Dmitry Jakobson and Frédéric Naud

Lower bounds for resonances of infinite-area Riemann surfaces
LOWER BOUNDS FOR RESONANCES OF INFINITE-AREA RIEMANN SURFACES

Dmitry Jakobson and Frédéric Naud

For infinite-area, geometrically finite surfaces \( X = \Gamma \setminus \mathbb{H}^2 \), we prove new omega lower bounds on the local density of resonances \( \mathcal{D}(z) \) when \( z \) lies in a logarithmic neighborhood of the real axis. These lower bounds involve the dimension \( \delta \) of the limit set of \( \Gamma \). The first bound is valid when \( \delta > \frac{1}{2} \) and shows logarithmic growth of the number \( \mathcal{D}(z) \) of resonances at high energy, that is, when \( |\text{Re}(z)| \to +\infty \). The second bound holds for \( \delta > \frac{3}{4} \) and if \( \Gamma \) is an infinite-index subgroup of certain arithmetic groups. In this case we obtain a polynomial lower bound. Both results are in favor of a conjecture of Guillopé and Zworski on the existence of a fractal Weyl law for resonances.

1. Introduction and results

Resonances arise in spectral theory on noncompact Riemannian manifolds when one tries to figure out what the natural replacement data should be for the missing eigenvalues of the Laplacian. The basic problem of the mathematical theory of resonances is to relate the resonances spectrum (which is a discrete set of complex numbers) to the geometry of the underlying manifold and its geodesic flow. In this paper we will focus on a particular setting where the spectral and scattering theory are already well developed: infinite-area surfaces with constant negative curvature. For a detailed account of the spectral theory of infinite-area surfaces, we refer the reader to [Borthwick 2007]. Let \( \mathbb{H}^2 \) be the hyperbolic plane endowed with its standard metric of constant Gaussian curvature \(-1\). Let \( \Gamma \) be a geometrically finite discrete group of isometries acting on \( \mathbb{H}^2 \). This means that \( \Gamma \) admits a finite sided polygonal fundamental domain in \( \mathbb{H}^2 \). We will require that \( \Gamma \) has no elliptic elements different from the identity and that the quotient \( \Gamma \setminus \mathbb{H}^2 \) is of infinite hyperbolic area. Under these assumptions, the quotient space \( X = \Gamma \setminus \mathbb{H}^2 \) is a nice Riemann surface whose geometry can be described as follows. The surface \( X \) can be decomposed into a finite area surface with geodesic boundary \( N \), called the Nielsen region, on which infinite-area ends \( F_i \) are glued: the funnels. We assume throughout that the number of funnels \( f \) is not zero. Each funnel \( F_i \) is isometric to a half-cylinder

\[ F_i = (\mathbb{R}/l_i\mathbb{Z})_\theta \times (\mathbb{R}^+)_t, \]

where \( l_i > 0 \), with the warped metric

\[ ds^2 = dt^2 + \cosh^2 t \, d\theta^2. \]
The Nielsen region $N$ is itself decomposed into a compact surface $K$ with geodesic and horocyclic boundary on which $c$ noncompact, finite area ends $C_i$ are glued: the cusps. A cusp $C_i$ is isometric to a half-cylinder

$$C_i = (\mathbb{R}/h_i\mathbb{Z})_\theta \times ([1, +\infty))_y,$$

where $h_i > 0$, endowed with the familiar Poincaré metric

$$ds^2 = d\theta^2 + \frac{dy^2}{y^2}.$$ 

Let $\Delta_X$ be the hyperbolic Laplacian on $X$. Its spectrum on $L^2(X)$ has been described by Lax and Phillips [1984a; 1984b; 1985]: $[\frac{1}{4}, +\infty)$ is the continuous spectrum, and there are no embedded eigenvalues. The rest of the spectrum is made of a (possibly empty) finite set of eigenvalues, starting at $\delta(1 - \delta)$, where $0 \leq \delta < 1$ is the Hausdorff dimension of the limit set of $\Gamma$. The fact that the bottom of the spectrum is related to the dimension $\delta$ was first pointed out by Patterson [1976] for convex cocompact groups (which amounts to saying that there are no cusps on $X$ or equivalently, no parabolic elements in $\Gamma$). This result was later extended for geometrically finite groups by Sullivan [1979; 1984].

The dimension $\delta$ has another important interpretation. Let $S_1X$ denote the unit tangent bundle; then the trapped set is defined as the set of points in $S_1X$ whose orbit under the geodesic flow remains (after projection on $X$) in the Nielsen region $N$ in the past and future. The Liouville measure of this set is always zero, but its Hausdorff dimension is actually $2\delta + 1$.

By the preceding description of the spectrum, the resolvent

$$R(\lambda) = (\Delta_X - \frac{1}{4} - \lambda^2)^{-1} : L^2(X) \to L^2(X),$$

is therefore well defined and analytic on the lower half-plane $\{\text{Im} \lambda < 0\}$ except at a possible finite set of poles corresponding to the finite point spectrum. Resonances are then defined as poles of the meromorphic continuation of

$$R(\lambda) : C^\infty_0(X) \to C^\infty(X)$$

to the whole complex plane. The set of poles is denoted by $\mathcal{R}_X$. This continuation is usually performed via the analytic Fredholm theorem after the construction of an adequate parametrix. The first result of this kind in the more general setting of asymptotically hyperbolic manifolds is due to Mazzeo and Melrose [1987]. A more precise parametrix for surfaces was constructed by Guillopé and Zworski [1995; 1997]; it allowed them to obtain global counting results for resonances of the following type. Let $N(R)$ be the number of resonances (counted with multiplicity) of modulus smaller than $R$. We have for all $R \geq 0$,

$$C^{-1}R^2 \leq N(R) \leq C + CR^2,$$

for some $C > 0$. Hence the set of resonances satisfies a quadratic growth law similar to the usual Weyl law for finite area surfaces. These bounds are actually valid for compact perturbations of the hyperbolic metric [Borthwick 2008], and in particular are not sensitive to the geometry of the trapped set. It is therefore necessary to examine finer properties of $\mathcal{R}_X$ to recover some geometrical information on $X$. The most natural thing to do is to look at resonances that are close to the real axis. Physically, these are the most relevant resonances, because they correspond to metastable states that live the longest (the imaginary part corresponding to the decay rate). In the case of Schottky groups (equivalently, convex cocompact quotients
in dimension 2), a “fractal” upper bound was obtained in [Zworski 1999; Guillopé et al. 2004], namely
\[
N_C(T) = O(T^{1+\delta}),
\]
where
\[
N_C(T) := \#\{z \in \mathbb{R}_X : \text{Im } z \leq C, |\text{Re } z| \leq T\}.
\]

The first proof of a geometric bound of this type involving fractal dimension is due to Sjöstrand [1990] for potential scattering. This upper bound, together with numerical experiments, has led to the following conjecture, known as the fractal Weyl law.

**Conjecture 1.1** (Guillopé–Zworski). There exist \( C > 0 \) and \( A > 0 \) such that for all \( T \) large enough,
\[
A^{-1} T^{1+\delta} \leq N_C(T) \leq A T^{1+\delta}.
\]
The only existing lower bound can be found in [Guillopé and Zworski 1999], where the authors show that for all \( \epsilon > 0 \), one can find \( C_\epsilon > 0 \) such that
\[
N_{C_\epsilon}(T) = O(T^{1-\epsilon}),
\]
where \( \Omega(\cdot) \) means being not a \( O(\cdot) \); in other words, one can find a sequence \( (T_i)_{i \in \mathbb{N}} \) with \( T_i \to \infty \) such that
\[
\lim_{i \to \infty} \frac{N_{C_\epsilon}(T_i)}{T_i^{1-\epsilon}} = +\infty.
\]
This is a frustrating lower bound: not only it does not involve \( \delta \) but it is not even optimal in the computable case of elementary groups where \( N_C(T) \) grows linearly. Guillopé et al. [2004] actually prove a stronger statement than (1). Let \( \mathcal{D}(z) \) be the number of resonances in the disc centered at \( z \) and radius one:
\[
\mathcal{D}(z) := \#\{\lambda \in \mathbb{R}_X : |\lambda - z| \leq 1\}.
\]
Then if \( \text{Im } z \leq C \), we have \( \mathcal{D}(z) = O(|\text{Re } z|^{\delta}) \), the implied constant depending solely on \( C \). A similar statement for semiclassical Schrödinger operators can be found in [Sjöstrand and Zworski 2007]. Note that if the Guillopé–Zworski conjecture holds, then by the box principle, for all \( \epsilon > 0 \), one can find a sequence \( (z_i) \) with \( |\text{Re } z_i| \to +\infty \) and \( \text{Im } z_i \leq C \) such that for all \( i \in \mathbb{N} \),
\[
\mathcal{D}(z_i) \geq |\text{Re } z_i|^{\delta-\epsilon}.
\]
To state our results, we need one more piece of notation. Let \( A > 0 \) and set
\[
W_A = \{\lambda \in \mathbb{C} : \text{Im } \lambda \leq A \log(1 + |\text{Re } \lambda|)\}.
\]
Guillopé and Zworski [1997] have shown that in logarithmic regions \( W_A \), the density of resonances grows at least linearly. We shall prove the following thing.

**Theorem 1.2.** Let \( \Gamma \) be a geometrically finite group as above. Assume that \( \delta > \frac{1}{2} \), and fix arbitrarily small \( \epsilon > 0 \) and \( A > 0 \). Then there exists a sequence \( (z_i)_{i \in \mathbb{N}} \) with \( z_i \in W_A \) and \( |\text{Re } z_i| \to +\infty \), such that for all \( i \geq 0 \),
\[
\mathcal{D}(z_i) \geq (\log |\text{Re } z_i|)^{(\delta-1/2)/\delta-\epsilon}.
\]
In other words, the local density \( \mathcal{D}(z) \) of resonances in logarithmic regions \( W_A \) is not bounded, and sensitive to the dimension of the trapped set. This implies in particular that the resonance set \( \mathbb{R}_X \cap W_A \)
is different from a lattice when $\delta > \frac{1}{2}$, which clearly could not follow from the existing lower bound in strips nor the global counting results. Building groups with $\delta > \frac{1}{2}$ is easy: if there is a parabolic element this is always the case and if one wants to consider only convex-cocompact groups, pinching a pair of pants will do it; see Section 4. We point out that the proof is based on Dirichlet box arguments, a technique that has proved useful to obtain lower bounds for the remainder in Weyl’s law on compact negatively curved manifolds; see [Jakobson et al. 2008; Jakobson and Polterovich 2007].

It is possible to obtain significantly better lower bounds that are closer to (2), by using infinite-index subgroups of arithmetic groups. Arithmetic groups are algebraically defined discrete groups of isometries of $\mathbb{H}^2$, the most celebrated being the modular group $\text{PSL}_2(\mathbb{Z})$. For more details on definitions and references, see Section 3. Our result is as follows.

**Theorem 1.3.** Let $\Gamma$ be a geometrically finite group as above, and assume that $\Gamma$ is an infinite-index subgroup of an arithmetic group $\Gamma_0$ derived from a quaternion algebra. Suppose $\delta > \frac{3}{4}$, and fix arbitrarily small $\epsilon > 0$ and $A > 0$. Then there exists a sequence $(z_k) \in W_A$ with $|\text{Re } z_k| \rightarrow +\infty$, such that for all $k \geq 0$,

$$\mathcal{D}(z_k) \geq |\text{Re } z_k|^{2\delta - \frac{3}{2} - \epsilon}.$$  

This improvement is based on the very special structure of closed geodesics on arithmetic surfaces: the set of lengths has high multiplicities and good separation (see Section 3 for more details). We point out that these techniques due to Selberg have been used recently by Anantharaman [2009] to obtain some results on the spectral deviations for the damped wave equation on compact arithmetic surfaces. This lower bound is clearly in favor of the Guillopé–Zworski conjecture, at least for the class of groups considered above. One may wonder at this point if Theorem 1.3 is not empty: Gamburd [2002] has shown in (see Section 4 for details) the existence of several geometrically finite subgroups $\Gamma$ of $\text{PSL}_2(\mathbb{Z})$ with dimension $\delta > \frac{3}{4}$. Another natural question is can we give a bound on the sequence $|\text{Re } z_k|$? We explain at the end of Section 3 how one can obtain a polynomial upper bound: for each $\epsilon > 0$ one can find an exponent $p_\epsilon > 0$ such that $|\text{Re } z_k| = O \left( k^{p_\epsilon} \right)$.

The lower bounds obtained above are to our knowledge the first examples in the literature which are related to the dimension of the trapped set, at least for fractal dimensions. Similar results should hold for higher dimensional convex-compact manifolds, by applying a similar strategy of proof based on the trace formula in [Guillarmou and Naud 2006].

The plan of the paper is as follows: in Section 2 we recall the necessary material for the proofs, including the wave trace formula which is at the basis of our results. We then prove Theorem 1.2 by a Dirichlet box-principle argument. Section 3 is devoted to the case of arithmetically built groups. The heart of the proof is based on a trick of Selberg and Hejhal on mean square estimates. This is where the high multiplicity and the separation play a key role. In Section 4 we discuss various examples of geometrically finite groups with $\delta$ large, and we construct an explicit family of convex cocompact subgroups of $\text{PSL}_2(\mathbb{Z})$ with $\delta > \frac{3}{4}$.

## 2. Wave trace and log lower bounds

In this section, we prove Theorem 1.2. Some of the technical estimates below will be of some use in the next section. We use the notation of the introduction. The constant $A > 0$ defining the logarithmic region $W_A$ is set once for all.
The variant of Selberg’s trace formula we need here is due to [Guillopé and Zworski 1999]. We denote by $\mathcal{P}$ the set of primitive closed geodesics on the surface $X = \Gamma \backslash \mathbb{H}^2$, and if $\gamma \in \mathcal{P}$, $l(\gamma)$ is the length. In the following, $c$ is the number of cusps, and $N$ is the Nielsen region. Let $\varphi \in C_0^\infty((0, +\infty))$, that is, a smooth function, compactly supported in $\mathbb{R}_+^*$. We have the identity

$$
\sum_{\lambda \in \mathbb{R}_X} \hat{\varphi}(-\lambda) = -\frac{\text{Vol}(N)}{4\pi} \int_0^{+\infty} \frac{\cosh(x/2)}{\sinh^2(x/2)} \varphi(x)dx + \frac{c}{2} \int_0^{+\infty} \frac{\cosh(x/2)}{\sinh(x/2)} \varphi(x)dx + \sum_{\gamma \in \mathcal{P}} \sum_{k \geq 1} \frac{l(\gamma)}{2\sinh(\lambda(\gamma)/2)} \varphi(l(\gamma)),
$$

where $\hat{\varphi}$ is the usual Fourier transform,

$$
\hat{\varphi}(\xi) = \int_{\mathbb{R}} \varphi(x)e^{-ix\xi}dx.
$$

We recall that $\mathbb{R}_X$ (except a possible finite number of term on the imaginary axis starting at $\lambda = i(\frac{1}{2} - \delta)$) is included in the upper half-plane. Note that we have omitted the main singular terms at $t = 0$ which are not relevant for our problem; see [Guillopé and Zworski 1999] for the formula in full detail. Proofs of Theorem 1.2 and 1.3 are based on the use of test functions of the form

$$
\varphi_{t,\alpha}(x) = e^{-itx} \varphi_0(x - \alpha),
$$

where $t > 0$, $\alpha > 0$ will be large and $\varphi_0 \in C_0^\infty(\mathbb{R})$ is a positive function, supported on the interval $[-1, +1]$ identical to 1 on $[-\frac{1}{2}, +\frac{1}{2}]$. The basic idea is to use the full-length spectrum (the set of lengths of closed geodesics) in the contribution from the geometric side instead of one single, closed primitive geodesic and its iterates as in the proof of Guillopé and Zworski [1999]. The price to pay for that is to lose positivity and deal with oscillating contributions. We start with some useful lemmas that consist mainly of brute force estimates. They will be used to control sums over resonances in the proof of Theorem 1.2 and 1.3. The reader can skip it for its first reading.

**Lemma 2.1.** For all $N \geq 0$, one can find $C_N > 0$ such that for all $\zeta \in \mathbb{C}$,

$$
|\hat{\varphi}_{t,\alpha}(\zeta)| \leq C_N \frac{e^{\alpha \Im \zeta + |\Im \zeta|}}{(1 + |t + \zeta|)^N}.
$$

**Proof.** Write $\hat{\varphi}_{t,\alpha}(\zeta) = e^{-i\alpha(t+\zeta)}\hat{\varphi}_0(t + \zeta)$, and integrate by parts $N$ times. Notice that while estimating $|\hat{\varphi}_0(u)|$ with $u \in \mathbb{C}$, there is an extra factor $e^{\Im u}$ coming out, which explain the presence of the (harmless) extra term $|\Im \zeta|$ in the exponents above. \hfill \Box

**Lemma 2.2.** Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be either $f(x) = (\log(1 + x))^\beta$ or $f(x) = x^\beta$ with $0 < \beta < 1$. Assume that $\mathcal{D}(z) = O(f(|\Re z|))$ for all $z \in W_A$ with $|\Re z|$ large enough. Then, for all $\alpha$, $t$ large and all $k \geq 0$, one has

$$
\left| \sum_{\lambda \in W_A \cap \mathbb{R}_X} \varphi_{\alpha,t}(-\lambda) \right| = O\left(\frac{e^{\alpha(\beta - 1/2)}}{t^k}\right) + O(f(t)),
$$

where the implied constants do not depend on $\alpha$, $t$. 

**Proof.** This follows from Lemma 2.1.
Proof. Let us assume that $\mathcal{D}(z) = O(f(|\Re z|))$ whenever $|\Re(z)| \geq p_0 \geq 1$ and $z \in W_A$. Let $t > 0$ be so large that $t > p_0 + 1$, assume that $\alpha > 1$. By absolute convergence one can write

$$\sum_{\lambda \in W_A \cap \mathbb{R}^X} \hat{\varphi}_{\alpha,t}(-\lambda) = \sum_{p \in \mathbb{Z}} \sum_{\lambda \in W_A \cap \mathbb{R}^X ; p \leq \Re \lambda \leq p+1} \hat{\varphi}_{\alpha,t}(-\lambda).$$

Let us set

$$S_p(\alpha, t) = \sum_{p \leq \Re \lambda \leq p+1} \hat{\varphi}_{\alpha,t}(-\lambda).$$

We split the above sum as

$$\sum_{\lambda \in W_A \cap \mathbb{R}^X} \hat{\varphi}_{\alpha,t}(-\lambda) = \sum_{p < -p_0} S_p(\alpha, t) + \sum_{-p_0 \leq p \leq p_0} S_p(\alpha, t) + \sum_{p > p_0} S_p(\alpha, t).$$

The middle term involves only finitely many resonances $\lambda \in W_A$, and they satisfy $\Im \lambda \geq \frac{1}{2} - \delta$. Therefore using Lemma 2.1, we have

$$\left| \sum_{-p_0 \leq p \leq p_0} S_p(\alpha, t) \right| \leq C_k \sum_{\lambda \in \mathbb{R}^X \cap W_A} e^{(-\alpha + 1)(1/2 - \delta)} \sum_{p \leq \Re \lambda \leq p_0} 1 = O\left(\frac{e^{\alpha (\delta - 1/2)}}{t^k}\right).$$

The first term can be estimated as

$$\left| \sum_{p < -p_0} S_p(\alpha, t) \right| \leq C_2 \sum_{p < -p_0} \frac{1}{(1 + |p - t| - 1)^2} \sum_{p \leq \Re \lambda \leq p + 1} e^{(-\alpha + 1) \Im \lambda},$$

while the last term is of size

$$\left| \sum_{p > p_0} S_p(\alpha, t) \right| \leq C_2 \sum_{p > p_0} \frac{\tilde{S}_p(\alpha)}{(1 + \min\{|p - t|, |p + 1 - t|\})^2},$$

where we have set

$$\tilde{S}_p(\alpha) = \sum_{p \leq \Re \lambda \leq p+1} e^{(-\alpha + 1) \Im \lambda}.$$

The following lemma will be convenient (this is where the hypothesis on $\mathcal{D}(z)$ is used).

**Lemma 2.3.** Under the hypothesis of Lemma 2.2, there exists a constant $M$, independent of $\alpha$, $p$ and such that for all $|p| \geq p_0$, we have

$$\tilde{S}_p(\alpha) \leq M f(|p|).$$

Let us postpone the proof of this result for a moment and show how to end the proof of Lemma 2.2. Clearly, using Lemma 2.3, the sum of the first and last terms is smaller than

$$C \sum_{p \in \mathbb{Z}} \frac{f(|p|)}{(1 + |p - t|)^2}.$$
for a constant $C > 0$ large enough. We can now write, denoting by $[t]$ the integer part of $t$,

$$\sum_{p \in \mathbb{Z}} \frac{f(|p|)}{(1 + |p-t|)^2} = \sum_{q \in \mathbb{Z}} \frac{f(|q| + [t])}{(1 + |q + [t] - t|)^2} \leq C' \sum_{q \in \mathbb{Z}} \frac{f(|q| + [t])}{(1 + |q|)^2},$$

again for a well chosen $C' > 0$ (we have used the fact that $f$ is increasing). To end the proof, simply write

$$\sum_{q \in \mathbb{Z}} \frac{f(|q| + [t])}{(1 + |q|)^2} = \sum_{|q| \leq [t]} \frac{f(|q| + [t])}{(1 + |q|)^2} + \sum_{|q| > [t]} \frac{f(|q| + [t])}{(1 + |q|)^2},$$

which yields

$$\sum_{q \in \mathbb{Z}} \frac{f(|q| + [t])}{(1 + |q|)^2} \leq f(2[t]) \sum_{q \in \mathbb{Z}} \frac{1}{(1 + |q|)^2} + \sum_{|q| > [t]} \frac{f(2|q|)}{(1 + |q|)^2}.$$ 

Since $f(2|q|) = O(|q|^{1-\epsilon})$, the second term is clearly bounded in $t$ and we get the upper bound of size $O(f(2t))$. It remains to prove Lemma 2.3. It will follow from a standard covering argument. It is enough to consider just the case $p > p_0$. We recall that for all $\lambda \in \mathbb{R}_X$, then for Re $\lambda \not= 0$ we have actually Im $\lambda \geq 0$ by definition. Let $A_p$ denote the set

$$A_p = \{z \in W_A : p \leq \text{Re} z \leq p + 1\},$$

let $D(z)$ denote the unit disc centered at $z \in \mathbb{C}$, and set

$$K(p) = \max \{ k \geq 0 : k\sqrt{3} \leq A \log(1 + p) \}.$$

For $1 \leq k \leq K(p)$, we define the rectangle $R(k)$ by

$$R(k) = \{ z \in A_p : (k - 1)\sqrt{3} \leq \text{Im} z \leq k\sqrt{3} \}.$$

Set $l = A \log(1 + p) - K(p)\sqrt{3} < \sqrt{3}$. One can check that, for $p$ large enough,

$$A_p \subset \left( \bigcup_{k=1}^{K(p)} R(k) \right) \cup D(p + \frac{1}{2} + i(K(p) + l/2)) \cup D(p + \frac{1}{2} + i(K(p) + l)).$$

Indeed,

$$A_p \setminus \left( \bigcup_{k=1}^{K(p)} R(k) \right)$$

is exactly the set

$$\{ z \in \mathbb{C} : p \leq \text{Re} z \leq p + 1 \text{ and } K(p)\sqrt{3} \leq \text{Im} z \leq A \log(1 + \text{Re} z) \},$$

which is clearly covered by the union of the two above discs as long as

$$A \log(1 + p + 1) - A \log(1 + p) = A \log \left(1 + \frac{1}{p+1}\right) \leq \frac{\sqrt{3}}{2}.$$
Note that for all $k = 1, \ldots, K(p)$, we have $R(k) \subset D(p + \frac{1}{2} + i\sqrt{3}(\frac{1}{2} + k - 1))$. We can now conclude by estimating

$$\tilde{S}_p(\alpha) = \sum_{\lambda \in \mathbb{R} \cap \partial X_{K(p)-1}} e^{(-\alpha+1) \text{Im} \lambda} \leq \sum_{j=0}^{K(p)-1} \mathbb{D}(p + \frac{1}{2} + i\sqrt{3}(\frac{1}{2} + j)) e^{(-\alpha+1)j\sqrt{3} + \mathbb{D}(p + \frac{1}{2} + i(K(p)) + \frac{1}{2}) + \mathbb{D}(p + \frac{1}{2} + i(K(p)) + \frac{1}{2})}.$$ 

Recalling that $\alpha > 1$ and $\mathbb{D}(z) \leq Cf(|\text{Re} z|)$ for all $z \in W_A$ with $|\text{Re} z| \geq p_0$, we thus obtain

$$\tilde{S}_p(\alpha) \leq 2Cf(p + \frac{1}{2}) + C \frac{f(p + \frac{1}{2})}{1 - e^{(-\alpha+1)\sqrt{3}}},$$

and therefore $\tilde{S}_p(\alpha) = O(f(p))$, uniformly in $\alpha$. □

Before we start the proof of Theorem 1.2, we need one more lemma, which is the key observation that motivates the definition of the region $W_A$.

**Lemma 2.4.** There exist some constants $a_0, C_0 > 0$, independent of $\alpha, t$ such that for all $\alpha \geq a_0$,

$$\left| \sum_{\lambda \in \mathbb{R} \setminus W_A} \overline{\varphi}_{\alpha, t}(-\lambda) \right| \leq C_0.$$

**Proof.** We assume first that $\alpha > 1$. If $\lambda \notin W_A$, then $\text{Im} \lambda \geq 0$ and

$$|\lambda|^2 = (\text{Re} \lambda)^2 + (\text{Im} \lambda)^2 \leq e^{(2/A) \text{Im} \lambda} + (\text{Im} \lambda)^2 \leq e^{(3/A) \text{Im} \lambda},$$

whenever $\text{Im} \lambda \geq C_A$ where $C_A$ is a large enough constant depending on $A$. We can assume in the sequel that $C_A \geq 1$. Using Lemma 2.1 with $N = 0$, we get

$$\left| \sum_{\lambda \in \mathbb{R} \setminus W_A} \overline{\varphi}_{\alpha, t}(-\lambda) \right| \leq C_0\#\{\lambda \in \mathbb{R} \setminus W_A : \text{Im} \lambda \leq C_A\} + \sum_{\lambda \in \mathbb{R} \setminus W_A \atop \text{Im} \lambda \geq C_A} \frac{1}{|\lambda|^{(a-1)2A/3}}.$$

The first term is clearly independent of $\alpha$ while the second can be bounded by the Stieltjes integral

$$\sum_{\lambda \in \mathbb{R} \setminus W_A \atop \text{Im} \lambda \geq C_A} \frac{1}{|\lambda|^{(a-1)2A/3}} \leq \int_{1}^{+\infty} u^{-(a-1)2A/3} dN(u),$$

where $N(u) = O(u^2)$ is the counting function for resonances in discs defined in Section 1. By integration by parts, the above integral is clearly convergent and bounded in $\alpha$ as long as

$$A(a-1) > 3.$$

The proof is complete. □
We can now start the proof of Theorem 1.2. Let’s test the trace formula (3) with the family $\varphi_{\alpha,t}$, where $\alpha$ is a large positive number:

$$\sum_{\lambda \in \mathbb{R}_x} \varphi_{\alpha,t}(-\lambda) = -\frac{\text{Vol}(N)}{4\pi} \int_{\alpha-1}^{\alpha+1} \cosh(x/2) \varphi_{\alpha,t}(x) \, dx + \frac{c}{2} \int_{\alpha-1}^{\alpha+1} \frac{\cosh(x/2)}{\sinh(x/2)} \varphi_{\alpha,t}(x) \, dx + \sum_{\alpha - 1 \leq kl(\gamma) \leq \alpha + 1} \frac{l(\gamma)}{2 \sinh(kl(\gamma)/2)} e^{-itkl(\gamma)} \varphi_0(kl(\gamma) - \alpha).$$

The first two terms on the right side are clearly bounded with respect to $\alpha$ and $t$. To get an appropriate control on the sum

$$\mathcal{G}_{\alpha,t} := \sum_{\alpha - 1 \leq kl(\gamma) \leq \alpha + 1} \frac{l(\gamma)}{2 \sinh(kl(\gamma)/2)} e^{-itkl(\gamma)} \varphi_0(kl(\gamma) - \alpha),$$

we will use the following lemma, also known as the Dirichlet box theorem.

**Lemma 2.5.** Let $\alpha_1, \ldots, \alpha_N \in \mathbb{R}$, and $D \in \mathbb{N}^*$. For all $Q \geq 2$ one can find an integer $q \in \{D, \ldots, DQ^N\}$ such that

$$\max_{1 \leq j \leq N} \|q\alpha_j\| \leq \frac{1}{Q},$$

where $\|x\| = \min_{n \in \mathbb{Z}} |x - n|$.

**Proof.** We use the box principle. Set

$$N_\alpha := |\{(k, l(\gamma)) \in \mathbb{N}^* \times \mathcal{P} : kl(\gamma) \in [\alpha - 1, \alpha + 1]\}|.$$

Fix a constant $\varepsilon_0 > 0$ and set $D_\alpha = [(4\pi)^\varepsilon_0 N_\alpha]$. By Lemma 2.5 with $Q = [4\pi]$, for all $\alpha \gg 1$, one can find $q_\alpha \in \{D_\alpha, \ldots, D_\alpha Q^N\}$ such that

$$\max_{\alpha - 1 \leq kl(\gamma) \leq \alpha + 1} \|q_\alpha k l(\gamma)\| \leq \frac{1}{Q}.$$  

Set $t_\alpha := 2\pi q_\alpha$, we have for all $\alpha - 1 \leq kl(\gamma) \leq \alpha + 1$,

$$|e^{it_\alpha kl(\gamma)} - 1| \leq \frac{2\pi}{Q} < \frac{2}{3}.$$

Hence we get

$$|\mathcal{G}_{\alpha,t_\alpha}| \geq \frac{1}{3} \left( \sum_{\alpha - 1 \leq kl(\gamma) \leq \alpha + 1} \frac{l(\gamma)}{2 \sinh(kl(\gamma)/2)} \varphi_0(kl(\gamma) - \alpha) \right) \geq C_0 e^{-\alpha/2} \sum_{\alpha - \frac{1}{2} \leq kl(\gamma) \leq \alpha + \frac{1}{2}} 1,$$

for a well chosen constant $C_0 > 0$. We now recall that by the prime geodesic theorem (see [Naud 2005] for a proof and references in the case of infinite-area surfaces), one has, as $T \to +\infty$,

$$\#\{(k, l(\gamma)) \in \mathbb{N}^* \times \mathcal{P} : kl(\gamma) \leq T\} = \frac{e^{\delta T}}{\delta T} (1 + o(1)).$$

This yields, for $\alpha$ large,

$$|\mathcal{G}_{\alpha,t_\alpha}| \geq C \frac{e^{(\delta - 1/2)\alpha}}{\alpha},$$
where $C_1$ is again a suitable constant. Using the prime geodesic theorem, one shows also that

$$C_1^{-1} e^{\delta \alpha} \leq N_\alpha \leq C_2 e^{\delta \alpha},$$

with $C_2 > 0$ and $\alpha$ large. We have therefore $\log \log t_\alpha \leq \delta \alpha + \text{constants}$, which can be more conveniently restated as

$$\log \log t_\alpha \leq (\delta + \epsilon) \alpha \quad \text{for all } \epsilon > 0 \text{ and } \alpha \text{ large.}$$

Similarly we get the lower bound

$$\log \log t_\alpha \geq (\delta - \epsilon) \alpha.$$

We can now conclude the proof. Assume that $\delta > \frac{1}{2}$. Suppose that for all $z \in W_A$ with $|\Re z| \geq R_0$, one has $\mathcal{D}(z) \leq (\log |\Re z|)^\beta$, where $\beta > 0$ will be determined later on. Then by Lemma 2.2 with $k = 1$, and Lemma 2.4, one gets as $\alpha \to +\infty$,

$$C_1 e^{(\delta - 1/2)\alpha} \leq |\mathcal{F}_{\alpha, t_\alpha}| \leq O(1) + O \left( \frac{e^{\alpha(\delta - 1/2)}}{t_\alpha} \right) + O((\log t_\alpha)^\beta).$$

Now recall that

$$\frac{\log \log t_\alpha}{\delta + \epsilon} \leq \alpha \leq \frac{\log \log t_\alpha}{\delta - \epsilon},$$

so that we have

$$\frac{C_1 (\delta + \epsilon)}{\log \log t_\alpha} (\log t_\alpha)^{(\delta - 1/2)/\delta + \epsilon)} \leq O(1) + O \left( \frac{(\log t_\alpha)^{(\delta - 1/2)/\delta - \epsilon)}}{t_\alpha} \right) + O((\log t_\alpha)^\beta).$$

We have a contradiction whenever $\beta < (\delta - 1/2)/(\delta + \epsilon)$. As a conclusion, for all $\epsilon > 0$ and all $R_0 \geq 0$ one can find $z \in W_A$ with $|\Re z| \geq R_0$ and $\mathcal{D}(z) > (\log |\Re z|)^{(\delta - 1/2)/(\delta)} - \epsilon$. This proves Theorem 1.2. ☐

### 3. Mean square lower bounds and arithmetic length spectrum

The goal of this section is to prove Theorem 1.3. First we need to recall a few basic facts about arithmetic group. Instead of detailing the construction of such groups, we refer the reader to the introductory book [Katok 1992], and will use a characterization of arithmetic groups derived from quaternion algebra due to Takeuchi [1975], which is all we need for this section.

We recall that a discrete group of isometries of the hyperbolic plane $\mathbb{H}^2$ can be viewed as a discrete subgroup of $\text{PSL}_2(\mathbb{R})$. If $M \in \text{PSL}_2(\mathbb{R})$ corresponds to a hyperbolic isometry, then $\text{Tr} M$ is related to the translation length $l$ of $M$ by the formula $2 \cosh(l/2) = |\text{Tr} M|$. Takeuchi’s result is as follows.

**Theorem 3.1** (Takeuchi). Let $\Gamma$ be a discrete, cofinite subgroup of $\text{PSL}_2(\mathbb{R})$. Set

$$\text{Tr} \Gamma := \{ \text{Tr} T : T \in \Gamma \}.$$

Then $\Gamma$ is derived from a quaternion algebra if and only if

1. the field $K = \mathbb{Q}(\text{Tr} \Gamma)$ is an algebraic field of finite degree and $\text{Tr} \Gamma$ is a subset of the ring of integers of $K$, and
2. for all embeddings $\varphi : K \to \mathbb{C}$, $\varphi \neq 1d$, the set $\varphi(\text{Tr} \gamma)$ is bounded in $\mathbb{C}$. 


For a proof of this characterization, see [Katok 1992; Takeuchi 1975]. Condition (2) has some strong implications on the structure of the trace set $\text{Tr} \Gamma$, as the next result shows. A similar statement can be found in [Luo and Sarnak 1994].

Lemma 3.2. Let $\Gamma_0$ be an arithmetic group derived from a quaternion algebra.

(1) There exists a constant $C_0 > 0$ depending solely on $\Gamma_0$ such that for all $x, x' \in \text{Tr} \Gamma_0$ with $x \neq x'$, 
\[ |x - x'| \geq C_0. \]

(2) There exists a constant $M_0$ depending only on $\Gamma_0$ such that for all $R$ large,
\[ \#\{x \in \text{Tr} \Gamma_0 : |x| \leq R\} \leq M_0 R. \]

Proof. Clearly (1) implies (2) by a box argument. Let us prove (1). The field $K = \mathbb{Q}(\text{Tr} \Gamma_0)$ is a totally real number field of degree say $n = [K : \mathbb{Q}]$. Let $\phi_1 = \text{Id}, \phi_2, \ldots, \phi_n$ be the $n$ distinct embeddings of $K$ into $\mathbb{C}$. The set $\text{Tr} \Gamma_0$ is a subset of the ring of integers $O_K$ of $K$. We denote by $N^K_{\mathbb{Q}}(\cdot)$ the norm on $K$.

We recall that if $x \in O_K$ then $N^K_{\mathbb{Q}}(x) \in \mathbb{Z}$. Let $x \neq x'$ belong to $\text{Tr} \Gamma_0$, we have
\[ 1 \leq |N^K_{\mathbb{Q}}(x - x')| = \prod_{i=1}^{n} |\phi_i(x - x')| \leq |x - x'| M^{n-1}, \]
where $M > 0$ is given by property (2) of Takeuchi’s characterization.

This important feature of the trace set was noticed by physicists working on quantum chaos [Bogomolny et al. 1997] and was clearly emphasized by Luo and Sarnak [1994] in their work on the number variance of arithmetic surfaces. Selberg and Hejhal [1976], when trying to obtain sharp lower bounds for the error term in Weyl’s law, had already noticed similar properties for some examples of cocompact arithmetic groups.

In the rest of this section we will work with a geometrically finite group $\Gamma$ as defined in Section 1, and we assume in addition that $\Gamma$ is an infinite-index subgroup of an arithmetic group $\Gamma_0$, derived from a quaternion algebra. The simplest examples of such groups $\Gamma$ that one can think of are finitely generated Schottky subgroups of $\text{PSL}_2(\mathbb{Z})$, but there are definitely many other examples, see the next section.

Given such a group $\Gamma$, one can define the length spectrum of $X = \Gamma \backslash \mathbb{H}^2$ by
\[ \mathcal{L}_\Gamma := \{kl(\gamma) : (k, \gamma) \in \mathbb{N}^* \times \mathcal{P}\}, \]
where as in the preceding section, $\mathcal{P}$ is the set of primitive closed geodesics. We have the following key properties.

Proposition 3.3. Let $\Gamma$ be a fuchsian group as above, then we have:

(1) Let $l_1, l_2 \in \mathcal{L}_\Gamma$ with $2 \cosh(l_i/2) = \text{Tr} M_i$, $i \in \{1, 2\}$, then
\[ |l_1 - l_2| \geq e^{-(\max(l_1, l_2))/2} |\text{Tr} M_1 - \text{Tr} M_2|. \]

(2) There exists a constant $C_1 > 0$ depending only on $\Gamma_0$ such that for all $\alpha$ large,
\[ \#\{l \in \mathcal{L}_\Gamma : \alpha - 1 \leq l \leq \alpha + 1\} \leq C_1 e^{\alpha/2}. \]
Proof. The set of closed geodesics on \( X = \Gamma \setminus \mathbb{H}^2 \) is in one-to-one correspondence with the set of conjugacy classes of hyperbolic elements in the fundamental group \( \Gamma \), each closed geodesic \( \gamma \) having its length \( l(\gamma) \) given by the formula
\[
2 \cosh(l(\gamma)/2) = |\text{Tr } T_{\gamma}|,
\]
where \( T_{\gamma} \in \Gamma \) is an hyperbolic isometry. The length spectrum \( \mathcal{L}_\Gamma \) is therefore in one-to-one correspondence with the trace set \( \text{Tr } \Gamma \) via the above formula (except for the conjugacy classes of parabolic elements with trace 2). Since we have \( \text{Tr } \Gamma \subset \text{Tr } \Gamma_0 \), we can use the preceding Lemma and crude bounds to prove estimate (2). To obtain the first lower bound (1), one simply writes (assuming \( l_2 > l_1 \)),
\[
l_2 - l_1 = 2 \int_{x_1}^{x_2} \frac{dt}{t} \geq 2 \frac{x_2 - x_1}{x_2},
\]
where we have
\[
x_i = e^{l_i/2} = \frac{1}{2} (\text{Tr } M_i + \sqrt{\text{Tr } M_i^2 - 4}).
\]
Clearly one gets
\[
x_2 - x_1 = \frac{1}{2} \int_{\text{Tr } M_1}^{\text{Tr } M_2} \left(1 + \frac{u}{\sqrt{u^2 - 4}}\right) du \geq \frac{1}{2} (\text{Tr } M_2 - \text{Tr } M_1),
\]
and the proof is done. \( \square \)

When compared with the prime geodesic theorem (see Section 2), estimate (2) in the proposition shows that whenever \( \delta > \frac{1}{2} \) there must be some exponentially large multiplicities in the length spectrum. This is the key observation of Selberg and Hejhal [Hejhal 1976, Section 18, Chapter 2] that will allow us to produce a better lower bound than in Section 2. More precisely:

**Proposition 3.4.** Let \( \Gamma \) be a group as above, \( \delta \) being the dimension of its limit set. Let \( \mathcal{S}_{a,t} \) be the sum defined by
\[
\mathcal{S}_{a,t} := \sum_{a-1 \leq kl(\gamma) \leq a+1} \frac{l(\gamma)}{2 \sinh(kl(\gamma)/2)} e^{-itkl(\gamma)} \varphi_0(kl(\gamma) - a).
\]
There exists a constant \( A > 0 \) such that for all \( T \) large, if one sets \( \alpha = 2 \log T - A \) then the integral \( \mathcal{S}(T) \) defined by
\[
\mathcal{S}(T) = \int_T^{3T} \left(1 - \frac{|t-2T|}{T}\right) |\mathcal{S}_{a,t}|^2 dt,
\]
enjoys the lower bound
\[
\mathcal{S}(T) \geq C_2 \frac{T^{1+4\delta-3}}{(\log T)^2},
\]
for some constant \( C_2 > 0 \) independent of \( T \).

Let us show how Theorem 1.3 follows from this lower bound. First we assume that for all \( z \in W_A \) with \( |\text{Re } z| \geq R_0 \), we have
\[
\mathcal{B}(z) \leq |\text{Re } z|^\beta,
\]
for some \( 0 < \beta < 1 \). Set \( \alpha = 2 \log T - A \), where \( A \) is given by the above proposition. We have
\[
C_2 \frac{T^{1+4\delta-3}}{(\log T)^2} \leq \mathcal{S}(T) \leq \int_T^{3T} |\mathcal{S}_{a,t}|^2 dt.
\]
By the trace formula (3) applied to \( \varphi_{a,t} \), and Lemma 2.2 with \( k = 2 \), Lemma 2.4, we have

\[
|s_{a,t}| \leq O(1) + O\left(\frac{T^{2\delta-1}}{T^2}\right) + O(t^\beta);
\]

therefore \( \int_T^{3T} |s_{a,t}|^2 dt = O(T^{2\beta+1}) \), which produces a contradiction whenever \( \beta < 2\delta - 3/2 \). This proves Theorem 1.3.

**Proof of Proposition 3.4.** We start with an elementary observation. For all \( \lambda \in \mathbb{R} \) and \( T > 0 \) set

\[
J(T, \lambda) = \int_T^{3T} \left(1 - \frac{|t - 2T|}{T}\right)e^{-i\lambda t} dt.
\]

**Lemma 3.5.** With the preceding notation, we have \( |J(T, \lambda)| \leq \frac{4}{\lambda^2 T} \) for all \( \lambda \neq 0 \), while \( J(T, 0) = T \).

**Proof:** This follows by direct computation.

At this point we need some more notation. If \( \ell \in \mathcal{L}_\Gamma \), we denote by \( \mu(\ell) \) the multiplicity of \( \ell \) as the length of a closed geodesic. If \( \ell \in \mathcal{L}_\Gamma \), then let \( \tilde{\ell} \) denote the primitive length of \( \ell \), that is, if \( \ell = k\ell(\gamma) \) with \( \gamma \) a primitive closed geodesic, then \( \tilde{\ell} = \ell(\gamma) \). Using this notation, we have

\[
\mathcal{J}(T) = \sum_{\ell, \ell' \in \mathcal{L}_\Gamma} \frac{\tilde{\ell}\ell' \mu(\ell) \mu(\ell')}{4 \sinh(\ell/2) \sinh(\ell'/2)} J(T, \ell - \ell')\varphi_0(\ell - \alpha)\varphi_0(\ell' - \alpha).
\]

We now set \( \mathcal{J}(T) = \mathcal{J}_1(T) + \mathcal{J}_2(T) \), where

\[
\mathcal{J}_1(T) = T \sum_{\ell \in \mathcal{L}_\Gamma} \frac{(\ell \mu(\ell))^2}{4 \sinh^2(\ell/2)} \varphi_0^2(\ell - \alpha)
\]

and

\[
\mathcal{J}_2(T) = \sum_{\ell, \ell' \in \mathcal{L}_\Gamma} \frac{\tilde{\ell}\ell' \mu(\ell) \mu(\ell')}{4 \sinh(\ell/2) \sinh(\ell'/2)} J(T, \ell - \ell')\varphi_0(\ell - \alpha)\varphi_0(\ell' - \alpha).
\]

By Lemma 3.5, we have

\[
|\mathcal{J}_2(T)| \leq \frac{4}{T} \sum_{\ell, \ell' \in \mathcal{L}_\Gamma} \frac{\tilde{\ell}\ell' \mu(\ell) \mu(\ell') \varphi_0(\ell - \alpha) \varphi_0(\ell' - \alpha)}{4 \sinh(\ell/2) \sinh(\ell'/2)(\ell - \ell')^2}.
\]

Using the inequality \( ab \leq \frac{1}{2}(a^2 + b^2) \) for all \( a, b \in \mathbb{R} \), we get by the symmetry of the summation

\[
|\mathcal{J}_2(T)| \leq \frac{4}{T} \sum_{\ell, \ell' \in \mathcal{L}_\Gamma} \frac{(\ell \mu(\ell))^2 \varphi_0^2(\ell - \alpha)}{4 \sinh(\ell/2) \sinh(\ell'/2)(\ell - \ell')^2}.
\]

Therefore, one can find a constant \( C > 0 \) such that, for all \( \alpha \) and \( T \) large,

\[
|\mathcal{J}_2(T)| \leq \frac{C e^{-\alpha}}{T} \sum_{\ell \in \mathcal{L}_\Gamma} (\ell \mu(\ell))^2 \varphi_0^2(\ell - \alpha) \sum_{\ell' \in \mathcal{L}_\Gamma \cap [\alpha - 1, \alpha + 1]} \frac{1}{(\ell - \ell')^2}.
\]
By Proposition 3.3(1), we can write \( x = 2 \cosh(\ell/2) \), where \( x \in \text{Tr } \Gamma \), and thus

\[
\sum_{\ell' \in \mathcal{F} \cap [\alpha - 1, \alpha + 1]} \frac{1}{(\ell - \ell')^2} \leq e^{\alpha + 1} \sum_{x' \in \text{Tr } \Gamma} \frac{1}{(x - x')^2}.
\]

We can now bound

\[
\sum_{x' \in \text{Tr } \Gamma} \frac{1}{(x - x')^2} \leq \int_{x_0}^{x} \frac{d\Pi_0(u)}{u^2} + \int_{x_0}^{+\infty} \frac{d\Pi_0(u)}{u^2},
\]

where \( \Pi_0 \) is the counting function for the trace set of the arithmetic group \( \Gamma_0 \) and the constant \( C_0 \) is given by Lemma 3.2. Using the fact that the growth \( \Pi_0(u) = O(u) \), two Stieltjes integration by parts show that there exists a constant \( \tilde{C}_0 \) depending only on \( \Gamma_0 \) such that for all \( x \in \text{Tr } \Gamma \),

\[
\sum_{x' \in \text{Tr } \Gamma} \frac{1}{(x - x')^2} \leq \tilde{C}_0.
\]

Going back to \( \mathcal{J}_2(T) \), we have obtained for \( T \) and \( \alpha \) large,

\[
|\mathcal{J}_2(T)| \leq \frac{C'}{T} \sum_{\ell \in \mathcal{F}} (\tilde{\ell} \mu(\ell))^2 \phi_0^2(\ell - \alpha).
\]

Recall that

\[
\mathcal{J}_1(T) = T \sum_{\ell \in \mathcal{F}} \frac{(\tilde{\ell} \mu(\ell))^2}{4 \sinh^2(\ell/2)} \phi_0^2(\ell - \alpha) \geq C'' e^{-a} T \sum_{\ell \in \mathcal{F}} (\tilde{\ell} \mu(\ell))^2 \phi_0^2(\ell - \alpha),
\]

again for \( \alpha \) large and some \( C'' > 0 \). Therefore \( |\mathcal{J}_2| \leq \frac{1}{2} \mathcal{J}_1 \) as long as

\[
e^{-\alpha} \leq \frac{1}{2} \frac{C''}{C'} T^2,
\]

which is definitely achieved if one sets \( \alpha = 2 \log T - A \), where \( A \gg 1 \). We have thus

\[
|\mathcal{J}(T)| \geq \frac{1}{2} |\mathcal{J}_1(T)| \geq C'' e^{-a} T \sum_{\ell \in \mathcal{F} \cap [\alpha - 1, \alpha + 1]} (\tilde{\ell} \mu(\ell))^2 \phi_0^2(\ell - \alpha) \geq \tilde{C}'' e^{-a} T \sum_{\ell \in \mathcal{F} \cap [\alpha - \frac{1}{2}, \alpha + \frac{1}{2}]} (\mu(\ell))^2,
\]

for some \( \tilde{C}'' > 0 \). By Schwarz inequality we get

\[
\left( \sum_{a - \frac{1}{2} \leq k \gamma \leq a + \frac{1}{2}} 1 \right)^2 = \left( \sum_{\ell \in \mathcal{F} \cap [a - \frac{1}{2}, a + \frac{1}{2}]} \mu(\ell) \right)^2 \leq \left( \sum_{\ell \in \mathcal{F} \cap [a - \frac{1}{2}, a + \frac{1}{2}]} (\mu(\ell))^2 \right) \left( \sum_{\ell \in \mathcal{F} \cap [a - \frac{1}{2}, a + \frac{1}{2}]} 1 \right).
\]

By Proposition 3.3(2),

\[
\sum_{\ell \in \mathcal{F} \cap [a - \frac{1}{2}, a + \frac{1}{2}]} 1 = O(e^{\alpha/2}),
\]

while the prime geodesic theorem yields

\[
\sum_{a - \frac{1}{2} \leq k \gamma \leq a + \frac{1}{2}} 1 \geq B e^{\delta \alpha} \frac{\alpha}{\alpha},
\]
where $B > 0$. Hence we have obtained

$$
\sum_{\ell \in \mathbb{H} \cap [\alpha - \frac{1}{2}, \alpha + \frac{1}{2}]} (\mu(\ell))^2 \geq B^2 \frac{e^{(2\delta - 1/2)\alpha}}{\alpha^2}.
$$

Going back to $\mathcal{J}(T)$ and recalling that $\alpha = 2 \log T - A$ we get

$$
|\mathcal{J}(T)| \geq B\left(\frac{T^{1+4\delta - 3}}{(\log T)^2}\right).
$$

The proof is now complete. \(\square\)

It is now time to indicate how to get upper bounds on the sequence $|\text{Re } z_k|$ as $k \to \infty$. First, we can notice that the above Proposition 3.4 still holds on shorter intervals. Indeed, pick any $0 < \rho < 1$ and set

$$
\mathcal{J}_\rho(T) = \int_{2T - T^\rho}^{2T + T^\rho} \left(1 - \frac{|t - 2T|}{T^\rho}\right)|\mathcal{J}_{a,t}|^2 dt,
$$

then one can show that taking $a = 2\rho \log T - A$, for some $A \gg 1$, there exists a constant $C_\rho > 0$ such that for $T$ large one has

$$
\mathcal{J}_\rho(T) \geq C_\rho \frac{T^{(4\delta - 3)\rho + \rho}}{(\log T)^2}.
$$

The assumption of Lemma 2.2 can be weakened: indeed to get the desired upper bound on $|\mathcal{J}_{a,t}| = O(t^\beta)$, it is enough to assume that

$$
\mathcal{D}(z) = O(|\text{Re } z|^\beta)
$$

for all $z \in W_A$ and $\text{Re } z \in [2T - t^\mu, 2T + t^\mu]$, for some $0 < \mu < 1$. These two minor modifications allow to obtain a more precise statement (by following the same line of proof). For all $\epsilon > 0$, one can find an exponent $1 > \rho_\epsilon > 0$ such that for all $T$ large, there exists $z \in W_A$ with the property

$$
\text{Re } z \in [2T - T^{\rho_\epsilon}, 2T + T^{\rho_\epsilon}] \text{ and } \mathcal{D}(z) \geq \text{Re } z^{2\delta - 3/2 - \epsilon}.
$$

Choose $1 > \mu_\epsilon > \rho_\epsilon$ and define by induction a sequence $(T_k)$ by $T_0 \gg 1$ and for all $k \geq 0$, $T_{k+1} = T_k + (T_k)^{\mu_\epsilon}$. For all $k \geq 0$, set

$$
I_k = [2T_k - (T_k)^{\rho_\epsilon}, 2T_k + (T_k)^{\rho_\epsilon}].
$$

For all $k \geq 0$, one can find $z_k \in W_A$ with

$$
\text{Re } z_k \in I_k \text{ and } \mathcal{D}(z_k) \geq (\text{Re } z_k)^{2\delta - 3/2 - \epsilon}.
$$

Moreover because $\mu_\epsilon > \rho_\epsilon$, we have $D(z_k) \cap D(z_{k+1}) = \emptyset$ for $k$ large. To obtain the leading behavior of $T_k$ as $k \to +\infty$, one can perform a change of variable $x_k = 1/T_k$ and consider the dynamical system on the real line given by

$$
f_{\mu_\epsilon}(x) = \frac{x}{1 + x^{1-\mu_\epsilon}}.
$$

Clearly 0 is a neutral fixed point for $f_{\mu_\epsilon}$ and for all $x_0 > 0$,

$$
x_k = f_{\mu_\epsilon}(x_0) > 0.$$
tends to 0 as $k \to +\infty$. Remark that since we have for $x \leq 1$,
\[ f_{\mu_\epsilon}(x) \leq \frac{x}{1+x}, \]
we get the crude upper bound
\[ x_k = O\left(\frac{1}{k}\right). \]
To obtain an asymptotic estimate, we set $u_k = (x_k)^\alpha$, where $\alpha$ will be determined later on. Writing
\[ u_N - u_0 = \sum_{k=0}^{N-1} f_{\mu}(x_k)^\alpha - x_k^\alpha, \]
and since we have the local expansion at $x = 0$
\[ f_{\mu}(x)^\alpha - x^\alpha = -\alpha x^{1-\mu+\alpha} + O(x^{2-2\mu+\alpha}), \]
the choice of $\alpha = \mu - 1$ yields as $N \to +\infty$,
\[ u_N = (1-\mu)N + O\left(\sum_{k=1}^{N} \frac{1}{k^{1-\mu}}\right) = (1-\mu)N + O(N^\mu). \]
Therefore
\[ \lim_{k \to \infty} (1-\mu_\epsilon)^{1/(1-\mu_\epsilon)k^{1/(1-\mu_\epsilon)}} x_k = 1. \]
Thus we have the polynomial bound $|\text{Re} \ z_k| = O(k^{1/(1-\epsilon \mu)})$. Clearly the exponent $p_\epsilon = \frac{1}{1-\mu_\epsilon}$ will tend to infinity as $\epsilon$ goes to 0.

4. Examples

In this section we discuss briefly examples of surfaces $X = \Gamma \backslash \mathbb{H}^2$ satisfying the assumptions of Theorems 1.2 and 1.3. We assume that the reader has some basic knowledge in fuchsian groups and hyperbolic geometry, for which we refer to [Katok 1992]. By the work of Patterson [1976], we know that if $X$ has at least one cusp, that is, if $\Gamma$ has at least one nontrivial parabolic element, then the dimension $\delta > \frac{1}{2}$. If one wants examples without cusps, then $\delta$ can be made arbitrarily close to 1 by “pinching” the geodesic boundary of Nielsen’s region. Let us explain what we mean. By [Patterson 1976] and the spectral analysis in [Lax and Phillips 1984a; 1984b; 1985], we have $\delta > \frac{1}{2}$ if and only if $\lambda_0(X) < \frac{1}{4}$, where $\lambda_0(X)$ is the bottom of the spectrum of the Laplacian $\Delta_X$. In that case $\lambda_0(X) = \delta(1-\delta)$. Hence to get $\delta > \frac{1}{2}$, it is enough to show that the Rayleigh quotient
\[ \lambda_0(X) = \inf_{f \neq 0} \frac{\int_X |\nabla f|^2 d\text{Vol}}{\int_X f^2 d\text{Vol}} < \frac{1}{4}, \]
where $f$ is an $L^2$ function on $X$ with an $L^2$ gradient $\nabla f$. Based on the above formula, Pignataro and Sullivan proved the following, where $\ell(X)$ denotes the maximum length of the closed geodesics which are the boundary of the Nielsen region of $X$ (the convex core):
Proposition 4.1 [Pignataro and Sullivan 1986]. There exists a constant $C(X) > 0$ depending only the topology of $X$ such that

$$\lambda_0(X) \leq C(X)\ell(X).$$

Therefore if $\ell(X)$ is small enough, one definitely has $\delta > \frac{1}{2}$. Applying the same strategy to find examples satisfying the hypothesis of Theorem 1.3 is harder. Indeed, the discreteness of arithmetic groups makes it difficult to perform deformations. What we are looking for are geometrically finite, infinite-index subgroups $\Gamma$ of arithmetic groups derived from quaternion algebras with $\delta(\Gamma) > \frac{3}{4}$. The easiest thing to do is to consider first $\text{PSL}_2(\mathbb{Z})$ and look at some of its subgroups.

Let us first consider the group $\Lambda_N$ obtained as

$$\Lambda_N := \langle g_0, g_1, \ldots, g_N \rangle,$$

where

$$g_0(z) = \frac{-1}{z} \simeq \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad g_k = \tau^k g_0 \tau^{-k}, \quad \tau(z) = z + 2 \simeq \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$ 

Let $D_j$, $j = 0, \ldots, N$ be the unit closed disc centered at $2j$. A fundamental domain for the action of $\Lambda_N$ on $\mathbb{H}^2$ is given by

$$\mathfrak{F} = \mathbb{H}^2 \setminus (D_0 \cup \cdots \cup D_N).$$

The group $\Lambda_N$ is therefore geometrically finite and has no parabolic elements, despite the presence of (false) “cusps” in the fundamental domain. The elliptic elements are the conjugacy classes of $g_0, \ldots, g_N$, which are of order 2. Up to a covering of order 2, we can get rid of them.

For $k = 1, \ldots, N$, set $h_k = g_0 g_k$, and consider the subgroup

$$\Gamma_N = \langle h_1, \ldots, h_N; h_1^{-1}, \ldots, h_N^{-1} \rangle,$$

then it is easy to see that $\Gamma_N$ is a subgroup of $\Lambda_N$ of index 2 and has no elliptic elements, hence a convex cocompact group. Because $\Gamma_N$ is of finite index the critical exponents $\delta(\Gamma_N)$ and $\delta(\Lambda_N)$ are the same: the critical exponent is defined as the infimum of positive real numbers $\sigma$ such that the Poincaré series

$$P(\sigma) := \sum_{\gamma \in \Gamma} e^{-\sigma d(i, \gamma i)}$$

are convergent. Here $d$ is the hyperbolic distance in the half-plane model. A classical result of Sullivan [1984] shows that for geometrically finite groups, the critical exponent is also the Hausdorff dimension of the limit set, hence $\Lambda_N$ and $\Gamma_N$ have same dimension for their limit set. The group $\Lambda_N$ is also considered in [Gamburd 2002], where he shows using a min-max argument and a suitable test function that $\delta(\Lambda_N)$ can be made as close to 1 as we want, provided $N$ is large enough (estimates are effective).

An alternative way to construct similar convex cocompact subgroups of $\text{PSL}_2(\mathbb{Z})$ with $\delta$ close to 1 is given in [Bourgain and Kontorovich 2010]. The idea is to start with the free subgroup $\Gamma(2) = \langle A, B, A^{-1}, B^{-1} \rangle$ generated by

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}.$$
Its commutator subgroup is a free, infinitely generated subgroup with critical exponent 1. Moreover it has no parabolic elements. This commutator subgroup contains finitely generated (hence convex cocompact) subgroups with critical exponent $\delta$ arbitrarily close to 1.

As a conclusion, we have found several examples of convex cocompact subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ with $\delta > \frac{3}{4}$. By a similar technique, one can produce several examples with cusps. In that direction, let us point out that the Hecke group $\Gamma_3$ generated by $g : z \mapsto -1/z$ and $h : z \mapsto z + 3$ is a good candidate: its Hausdorff dimension was estimated by Phillips and Sarnak [1985] to be $\delta = 0.753 \pm 0.003$. Can one prove (or disprove) rigorously that $\delta > 0.75$?

It would be interesting in itself to find similar constructions for arithmetic groups that were not considered in this paper. In a sequel, the authors plan to address the case of arithmetic groups derived from quaternion division algebras (which are cocompact surface groups). It would also be interesting to consider groups acting on higher-dimensional hyperbolic spaces, for example arithmetic Kleinian groups.

Acknowledgments

Both authors thank the Banff International Research Station, where part of this work was done, for its hospitality. This paper benefited from interesting discussion with Nalini Anantharaman. We also thank Iosif Polterovich and Julie Rowlett for the stimulating discussions that led to this paper.

References


Received 24 Sep 2009. Accepted 10 Feb 2010.

DMITRY JAKOBSON: jakobson@math.mcgill.ca
Department of Mathematics and Statistics, McGill University, Montreal, QC H3A 2K6, Canada

FRÉDÉRIC NAUD: frederic.naud@univ-avignon.fr
Laboratoire d’Analyse Non-linéaire et Géométrie (EA 2151), Université d’Avignon et des pays de Vaucluse, 84018 Avignon, France
Polynomials with no zeros on the bidisk
Greg Knese

Local wellposedness for the 2+1-dimensional monopole equation
Magdalena Czubak

Regularity of almost periodic modulo scaling solutions for mass-critical NLS and applications
Dong Li and Xiaoyi Zhang

Estimées des noyaux de Green et de la chaleur sur les espaces symétriques
Gilles Carron

Lower bounds for resonances of infinite-area Riemann surfaces
Dmitry Jakobson and Frédéric Naud