lemma 4.3. *Let h be as above. Then*

$$\begin{aligned} \sum_n h(n)|v_j(n)|^2 &= \frac{12}{\pi^2}N + r(t_j, N), \\ \sum_{t_j \leq T} |r(t_j, N)| &\ll T^2 N^{1/2} \log^2 T. \end{aligned}$$

Proof. See [Luo and Sarnak 1995, p. 233]. □

Lemma 4.4. *Let $A, T > 2$. Then, for any $\varepsilon > 0$,*

$$\frac{1}{A} \int_A^{2A} \left| \sum_{t_j} X^{it_j} e^{-t_j/T} \right|^2 dX \ll_{\varepsilon} A^{1/4} T^2 (AT)^{\varepsilon}.$$

Proof. By the previous lemma

$$\begin{aligned} \frac{1}{N} \sum_n h(n) \left(\sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right) &= \frac{1}{N} \sum_{t_j} \left(\sum_n |v_j(n)|^2 h(n) \right) \hat{\phi}(t_j) \\ &= \frac{12}{\pi^2} \sum_{t_j} \hat{\phi}(t_j) + \frac{1}{N} \sum_{t_j} r(t_j, N) \hat{\phi}(t_j) \\ &= \frac{12}{\pi^2} \sum_{t_j} \hat{\phi}(t_j) + O(T^2 N^{-1/2} \log^2 T), \end{aligned} \tag{4-8}$$

where the last step can be obtained by $|\hat{\phi}(t_j)| \ll e^{-t_j/T}$ and partial summation. Note now that from (4-3) it follows that

$$\sum_{t_j} \hat{\phi}(t_j) = \sum_{t_j} X^{it_j} e^{-t_j/T} + O(1). \tag{4-9}$$

Combining (4-8) and (4-9) we deduce that

$$\sum_{t_j} X^{it_j} e^{-t_j/T} \ll \frac{1}{N} \sum_n h(n) \left(\sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right) + T^2 N^{-1/2} \log^2 T.$$

Taking absolute value, squaring, and integrating over X , we infer that

$$\frac{1}{A} \int_A^{2A} \left| \sum_{t_j} X^{it_j} e^{-t_j/T} \right|^2 dX \ll \frac{1}{A} \int_A^{2A} \left| \frac{1}{N} \sum_n h(n) \left(\sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right) \right|^2 dX + T^4 N^{-1} \log^4 T.$$

Moreover, by Cauchy–Schwarz and the properties of h , we have

$$\begin{aligned} \left| \frac{1}{N} \sum_n h(n) \left(\sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right) \right|^2 &\ll \frac{1}{N^2} \left(\sum_{n=N}^{2N} |h(n)|^2 \right) \left(\sum_{n=N}^{2N} \left| \sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right|^2 \right) \\ &\ll \frac{1}{N} \sum_{n=N}^{2N} \left| \sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right|^2 \end{aligned}$$

Finally, we obtain

$$\begin{aligned} \frac{1}{A} \int_A^{2A} \left| \sum_{t_j} X^{it_j} e^{-t_j/T} \right|^2 dX &\ll_\varepsilon \frac{1}{N} \sum_{n=N}^{2N} \left(\frac{1}{A} \int_A^{2A} \left| \sum_{t_j} |v_j(n)|^2 \hat{\phi}(t_j) \right|^2 dX \right) + \frac{T^{4+\varepsilon}}{N} \\ &\ll_\varepsilon (NA^{1/2} + T^2)(ANT)^\varepsilon + \frac{T^{4+\varepsilon}}{N}, \end{aligned}$$

where the last inequality follows from Lemma 4.2. Choose $N = A^{-1/4}T^2$ to complete the proof. \square

The next step is to replace the smoothed sum $\sum X^{it_j} e^{-t_j/T}$ with the truncated one. This gives us a corresponding mean square estimate for $R(X, T)$.

Proposition 4.5. *Let $A, T > 2$, and let $R(X, T)$ be as in (4-2). Then, for any $\varepsilon > 0$,*

$$\frac{1}{A} \int_A^{2A} |R(X, T)|^2 dX \ll_\varepsilon A^{1/4} T^2 (AT)^\varepsilon.$$

Proof. We start by choosing a smooth function g that approximates the characteristic function of $[1, T]$. Let g be a smooth function supported on $[\frac{1}{2}, T + \frac{1}{2}]$ such that $0 \leq g(\xi) \leq 1$, and $g(\xi) = 1$ when $\xi \in [1, T]$. By the strong Weyl’s law we know that $|\{t_j : U \leq t_j \leq U + 1\}| \ll U$, and so we have

$$R(X, T) = \sum_{t_j} g(t_j) X^{it_j} + O(T).$$

Define $\hat{g}(\xi)$ to be the Fourier transform of $g(\xi) \exp(\xi/T)$. By [Luo and Sarnak 1995, pp. 235–236] we have that

$$\begin{aligned} \sum_{t_j} g(t_j) X^{it_j} &= \int_{-1}^1 \hat{g}(\xi) \left(\sum_{t_j} (X e^{-2\pi\xi})^{it_j} e^{-t_j/T} \right) d\xi + O \left(\sum_{t_j} \frac{e^{-t_j/T}}{t_j} + \frac{e^{-t_j/T}}{|T - t_j| + 1} + \frac{\log(T + t_j) e^{-t_j/T}}{T} \right), \end{aligned}$$

where the error term can be bounded by $O(T \log T)$. Also note the estimate

$$\hat{g}(\xi) \ll \min \left(T, \frac{1}{|x|} \right). \tag{4-10}$$

Defining

$$k(X, T, \xi) := \sum_{t_j} (X e^{-2\pi\xi})^{it_j} e^{-t_j/T},$$

we deduce that

$$\begin{aligned} & \frac{1}{A} \int_A^{2A} |R(X, T)|^2 dX \\ & \ll \frac{1}{A} \int_A^{2A} \left| \int_{-1}^1 \hat{g}(\xi) k(X, T, \xi) d\xi \right|^2 dX + T^2 \log^2 T \\ & \ll \frac{1}{A} \int_A^{2A} \left| \int_{|\xi| \leq \delta} \hat{g}(\xi) k(X, T, \xi) d\xi \right|^2 dX + \frac{1}{A} \int_A^{2A} \left| \int_{\delta < |\xi| \leq 1} \hat{g}(\xi) k(X, T, \xi) d\xi \right|^2 dX + T^2 \log^2 T. \end{aligned}$$

The first term in the sum above we bound, using Lemma 4.4 and the first bound in (4-10), by

$$\int_{|\xi| \leq \delta} |\hat{g}(\xi)|^2 \int_{|\xi| \leq \delta} \frac{1}{A} \int_A^{2A} |k(X, T, \xi)|^2 dXd\xi \ll_\epsilon A^{1/4} T^4 (AT)^\epsilon \delta^2,$$

and the second term, using again Lemma 4.4 and the second bound in (4-10), by

$$\int_{\delta < |\xi| \leq 1} |\hat{g}(\xi)|^2 |\xi| \int_{\delta < |\xi| \leq 1} \frac{1}{|\xi|} \frac{1}{A} \int_A^{2A} |k(X, T, \xi)|^2 dXd\xi \ll_\epsilon A^{1/4} T^2 (AT)^\epsilon \log(\delta^{-1}).$$

On taking $\delta = T^{-2}$ we obtain the claim. □

We are now able to conclude the proof of Theorem 1.4. Let $2 < T \leq A^{1/2}(\log A)^{-2}$. By partial summation,

$$\sum_{t_j \leq T} \frac{X^{1/2+it_j}}{\frac{1}{2} + it_j} = \frac{R(X, T)X^{1/2}}{\frac{1}{2} + iT} + iX^{1/2} \int_1^T \frac{R(X, U)}{(\frac{1}{2} + iU)^2} dU.$$

Therefore,

$$\frac{1}{A} \int_A^{2A} \left| \sum_{t_j \leq T} \frac{X^{1/2+it_j}}{\frac{1}{2} + it_j} \right|^2 dX \ll \frac{1}{A} \int_A^{2A} \left| \frac{R(X, T)X^{1/2}}{\frac{1}{2} + iT} \right|^2 dX + \frac{1}{A} \int_A^{2A} \left| X^{1/2} \int_1^T \frac{R(X, U)}{(\frac{1}{2} + iU)^2} dU \right|^2 dX.$$

The first term on the right hand side is bounded, in view of Proposition 4.5, by $O(A^{5/4+\epsilon})$. The second term can be bounded by using Cauchy–Schwarz inequality and again Proposition 4.5, giving

$$A(\log T) \int_1^T \frac{1}{U^3} \left(\frac{1}{A} \int_A^{2A} |R(X, U)|^2 dX \right) dU \ll_\epsilon A^{5/4+\epsilon} \int_1^T U^{-1} dU \ll_\epsilon A^{5/4+\epsilon}.$$

Inserting these bounds in (4-1) we obtain

$$\frac{1}{A} \int_A^{2A} |\psi_G(X) - X|^2 dX \ll_\epsilon A^{5/4+\epsilon} + \frac{A^2}{T^2} \log^2 A.$$

Choosing $T = A^{3/8}$ above concludes the proof of Theorem 1.4.

Weighted sums of Kloosterman sums. In the remainder of the section we prove Lemma 4.1. Recall that we want to estimate, on average, the following sum of Kloostermann sums

$$\frac{1}{A} \int_A^{2A} |S_n(\phi_{X,T})|^2 dX,$$

where $S_n(\psi)$ is the weighted sum of Kloosterman sums defined in (4-7), i.e.,

$$S_n(\psi) = \sum_{c=1}^{\infty} \frac{S(n, n, c)}{c} \psi\left(\frac{4\pi n}{c}\right),$$

with n a (large) positive integer. The function $\phi_{X,T}$ carries some oscillation in X , that we exploit when integrating over $X \in [A, 2A]$.

Lemma 4.6. *Let $A, T > 2$, and let z_1, z_2 be positive real numbers. Then*

$$\frac{1}{A} \int_A^{2A} \phi_{X,T}(z_1) \overline{\phi_{X,T}(z_2)} dX \ll z_1 z_2 A \exp\left(-\frac{A^{1/2}(z_1 + z_2)}{T}\right)$$

and

$$\frac{1}{A} \int_A^{2A} \phi_{X,T}(z_1) \overline{\phi_{X,T}(z_2)} dX \ll \frac{z_1 z_2 A^{1/2}}{|z_1 - z_2|}.$$

For the last bound we assume $z_1 \neq z_2$.

Proof. Inserting the definition of $\phi_{X,T}$ we can write the integral as

$$\frac{z_1 z_2}{\pi^2} \int_A^{2A} |\sinh \beta|^2 \exp(i \cosh \beta z_1 + i \overline{\cosh \beta z_2}) dX. \tag{4-11}$$

Bounding the integrand uniformly using

$$|\sinh \beta|^2 \ll X \quad \text{and} \quad \exp(i \cosh \beta z_1 + i \overline{\cosh \beta z_2}) \ll \exp\left(-\frac{A^{1/2}(z_1 + z_2)}{T}\right),$$

we obtain the first bound

$$\frac{1}{A} \int_A^{2A} \phi_{X,T}(z_1) \overline{\phi_{X,T}(z_2)} dX \ll z_1 z_2 A \exp\left(-\frac{A^{1/2}(z_1 + z_2)}{T}\right).$$

To obtain the second bound we use integration by parts in (4-11) to get

$$\frac{z_1 z_2}{\pi^2} f(X) \exp(i \cosh \beta z_1 + i \overline{\cosh \beta z_2}) \Big|_A^{2A} + \frac{z_1 z_2}{\pi^2} \int_A^{2A} f'(X) \exp(i \cosh \beta z_1 + i \overline{\cosh \beta z_2}) dX,$$

where

$$f(X) = \frac{2X |\sinh \beta|^2}{i \sinh \beta z_1 + i \overline{\sinh \beta z_2}}.$$

We can bound these terms (up to a constant) by

$$\frac{z_1 z_2 A^{3/2}}{|z_1 - z_2|},$$

which gives us the second bound

$$\frac{1}{A} \int_A^{2A} \phi_{X,T}(z_1) \overline{\phi_{X,T}(z_2)} dX \ll \frac{z_1 z_2 A^{1/2}}{|z_1 - z_2|}.$$

This proves the lemma. □

The lemma above will provide the necessary bounds to establish Lemma 4.1.

Proof of Lemma 4.1. We expand the square in the integrand and exchange the integral with the sums, so that we can rewrite the integral as

$$\sum_{c_1, c_2=1}^{\infty} \frac{S(n, n, c_1) \overline{S(n, n, c_2)}}{c_1 c_2} \frac{1}{A} \int_A^{2A} \phi_{X,T}\left(\frac{4\pi n}{c_1}\right) \overline{\phi_{X,T}\left(\frac{4\pi n}{c_2}\right)} dX. \tag{4-12}$$

We now split the sum in (4-12) into two sums Σ_d and Σ_{nd} , where Σ_d is the sum over the diagonal terms $c_1 = c_2$, and Σ_{nd} is the sum over the terms $c_1 \neq c_2$. We shall make use of the Weil bound on Kloosterman sums throughout the proof, namely

$$|S(n, n, c)| \leq (n, c)^{1/2} c^{1/2} d(c). \tag{4-13}$$

Moreover, we have

$$\sum_{c \leq x} (n, c) d^2(c) \ll x \log^3 x d(n).$$

We bound the diagonal terms using the first bound in Lemma 4.6 (with $z_1 = z_2 = (4\pi n)/c$), obtaining

$$\Sigma_d \ll n^2 A \sum_c \frac{(n, c) d^2(c)}{c^3} \exp\left(-\frac{A^{1/2} n}{Tc}\right) \ll T^2 \log^3(An) d(n).$$

To bound the nondiagonal terms we interpolate the two bounds in Lemma 4.6 to get, for $0 < \lambda < 1$,

$$\frac{1}{A} \int_A^{2A} \phi_{X,T}\left(\frac{4\pi n}{c_1}\right) \overline{\phi_{X,T}\left(\frac{4\pi n}{c_2}\right)} dX \ll \left(\frac{n^2 A}{c_1 c_2}\right)^{1-\lambda} \left(\frac{n A^{1/2}}{|c_1 - c_2|}\right)^\lambda.$$

Therefore,

$$\begin{aligned} \Sigma_{nd} &\ll \sum_{c_1 \neq c_2=1}^{\infty} \frac{|S(n, n, c_1) S(n, n, c_2)|}{c_1 c_2} \left(\frac{n^2 A}{c_1 c_2}\right)^{1-\lambda} \left(\frac{n A^{1/2}}{|c_1 - c_2|}\right)^\lambda \\ &\ll_{\lambda} (n A^{1/2})^{2-\lambda} \left(\sum_{c=1}^{\infty} \frac{|S(n, n, c)|^{2/(2-\lambda)}}{c^2}\right)^{2-\lambda}, \end{aligned}$$

where the last inequality follows from the Hardy–Littlewood–Pólya inequality (1-8). Applying the Weil bound (4-13) we obtain

$$\begin{aligned}\Sigma_{nd} &\ll_{\lambda} (nA^{1/2})^{2-\lambda} \left(\sum_{c=1}^{\infty} \frac{((n, c)^{1/2} d(c) \sqrt{c})^{2/(2-\lambda)}}{c^2} \right)^{2-\lambda} \\ &\ll_{\lambda} (nA^{1/2})^{2-\lambda} \left(\sum_{c=1}^{\infty} \frac{(n, c) d^2(c)}{c^{1+(1-\lambda)/(2-\lambda)}} \right)^{2-\lambda} \\ &\ll_{\lambda} (nA^{1/2} d(n))^{2-\lambda}.\end{aligned}$$

Pick $\lambda = 1 - \varepsilon$ and note that $d(n) \ll_{\varepsilon} n^{\varepsilon}$ to finish the proof. \square

References

- [Avdispahić 2017] M. Avdispahić, “Prime geodesic theorem of Gallagher type”, preprint, 2017. arXiv
- [Avdispahić 2018] M. Avdispahić, “On Koyama’s refinement of the prime geodesic theorem”, *Proc. Japan Acad. Ser. A Math. Sci.* **94**:3 (2018), 21–24. MR
- [Cai 2002] Y. Cai, “Prime geodesic theorem”, *J. Théor. Nombres Bordeaux* **14**:1 (2002), 59–72. MR Zbl
- [Carneiro and Vaaler 2010] E. Carneiro and J. D. Vaaler, “Some extremal functions in Fourier analysis, II”, *Trans. Amer. Math. Soc.* **362**:11 (2010), 5803–5843. MR Zbl
- [Cramér 1922] H. Cramér, “Ein Mittelwertsatz in der Primzahltheorie”, *Math. Z.* **12**:1 (1922), 147–153. MR JFM
- [Gallagher 1970] P. X. Gallagher, “A large sieve density estimate near $\sigma = 1$ ”, *Invent. Math.* **11**:4 (1970), 329–339. MR Zbl
- [Hardy et al. 1934] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge Univ. Press, 1934. JFM
- [Hejhal 1983] D. A. Hejhal, *The Selberg trace formula for $\mathrm{PSL}(2, \mathbb{R})$, II*, Lecture Notes in Math. **1001**, Springer, 1983. MR Zbl
- [Huber 1961] H. Huber, “Zur analytischen Theorie hyperbolischer Raumformen und Bewegungsgruppen, II”, *Math. Ann.* **142**:4 (1961), 385–398. Appendix in **143**:5 (1961), 463–464. MR Zbl
- [Ivić 1985] A. Ivić, *The Riemann zeta-function*, Wiley, New York, 1985. MR Zbl
- [Iwaniec 1984a] H. Iwaniec, “Nonholomorphic modular forms and their applications”, pp. 157–196 in *Modular forms* (Durham, 1983), edited by R. A. Rankin, Horwood, Chichester, UK, 1984. MR Zbl
- [Iwaniec 1984b] H. Iwaniec, “Prime geodesic theorem”, *J. Reine Angew. Math.* **349** (1984), 136–159. MR Zbl
- [Iwaniec 1995] H. Iwaniec, *Introduction to the spectral theory of automorphic forms*, Revista Matemática Iberoamericana, Madrid, 1995. MR Zbl
- [Koyama 1998] S.-y. Koyama, “Prime geodesic theorem for arithmetic compact surfaces”, *Int. Math. Res. Not.* **1998**:8 (1998), 383–388. MR Zbl
- [Koyama 2016] S.-y. Koyama, “Refinement of prime geodesic theorem”, *Proc. Japan Acad. Ser. A Math. Sci.* **92**:7 (2016), 77–81. MR Zbl
- [Kuznetsov 1980] N. V. Kuznetsov, “Pettersson’s conjecture for cusp forms of weight zero and Linnik’s conjecture: sums of Kloosterman sums”, *Mat. Sb. (N.S.)* **111(153)**:3 (1980), 334–383. In Russian; translation in *Math. USSR-Sb.* **39**:3 (1981), 299–342. MR Zbl
- [Li and Villavert 2011] C. Li and J. Villavert, “An extension of the Hardy–Littlewood–Pólya inequality”, *Acta Math. Sci. Ser. B Engl. Ed.* **31**:6 (2011), 2285–2288. MR Zbl
- [Luo and Sarnak 1995] W. Z. Luo and P. Sarnak, “Quantum ergodicity of eigenfunctions on $\mathrm{PSL}_2(\mathbb{Z}) \backslash \mathbb{H}^2$ ”, *Inst. Hautes Études Sci. Publ. Math.* **81** (1995), 207–237. MR Zbl
- [Luo et al. 1995] W. Luo, Z. Rudnick, and P. Sarnak, “On Selberg’s eigenvalue conjecture”, *Geom. Funct. Anal.* **5**:2 (1995), 387–401. MR Zbl

- [Petridis and Risager 2017] Y. N. Petridis and M. S. Risager, “Local average in hyperbolic lattice point counting”, *Math. Z.* **285**:3-4 (2017), 1319–1344. With an appendix by N. Laaksonen. MR Zbl
- [Phillips and Rudnick 1994] R. Phillips and Z. Rudnick, “The circle problem in the hyperbolic plane”, *J. Funct. Anal.* **121**:1 (1994), 78–116. MR Zbl
- [Rubinstein and Sarnak 1994] M. Rubinstein and P. Sarnak, “Chebyshev’s bias”, *Experiment. Math.* **3**:3 (1994), 173–197. MR Zbl
- [Sarnak 1980] P. C. Sarnak, *Prime geodesic theorems*, Ph.D. thesis, Stanford University, 1980, Available at <https://tinyurl.com/sarurl>.
- [Sarnak 1982] P. Sarnak, “Class numbers of indefinite binary quadratic forms”, *J. Number Theory* **15**:2 (1982), 229–247. Correction in **16**:2 (1983), 284. MR Zbl
- [Soundararajan and Young 2013] K. Soundararajan and M. P. Young, “The prime geodesic theorem”, *J. Reine Angew. Math.* **676** (2013), 105–120. MR Zbl
- [Venkov 1990] A. B. Venkov, *Spectral theory of automorphic functions and its applications*, Math. and Its Appl. (Soviet Series) **51**, Kluwer, Dordrecht, 1990. MR Zbl
- [Wintner 1935] A. Wintner, “On the asymptotic distribution of the remainder term of the prime-number theorem”, *Amer. J. Math.* **57**:3 (1935), 534–538. MR Zbl

Communicated by Peter Sarnak

Received 2017-05-23 Revised 2017-10-26 Accepted 2017-12-30

cherubini@mpim-bonn.mpg.de

Max-Planck-Institut für Mathematik, Bonn, Germany

guerreiro@mpim-bonn.mpg.de

Max-Planck-Institut für Mathematik, Bonn, Germany

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR

Bjorn Poonen
Massachusetts Institute of Technology
Cambridge, USA

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Richard E. Borcherds	University of California, Berkeley, USA	Martin Olsson	University of California, Berkeley, USA
J-L. Colliot-Thélène	CNRS, Université Paris-Sud, France	Raman Parimala	Emory University, USA
Brian D. Conrad	Stanford University, USA	Jonathan Pila	University of Oxford, UK
Samit Dasgupta	University of California, Santa Cruz, USA	Anand Pillay	University of Notre Dame, USA
Hélène Esnault	Freie Universität Berlin, Germany	Michael Rapoport	Universität Bonn, Germany
Gavril Farkas	Humboldt Universität zu Berlin, Germany	Victor Reiner	University of Minnesota, USA
Hubert Flenner	Ruhr-Universität, Germany	Peter Sarnak	Princeton University, USA
Sergey Fomin	University of Michigan, USA	Joseph H. Silverman	Brown University, USA
Edward Frenkel	University of California, Berkeley, USA	Michael Singer	North Carolina State University, USA
Andrew Granville	Université de Montréal, Canada	Christopher Skinner	Princeton University, USA
Joseph Gubeladze	San Francisco State University, USA	Vasudevan Srinivas	Tata Inst. of Fund. Research, India
Roger Heath-Brown	Oxford University, UK	J. Toby Stafford	University of Michigan, USA
Craig Huneke	University of Virginia, USA	Pham Huu Tiep	University of Arizona, USA
Kiran S. Kedlaya	Univ. of California, San Diego, USA	Ravi Vakil	Stanford University, USA
János Kollár	Princeton University, USA	Michel van den Bergh	Hasselt University, Belgium
Philippe Michel	École Polytechnique Fédérale de Lausanne	Marie-France Vignéras	Université Paris VII, France
Susan Montgomery	University of Southern California, USA	Kei-Ichi Watanabe	Nihon University, Japan
Shigefumi Mori	RIMS, Kyoto University, Japan	Shou-Wu Zhang	Princeton University, USA

PRODUCTION

production@msp.org
Silvio Levy, Scientific Editor


See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2018 is US \$340/year for the electronic version, and \$535/year (+\$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW[®] from MSP.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2018 Mathematical Sciences Publishers

Algebra & Number Theory

Volume 12 No. 3 2018

Pseudo-exponential maps, variants, and quasiminimality MARTIN BAYS and JONATHAN KIRBY	493
On faithfulness of the lifting for Hopf algebras and fusion categories PAVEL ETINGOF	551
Mean square in the prime geodesic theorem GIACOMO CHERUBINI and JOÃO GUERREIRO	571
Elliptic quantum groups and Baxter relations HUAFENG ZHANG	599
Differential forms in positive characteristic, II: cdh-descent via functorial Riemann–Zariski spaces ANNETTE HUBER and SHANE KELLY	649
Nilpotence order growth of recursion operators in characteristic p ANNA MEDVEDOVSKY	693
Algebraic de Rham theory for weakly holomorphic modular forms of level one FRANCIS BROWN and RICHARD HAIN	723



1937-0652(2018)12:3;1-E