Journal of Number Theory 227 (2021) 235–264

Contents lists available at ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

General Section

Joint universality theorem of Selberg zeta functions for principal congruence subgroups

Hidehiko Mishou

Tokyo Denki University, Department of Mathematics and Technology, School of System Design, 5 Senju-Asahi-cho, Adachi-ku, Tokyo, 120-8551, Japan

A R T I C L E I N F O

Article history:
Received 25 October 2020
Received in revised form 25 March 2021
Accepted 25 March 2021
Available online 20 April 2021
Communicated by S.J. Miller

MSC:
11M36

Keywords:
Selberg zeta functions
Value distribution
Congruence subgroups
Prime geodesic theorem

A B S T R A C T

In this paper, we investigate the joint functional distribution of Selberg zeta functions for principal congruence subgroups. We prove that the joint universality theorem for these zeta functions holds in the strip $0.85 < \sigma < 1$. As a corollary, we obtain the functional independence for the zeta functions.

1. Introduction

As usual, let $s = \sigma + it$ be a complex valuable and $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ be the set of integers, rational numbers, real numbers, and complex numbers respectively.

The investigation of the value distribution of the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p:\text{prime}} \left(1 - \frac{1}{p^s}\right)^{-1} \quad (\sigma > 1)$$

E-mail address: h_mishou@mail.dendai.ac.jp.

https://doi.org/10.1016/j.jnt.2021.03.009
0022-314X/

was started in 1910s. Bohr and Courant [2] showed that for any fixed $1/2 < \sigma_0 \leq 1$ the set

$$\{\zeta(\sigma_0 + it) \in \mathbb{C} \mid t \in \mathbb{R}\}$$

is dense in $\mathbb{C}$. In 1972, Voronin [13] extended this denseness result to the multi-dimensional space. He showed that for any fixed $1/2 < \sigma_0 \leq 1$ and any positive integer r the set

$$\left\{\left(\zeta(\sigma_0 + it), \zeta'(\sigma_0 + it), \ldots, \zeta^{(r-1)}(\sigma_0 + it)\right) \in \mathbb{C}^r \mid t \in \mathbb{R}\right\}$$

is dense in $\mathbb{C}^r$. Further, Voronin [14] extended it to the functional space and obtained the remarkable universality theorem. To state it in a modern form which was established by Bagchi [1], we define a probability measure on $\mathbb{R}$. Let μ be the Lebesgue measure on $\mathbb{R}$. For $T > 0$ define

$$\nu_T(\cdots) = \frac{1}{T}\mu\{\tau \in [0, T] \ : \cdots\},$$

where in place of dots we write some conditions satisfied by a real number τ.

Theorem 1 *(Voronin [14]). Let K be a compact subset of the strip $\frac{1}{2} < \sigma < 1$ with connected complement and $f(s)$ be a non-vanishing and continuous function on K which is analytic in the interior of K. Then for any small positive number ε we have*

$$\liminf_{T \to \infty} \nu_T(\max_{s \in K} |\zeta(s + i\tau) - f(s)| < \varepsilon) > 0.$$

Roughly speaking this theorem asserts that any analytic function can be uniformly approximated by suitable vertical translation of $\zeta(s)$. We call such analytic property of a function *universality.*

In general, a Dirichlet series is called arithmetic zeta function if it has the Euler product expression over prime numbers and some analytic properties. In the proof of Theorem 1, the fact that the set $\{\log p \mid p\text{:prime}\}$ of Dirichlet exponents of $\log \zeta(s)$ is linearly independent over $\mathbb{Q}$ plays an essential role. After Theorem 1, several mathematicians proved universality property for many types of arithmetic zeta functions (see Steuding [12, Section 1.6]).

Let Γ be a discrete subgroup of $PSL_2(\mathbb{R})$ such that $\mathrm{vol}(H/\Gamma) < \infty$. The Selberg zeta function for Γ is defined by

$$Z_\Gamma(s) = \prod_{\{P\}} \prod_{k=0}^{\infty} \left(1 - N(P)^{-s-k}\right), \tag{1.1}$$

where $\{P\}$ runs through all primitive hyperbolic conjugacy classes of Γ and $N(P)$ denotes the norm of P (we give the definitions in §2.). This product converges absolutely in

$\sigma > 1$ and has a meromorphic continuation to the whole s-plane. Selberg [10] showed that the Selberg zeta function bears a strong resemblance to the Riemann zeta function. Especially $Z_\Gamma(s)$ satisfies the Riemann hypothesis for some groups Γ.

In 2013, Drungilas, Garunkštis, and Kačėnas [3] proved that when $\Gamma = PSL_2(\mathbb{Z})$, the Selberg zeta function behaves similarly to the Riemann zeta function regarding non-zero value-distribution. Namely, $Z_\Gamma(s)$ has the universality property.

Theorem 2 *(Drungilas, Garunkštis, and Kačėnas [3]). Let* Γ *be the full modular group* $PSL_2(\mathbb{Z})$*. Let* α *be a positive constant for which the prime geodesic theorem*

$$\pi_\Gamma(x) := \sum_{\substack{\{P\} \\ N(P) \le x}} 1 = \mathrm{Li}(x) + O(x^\alpha)$$

holds for Γ*. Let* K *be a compact subset of the strip* $\frac{\alpha+1}{2} < \sigma < 1$ *with connected complement and* $f(s)$ *be a non-vanishing and continuous function on* K *which is analytic in the interior of* K*. Then for any small positive number* ε*, we have*

$$\liminf_{T \to \infty} \nu_T\left(\max_{s \in K} |Z_\Gamma(s + i\tau) - f(s)| < \varepsilon\right) > 0.$$

At present, the best estimate of the error term of the prime geodesic theorem is $O(x^{\frac{25}{36}+\varepsilon})$ due to Soundararajan and Young [11]. Therefore the universality for $Z_\Gamma(s)$ holds in the strip $\frac{61}{72} < \sigma < 1$.

The proof of Theorem 2 is similar to the proof of Theorem 1 in several points. Let D be the set of positive discriminants. For $d \in D$, let $\varepsilon(d)$ and $h(d)$ be the fundamental unit and the class number of the quadratic field $\mathbb{Q}(\sqrt{d})$ respectively (we state the definitions of them in §2.1.). Then $Z_\Gamma(s)$ for $\Gamma = PSL_2(\mathbb{Z})$ has the expression

$$Z_\Gamma(s) = \prod_{d \in D} \prod_{k=0}^{\infty} \left(1 - \varepsilon(d)^{-2s-2k}\right)^{h(d)}.$$

Let D^* be a subset of D consisting of positive fundamental discriminants. The authors proved Theorem 2 by using the fact that the set $\{\log \varepsilon(d) : d \in D^*\}$ is linearly independent over $\mathbb{Q}$.

In [3], the authors predicted that the universal property of the Selberg zeta function also holds for other types of discrete subgroups of $PSL_2(\mathbb{R})$. One of them is the principal congruence subgroup. For a positive integer N, the principal congruence subgroup of level N is defined by

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

A subgroup Δ of $SL_2(\mathbb{Z})$ is called a *congruence subgroup* if $\Gamma(N) \subset \Delta$ for some $N \ge 1$. Let $\bar{\Gamma}(N)$ be the image of $\Gamma(N)$ in $PSL_2(\mathbb{R})$. Namely $\bar{\Gamma}(1) = PSL_2(\mathbb{Z})$.

Now we state our main result.

Theorem 3. *Let $\alpha > 0$. Suppose that the prime geodesic theorem*

$$\pi_\Gamma(x) = \sum_{\substack{\{P\} \\ N(P) \le x}} 1 = \mathrm{Li}(x) + O(x^\alpha) \tag{1.2}$$

holds for all congruence subgroups Γ. Let r be a positive integer. Let $N_0 = 1$ and $N_1, \ldots, N_r$ be positive integers which are relatively prime each other and $N_j \geq 3$. In the following, we put

$$Z_j(s) = Z_{\bar{\Gamma}(N_j)}(s).$$

For each $j = 0 \ldots r$, let K_j be a compact subset of the strip $\frac{\alpha+1}{2} < \sigma < 1$ with connected complement and $f_j(s)$ be a non-vanishing and continuous function on K_j which is analytic in the interior of K_j. Then for any small positive number ε, we have

$$\liminf_{T \to \infty} \nu_T \left(\max_{0 \le j \le r} \max_{s \in K_j} |Z_j(s + i\tau) - f_j(s)| < \varepsilon \right) > 0.$$

This theorem asserts that for a set of Selberg zeta functions associated with principal congruence subgroups with distinct levels, the universal property for each zeta function hold simultaneously. We call this type of property for a collection of functions *joint universality.*

Luo, Rudnick, and Sarnak [7] proved that the remainder term of (1.2) is $O(x^{\frac{7}{10}})$ for all congruence subgroups, which is the best estimate at present. Therefore the joint universality for $Z_j(s)$'s holds in the strip $\frac{17}{20} < \sigma < 1$. Similarly to the prime number theorem, it is expected that the error term of (1.2) is $O(x^{\frac{1}{2}+\varepsilon})$. Therefore we expect that the joint universality for $Z_j(s)$'s will hold in $\frac{3}{4} < \sigma < 1$.

The following corollaries are typical and simple consequences of the joint universality theorem. They easily follow from Theorem 3 in the similar way as the proofs of the corollaries in Mishou [8].

Corollary 1. *Let r be a non-negative integer. Suppose that integers $N_0, \ldots, N_r$ satisfy the assumption in Theorem 3. Let σ_0 be a real number with $\frac{\alpha+1}{2} < \sigma_0 < 1$ and M be a positive integer. Then the set*

$$\Big\{ \Big(Z_0(\sigma_0 + it), \cdots Z_r(\sigma_0 + it), \cdots, Z_0^{(M-1)}(\sigma_0 + it), \ldots,$$
$$Z_r^{(M-1)}(\sigma_0 + it) \Big) \in \mathbb{C}^{(r+1)M} \;\Big|\; t \in \mathbb{R} \Big\}$$

is dense in $\mathbb{C}^{(r+1)M}$.

Proof. See the proof of [8, Corollary 4]. □

Corollary 2. *Let r be a non-negative integer. Suppose that integers $N_0, \ldots, N_r$ satisfy the assumption in Theorem 3. Let M be a positive integer. If continuous functions $f_l : \mathbb{C}^{(r+1)M} \to \mathbb{C}$ $(0 \leq l \leq L)$ satisfy*

$$\sum_{l=0}^{L} s^l f_l(Z_0(s), \cdots, Z_r(s), \ldots, Z_0^{(M-1)}(s), \cdots, Z_r^{(M-1)}(s)) \equiv 0$$

for all $s \in \mathbb{C}$, then $f_l \equiv 0$ $(0 \leq l \leq L)$.

Proof. See the proof of [8, Corollary 5]. □

Corollary 3. *Let r be a positive integer. Suppose that integers $N_0, \ldots, N_r$ satisfy the assumption in Theorem 3. Let $a_0, \ldots, a_r$ be non-zero complex numbers. Put*

$$Z(s) = \sum_{j=0}^{r} a_j Z_j(s).$$

Then $Z(s)$ is strongly universal in the strip $\frac{\alpha+1}{2} < \sigma < 1$. Namely, let K be a compact subset of the strip with connected complement and $h(s)$ be a continuous function on K which is analytic in the interior of K. Remark that the function $h(s)$ is allowed to have zeros on K. For any $\varepsilon > 0$

$$\liminf_{T \to \infty} \nu_T \left(\max_{s \in K} |Z(s + i\tau) - h(s)| < \varepsilon \right) > 0.$$

Proof. See the proof of [8, Corollary 3]. □

Corollary 4. *Let r be a positive integer. Suppose that integers $N_0, \ldots, N_r$ satisfy the assumption in Theorem 3. Let $a_0, \ldots, a_r$ be non-zero complex numbers. For real numbers σ_1, σ_2 with $\frac{\alpha+1}{2} < \sigma_1 < \sigma_2 < 1$ and $T \geq 2$, denote by $N(\sigma_1, \sigma_2, T)$ the number of zeros of the function*

$$Z(s) = \sum_{j=0}^{r} a_j Z_j(s)$$

in the rectangle $\{s = \sigma + it \mid \sigma_1 < \sigma < \sigma_2,\ 0 \leq t \leq T\}$. Then we have for sufficiently large T

$$N(\sigma_1, \sigma_2, T) \gg T.$$

Proof. See the proof of [8, Corollary 6]. □

The construction of the paper is as follows. In Section 2, we explain the connection between primitive hyperbolic elements of $\Gamma(N)$ and primitive indefinite quadratic forms, which are given in Sarnak [9]. In Section 3, we quote some results in [3], which are basic tools to prove universality for general Dirichlet series. In Section 4, we represent the logarithm of the Selberg zeta function by a sum of two Dirichlet series over the set of fundamental discriminants. In the last section, we prove Theorem 3.

2. Connection between indefinite quadratic forms and hyperbolic elements of principal congruence subgroups

In this section, we explain the connection between primitive hyperbolic elements of principal congruence subgroups and primitive indefinite quadratic forms. By applying this connection, we will obtain the infinite product expression of $Z_{\bar{\Gamma}(N)}(s)$ over the set of positive discriminants and the prime geodesic theorem for congruence subgroups. We quote many notions and results from Sarnak [9] without proofs.

2.1. Primitive indefinite quadratic forms

Let a, b, c be integers such that $(a, b, c) = 1$ and the discriminant $d = b^2 - 4ac$ is positive. Then the quadratic form

$$Q(x, y) = ax^2 + bxy + cy^2,$$

is primitive and indefinite. We will use $[a, b, c]$ to denote such a form. Let D be the set of positive discriminants, then

$$D = \{d \in \mathbb{Z}_{>0} \mid d \equiv 0 \text{ or } 1 \pmod 4, \quad d \text{ is not square}\}.$$

Two forms $Q = [a, b, c]$ and $Q' = [a', b', c']$ are called equivalent in the narrow sense, if there exists $\gamma = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in SL_2(\mathbb{Z})$ such that

$$Q'(x, y) = Q(px + qy, rx + sy),$$

in other words,

$$\begin{pmatrix} a' & \frac{b'}{2} \\ \frac{b'}{2} & c' \end{pmatrix} = \gamma^t \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c' \end{pmatrix} \gamma.$$

We denote the equivalent relation by $Q \sim Q'$ or $[a, b, c] \sim [a', b', c']$. The following properties are well known. For more details, see Landau [6].

(1) If $[a, b, c]$ and $[a', b', c']$ are equivalent, then they have the same discriminant.

(2) For any $d \in D$, the number of equivalent classes of primitive indefinite quadratic forms with discriminant d is finite. We call the number of classes the class number. We denote it $h(d)$.

(3) The group of automorphism of $Q = [a, b, c]$ is defined by

$$G_Q = \left\{ \gamma \in SL_2(\mathbb{Z}) \;\middle|\; \gamma^t \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \gamma = \begin{pmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{pmatrix} \right\}.$$

Then G_Q is an infinite cyclic group $\{\pm M_Q^n : n \in \mathbb{Z}\}$ with a generator

$$M_Q = \begin{pmatrix} \dfrac{t(d) - bu(d)}{2} & -cu(d) \\ au(d) & \dfrac{t(d) + bu(d)}{2} \end{pmatrix}, \tag{2.1}$$

where $(t(d), u(d)) \in \mathbb{Z}^2_{>0}$ is the fundamental solution of the Pell equation $t^2 - du^2 = 4$. We call M_Q the fundamental automorphism of the form Q.

2.2. Primitive hyperbolic elements

The group $PSL_2(\mathbb{R}) = SL_2(\mathbb{R})/\{\pm I\}$ acts on the upper half-plane H by a linear fractional transformation

$$\gamma z = \frac{pz + q}{rz + s} \quad \left(z \in H,\ \gamma = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in PSL_2(\mathbb{R}) \right).$$

Let Γ be a discrete subgroup of $PSL_2(\mathbb{R})$. An element $P \in \Gamma$ is called hyperbolic if its fixed points are two different real numbers. If P is hyperbolic, the eigenvalues of P are $\alpha > 1$ and $0 < \alpha^{-1} < 1$. We define the norm of P by

$$N(P) = \alpha^2.$$

If P and P' are Γ-equivalent, that is, there is an element $\gamma \in \Gamma$ such that $P' = \gamma^{-1} P \gamma$, then $N(P) = N(P')$. A power of a hyperbolic element is also hyperbolic. We say that a hyperbolic element $P \in \Gamma$ is primitive if it is not a non-trivial power of another hyperbolic element of Γ.

Let $M_Q \in SL_2(\mathbb{Z})$ be the fundamental automorphism of the form Q and $\bar{M}_Q$ be the image of M_Q in $PSL_2(\mathbb{Z})$. By (2.1) it is easily showed that $\bar{M}_Q$ is a primitive hyperbolic element of $PSL_2(\mathbb{Z})$ and its eigenvalues are

$$\varepsilon(d) = \frac{t(d) + u(d)\sqrt{d}}{2} > 1 \tag{2.2}$$

and $0 < \varepsilon(d)^{-1} < 1$. Thus we have $N(\bar{M}_Q) = \varepsilon(d)^2$. As usual, we call $\varepsilon(d)$ the fundamental unit of discriminant d. Now we state the connection between fundamental infinite quadratic forms and primitive hyperbolic elements of $PSL_2(\mathbb{Z})$.

Proposition 1. *Define the map ϕ by $\phi(Q) = \bar{M}_Q$. Then*

(1) *ϕ is a one-to-one map from the set of all primitive indefinite quadratic forms onto the set of all primitive hyperbolic elements of $PSL_2(\mathbb{Z})$.*
(2) *ϕ commutes with the action of $PSL_2(\mathbb{Z})$, so that*

$$Q \sim Q' \iff \bar{M}_Q \sim \bar{M}_{Q'}.$$

(3) *The norms of the conjugacy classes of primitive hyperbolic elements of $PSL_2(\mathbb{Z})$ are the numbers $\varepsilon(d)^2$ where $d \in D$, with multiplicity $h(d)$.*

Proof. See [9, Proposition 1.4 and Corollary 1.5]. □

By the proposition, the prime geodesic theorem (1.2) for $\Gamma = PSL_2(\mathbb{Z})$ can be rewritten as

$$\pi_\Gamma(x) = \sum_{\substack{d \in D \\ \varepsilon(d)^2 \le x}} h(d) = \mathrm{Li}(x) + O(x^\alpha). \tag{2.3}$$

The Selberg zeta function (1.1) is also rewritten as

$$Z_\Gamma(s) = \prod_{d \in D} \prod_{k=0}^{\infty} \left(1 - \varepsilon(d)^{-2s-2k}\right)^{h(d)} \tag{2.4}$$

when $\Gamma = PSL_2(\mathbb{Z})$.

2.3. Principal congruence subgroups

For a positive integer N, the principal congruence subgroup of level N is defined by

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

Let $\bar{\Gamma}(N)$ be the image of $\Gamma(N)$ in $PSL_2(\mathbb{Z})$. The following basic property of the principal congruence subgroup is well-known.

Lemma 1.

(1) *For any integer $N \geq 1$, the group $\Gamma(N)$ is a normal subgroup of $SL_2(\mathbb{Z})$ and*

$$\nu^*(N) = [SL_2(\mathbb{Z}) : \Gamma(N)] = N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right).$$

(2) *If $N \geq 3$, then $-I \notin \Gamma(N)$. Namely, $\bar{\Gamma}(N)$ is a normal subgroup of $PSL_2(\mathbb{Z})$ and*

$$\nu(N) = [PSL_2(\mathbb{Z}) : \overline{\Gamma}(N)] = \frac{1}{2}\nu^*(N).$$

Proof. See, for instance, [5, Section 11.5]. □

Remark that when $(M, N) = 1$,

$$\Gamma(M) \cap \Gamma(N) = \Gamma(MN) \tag{2.5}$$

and

$$\nu^*(M)\nu^*(N) = \nu^*(MN) \tag{2.6}$$

hold. In contrast, concerning $\bar{\Gamma}(N)$ we have the following lemma.

Lemma 2.

(1) *Let M and N be positive integers with $M, N \geq 3$. If $(M, N) = 1$, then $\bar{\Gamma}(MN)$ is a subgroup of $\bar{\Gamma}(M) \cap \bar{\Gamma}(N)$ satisfying*

$$[\bar{\Gamma}(M) \cap \bar{\Gamma}(N) : \bar{\Gamma}(MN)] = 2.$$

(2) *Let $N_1, \ldots, N_r$ be positive integers which are relatively prime each other and $N_j \geq 3$. Put $N = \prod_{j=1}^{r} N_j$ and*

$$\bar{\Gamma}'(N) = \bigcap_{j=1}^{r} \bar{\Gamma}(N_j).$$

Then $\bar{\Gamma}(N)$ is a subgroup of $\bar{\Gamma}'(N)$ satisfying

$$[\bar{\Gamma}'(N) : \bar{\Gamma}(N)] = 2^{r-1}.$$

Namely, $\bar{\Gamma}'(N)$ is a normal subgroup of $PSL_2(\mathbb{Z})$ satisfying

$$[PSL_2(\mathbb{Z}) : \overline{\Gamma}'(N)] = \prod_{j=1}^{r} \frac{\nu^*(N_j)}{2} = \frac{1}{2^r}\nu^*(N).$$

Proof. First we prove that

$$\Gamma(M)\Gamma(N) = SL_2(\mathbb{Z}) \tag{2.7}$$

when $(M, N) = 1$. By the second isomorphism theorem and (2.5),

$$\Gamma(M)\Gamma(N)/\Gamma(N) \cong \Gamma(M)/(\Gamma(M) \cap \Gamma(N)) = \Gamma(M)/\Gamma(MN).$$

Therefore, by (2.6),

$$\begin{aligned}[\Gamma(M)\Gamma(N) : \Gamma(N)] &= [\Gamma(M) : \Gamma(MN)] \\ &= \frac{\nu^*(MN)}{\nu^*(M)} = \nu^*(N) = [SL_2(\mathbb{Z}) : \Gamma(N)],\end{aligned}$$

which means (2.7).

Next, we consider the similar problem in $PSL_2(\mathbb{R})$. By the second isomorphism theorem,

$$\bar{\Gamma}(M)/(\bar{\Gamma}(M) \cap \bar{\Gamma}(N)) \cong \bar{\Gamma}(M)\bar{\Gamma}(N)/\bar{\Gamma}(N).$$

By (2.7), we have $\bar{\Gamma}(M)\bar{\Gamma}(N) \supset \overline{\Gamma(M)\Gamma(N)} = PSL_2(\mathbb{Z})$. Therefore

$$[\bar{\Gamma}(M) : \bar{\Gamma}(M) \cap \bar{\Gamma}(N)] = [PSL_2(\mathbb{Z}) : \bar{\Gamma}(N)] = \frac{1}{2}\nu^*(N).$$

Now we have

$$\begin{aligned}\frac{1}{2}\nu^*(MN) &= [PSL_2(\mathbb{Z}) : \bar{\Gamma}(MN)] \\ &= [PSL_2(\mathbb{Z}) : \bar{\Gamma}(M)][\bar{\Gamma}(M) : \bar{\Gamma}(M) \cap \bar{\Gamma}(N)][\bar{\Gamma}(M) \cap \bar{\Gamma}(N) : \bar{\Gamma}(MN)] \\ &= \frac{1}{2}\nu^*(M)\frac{1}{2}\nu^*(N)[\bar{\Gamma}(M) \cap \bar{\Gamma}(N) : \bar{\Gamma}(MN)].\end{aligned}$$

Therefore, by (2.6),

$$[\bar{\Gamma}(M) \cap \bar{\Gamma}(N) : \bar{\Gamma}(MN)] = 2.$$

This completes the proof of the first assertion of the lemma. The second assertion easily follows from the first assertion. □

For a positive integer N, define the sets of discriminants

$$D'_N := \{d \in D \mid N|u(d)\} \tag{2.8}$$

and

$$D_N := \{d \in D'_N \mid t(d) \equiv \pm 2 \pmod N\}, \tag{2.9}$$

of D. It is obvious that when $(M, N) = 1$,

$$D'_M \cap D'_N = D'_{MN} \tag{2.10}$$

and $D_M \cap D_N \supsetneq D_{MN}$ hold. Now we prove the following lemma.

Lemma 3. *For a primitive hyperbolic element $P \in PSL_2(\mathbb{Z})$, denote by d_P the discriminant of the primitive quadratic form which is corresponding to P by Proposition 1.*

(1) *If N is a positive integer with $N \geq 3$,*

$$P \in \bar{\Gamma}(N) \iff d_P \in D_N. \tag{2.11}$$

(2) *If N is a prime number with $N \geq 3$, then $D'_N = D_N$. Therefore*

$$P \in \bar{\Gamma}(N) \iff d_P \in D'_N. \tag{2.12}$$

(3) *Let $N_1, \dots N_r$ be positive integers which are relatively prime each other and $N_j \geq 3$. Put $N = N_1 \cdots N_r$ and $\bar{\Gamma}'(N)$ is the subgroup of $PSL_2(\mathbb{R})$ given in Lemma 2. Then*

$$P \in \bar{\Gamma}'(N) \iff d_P \in \bigcap_{j=1}^{r} D_{N_j}. \tag{2.13}$$

In particular, if $N_1, \dots N_r$ are distinct prime numbers with $N_j \geq 3$, then

$$P \in \bar{\Gamma}'(N) \iff d_P \in D'_N. \tag{2.14}$$

Proof. The second assertion has already been obtained in [9, Proposition 3.3], however, we give the proof of all the assertions. By Proposition 1, the primitive hyperbolic element P may be identified with the fundamental automorphism

$$M_Q = \begin{pmatrix} \dfrac{t(d) - bu(d)}{2} & -cu(d) \\ au(d) & \dfrac{t(d) + bu(d)}{2} \end{pmatrix}$$

of the form $Q = [a, b, c]$ with discriminant d. $\bar{M}_Q \in \bar{\Gamma}(N)$ if and only if

$$\frac{t(d) - bu(d)}{2} \equiv \frac{t(d) + bu(d)}{2} \equiv \pm 1, \quad -cu(d) \equiv au(d) \equiv 0 \pmod N.$$

Since $(a, b, c) = 1$, we have

$$u(d) \equiv 0, \quad t(d) \equiv \pm 2 \pmod N.$$

This means $d \in D_N$.

Next we assume that N is a prime number with $N \geq 3$. If $d \in D'_N$ then $u(d) \equiv 0 \pmod{N}$. Since $(t(d), u(d))$ satisfies the Pell equation $t^2 - u^2 d = 4$, we have

$$t(d)^2 \equiv 4 \pmod{N}.$$

When N is a prime, this means $t(d) \equiv \pm 2 \pmod{N}$, that is, $d \in D_N$. We obtain the second assertion.

The relation (2.13) easily follows from the definition of $\bar{\Gamma}'(N)$ in Lemma 2 and the first assertion. Further if $N_1, \ldots, N_r$ are distinct primes with $N_j \geq 3$, by the second assertion and (2.10) we have,

$$\bigcap_{j=1}^{r} D_{N_j} = \bigcap_{j=1}^{r} D'_{N_j} = D'_{\prod N_j} = D'_N.$$

Therefore (2.14) is obtained as the special case of (2.13). $\square$

The Selberg zeta function for $\bar{\Gamma}(N)$ is defined by

$$Z_{\bar{\Gamma}(N)}(s) = \prod_{\{P_N\}_N} \prod_{k=0}^{\infty} \left(1 - N(P_N)^{-s-k}\right),$$

where P_N denotes the primitive hyperbolic element of $\bar{\Gamma}(N)$ and $\{P_N\}_N$ denotes its $\bar{\Gamma}(N)$-conjugacy class. By Lemma 1, the $PSL_2(\mathbb{Z})$-conjugacy class $\{P_N\}$ of P_N is divided into $\bar{\Gamma}(N)$-conjugacy classes $\{\gamma_j^{-1} P_N \gamma_j\}_N$, where $\gamma_1 \ldots \gamma_{\nu(N)}$ are representatives of $PSL_2(\mathbb{Z}) / \bar{\Gamma}(N)$ and $\nu(N)$ is the constant in Lemma 1. By Proposition 1, Lemma 3 and the above argument, we have the following expression

$$Z_{\bar{\Gamma}(N)}(s) = \prod_{d \in D_N} \prod_{k=0}^{\infty} \left(1 - \varepsilon(d)^{-2s-2k}\right)^{h(d)\nu(N)} \quad (\sigma > 1). \tag{2.15}$$

Next we will obtain the prime geodesic theorem of the type (2.3) for the congruence subgroups. Let Γ be a discrete subgroup of $PSL_2(\mathbb{R})$ and $M = H/\Gamma$ be its Riemann surface. Let $CP(\Gamma)$ denote the set of oriented closed geodesics on M. For $\gamma \in CP(\Gamma)$ let $\tau(\gamma)$ denote the length of γ. As we know, each closed geodesic $\gamma \in CP(\Gamma)$ is associated with a Γ-conjugacy class $\{P_\gamma\}_\Gamma$ of the primitive hyperbolic element $P_\gamma \in \Gamma$ and we have $\tau(\gamma) = \log N(P_\gamma)$. Thus the function $\pi_\Gamma(x)$ is represented as follows

$$\pi_\Gamma(x) = \sharp \left\{\gamma \in CP(\Gamma) : \ \tau(\gamma) \leq \log x\right\}. \tag{2.16}$$

Now we assume that $N_1, \ldots, N_r$ are positive integers which are relatively prime each other and $N_j \geq 3$. Put $N = \prod_{j=1}^{r} N_j$. By Lemma 1 and Lemma 2, the subgroup

$$\Gamma_1 = \bar{\Gamma}(N) \quad \text{or} \quad \bar{\Gamma}'(N)$$

is a normal subgroup of $\Gamma_0 = PSL_2(\mathbb{Z})$. Geometrically speaking, this means that the surface $M_1 = H/\Gamma_1$ is a finite regular cover of the surface $M_0 = H/\Gamma_0$ and the group of cover transformations G is isomorphic to Γ_0/Γ_1. As usual, we call an oriented primitive closed geodesic on the surface M a *prime* in M. Let $\pi : M_1 \to M_0$ be the natural projection. We say that a prime $\tilde{\gamma}$ in M_1 lies over a prime γ in M_0 if $\pi(\tilde{\gamma}) = \gamma$. Now we quote the following result.

Lemma 4. *For a prime $\gamma \in CP(\Gamma_0)$, let m be the order of $P_\gamma \in \Gamma_0$ in the group $G = \Gamma_0/\Gamma_1$ and $k = \frac{|G|}{m}$. Then γ is decomposed into the primes $\tilde{\gamma}_1, \ldots, \tilde{\gamma}_k \in CP(\Gamma_1)$ such that $\tau(\tilde{\gamma}_j) = m\tau(\gamma)$ for all $1 \le j \le k$. Especially, $P_\gamma \in \Gamma_1$ if and only if γ splits completely in M_1, namely, $\tau(\tilde{\gamma}_j) = \tau(\gamma)$ for all $1 \le j \le |G|$.*

Proof. This is [9, Proposition 2.3]. □

By (2.16) and the above lemma,

$$\begin{aligned}
\pi_{\Gamma_1}(x) &= \sharp \left\{ \tilde{\gamma} \in CP(\Gamma_1) : \begin{array}{l} \tau(\tilde{\gamma}) \le x \\ \tau(\tilde{\gamma}) = \tau(\gamma) \end{array} \right\} \\
&\quad + \sharp \left\{ \tilde{\gamma} \in CP(\Gamma_1) : \begin{array}{l} \tau(\tilde{\gamma}) \le x \\ \tau(\tilde{\gamma}) = m\tau(\gamma) \ (m \ge 2) \end{array} \right\} \\
&= |G| \sharp \left\{ \gamma \in CP(\Gamma_0) : \begin{array}{l} \tau(\gamma) \le x \\ P_\gamma \in \Gamma_1 \end{array} \right\} + O(\pi_{\Gamma_1}(\sqrt{x})).
\end{aligned}$$

Thus we have

$$\sum_{\substack{\{P\}_{\Gamma_0} \\ P \in \Gamma_1,\ N(P) \le x}} 1 = \frac{1}{|G|} \pi_{\Gamma_1}(x) + O(\pi_{\Gamma_1}(\sqrt{x})).$$

For the group $\Gamma_1 = \bar{\Gamma}(N)$ or $\bar{\Gamma}'(N)$, let D_{Γ_1} be the associated set of discriminants by Lemma 3. Then by (1.2), Proposition 1 and Lemma 3, we have

$$\sum_{\substack{d \in D_{\Gamma_1} \\ \varepsilon(d)^2 \le x}} h(d) = \frac{1}{[PSL_2(\mathbb{Z}) : \Gamma_1]} Li(x) + O(x^\alpha).$$

Namely, by Lemma 1 and Lemma 2, we obtain the following prime geodesic theorem.

Proposition 2.

(1) *Let N be a positive integer with $N \geq 3$. Then*

$$\sum_{\substack{d \in D_N \\ \varepsilon(d)^2 \leq x}} h(d) = \mu(N)\mathrm{Li}(x) + O(x^\alpha) \tag{2.17}$$

holds, where

$$\mu(N) = \frac{2}{\nu^*(N)} = \frac{1}{\nu(N)}$$

and $\nu^*(N)$ and $\nu(N)$ are the constants in Lemma 1.

(2) *Let $N_1, \dots, N_r$ be positive integers which are relatively prime each other and $N_j \geq 3$. Then*

$$\sum_{\substack{d \in \bigcap_{j=1}^r D_{N_j} \\ \varepsilon(d)^2 \leq x}} h(d) = \left(\prod_{j=1}^r \mu(N_j)\right) \mathrm{Li}(x) + O(x^\alpha). \tag{2.18}$$

In particular, if $N_1, \dots, N_r$ are distinct primes with $N_j \geq 3$,

$$\sum_{\substack{d \in D'_N \\ \varepsilon(d)^2 \leq x}} h(d) = \mu'(N)\mathrm{Li}(x) + O(x^\alpha) \tag{2.19}$$

holds, where

$$\mu'(N) = \prod_{j=1}^r \frac{2}{\nu^*(N_j)} = \frac{2^r}{\nu^*(N)}.$$

The formulas (2.17) and (2.18) play essential roles in the proof of Theorem 3. Remark that when N is a prime number with $N \geq 3$, (2.17) and (2.19) mean the same formula

$$\sum_{\substack{N | u(d) \\ \varepsilon(d)^2 \leq x}} h(d) = \frac{2}{N(N^2-1)}\mathrm{Li}(x) + O(x^\alpha),$$

which was [9, Theorem 3.4].

3. Basic tools for to prove universality

Let us consider a general Dirichlet series of the form

$$\sum_{\lambda \in \Lambda} \frac{a_\lambda}{e^{\lambda s}}, \tag{3.1}$$

where $\Lambda = \{\lambda\}$ is a monotone increasing sequence of positive real numbers tending to infinity and $a_\lambda \in \mathbb{C}$. In this section, we quote three propositions in Drungilas, Garunkštis, and Kačėnas [3], which are basic tools to prove the universality for the general Dirichlet series.

Proposition 3. *Suppose that the general Dirichlet series (3.1) satisfies the following conditions:*

(1) *(3.1) converges on some half-plane,*
(2) *(3.1) has an analytic continuation to a function $L(s)$ which is meromorphic for $\sigma > \sigma_1 > 1/2$,*
(3) *$L(s)$ is polynomial growth in $\sigma > \sigma_1$,*
(4) *$L(s)$ satisfies a mean value estimate*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T |L(\sigma + it)|^2 dt < \infty \quad (\sigma > \sigma_1).$$

Let K be a compact subset of $\sigma_1 < \sigma < 1$ with connected complement. Then for any small positive real numbers ε and ε_1 there exists a large positive real number Q_0 such that if $Q > Q_0$,

$$\mu\left(\left\{\tau \in [0,T] : \max_{s\in K}\left|L(s+i\tau) - \sum_{\substack{\lambda\in\Lambda \\ e^\lambda < Q}} \frac{a_\lambda}{e^{\lambda(s+i\tau)}}\right| < \varepsilon\right\}\right) > (1-\varepsilon_1)T.$$

Proof. This is [3, Proposition 2.2]. □

This proposition means that the Dirichlet series satisfying the mean value estimate can be uniformly approximated by the Dirichlet polynomial with sufficient length for almost all vertical translation.

Definition 1. For $x > 0$ put

$$N(x) := \sum_{\lambda \le x} |a_\lambda|.$$

We say that the series (3.1) satisfies the packing condition if for any $c > 0$ and $\varepsilon > 0$,

$$\left| N\left(x \pm \frac{c}{x^2}\right) - N(x) \right| \gg e^{(1-\varepsilon)x}. \tag{3.2}$$

The following proposition is called the general denseness lemma, which follows from some results in functional analysis such as the theorem of Hahn-Banach.

Proposition 4. *Suppose that the general Dirichlet series (3.1) satisfies the packing condition (3.2). Let K be a compact subset of $\frac{1}{2} < \sigma_1 < \sigma < \sigma_2 < 1$ with connected complement and $f(s)$ be a continuous function on K which is analytic in the interior of K. Then for any $\mu > 0$ there exists a positive constant $\rho_0 = \rho_0(\sigma_1, \sigma_2, K, f, \mu)$ such that for any $\rho > \rho_0$ there is a sequence $\theta_\lambda \in [0, 1)$ for which*

$$\max_{s \in K} \left| f(s) - \sum_{\substack{\lambda \in \Lambda \\ \mu < e^\lambda \le \rho}} \frac{a_\lambda e(\theta_\lambda)}{e^{\lambda s}} \right| \le C_K \sum_{\substack{\lambda \in \Lambda \\ \mu < e^\lambda \le \rho}} \frac{|a_\lambda|^2}{e^{2\lambda\sigma_1}},$$

where $e(x) = e^{2\pi i x}$ and C_K is the constant depends only on σ_1, σ_2, K and Λ.

Proof. This is [3, Proposition 2.3]. □

As we stated in §1, in the proof of the universality theorem for a function given by a Dirichlet series, the linearly independence of the set of Dirichlet exponents over $\mathbb{Q}$ plays an essential role. We need this condition to apply the following proposition.

Proposition 5. *Suppose that the general Dirichlet series (3.1) satisfies the following conditions:*

(1) *the set Λ is linearly independent over $\mathbb{Q}$,*
(2) *the series*

$$\sum_{\lambda \in \Lambda} \frac{|a_\lambda|^2}{e^{2\lambda\sigma}}$$

converges for $\sigma > \alpha$.

For numbers $\theta_\lambda \in [0, 1)$ $(\lambda \in \Lambda)$, $0 < \mu < \rho$ and $0 < \delta < 1/2$, we consider

$$S_T = S_T(\delta, \mu, \rho) = \left\{ \tau \in [0, T] : \left\| -\frac{\tau\lambda}{2\pi} - \theta_\lambda \right\| < \delta \quad (\mu < e^\lambda \le \rho) \right\},$$

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

(1) *We have*

$$\lim_{T\to\infty}\frac{\mu(S_T)}{T}=(2\delta)^M,$$

where $M=\sharp\{\lambda\in\Lambda:\mu<e^{\lambda}\le\rho\}$

(2) *Let* K *be a compact subset of* $\alpha<\sigma<1$ *and* Q *be a positive real number with* $Q>\rho$. *Denote by* S'_T *the set of* $\tau\in S_T$ *satisfying*

$$\max_{s\in K}\left|\sum_{\substack{\rho<e^{\lambda}\le Q\\ \lambda\in\Lambda}}\frac{a_\lambda}{e^{\lambda(s+i\tau)}}\right|\le C'_K\left(\sum_{\substack{e^{\lambda}>\rho\\ \lambda\in\Lambda}}\frac{|a_\lambda|^2}{e^{2\lambda\sigma_1}}\right)^{\frac{1}{4}},$$

where σ_1 *is a real number satisfying* $\alpha<\sigma_1<\min\{\Re s:s\in K\}$ *and* $C'_K=C'_K(\sigma_1)$ *be a positive constant depending only on* K *and* σ_1. *Then we have*

$$\lim_{T\to\infty}\frac{\mu(S'_T)}{T}>\frac{1}{2}(2\delta)^M.$$

Proof. This is a combination of [3, Lemma 2.9] (Generalized Kronecker's theorem) and [3, Proposition 2.8]. □

4. Approximation by sum of Dirichlet series over fundamental discriminants

In this section, we express the logarithm of the Selberg zeta function by a sum of two Dirichlet series over the set of fundamental discriminants.

4.1. Approximation by a Dirichlet series over fundamental discriminants

Define the set of positive fundamental discriminants by

$$\begin{aligned}D^* = &\{d\in\mathbb{Z}_{>0}:d\equiv 1 \mod 4,\ d \text{ is square-free}\}\\ &\cup\{d=4m:m\in\mathbb{Z}_{>0},\ m\equiv 2,3 \mod 4,\ m \text{ is square-free}\}.\end{aligned}$$

Drungilas, Garunkštis, and Kačėnas [3] proved the following proposition.

Proposition 6. *For* $d\in D$, *let* $\varepsilon(d)$ *be the fundamental unit given by (2.2). Then the set*

$$\{\log\varepsilon(d):d\in D^*\}$$

is linearly independent over $\mathbb{Q}$.

Proof. This is [3, Proposition 3.1]. □

The authors also proved that the logarithm of the Selberg zeta function for $PSL_2(\mathbb{Z})$ is approximated by the Dirichlet series over the fundamental discriminants. According to their method, we will represent the logarithms of $Z_j(s)$ by Dirichlet series over a subset of D^*. Let N be a positive integer which is square-free. For a subset D_N given in (2.9), define

$$D_N^* = D_N \cap D^*.$$

Let $\nu(N)$ be the constant in Lemma 1 when $N \geq 3$. When $N = 1$, put $\nu(1) = 1$. For $d \in D^*$, put $b(d) = \sharp\{n \in \mathbb{N} : n|u(d)\}$. Define a non-decreasing sequence $\{y_n\}$ of positive real numbers by

$$\{y_n\} = \{\varepsilon(d)^m : d \in D^*,\ m \in \mathbb{N}\}.$$

Proposition 7. *For each $j = 0 \dots r$ there exist a sequence $\{a_n^{(j)}\}$ of complex numbers such that for $\sigma > 1$,*

$$\log Z_j(s) = -\nu(N_j) \sum_{d \in D_{N_j}^*} \frac{h(d)b(d)}{\varepsilon_d^{2s}} + \sum_{n=1}^{\infty} \frac{a_n^{(j)}}{y_n^s}, \tag{4.1}$$

where the second series converges absolutely for $\sigma > 1/2$.

Proof. By (2.15),

$$\begin{aligned} Z_j(s) &= \prod_{d \in D_{N_j}} \left(1 - \varepsilon(d)^{-2s}\right)^{h(d)\nu(N_j)} Z_j(s+1) \\ &=: R_j(s) Z_j(s+1), \end{aligned}$$

where the function $R_j(s)$ is called the Ruelle zeta function. By taking logarithm of both sides, we have

$$\log Z_j(s) = \log R_j(s) + \log Z_j(s+1), \tag{4.2}$$

where the second term is absolutely convergent for $\sigma > 1/2$. By the Tailor expansion of $\log(1-x)$,

$$\begin{aligned} \log R_j(s) &= \sum_{d \in D_{N_j}} \nu(N_j) h(d) \log\left(1 - \varepsilon(d)^{-2s}\right) \qquad (4.3) \\ &= -\nu(N_j) \sum_{d \in D_{N_j}} \sum_{k=1}^{\infty} \frac{h(d)}{k\varepsilon(d)^{2ks}} \end{aligned}$$

$$= -\nu(N_j)\left\{\sum_{d\in D_{N_j}}\frac{h(d)}{\varepsilon(d)^{2s}} + \sum_{d\in D_{N_j}}\sum_{k=2}^{\infty}\frac{h(d)}{k\varepsilon(d)^{2ks}}\right\},$$

where the second double series converges absolutely for $\sigma > 1/2$.

Now we rewrite the first sum

$$\sum_{d\in D_{N_j}}\frac{h(d)}{\varepsilon(d)^{2s}}$$

to the sum over the set $D^*_{N_j}$. By definitions of D_{N_j} and $D^*_{N_j}$,

$$D_{N_j}\setminus D^*_{N_j} = \{d' = dn^2 : d\in D^*_{N_j},\ n>1\}.$$

Put $E^* = \{\varepsilon(d) : e\in D^*\}$. Assume that the discriminant $d' = dn^2 \in D_{N_j}\setminus D^*_{N_j}$ satisfies $\varepsilon(d')\in E^*$. Then there is some $d_1\in D^*$ satisfying $\varepsilon(d') = \varepsilon(d_1)$. This means that

$$\frac{t(dn^2)+u(dn^2)\sqrt{dn^2}}{2} = \frac{t(d_1)+u(d_1)\sqrt{d_1}}{2}$$

and that

$$d = d_1 \quad\text{and}\quad n|u(d).$$

Therefore, if we put $b(d) = \sharp\{n\in\mathbb{N} : n|u(d)\}$, we have

$$\begin{aligned}\sum_{d\in D_{N_j}}\frac{h(d)}{\varepsilon(d)^{2s}} &= \sum_{d\in D^*_{N_j}}\frac{h(d)}{\varepsilon(d)^{2s}} + \sum_{d\in D_{N_j}\setminus D^*_{N_j}}\frac{h(d)}{\varepsilon(d)^{2s}} \\ &= \sum_{d\in D^*_{N_j}}\frac{h(d)b(d)}{\varepsilon(d)^{2s}} + \sum_{d\in D^*_{N_j}}\sum_{\substack{n\geq 2\\ \varepsilon(dn^2)\notin E^*}}\frac{h(d)}{\varepsilon(d)^{2s}}.\end{aligned} \tag{4.4}$$

By the property of fundamental units (see, for instance, Davenport [4, Chapter 6]), we have

$$\{\varepsilon(dn^2) : d\in D^*_{N_j},\ n\geq 2,\ \varepsilon(dn^2)\notin E^*\} \subset \{\varepsilon(d)^m : d\in D^*_{N_j},\ m\geq 2\}.$$

If $\varepsilon(dn^2) = \varepsilon(d)^m$ holds, then

$$\frac{t(dn^2)+u(dn^2)\sqrt{dn^2}}{2} = \left(\frac{t(d)+u(d)\sqrt{d}}{2}\right)^m = \frac{t_m(d)+u_m(d)\sqrt{d}}{2}$$

for some integers $t_m(d), u_m(d)$. This means that for fixed d and m

$$\sharp\{n \in \mathbb{N} : \varepsilon(dn^2) = \varepsilon(d)^m\} \le \sum_{n | t_m(d)} 1 \ll_\varepsilon \varepsilon(d)^{m\varepsilon}.$$

By this and the trivial estimate $h(d) < d$, the second series in (4.4) is estimated by

$$\ll_\varepsilon \sum_{d \in D^*} \sum_{m \ge 2} \frac{h(d)}{\varepsilon(d)^{(2\sigma - \varepsilon)m}} \ll \sum_{d \in D^*} \frac{1}{\varepsilon(d)^{4\sigma - 1 - \varepsilon}}.$$

Therefore the second series in (4.4) is absolutely convergent for $\sigma > 1/2$. Combining (4.2) with (4.3) and (4.4), we obtain (4.1). □

In the following, we put

$$L_j(s) = \sum_{d \in D^*_{N_j}} \frac{h(d)b(d)}{\varepsilon(d)^{2s}} \tag{4.5}$$

for each $j = 0, \ldots, r$.

4.2. Outline of the proof of the universality for individual zeta functions

In [3], the authors proved that the series

$$L_0(s) = \sum_{d \in D^*} \frac{h(d)b(d)}{\varepsilon(d)^{2s}}$$

for $PSL_2(\mathbb{Z})$ has the following analytic property.

Proposition 8. *Let the function $r(s)$ be an analytic continuation to $\sigma > 1/2$, $t > 0$, of the Dirichlet series*

$$\sum_{d \in D^*} \frac{h(d)b(d)}{\varepsilon(d)^{2s}}.$$

Then $r(s)$ is of polynomial growth for $\sigma > \alpha$, where α is the same constant in the prime geodesic theorem (2.3), and

$$\lim_{T \to \infty} \frac{1}{T} \int_1^T |r(\sigma + it)|^2 dt = \sum_{d \in D^*} \frac{h(d)^2 b(d)^2}{\varepsilon(d)^4 \sigma}$$

for $\sigma > \frac{\alpha+1}{2}$.

Proof. This is [3, Proposition 3.4]. □

This proposition follows from the prime geodesic theorem (2.3) for $PSL_2(\mathbb{Z})$ and the next lemma.

Lemma 5. *Define a non-decreasing sequence* $\{x_n\}$ *of positive real numbers by*

$$\{\varepsilon(d)^2 : d \in D^*\} \cup \mathbb{N} = \{x_1 < x_2 < x_3 \cdots\}.$$

Then the inequality

$$x_{n+1} - x_n \geq \frac{1}{x_n + 1}$$

holds for each $n > 3$.

Proof. This is [3, Lemma 3.7]. □

By the prime geodesic theorem (2.3) for $\Gamma = PSL_2(\mathbb{Z})$, we have

$$\left|\pi_\Gamma\left(\exp\left(x + \frac{c}{x^2}\right)\right) - \pi_\Gamma(e^x)\right| \gg \frac{e^x}{x^3}$$

for any positive constant c. This means that the packing condition (3.2) holds for the series $L_0(s)$. Therefore, by Proposition 7 and Proposition 8, the series $L_0(s)$ satisfies all the assumptions of Propositions 3-5 in the strip $\frac{\alpha+1}{2} < \sigma < 1$. Theorem 2 follows from these propositions and Proposition 6.

Now we have the prime geodesic theorem (2.17) for principal congruence subgroups, so the packing condition (3.2) and Proposition 8 also hold even if we replace D^* by a subset D_N^*. From this and Proposition 7, as Drungilas, Garunkštis, and Kačėnas [3] predicted, the ordinary universality theorem for each $Z_j(s)$ holds in the strip $\frac{\alpha+1}{2} < \sigma < 1$.

4.3. Definition of $L_j^{(1)}(s)$ and $L_j^{(2)}(s)$

To prove Theorem 3, we divide the series

$$L_j(s) = \sum_{d \in D^*_{N_j}} \frac{h(d)b(d)}{\varepsilon(d)^{2s}}$$

into two sub series.

Definition 2. For each $j = 0 \ldots r$, we define subsets $D_j^{(1)}$ and $D_j^{(2)}$ of $D^*_{N_j}$ as follows.

(1) For $j = 0$,

$$D_0^{(2)} = \bigcup_{i=1}^{r} D^*_{N_i}, \quad D_0^{(1)} = D^* \setminus D_0^{(2)}.$$

(2) For $j = 1 \dots r$,

$$D_j^{(2)} = \bigcup_{\substack{i=1 \\ i \neq j}}^{r} \left(D^*_{N_j} \cap D^*_{N_i} \right), \quad D_j^{(1)} = D^*_{N_j} \setminus D_j^{(2)}.$$

For instance, when $r = 1$ we have

$$D_0^{(1)} = D^* \setminus D^*_{N_1}, \quad D_0^{(2)} = D_1^{(1)} = D^*_{N_1}, \quad D_1^{(2)} = \emptyset.$$

When $r = 2$ we have

$$\begin{aligned} D_0^{(1)} &= D^* \setminus \left(D^*_{N_1} \cup D^*_{N_2} \right), & D_0^{(2)} &= D^*_{N_1} \cup D^*_{N_2}, \\ D_1^{(1)} &= D^*_{N_1} \setminus (D^*_{N_1} \cap D^*_{N_2}), & D_1^{(2)} &= D^*_{N_1} \cap D^*_{N_2}, \\ D_2^{(1)} &= D^*_{N_2} \setminus (D^*_{N_1} \cap D^*_{N_2}), & D_2^{(2)} &= D^*_{N_1} \cap D^*_{N_2}. \end{aligned}$$

It is clear that $D_j^{(1)} \cup D_j^{(2)} = D^*_{N_j}$ for each $j = 0, \dots, r$ and that all $D_j^{(1)}$ $(j = 0, \dots, r)$ and $\bigcup_{j=0}^{r} D_j^{(2)}$ are disjoint. Now we prove that each $D_j^{(1)}$ has a positive density.

Proposition 9. *For $x > 0$,*

$$\sum_{\substack{d \in D_j^{(1)} \\ \varepsilon(d)^2 \leq x}} h(d) b(d) = C_j \mathrm{Li}(x) + O(x^\alpha), \tag{4.6}$$

where

$$C_j = \begin{cases} \prod_{i=1}^{r} (1 - \mu(N_i)) & (j = 0), \\ \mu(N_j) \prod_{\substack{i=1 \\ i \neq j}}^{r} (1 - \mu(N_i)) & (j = 1 \dots r), \end{cases}$$

and $\mu(N)$ is the constant in Proposition 2.

Proof. First recall the prime geodesic theorem (2.3) for $PSL_2(\mathbb{Z})$

$$\sum_{\substack{d \in D \\ \varepsilon(d)^2 \leq x}} h(d) = \mathrm{Li}(x) + O(x^\alpha).$$

By the argument in §4.1, the left-hand side is rewritten as follows.

$$\sum_{\substack{d \in D^* \\ \varepsilon(d)^2 \leq x}} h(d) b(d) + \sum_{d \in D^*} \sum_{\substack{n \geq 2 \\ \varepsilon(dn^2) \leq x \\ \varepsilon(dn^2) \notin E}} h(d).$$

The second term is estimated by

$$\ll \sum_{d \in D^*} \sum_{\substack{m=2 \\ \varepsilon(d)^{2m} \leq x}} h(d)\varepsilon(d)^{m\varepsilon} \ll x^{\varepsilon} \sum_{\substack{d \in D \\ \varepsilon(d) \leq \sqrt{x}}} h(d) \ll x^{\alpha}.$$

Therefore we have for $j = 0$,

$$\sum_{\substack{d \in D^* \\ \varepsilon(d)^2 \leq x}} h(d)b(d) = \mathrm{Li}(x) + O(x^{\alpha}). \tag{4.7}$$

By the first assertion of Proposition 2, we also have for $j = 1 \dots r$,

$$\sum_{\substack{d \in D^*_{N_j} \\ \varepsilon(d)^2 \leq x}} h(d)b(d) = \mu(N_j)\mathrm{Li}(x) + O(x^{\alpha}). \tag{4.8}$$

Further, by the second assertion of Proposition 2,

$$\sum_{\substack{d \in D^*_I \\ \varepsilon(d)^2 \leq x}} h(d)b(d) = \left(\prod_{i \in I} \mu(N_i)\right) \mathrm{Li}(x) + O(x^{\alpha}), \tag{4.9}$$

where $D^*_I = \bigcap_{i \in I} D^*_{N_i}$ for a subset $I \subset \{1, 2, \dots, r\}$. Now we prove the proposition. First we consider the case $j = 0$. By the definition of $D^{(1)}_0$, (4.7) and (4.9),

$$\begin{aligned} \sum_{\substack{d \in D^{(1)}_0 \\ \varepsilon(d)^2 \leq x}} h(d)b(d) &= \left\{ \sum_{\substack{d \in D^* \\ \varepsilon(d)^2 \leq x}} + \sum_{I \subset \{1,\dots,r\}} (-1)^{|I|} \sum_{\substack{d \in D^*_I \\ \varepsilon(d)^2 \leq x}} \right\} h(d)b(d) \\ &= C_0 \mathrm{Li}(x) + O(x^{\alpha}), \end{aligned}$$

where

$$\begin{aligned} C_0 &= 1 + \sum_{I \subset \{1,\dots,r\}} (-1)^{|I|} \prod_{i \in I} \mu(N_i) \\ &= \prod_{i=1}^{r} (1 - \mu(N_i)). \end{aligned}$$

Next we consider the case $j = 1, \dots, r$. By the definition of $D^{(1)}_j$, (4.8) and (4.9),

$$\sum_{\substack{d\in D_j^{(1)}\\ \varepsilon(d)^2\le x}} h(d)b(d) = \left\{ \sum_{\substack{d\in D_{N_j}^*\\ \varepsilon(d)^2\le x}} + \sum_{\substack{I\subset\{1,\dots,r\}\\ j\notin I}} (-1)^{|I|} \sum_{\substack{d\in D_{N_j}^*\cap D_I^*\\ \varepsilon(d)^2\le x}} \right\} h(d)b(d)$$
$$= C_j \mathrm{Li}(x) + O(x^\alpha),$$

where

$$C_j = \mu(N_j) + \sum_{\substack{I\subset\{1,\dots,r\}\\ j\notin I}} (-1)^{|I|}\mu(N_j)\prod_{i\in I}\mu(N_i)$$
$$= \mu(N_j)\prod_{\substack{i=1\\ i\neq j}}^{r}(1-\mu(N_i)). \quad \square$$

By (4.7)–(4.8) we also have the prime geodesic theorems for $D_j^{(2)}$

$$\sum_{\substack{d\in D_j^{(2)}\\ \varepsilon(d)^2\le x}} h(d)b(d) = C_j^{(2)}\mathrm{Li}(x) + O(x^\alpha), \tag{4.10}$$

where $C_j^{(2)} \geq 0$ for each $j = 0, \dots, r$. We divide $L_j(s)$ into two sub series

$$L_j(s) = \sum_{d\in D_j^{(1)}} \frac{h(d)b(d)}{\varepsilon(d)^{2s}} + \sum_{d\in D_j^{(2)}} \frac{h(d)b(d)}{\varepsilon(d)^{2s}}$$
$$=: L_j^{(1)}(s) + L_j^{(2)}(s).$$

Now we have the prime geodesic theorems (4.6) and (4.10), which are corresponding to each series $L_j^{(1)}(s)$ and $L_j^{(2)}(s)$. By the same argument in the previous section, for each $j = 0, \dots, r$, the following statements hold:

(1) $L_j(s)$ satisfies the assumption of Proposition 3.
(2) $L_j^{(1)}(s)$ satisfies the assumption of Proposition 4,
(3) $L_j^{(1)}(s)$ and $L_j^{(2)}(s)$ satisfy the assumption of Proposition 5.

5. Proof of Theorem 3

To simplify the proof of Theorem 3, let us define some symbols. For a Dirichlet series $L(s) = \sum_{\lambda\in\Lambda} \frac{a_\lambda}{e^{\lambda s}}$ and positive numbers $X < Y$, we put

$$L_X(s)=\sum_{\substack{\lambda\in\Lambda\\ e^\lambda\le X}}\frac{a_\lambda}{e^{\lambda s}},\quad L_{X<Y}(s)=\sum_{\substack{\lambda\in\Lambda\\ X<e^\lambda\le Y}}\frac{a_\lambda}{e^{\lambda s}}.$$

Assume that compact subsets K_j and functions $f_j(s)$ satisfy the assumption in Theorem 3. Let σ_1 be a real number with $\frac{\alpha+1}{2}<\sigma_1<\min\{\Re s : s\in\cup_{j=0}^r K_j\}$ and σ_2 be a real number with $\max\{\Re s : s\in\cup_{j=0}^r K_j\}<\sigma_2<1$. We will show that for any small positive number ε and any sufficiently large number T there exists a subset C_T of the interval $[0,T]$ with positive density such that for any $\tau\in C_T$

$$I_j:=\max_{s\in K_j}|\log Z_j(s+i\tau)-\log f_j(s)|<\varepsilon \tag{5.1}$$

holds for each $j=0,\dots,r$. By Proposition 7,

$$\log Z_j(s)=-\nu(N_j)L_j(s)+\sum_{n=1}^{\infty}\frac{a_n^{(j)}}{y_n^s} \tag{5.2}$$

holds for $\sigma>1$, where $L_j(s)$ is given by (4.5), y_n is a non-decreasing sequence given by

$$\{y_n\}=\{\varepsilon(d)^m : d\in D^*,\ m\in\mathbb{N}\}$$

and the second series converges absolutely for $\sigma>1/2$. Fix a sufficiently large positive number ρ_3 such that

$$\max_{s\in K_j}\left|\sum_{y_n>\rho_3}\frac{a_n^{(j)}}{y_n^s}\right|<\frac{\varepsilon}{2} \tag{5.3}$$

holds for each $j=0,\dots,r$. Put

$$L_{j,\rho_3}^{(3)}(s)=-\frac{1}{\nu(N_j)}\sum_{y_n\le\rho_3}\frac{a_n^{(j)}}{y_n^s}$$

and

$$g_j(s):=-\frac{1}{\nu(N_j)}\log f_j(s).$$

Then, by (5.1)–(5.3),

$$I_j\le\nu(N_j)I_j'+\frac{\varepsilon}{2} \tag{5.4}$$

holds for all $\tau\in\mathbb{R}$, where we put

$$I_j'=\max_{s\in K_j}\left|L_j(s+i\tau)+L_{j,\rho_3}^{(3)}(s+i\tau)-g_j(s)\right|. \tag{5.5}$$

Let us divide I'_j into seven parts. Recall that all $L_j(s)$ satisfy the assumption in Proposition 3. We may say that $L_j(s)$ can be replaced by the Dirichlet polynomial $L_{j,X}(s)$ if $X > 0$ is a sufficiently large number. Fix a sufficiently large positive number ρ_2 satisfying $\rho_2 > \rho_3$ and

$$C'_{K_j} \left(\sum_{\substack{d \in D_j^{(2)} \\ \varepsilon(d)^2 > \rho_2}} \frac{|h(d)b(d)|^2}{\varepsilon(d)^{4\sigma_1}} \right)^{\frac{1}{4}} < \frac{\varepsilon}{14\nu(N_j)} \tag{5.6}$$

for each $j = 0, \ldots, r$, where $C'_K = C'_K(\sigma_1)$ is a positive constant in Proposition 5. Assume that $X > \rho_2$. By (5.5) and the Cauchy-Schwarz inequality, we have

$$I'_j \leq I''_j + \sum_{k=4}^{7} I_j^{(k)}, \tag{5.7}$$

where

$$\begin{aligned} I''_j &= \max_{s \in K_j} |L_{j,X}^{(1)}(s + i\tau) + L_{j,\rho_2}^{(2)}(s) + L_{j,\rho_3}^{(3)}(s) - g_j(s)|, \qquad (5.8) \\ I_j^{(4)} &= \max_{s \in K_j} |L_{j,\rho_2}^{(2)}(s + i\tau) - L_{j,\rho_2}^{(2)}(s)|, \\ I_j^{(5)} &= \max_{s \in K_j} |L_{j,\rho_2 < X}^{(2)}(s + i\tau)|, \\ I_j^{(6)} &= \max_{s \in K_j} |L_j(s + i\tau) - L_{j,X}(s + i\tau)|, \end{aligned}$$

and

$$I_j^{(7)} = \max_{s \in K_j} |L_{j,\rho_3}^{(3)}(s + i\tau) - L_{j,\rho_3}^{(3)}(s)|.$$

Next we apply Proposition 4 for each series $L_j^{(1)}(s)$. Fix a sufficiently large number $\mu_1 > 0$ satisfying $\mu_1 > \rho_3$ and

$$C_{K_j} \sum_{\substack{d \in D_j^{(1)} \\ \varepsilon(d)^2 > \mu_1}} \frac{|h(d)b(d)|^2}{\varepsilon(d)^{4\sigma_1}} < \frac{\varepsilon}{14\nu(N_j)}$$

for each $j = 0, \ldots, r$, where C_K is the positive constant in Proposition 4. By this and Proposition 4, there exists a positive constant $\rho_0 > \mu_1$ such that if $\rho_1 > \rho_0$ there exist numbers

$$\theta_d^{(j)} \in [0, 1) \quad (d \in D_j^{(1)},\ \mu_1 < \varepsilon(d)^2 \leq \rho_1)$$

such that

$$\max_{s\in K_j}\left|\sum_{\substack{d\in D_j^{(1)}\\ \mu_1<\varepsilon(d)^2\le\rho_1}}\frac{h(d)b(d)e(\theta_d^{(j)})}{\varepsilon(d)^{2s}}+L_{j,\mu_1}^{(1)}(s)+L_{j,\rho_2}^{(2)}(s)+L_{j,\rho_3}^{(3)}(s)-g_j(s)\right|<\frac{\varepsilon}{28\nu(N_j)} \tag{5.9}$$

holds for each $j=0,\dots,r$. Here we may assume that ρ_1 also satisfies

$$C'_{K_j}\left(\sum_{\substack{d\in D_j^{(1)}\\ \varepsilon(d)^2>\rho_1}}\frac{|h(d)b(d)|^2}{\varepsilon(d)^{4\sigma_1}}\right)^{\frac{1}{4}}<\frac{\varepsilon}{14\nu(N_j)} \tag{5.10}$$

for each $j=0,\dots,r$. Let $X>\max\{\rho_1,\rho_2\}$. By (5.8) and the Cauchy-Schwarz inequality,

$$I''_j\le\sum_{k=1}^{3}I_j^{(k)}, \tag{5.11}$$

where

$$I_j^{(1)}=\max_{s\in K_j}|L_{j,\mu_1<\rho_1}^{(1)}(s+i\tau)+L_{j,\mu_1}^{(1)}(s)+L_{j,\rho_2}^{(2)}(s)-L_{j,\rho_3}^{(3)}(s)|, \tag{5.12}$$

$$I_j^{(2)}=\max_{s\in K_j}|L_{j,\mu_1}^{(1)}(s+i\tau)-L_{j,\mu_1}(s)|,$$

and

$$I_j^{(3)}=\max_{s\in K_j}|L_{j,\rho_1<X}^{(1)}(s+i\tau)|.$$

Combining (5.7) with (5.11), we have

$$I'_j\le\sum_{k=1}^{7}I_j^{(k)}. \tag{5.13}$$

Now we define a subset of the interval $[0,T]$. Put

$$D_{\rho_1,\rho_2}:=\bigcup_{k=1}^{2}\left\{d\in\bigcup_{j=0}^{r}D_j^{(k)}:\varepsilon(d)^2\le\rho_k\right\}.$$

Since all sets $D_j^{(1)}$ $(j=0,\dots,r)$ and $\bigcup_{j=0}^{r}D_j^{(2)}$ are disjoint and the relation $\rho_3<\rho_2,\ \rho_3<\mu_1<\rho_1$, we can define the numbers θ_d $(d\in D_{\rho_1,\rho_2})$ as follows

$$\theta_d = \begin{cases} \theta_d^{(j)} & (\mu_1 < \varepsilon(d)^2 \le \rho_1 \text{ and } d \in D_j^{(1)} \text{ for } j = 0, \dots, r), \\ 0 & (\varepsilon(d)^2 \le \mu_1 \text{ and } d \in \bigcup_{j=0}^{r} D_j^{(1)}), \\ 0 & (\varepsilon(d)^2 \le \rho_2 \text{ and } d \in \bigcup_{j=0}^{r} D_j^{(2)}). \end{cases}$$

For a real number δ with $0 < \delta < 1/2$, define

$$S_T(\delta) = \left\{ \tau \in [0, T] : \left\| -\frac{\tau \log \varepsilon(d)}{\pi} - \theta_d \right\| < \delta \text{ for all } d \in D_{\rho_1,\rho_2} \right\}.$$

By the continuity of the Dirichlet polynomials, if we fix a sufficiently small $\delta > 0$, we have for $\tau \in S_T(\delta)$

$$\max_{s \in K_j} \left| \sum_{\substack{d \in D_j^{(1)} \\ \mu_1 < \varepsilon(d)^2 \le \rho_1}} \frac{h(d)b(d)(\varepsilon(d)^{-2i\tau} - e(\theta_d))}{\varepsilon(d)^{2s}} \right| < \frac{\varepsilon}{28\nu(N_j)}. \tag{5.14}$$

By the definition of θ_d, (5.9) and (5.12), we have

$$I_j^{(1)} < \frac{\varepsilon}{14\nu(N_j)} \tag{5.15}$$

for $\tau \in S_T(\delta)$. Similarly we have

$$I_j^{(2)},\ I_j^{(4)},\ I_j^{(7)} < \frac{\varepsilon}{14\nu(N_j)} \tag{5.16}$$

for any $\tau \in S_T(\delta)$. By Proposition 5-(1), the set $S_T(\delta)$ have the positive density,

$$\lim_{T \to \infty} \frac{\mu(S_T(\delta))}{T} = (2\delta)^{\sharp D_{\rho_1,\rho_2}} =: M_{\rho_1,\rho_2}, \tag{5.17}$$

where the number $\sharp D_{\rho_1,\rho_2}$ is computable according to the argument in [9, Chapter 4]. By Proposition 5-(2), (5.6) and (5.10), there exists a subset A_T of $S_T(\delta)$ such that

$$\lim_{T \to \infty} \frac{\mu(A_T)}{T} > \frac{1}{2} M_{\rho_1,\rho_2} \tag{5.18}$$

and that we have for any $\tau \in A_T$,

$$\max_{s \in K_j} \left| \sum_{\substack{d \in D_j^{(k)} \\ \rho_k < \varepsilon(d)^2 \le X}} \frac{h(d)b(d)}{\varepsilon(d)^{2(s+i\tau)}} \right| < \frac{\varepsilon}{14\nu(N_j)} \quad (k = 1, 2),$$

that is

$$I_j^{(3)},\ I_j^{(5)} < \frac{\varepsilon}{14\nu(N_j)}. \tag{5.19}$$

Now we apply Proposition 3 for the series $L_j(s)$, $\varepsilon = \frac{\varepsilon}{14\nu(N_j)}$, and $\varepsilon_1 = \frac{1}{4}M_{\rho_1,\rho_2}$. Then there is a positive constant X_0 such that if $X > X_0$ the set B_T of real numbers $\tau \in [0, T]$ for which

$$\max_{s\in K_j} \left| L_j(s+i\tau) - \sum_{\substack{d\in D_j \\ \rho_\varepsilon(d)^2 \le X}} \frac{h(d)b(d)}{\varepsilon(d)^{2(s+i\tau)}} \right| < \frac{\varepsilon}{14\nu(N_j)} \quad ,$$

that is

$$I_j^{(6)} < \frac{\varepsilon}{14\nu(N_j)} \tag{5.20}$$

has a positive density

$$\lim_{T\to\infty} \frac{\mu(B_T)}{T} > 1 - \frac{1}{4}M_{\rho_1,\rho_2}. \tag{5.21}$$

Now we fix a positive real number X satisfying $X > \max\{\rho_1, \rho_2, X_0\}$. Put

$$C_T = A_T \cap B_T.$$

By (5.18) and (5.21), we have

$$\lim_{T\to\infty} \frac{\mu(C_T)}{T} > \frac{1}{4}M_{\rho_1,\rho_2} > 0.$$

For any $\tau \in C_T$, the inequalities (5.15), (5.16), (5.19) and (5.20) hold. By (5.4) and (5.13) we have

$$I_j = \max_{s\in K_j} |\log Z_j(s+i\tau) - \log f_j(s)| < \varepsilon,$$

for any $\tau \in C_T$. This completes the proof of Theorem 3. $\square$

Acknowledgments

The author would like to express his appreciation to the referee for the comments and the suggestion. This work was supported by JSPS Grant-in-Aid for Scientific Research (B) 18H01111 and JSPS Grant-in-Aid for Scientific Research (C) 19K03455.

References

[1] B. Bagchi, The statistical behavior and universality properties of the Riemann zeta-function and other allied Dirichlet series, Ph.D. Thesis, Indian Statistical Institute, Calcutta, 1981.
[2] H. Bohr, R. Courant, Neue Anwendungen der Theorie der Diophantischen auf die Riemannsche Zetafunktion, J. Reine Angew. Math. 144 (1914) 249–274.
[3] P. Drungilas, R. Garunkštis, A. Kačėnas, Universality of the Selberg zeta-function for the modular group, Forum Math. 25 (3) (2013) 533–564.
[4] H. Davenport, Multiplicative Number Theory, 3rd ed., Graduate Texts in Mathematics, vol. 74, Springer, 2000.
[5] D.A. Hejhal, The Selberg Trace Formula for $PSL_2(\mathbb{R})$, Vol. 2, Lecture Notes in Mathematics, vol. 1001, Springer-Verlag, 1983.
[6] E. Landau, Elementary Number Theory, Chelsea, New York, 1958.
[7] W. Luo, Z. Rudnick, P. Sarnak, On Selberg's eigenvalue conjecture, Geom. Funct. Anal. 5 (1995) 387–401.
[8] H. Mishou, Joint universality theorems for pairs of automorphic zeta functions, Math. Z. 277 (2014) 1113–1154.
[9] P. Sarnak, Class number of indefinite binary quadratic forms, J. Number Theory 15 (2) (1982) 229–247.
[10] 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.
[11] K. Soundararajan, M.P. Young, The prime geodesic theorem, J. Reine Angew. Math. 676 (2013) 105–120.
[12] J. Steuding, Value-Distribution of L-Functions, Lecture Notes in Math., vol. 1877, Springer, 2007.
[13] S.M. Voronin, The distribution of the non-zero values of the Riemann zeta-function, Tr. Mat. Inst. Steklova 128 (1972) 131–150 (in Russian), Proc. Steklov Inst. Math. 128 (1972) 153–175.
[14] S.M. Voronin, Theorem on the universality of the Riemann zeta function, Izv. Acad. Nauk SSSR Ser. Mat. 39 (1975) 475–486 (in Russian), Math. USSR, Izv. 9 (1975) 443–453.