# The large sieve in Riemann surfaces 

by<br>Fernando Chamizo (Madrid)

1. Introduction. In analytic number theory it is sometimes needed to get non-trivial estimates for sums of exponential or character sums whose coefficients are only controlled in average. Under the general denomination of "large sieve" there are several inequalities for this purpose; perhaps one of the most representative examples is the following result (see Th. 4 of [Bo]):

Given $x_{1}, \ldots, x_{R} \in \mathbb{R} / \mathbb{Z}$, if $\left\|x_{\nu}-x_{\mu}\right\|>\delta>0$ for $\nu \neq \mu$, then

$$
\begin{equation*}
\sum_{\nu=1}^{R}\left|\sum_{n \leq N} a_{n} e^{2 \pi i n x_{\nu}}\right|^{2} \leq\left(N+\delta^{-1}\right) \sum_{n \leq N}\left|a_{n}\right|^{2} \tag{1.1}
\end{equation*}
$$

(Here $\|\cdot\|$ is the distance function in $\mathbb{R} / \mathbb{Z}$, i.e. $\|x\|=\min _{n \in \mathbb{Z}}|x-n|$.)
In the last years, the harmonic analysis in the upper half-plane $\mathbb{H}=\left\{x+i y: x \in \mathbb{R}, y \in \mathbb{R}^{+}\right\}$has entered broadly in analytic number theory giving different perspectives of old problems and opening new questions (see the surveys [Iw1] and [Iw2]). One of the main interests is the study of harmonic analysis in the Riemann surfaces $\Gamma \backslash \mathbb{H}$ where $\Gamma$ is a Fuchsian group of the first kind, specially a congruence group. Keeping this idea in mind one can ask first, what is the analogue of (1.1) with $\Gamma \backslash \mathbb{H}$ instead of $\mathbb{R} / \mathbb{Z}$ and secondly, what are its consequences.

In this paper we focus on the first objective leaving some of the applications for a second part of this work. Moreover, we show that under a compactness condition the result extends to Riemannian manifolds.

Acknowledgements. This work is a part of my Doctoral Thesis. I owe my gratitude to my supervisor Professor A. Córdoba for encouragement and his confidence in me along these years. I am also deeply indebted to Professor H. Iwaniec for very helpful numerous conversations, his suggestions and the privilege of attending his lectures. The final version was completed during a visit to Rutgers University. I am grateful for the support and hospitality. Finally, I want to thank specially for the encouraging help given by E. Valenti.
2. Review on Selberg's theory and main results. Before stating our main results we give a brief overview of harmonic analysis in Riemann surfaces in order to fix notation and quote some results. The main contributor to the theory was Selberg (see [Se]), but he did not publish the proofs of most of his results. A self-contained introduction with complete proofs can be found in [Iw3].

We consider the upper half-plane $\mathbb{H}$ endowed with the Poincaré metric, which induces the element of volume $d \mu$ and the distance function $\varrho$ given by

$$
\begin{gathered}
d \mu(x+i y)=\frac{d x d y}{y^{2}} \\
\varrho(z, w)=\operatorname{arccosh}(1+2 u(z, w)) \quad \text { where } u(z, w)=\frac{|z-w|^{2}}{4 \operatorname{Im} z \operatorname{Im} w}
\end{gathered}
$$

The direct isometries of $\mathbb{H}$ are the real Möbius transformations, which can be identified with $\mathrm{PSL}_{2}(\mathbb{R})$ in the usual way. They leave invariant the Laplace-Beltrami operator in $\mathbb{H}$,

$$
\Delta=y^{2}\left(\frac{\partial^{2}}{\partial x^{2}}+\frac{\partial^{2}}{\partial y^{2}}\right)=-(z-\bar{z})^{2} \frac{\partial}{\partial z} \frac{\partial}{\partial \bar{z}}
$$

If $\Gamma$ is a Fuchsian group of the first kind then $\Gamma \backslash \mathbb{H}$ has finite volume and a structure of Riemann surface. The distance function in $\mathbb{H}$ induces a distance in $\Gamma \backslash \mathbb{H}$ given by

$$
d(z, w)=\inf _{\gamma \in \Gamma} \varrho(\gamma z, w)
$$

here and subsequently we identify each point in $\Gamma \backslash \mathbb{H}$ with one of its representatives in $\mathbb{H}$.

A point in $\mathbb{R} \cup\{\infty\}$ which is the only fixed point of some element of $\Gamma$ is called a cusp. The cusps of $\Gamma \backslash \mathbb{H}$ are in some sense the points at infinity of the Riemann surface $\Gamma \backslash \mathbb{H}$; with this idea in mind, it is convenient to define for each cusp $\mathfrak{a}$ an element $\sigma_{\mathfrak{a}}$ of $\operatorname{PSL}_{2}(\mathbb{R})$ such that

$$
\sigma_{\mathfrak{a}} \infty=\mathfrak{a} \quad \text { and } \quad \sigma_{\mathfrak{a}}^{-1} \Gamma_{\mathfrak{a}} \sigma_{\mathfrak{a}}=\left\{\left(\begin{array}{cc}
1 & n \\
0 & 1
\end{array}\right) /\{ \pm I\}: n \in \mathbb{Z}\right\}
$$

with $\Gamma_{\mathfrak{a}}=\{\gamma \in \Gamma: \gamma \mathfrak{a}=\mathfrak{a}\}$.
We also define the function

$$
y_{\Gamma}(z)=\max _{\mathfrak{a}} \max _{\gamma \in \Gamma} \operatorname{Im} \sigma_{\mathfrak{a}}^{-1} \gamma z
$$

which measures, so to speak, how close a point $z \in \Gamma \backslash \mathbb{H}$ is to the cusps.
The smooth functions in $\Gamma \backslash \mathbb{H}$ having exponential decay to zero at every cusp and being eigenfunctions of $\Delta$ are called Maaß cusp forms. They belong
to $L^{2}(\Gamma \backslash \mathbb{H})$, the space of functions such that $\langle f, f\rangle$ is finite, where

$$
\langle f, g\rangle=\int_{\Gamma \backslash \mathbb{H}} f(z) \overline{g(z)} d \mu(z)
$$

Other important functions are the Eisenstein series, defined for each cusp $\mathfrak{a}$ and $s \in \mathbb{C}$ with $\operatorname{Re} s>1$ by

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

It can be proved by quite advanced methods that $E_{\mathfrak{a}}(z, s)$ has a meromorphic continuation in $\mathbb{C}$ as a function of $s$; moreover, its poles in $\operatorname{Re} s>1 / 2$ are simple, real and the residues are eigenfunctions of $\Delta$ in $L^{2}(\Gamma \backslash \mathbb{H})$.

If $\left\{u_{j}(z)\right\}$ is a complete orthonormal system in the space generated by Maaß cusp forms and residues of Eisenstein series, then the spectral theorem asserts that every function $f$ in $L^{2}(\Gamma \backslash \mathbb{H})$ can be expanded as

$$
f(z)=\sum_{j}\left\langle f, u_{j}\right\rangle u_{j}(z)+\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-\infty}^{\infty}\left\langle f, E_{\mathfrak{a}}(\cdot, 1 / 2+i t)\right\rangle E_{\mathfrak{a}}(z, 1 / 2+i t) d t
$$

A consequence of this is the following expansion for automorphic kernels created by the method of images:

$$
\begin{aligned}
\sum_{\gamma \in \Gamma} k(u(\gamma z, w))= & \sum_{j} h\left(t_{j}\right) u_{j}(z) \overline{u_{j}(w)} \\
& +\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} h(t) E_{\mathfrak{a}}(z, 1 / 2+i t) \overline{E_{\mathfrak{a}}(w, 1 / 2+i t)} d t
\end{aligned}
$$

where $1 / 4+t_{j}^{2}$ is the eigenvalue corresponding to $u_{j}$ and $h$ is the Selberg-Harish-Chandra transform of $k$, an even function defined by

$$
\begin{aligned}
h(t) & =\int_{\mathbb{H}} k(u(z, i))(\operatorname{Im} z)^{1 / 2+i t} d \mu(z) \\
& =2 \int_{0}^{\infty} \int_{0}^{\infty} k\left(\frac{x^{2}+(y-1)^{2}}{4 y}\right) y^{-3 / 2+i t} d x d y
\end{aligned}
$$

It is useful to invert this transform in order to recover the left hand side from the right hand side in the above expansion of automorphic kernels. The corresponding formula for the Selberg-Harish-Chandra inverse transform is

$$
k\left(\frac{\cosh u-1}{2}\right)=\frac{\sqrt{2}}{4 \pi^{2}} \int_{u}^{\infty} \int_{-\infty}^{\infty} \frac{t h(t) \sin (w t)}{\sqrt{\cosh w-\cosh u}} d w d t
$$

Now we state our main results.
By the spectral theorem we can consider the following to be the analogue of (1.1) in $\Gamma \backslash \mathbb{H}$.

Theorem 2.1. Given $T>1$ and $z_{1}, \ldots, z_{R} \in \Gamma \backslash \mathbb{H}$, if $d\left(z_{\nu}, z_{\mu}\right)>\delta>0$ for $\nu \neq \mu$, then
$\sum_{\nu=1}^{R}\left|\sum_{\left|t_{j}\right| \leq T} a_{j} u_{j}\left(z_{\nu}\right)+\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-T}^{T} a_{\mathfrak{a}}(t) E_{\mathfrak{a}}\left(z_{\nu}, 1 / 2+i t\right) d t\right|^{2} \ll\left(T^{2}+\delta^{-2}\right)\|a\|_{*}^{2}$
where

$$
\|a\|_{*}=\left(\sum_{\left|t_{j}\right| \leq T}\left|a_{j}\right|^{2}+\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-T}^{T}\left|a_{\mathfrak{a}}(t)\right|^{2} d t\right)^{1 / 2}
$$

and the " $<$ " constant depends on $\Gamma$ and $\max y_{\Gamma}\left(z_{\nu}\right)$.
Remark. Note that by Parseval's identity in $\Gamma \backslash \mathbb{H}$, the theorem is best possible when $\delta$ is small. The dependence of the constant on $\max y_{\Gamma}\left(z_{\nu}\right)$ is given more precisely in [Ch].

We can also ask about cancellation considering exponential sums over the eigenvalues instead of sums of eigenfunctions (both concepts coincide in $\mathbb{R} / \mathbb{Z})$. In the following result we study this kind of exponential sums but weighted with the eigenfunctions $u_{j}(z)$, which allows a considerable simplification in the proof (avoiding the use of Selberg's trace formula).

Theorem 2.2. Given $z \in \Gamma \backslash \mathbb{H}, T, X>1$ and $x_{1}, \ldots, x_{R} \in[X, 2 X]$, if $\left|x_{\nu}-x_{\mu}\right|>\delta>0$ for $\nu \neq \mu$, then

$$
\begin{aligned}
& \sum_{\nu=1}^{R}\left|\sum_{\left|t_{j}\right| \leq T} a_{j} x_{\nu}^{i t_{j}} u_{j}(z)+\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-T}^{T} a_{\mathfrak{a}}(t) x_{\nu}^{i t} E_{\mathfrak{a}}(z, 1 / 2+i t) d t\right|^{2} \\
&<\left(T^{2}+X T \delta^{-1}\right)\|a\|_{*}^{2}
\end{aligned}
$$

where $\|a\|_{*}$ is as in Theorem 2.1 and the " $\ll$ " constant depends on $\Gamma$ and $y_{\Gamma}(z)$.

If $\Gamma$ is a co-compact group, i.e. if $\Gamma \backslash \mathbb{H}$ is compact, the spectral theorem becomes much simpler. In general, if $M$ is a Riemannian manifold of dimension $n$, some classical results (see $\S 8.12$ in [Gi-Tr]) assure that the eigenfunctions $\phi_{1}, \phi_{2}, \ldots$ of the Laplace-Beltrami operator in $M$,

$$
\Delta=\frac{1}{\sqrt{g}} \sum_{i=1}^{n} \frac{\partial}{\partial x^{i}}\left(\sqrt{g} \sum_{k=1}^{n} g^{i k} \frac{\partial}{\partial x^{k}}\right),
$$

can be chosen to be a complete orthonormal system in the space of square integrable functions and the corresponding eigenvalues, $\mu_{1}, \mu_{2}, \ldots$, are nonpositive. We shall write $\mu_{i}=-\lambda_{i}$.

In $M$ the natural distance function $d$ is given by the minimal length of a path joinning two points; by the Hopf-Rinow theorem this minimal length
is actually attained by some geodesic. Note that this definition of $d$ agrees with that for $\Gamma \backslash \mathbb{H}$.

In this general setting the following large sieve inequality holds, which will be proved using some properties of the heat equation:

Theorem 2.3. Given $\Lambda>1$ and $x_{1}, \ldots, x_{R} \in M$, if $d\left(x_{\nu}, x_{\mu}\right)>\delta>0$ for $\nu \neq \mu$, then

$$
\sum_{\nu=1}^{R}\left|\sum_{\sqrt{\lambda_{j}} \leq \Lambda} a_{j} \phi_{j}\left(x_{\nu}\right)\right|^{2} \ll\left(\Lambda^{n}+\delta^{-n}\right) \sum_{\sqrt{\lambda_{j}} \leq \Lambda}\left|a_{j}\right|^{2}
$$

Remark. Note that this result gives (1.1) up to a constant upon taking $M=\mathbb{R} / \mathbb{Z}$, and Theorem 2.1 for $\Gamma$ co-compact upon taking $M=\Gamma \backslash \mathbb{H}$.
3. Proofs of the main results. In the proofs of Theorems 2.1 and 2.2 we shall use two lemmas; the first one gives estimates for functions having a certain type of Selberg-Harish-Chandra transform and the second one is a variant of a standard elementary inequality used in large sieve theory (see for instance (1.18) in $[\mathrm{Mo}]$ ). The proof of these results will be given at the end of this section.

Lemma 3.1. Given $T>1$ and $r>0$, let $k_{1}$ and $k_{2}$ be the Selberg-HarishChandra inverse transform of

$$
h_{1}(t)=e^{-t^{2} /\left(4 T^{2}\right)} \quad \text { and } \quad h_{2}(t)=e^{-t^{2} /\left(4 T^{2}\right)} \cos (r t)
$$

respectively. Then
(a) $k_{1}(t)$ is decreasing in $t>0$ and for every $u$,

$$
k_{1}((\cosh u-1) / 2) \ll T^{2} e^{-T^{2} u^{2}}
$$

(b) There exists an absolute positive constant $C$ such that
$k_{2}((\cosh u-1) / 2) \ll T^{2} e^{-C T^{2}(u-r)^{2}}$ for $u \geq r$ and $k_{2}(0) \ll \min \left(T^{2}, r^{-2}\right)$.
Lemma 3.2. Let $\vec{b}=\left(b_{1}, \ldots, b_{R}\right)$ be a unit complex vector and let $M=\left(m_{\nu \mu}\right)$ be an $R \times R$ complex matrix such that $\left|m_{\nu \mu}\right|=\left|m_{\mu \nu}\right|$. Then

$$
|\vec{b} \cdot M \vec{b}|=\left|\sum_{\nu, \mu=1}^{R} b_{\nu} \bar{b}_{\mu} m_{\nu \mu}\right| \leq \max _{\nu} \sum_{\mu=1}^{R}\left|m_{\nu \mu}\right|
$$

Proof of Theorem 2.1. Let $S$ be the left hand side of the conclusion. By duality there exists a unit complex vector, say $\vec{b}=\left(b_{1}, \ldots, b_{R}\right)$, such that

$$
S=\left(\sum_{\nu=1}^{R} b_{\nu}\left(\sum_{\left|t_{j}\right| \leq T} a_{j} u_{j}\left(z_{\nu}\right)+\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-T}^{T} a_{\mathfrak{a}}(t) E_{\mathfrak{a}}\left(z_{\nu}, 1 / 2+i t\right) d t\right)\right)^{2}
$$

Interchanging the summation and by Cauchy's inequality in $\mathbb{C} \times \ldots \times \mathbb{C} \times$ $L^{2}(\mathbb{R}) \times \ldots \times L^{2}(\mathbb{R})$ we have

$$
\begin{equation*}
S \leq\|a\|_{*}^{2} \widetilde{S} \tag{3.1}
\end{equation*}
$$

where $\widetilde{S}$ is the "dual sum"

$$
\widetilde{S}=\sum_{\left|t_{j}\right| \leq T}\left|\sum_{\nu=1}^{R} b_{\nu} u_{j}\left(z_{\nu}\right)\right|^{2}+\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-T}^{T}\left|\sum_{\nu=1}^{R} b_{\nu} E_{\mathfrak{a}}\left(z_{\nu}, 1 / 2+i t\right)\right|^{2} d t
$$

By positivity we can smooth the summation and the integration writing

$$
\begin{aligned}
\widetilde{S} \ll & \sum_{j} e^{-t_{j}^{2} /\left(4 T^{2}\right)}\left|\sum_{\nu=1}^{R} b_{\nu} u_{j}\left(z_{\nu}\right)\right|^{2} \\
& +\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} e^{-t^{2} /\left(4 T^{2}\right)}\left|\sum_{\nu=1}^{R} b_{\nu} E_{\mathfrak{a}}\left(z_{\nu}, 1 / 2+i t\right)\right|^{2} d t
\end{aligned}
$$

Opening the squares and interchanging the order of summation, by Lemma 3.2 and (3.1) we obtain

$$
\begin{equation*}
S \ll\|a\|_{*}^{2} \max _{\nu} \sum_{\mu=1}^{R}\left|S_{\nu \mu}\right| \tag{3.2}
\end{equation*}
$$

where

$$
\begin{aligned}
S_{\nu \mu}= & \sum_{j} e^{-t_{j}^{2} /\left(4 T^{2}\right)} u_{j}\left(z_{\nu}\right) \overline{u_{j}\left(z_{\mu}\right)} \\
& +\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} e^{-t^{2} /\left(4 T^{2}\right)} E_{\mathfrak{a}}\left(z_{\nu}, 1 / 2+i t\right) \overline{E_{\mathfrak{a}}\left(z_{\mu}, 1 / 2+i t\right)} d t
\end{aligned}
$$

and then the problem reduces to giving a suitable estimate for $S_{\nu \mu}$.
Considering $S_{\nu \mu}$ as the spectral expansion of an automorphic kernel (see Section 2) and using Lemma 3.1(a) for the computations, we have

$$
\begin{equation*}
S_{\nu \mu} \ll T^{2} \sum_{\gamma \in \Gamma} e^{-T^{2} \varrho\left(\gamma z_{\nu}, z_{\mu}\right)} \tag{3.3}
\end{equation*}
$$

By a trivial estimate for the so-called hyperbolic circle problem (see Prop. I.2.2 of [He] and Lemma 2.11 of [Iw3]) one deduces

$$
\log \left(1+\#\left\{\gamma \in \Gamma: r<\varrho\left(\gamma z_{\nu}, z_{\mu}\right) \leq 2 r\right\}\right) \ll r+1 \quad \text { for } r>0
$$

where the involved constant depends on $y_{\Gamma}\left(z_{\nu}\right), y_{\Gamma}\left(z_{\mu}\right)$ and $\Gamma$. Hence summing by parts in (3.3) we have

$$
\begin{equation*}
S_{\nu \mu} \ll T^{2} e^{-C T^{2} d^{2}\left(z_{\nu}, z_{\mu}\right)} \tag{3.4}
\end{equation*}
$$

for some absolute positive constant $C$.

Now if we consider a circle of radius $r, 0<r \ll 1$, and center $z_{\nu}$, we can estimate the number of $z_{\mu}$ 's inside it by the spacing condition, getting

$$
\begin{equation*}
\#\left\{z_{\mu}: d\left(z_{\nu}, z_{\mu}\right)<r\right\} \ll 1+r^{2} \delta^{-2} \tag{3.5}
\end{equation*}
$$

Note that one can always find a circle of radius $O(1)$ (depending on $\max _{\nu} y_{\Gamma}\left(z_{\nu}\right)$ and $\Gamma$ ) containing all of the $z_{\nu}$ 's, hence (3.5) holds true for every $r>0$.

Finally, by partial summation we conclude from (3.4) and (3.5) that

$$
\sum_{\mu=1}^{R}\left|S_{\nu \mu}\right| \ll T^{2}+\delta^{-2}
$$

and substituting in (3.2) gives the theorem.
Proof of Theorem 2.2. The beginning of the proof is completely analogous to that of Theorem 2.1 replacing $u_{j}\left(z_{\nu}\right)$ and $E_{\mathfrak{a}}\left(z_{\nu}, 1 / 2+i t\right)$ by $x_{\nu}^{i t_{j}} u_{j}(z)$ and $x_{\nu}^{i t} E_{\mathfrak{a}}(z, 1 / 2+i t)$ respectively, so if $S$ denotes the left hand side of the conclusion one proves (3.2), i.e.

$$
\begin{equation*}
S \ll\|a\|_{*}^{2} \max _{\nu} \sum_{\mu=1}^{R}\left|S_{\nu \mu}\right| \tag{3.6}
\end{equation*}
$$

but in this case

$$
\begin{aligned}
S_{\nu \mu}= & \sum_{j} e^{-t_{j}^{2} /\left(4 T^{2}\right)} \cos \left(r_{\nu \mu} t_{j}\right)\left|u_{j}(z)\right|^{2} \\
& +\frac{1}{4 \pi} \sum_{\mathfrak{a}} \int_{-\infty}^{\infty} e^{-t^{2} /\left(4 T^{2}\right)} \cos \left(r_{\nu \mu} t\right)\left|E_{\mathfrak{a}}(z, 1 / 2+i t)\right|^{2}
\end{aligned}
$$

where

$$
r_{\nu \mu}=\left|\log \left(x_{\nu} / x_{\mu}\right)\right|
$$

Upon a subdivision of the set $x_{1}, \ldots, x_{R}$ into a finite number of subsets depending on $y_{\Gamma}(z)$ we can always assume that $\varrho(\gamma z, z)>r_{\nu \mu} / 2$ for every $\nu$ and $\mu$ when $\gamma \neq \mathrm{Id}$. Then by the spectral expansion of automorphic kernels and Lemma 3.1(b) we deduce

$$
S_{\nu \mu} \ll \min \left(T^{2}, r_{\nu \mu}^{-2}\right)+T^{2} \sum_{\gamma \neq \mathrm{Id}} e^{-C T^{2}\left(\varrho(\gamma z, z)-r_{\nu \mu}\right)^{2}}
$$

for some absolute positive constant $C$. The trivial estimate for the hyperbolic circle problem mentioned in the proof of Theorem 2.1 proves that the above series converges quickly and hence

$$
S_{\nu \mu} \ll \min \left(T^{2}, r_{\nu \mu}^{-2}\right)
$$

Finally, by the definition of $r_{\nu \mu}$ and the spacing condition we have

$$
\sum_{\mu=1}^{R}\left|S_{\nu \mu}\right| \ll \sum_{\mu=1}^{R} \min \left(T^{2}, X^{2}\left|x_{\nu}-x_{\mu}\right|^{-2}\right) \ll T^{2}+X T \delta^{-1}
$$

and substituting in (3.6) completes the proof.
Proof of Theorem 2.3. The left hand side, say $S$, is formally similar to that of Theorem 2.1 without the contribution of the Eisenstein series, hence the first algebraic manipulations can be done in the same way giving (compare with (3.2))

$$
\begin{equation*}
S \ll\left(\sum_{\sqrt{\lambda_{j}} \leq \Lambda}\left|a_{j}\right|^{2}\right) \max _{\nu} \sum_{\mu=1}^{R}\left|S_{\nu \mu}\right| \tag{3.7}
\end{equation*}
$$

where

$$
S_{\nu \mu}=\sum_{j} e^{-\lambda_{j} / \Lambda^{2}} \phi_{j}\left(x_{\nu}\right) \overline{\phi_{j}\left(x_{\mu}\right)} .
$$

Now observe that $S_{\nu \mu}=u\left(x_{\nu}, x_{\mu}, \Lambda^{-2}\right)$ where $u$ is the fundamental solution of the heat equation in $M$, i.e. $u$ solves

$$
\left\{\begin{array}{l}
L u(x, y, t)=0,  \tag{3.8}\\
u(x, y, 0)=\delta(x-y),
\end{array} \quad x, y \in M, t>0,\right.
$$

with $L=-\partial / \partial t+\Delta_{x}$ and $\delta=$ Dirac's delta.
The asymptotic expansion of $u$ and of other related functions is a well known problem with interesting applications in geometry and differential topology. Here we only need an upper bound, so the arguments are much simpler. The key idea is that the solution of (3.8) can be constructed from a kind of local approximate solution (a parametrix) by an iterative process. We shall quote the required results from Chapter III of [ $\mathrm{Be}-\mathrm{Ga}-\mathrm{Ma}$ ].

Given $k>n / 2+2$, by E.III. 2 and E.III. 3 of [Be-Ga-Ma] there exist $r_{0}>0$ and functions $u_{1}, \ldots, u_{k} \in C^{\infty}(M \times M)$ and $\eta \in C^{\infty}(\mathbb{R})$ with $\operatorname{supp} \eta \subset[-2,2],\left.\eta\right|_{[-1,1]} \equiv 1$, such that

$$
\begin{aligned}
& H_{k}(x, y, t)=(4 \pi t)^{-n / 2} \\
& \quad \times e^{-d^{2}(x, y) /(4 t)}\left(u_{0}(x, y)+t u_{1}(x, y)+\ldots+t^{k} u_{k}(x, y)\right) \eta\left(r_{0}^{-1} d(x, y)\right)
\end{aligned}
$$

satisfies

$$
\lim _{t \rightarrow 0^{+}} H_{k}(x, y, t)=\delta(x-y), \quad L H_{k} \in C^{l}\left(M \times M \times \mathbb{R}^{+}\right) \quad \text { for } l<k-n / 2
$$

and

$$
L H_{k}(x, y, t)=(4 \pi t)^{-n / 2} t^{k} e^{-d^{2}(x, y) /(4 t)} \Delta_{x} u_{k} \quad \text { if } d(x, y)<r_{0} .
$$

If we define the convolution of two functions $A, B \in C^{\infty}\left(M \times M \times \mathbb{R}^{+}\right)$ by

$$
A * B(x, y, t)=\int_{0}^{t} \int_{M} A(x, z, \tau) B(z, y, t-\tau) d \operatorname{Vol} d \tau
$$

then the solution $u$ of (3.8) is given by the formula (see E.III. 8 of [Be-Ga-Ma])

$$
u=H_{k}-Q_{k} * H_{k} \quad \text { where } \quad Q_{k}=\sum_{j=1}^{\infty}(-1)^{j+1}\left(L H_{k}\right)^{* j}
$$

It is not difficult to prove that $Q_{k}$ is less than a positive power of $t$ (see E.III. 6 of [Be-Ga-Ma]); then recalling the definition of $H_{k}$ one deduces

$$
u(x, y, t) \ll t^{-n / 2} e^{-C d^{2}(x, y) / t}
$$

for some positive constant $C$ and small values of $t$. Hence

$$
S_{\nu \mu} \ll \Lambda^{n} e^{-C \Lambda^{2} d^{2}\left(x_{\nu}, x_{\mu}\right)}
$$

By the spacing condition

$$
\#\left\{x_{\mu}: d\left(x_{\nu}, x_{\mu}\right)<r\right\} \ll 1+\delta^{-n} \min \left(r^{n}, \operatorname{diam}^{n}(M)\right)
$$

where $\operatorname{diam}(M)$ is the diameter of $M$. Then partial summation gives

$$
\sum_{\mu=1}^{R}\left|S_{\nu \mu}\right| \ll \Lambda^{n}+\delta^{-n}
$$

and by (3.7) this proves the theorem.
We finish this section with the proofs of Lemmas 3.1 and 3.2.
Proof of Lemma 3.1. Using Parseval's identity one can simplify the formula for the Selberg-Harish-Chandra inverse transform at the special value $u=0$ (see Appendix of $[\mathrm{Ku}]$ ) getting

$$
k(0)=\frac{1}{4 \pi} \int_{-\infty}^{\infty} t \tanh (\pi t) h(t) d t
$$

Hence $k_{1}(0), k_{2}(0) \ll T^{2}$ and integrating by parts twice gives $k_{2}(0) \ll r^{-2}$. For non-zero values we consider separately (a) and (b).
(a) Define $f(w)=w e^{-w^{2} T^{2}} / \sinh w$. Then

$$
k_{1}\left(\frac{\cosh u-1}{2}\right)=\frac{T^{3} \sqrt{2}}{\pi^{3 / 2}} \int_{u}^{\infty} f(w) \frac{\sinh w}{\sqrt{\cosh w-\cosh u}} d w .
$$

Integrating by parts and differentiating the result with respect to $u$ we deduce that $k_{1}$ is decreasing on the positive axis and then the bound
$k_{1}(0) \ll T^{2}$ proves (a) for $u T \ll 1$. If $u T \gg 1$ we write

$$
k_{1}\left(\frac{\cosh u-1}{2}\right)=\frac{T^{3} \sqrt{2}}{\pi^{3 / 2}} \int_{u}^{\alpha}+\frac{T^{3} \sqrt{2}}{\pi^{3 / 2}} \int_{\alpha}^{\infty}=I_{1}+I_{2}
$$

where $\alpha=u+u^{-1} T^{-2}$. It is plain that

$$
I_{1} \ll T^{3} f(u) \int_{u}^{\alpha} \frac{\sinh w}{\sqrt{\cosh w-\cosh u}} d w \ll T^{2} e^{-T^{2} u^{2}}
$$

and

$$
I_{2} \ll \frac{T^{3}}{\sqrt{\cosh \alpha-\cosh u}} \int_{\alpha}^{\infty} w e^{-w^{2} T^{2}} d w \ll T^{2} e^{-T^{2} u^{2}}
$$

Hence (a) follows.
(b) In this case we have

$$
k_{2}\left(\frac{\cosh u-1}{2}\right)=\frac{T^{3}}{\pi^{3 / 2} \sqrt{2}} \int_{u}^{\infty} \frac{g(w)}{\sqrt{\cosh w-\cosh u}} d w
$$

with

$$
g(w)=(w+r) e^{-(w+r)^{2} T^{2}}+(w-r) e^{-(w-r)^{2} T^{2}}
$$

If $u \geq r$ we can also assume $r \geq T^{-1}$ because otherwise $g(w) \ll w e^{-C w^{2} T^{2}}$ for some $C>0$ and the proof of (a) applies. Under this assumption we divide the range of integration

$$
k_{2}\left(\frac{\cosh u-1}{2}\right)=\frac{T^{3}}{\pi^{3 / 2} \sqrt{2}} \int_{u}^{\alpha}+\frac{T^{3}}{\pi^{3 / 2} \sqrt{2}} \int_{\alpha}^{\infty}=I_{1}+I_{2}
$$

where $\alpha=u+T^{-1}$. Each of these integrals is bounded as in (a), giving

$$
I_{1} \ll T^{-1} e^{-C(u-r)^{2} T^{2}} \int_{u}^{\alpha} \frac{1}{\sqrt{\cosh w-\cosh u}} d w \ll u^{-1 / 2} T^{-3 / 2} e^{-C(u-r)^{2} T^{2}}
$$

and

$$
I_{2} \ll \frac{1}{\sqrt{\cosh \alpha-\cosh u}} \int_{\alpha}^{\infty} g(w) d w \ll u^{-1 / 2} T^{-3 / 2} e^{-C(u-r)^{2} T^{2}}
$$

Under our hypothesis we have $u \geq T^{-1}$ and then (b) follows.
Proof of Lemma 3.2. The proof reduces to the elementary inequalities

$$
\begin{aligned}
\left|\sum_{\nu, \mu=1}^{R} b_{\nu} \bar{b}_{\mu} m_{\nu \mu}\right| & \leq \frac{1}{2} \sum_{\nu=1}^{R}\left|b_{\nu}\right|^{2} \sum_{\mu=1}^{R}\left|m_{\nu \mu}\right|+\frac{1}{2} \sum_{\mu=1}^{R}\left|b_{\mu}\right|^{2} \sum_{\nu=1}^{R}\left|m_{\nu \mu}\right| \\
& \leq\left(\sum_{\nu=1}^{R}\left|b_{\nu}\right|^{2}\right) \max _{\nu} \sum_{\mu=1}^{R}\left|m_{\nu \mu}\right|
\end{aligned}
$$

and the term in parenthesis can be omitted because $\vec{b}$ is a unit vector.

## References

[Be-Ga-Ma] M. Berger, P. Gauduchon et E. Mazet, Le Spectre d'une Variété Riemannienne, Lecture Notes in Math. 194, Springer, 1971.
[Bo] E. Bombieri, Le grand crible dans la théorie analytique des nombres, Soc. Math. France, Astérisque 18 (1974).
[Ch] F. Chamizo, Topics in Analytic Number Theory, Doctoral Thesis, Universidad Autónoma de Madrid, 1994 (in Spanish).
[Gi-Tr] D. Gilbarg and N. S. Trudinger, Elliptic Partial Differential Equations of Second Order, 2nd ed., Springer, 1983.
[He] D. A. Hejhal, The Selberg Trace Formula for $P S L_{2}(\mathbb{R})$, Vol. I, Lecture Notes in Math. 548, Springer, 1976.
[Iw1] H. Iwaniec, Non-holomorphic modular forms and their applications, in: Modular Forms, Ellis Horwood Series of Halsted Press, New York, 1984, 157-196.
[Iw2] - , Spectral theory of automorphic forms and recent developments in analytic number theory, Proc. I.C.M. Berkeley, 1986.
[Iw3] -, Introduction to the Spectral Theory of Automorphic Forms, Bibl. Rev. Mat. Iberoamericana, Madrid, 1995.
[Ku] T. Kubota, Elementary Theory of Eisenstein Series, Wiley, 1973.
[Mo] H. L. Montgomery, Topics in Multiplicative Number Theory, Lecture Notes in Math. 227, Springer, 1971.
[Se] A. Selberg, Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with applications to Dirichlet series, J. Indian Math. Soc. 20 (1956), 47-87.

Departamento de Matemáticas
Facultad de Ciencias
Universidad Autónoma de Madrid
28049 Madrid, Spain
E-mail: fchamizo@ccuam3.sdi.uam.es

