

Selberg's zeta function and the spectral geometry of geometrically finite hyperbolic surfaces

David Borthwick*, Chris Judge and Peter A. Perry†

Abstract. For hyperbolic Riemann surfaces of finite geometry, we study Selberg's zeta function and its relation to the relative scattering phase and the resonances of the Laplacian. As an application we show that the conjugacy class of a finitely generated, torsion-free, discrete subgroup of $SL(2, \mathbb{R})$ is determined by its trace spectrum up to finitely many possibilities, thus generalizing results of McKean [20] and Müller [23] to groups which are not necessarily cofinite.

Mathematics Subject Classification (2000). Primary 58J50, 35P25; Secondary 47A40.

1. Introduction

A geometrically finite hyperbolic surface is a topologically finite, complete Riemannian surface of constant curvature -1 . In this paper we study the Selberg zeta function, $Z_X(s)$ (see §4 for the definition), associated to the length spectrum of a geometrically finite surface X . When X has infinite area, the discrete spectrum of the positive Laplacian Δ_X will be finite and possibly empty. The appropriate spectral invariants at this level of generality are the *scattering resonances* or simply *resonances*. These are the poles of the meromorphic continuation of the resolvent $(\Delta_X - s(1 - s))^{-1}$, with multiplicities (see §3 for the definition).

Building upon our previous results and the work of Guillopé and Zworski [10], we exhibit an explicit connection (Theorem 4.1) between Selberg's zeta function and a Hadamard product over the resonance set. This yields the following:

Theorem 1.1. *Let n_C denote the number of cusps of X , and let χ denote the Euler characteristic of X . The Selberg zeta function Z_X extends to a meromorphic function of order two, with a divisor that can be divided into spectral and topological components: The spectral zeros of Z_X are given by the resonance set \mathcal{R}_X (with multiplicities). In addition, $Z_X(s)$ has topological zeros at $s = -k$ for $k \in \mathbb{N}_0$ of order $(2k + 1) \cdot (-\chi)$, and topological poles of order n_C at $s \in \frac{1}{2} - \mathbb{N}_0$.*

*Supported in part by NSF grant DMS-0204985.

†Supported in part by NSF grant DMS-0100829.

In the full geometrically finite context, the meromorphic continuation of Z_X to \mathbb{C} was proven by Guillopé [7]. If X has finite area, then meromorphic continuation with the order bound and divisor as given here can be deduced from the Selberg trace formula (see [12]). In the convex co-compact case ($n_C = 0$) Patterson–Perry [26] established the order bound, using thermodynamic formalism of Ruelle and Fried, and also computed the divisor. The thermodynamic formalism does not extend to the case $n_C > 0$. To prove that Theorem 1.1 holds for all geometrically finite X requires a different approach from these previous results.

As in the compact case, one can exploit knowledge of the divisor of the zeta function to link the resonance set to the length spectrum.

Corollary 1.2. *The resonance set determines the length spectrum of X , the Euler characteristic, χ , and the number of cusps, n_C . The length spectrum determines χ and n_C up to a finite number of possibilities. The length spectrum, χ , and n_C , together determine the resonance set.*

At present, we do not know whether the length spectrum (by way of Z_X) determines χ and n_C when X has infinite area and $n_C > 0$. That is, we cannot rule out the possibility that the resonance set contributes to the multiplicity of each of the zeros of Z_X that lie in $1/2 - \mathbb{N}_0/2$. However, if X has finite area, the resonances are confined to a vertical strip, and hence for sufficiently negative $k \in -\mathbb{N}_0$, the multiplicity of the zero of Z_X equals $(2k + 1) \cdot (-\chi)$. And in the convex co-compact case (infinite area with no cusps), Martin Olbrich has pointed out to us that Corollary 6.9 of [24], in conjunction with the theory of [4], shows that the multiplicity of resonances at $-\mathbb{N}$ is either zero (for non-elementary groups) or two (for elementary groups). Thus, in the convex co-compact case χ can be recovered as $(\text{ord}_{s=-k} Z_X - \text{ord}_{s=-k-1} Z_X)/2$ for any $k \in \mathbb{N}$.¹

By combining Corollary 1.2 with methods of Teichmüller theory we obtain the following application:

Theorem 1.3. *Let X be a complete, geometrically finite hyperbolic surface of infinite area. Then the length spectrum of X determines X up to finitely many possibilities. In particular, the resonance set determines X up to finitely many possibilities.*

Our proof of the first claim in Theorem 1.3 requires Theorem 1.1. Indeed, we do not know how to obtain bounds on the Euler characteristic without using scattering theory. Examples of non-isometric infinite-area hyperbolic surfaces with the same resonance set (and hence the same length spectrum) were given by Brooks–Davidovich [3].

¹The possibility of resonances overlapping topological zeros was overlooked in [1], but Olbrich’s result implies that Theorem 3.1 of that paper is correct as stated.

If X is connected, then there exists a finitely generated, torsion free, discrete subgroup, Γ , of $SL(2, \mathbb{R})$, unique up to conjugation, such that X is the quotient of the hyperbolic upper half-plane, \mathbb{H} , by Γ acting as Möbius transformations. Conversely, any such quotient is a geometrically finite hyperbolic surface. Moreover, the length spectrum of X equals (twice) the set $\{\cosh(\text{trace}(\gamma)) \mid \gamma \in \Gamma\}$ (including multiplicities). Thus, from Theorem 1.3 we have

Corollary 1.4. *Let Γ be a finitely generated, torsion-free, discrete subgroup of $SL(2, \mathbb{R})$ of co-infinite area. Then the set of traces $\{\text{trace}(\gamma) \mid \gamma \in \Gamma\}$ (including multiplicities) determines the conjugacy class of Γ up to finitely many possibilities.*

Theorem 1.3 is due to McKean [20] in the case that X is compact and to Müller [23] in the case that X has finite area. Corollary 1.4 is also due to McKean [20] in the case that Γ is cocompact.

Theorem 1.1 is proven by linking the zeta function to the scattering theory of Δ_X . For geometrically finite X scattering theory can be set up (following the approach of Guillopé and Zworski [10], [11]) using a decomposition of the form

$$X = Z \sqcup (C_1 \sqcup \dots \sqcup C_{n_C}) \sqcup (F_1 \sqcup \dots \sqcup F_{n_F}), \tag{1.1}$$

where Z is a compact surface, each F_i , $1 \leq i \leq n_F$, is a funnel, and each C_j , $1 \leq j \leq n_C$, is a cusp (end). This is illustrated in Figure 1. (For an exact description of funnels and cusps, see §2.) The boundary of Z consists of n_F closed geodesics (uniquely determined) and n_C horocycles (the choice of which is not unique) along which Z is glued to the funnel and cusp ends, respectively. The ideal boundary $\partial_\infty X$ is the disjoint union of n_C ideal points and n_F circles.

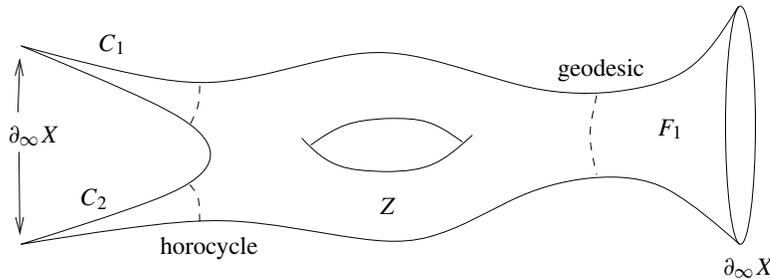


Figure 1. Decomposition of the surface X .

The scattering operator $S_X(s)$, described in greater detail in §3, can be viewed as a map from $C^\infty(\partial_\infty X)$ to itself, where a “smooth function” on $\partial_\infty X$ is interpreted as an element of $C^{n_C} \oplus (\bigoplus_{j=1}^{n_F} C^\infty(S^1))$. Because the funnels are isometric to standard

models (see §2), one can define relative scattering from the disjoint union

$$Y = F_1 \cup F_2 \cup \cdots \cup F_{n_F}$$

to X . On each funnel end, the Laplacian Δ_{F_j} carries Dirichlet boundary conditions on the closed geodesic which joins F_j to Z . The scattering operator for $\bigoplus_{j=1}^{n_F} \Delta_{F_j}$, denoted $S_Y(s)$, is a direct sum of scattering operators for the Dirichlet Laplacians Δ_{F_j} , and acts on $\bigoplus_{j=1}^{n_F} C^\infty(S^1)$. If we set

$$S_0(s) = \left(\bigoplus_{j=1}^{nc} \mathbf{1} \right) \oplus S_Y(s)$$

(where $\mathbf{1}$ is the identity map on \mathbb{C}) then $S_0(s)$ acts on $C^\infty(\partial_\infty X)$ and the relative scattering operator

$$S_{\text{rel}}(s) = S_X(s) S_0(s)^{-1}$$

is determinant class. Thus one has a relative scattering determinant:

$$\tau_X(s) = \det(S_{\text{rel}}(s)).$$

In [10], Guillopé and Zworski show that $\tau_X(s)$ extends to the quotient of analytic functions of order at most four and compute its divisor in terms of the resonance set. Our proof of Theorem 1.1 in §4 involves establishing the connection between $Z_X(s)$ and $\tau_X(s)$, and then exploiting this information about $\tau_X(s)$. In the course of the proof we will improve the order bound on $\tau_X(s)$ from four to two.

Acknowledgments. We are grateful to Reza Chamanara and Jeffrey McGowan for helpful conversations, and to Martin Olbrich for point out to us his results in [24]. We also thank Peter Sarnak for encouraging us to remove the unnecessary geometric assumptions used in [1].

2. Regularized traces and model problems

We start with X a geometrically finite hyperbolic surface decomposed into a compact core Z plus cusps C_i and funnels F_j as in (1.1). The cusp and funnel ends are isometric to standard models which we now describe. Here and in what follows, we use the upper half-plane model of hyperbolic space: $\mathbb{H} = \{z = x + iy \mid y > 0\}$, with metric

$$ds^2 = y^{-2}(dx^2 + dy^2).$$

For $\ell > 0$, denote by Γ_ℓ the abelian discrete group generated by the hyperbolic isometry $z \mapsto e^\ell \cdot z$. The quotient $\Gamma_\ell \backslash \mathbb{H}$ is the *hyperbolic cylinder* with a single

primitive closed geodesic of length ℓ . It is well-known and easy to see that this cylinder is isometric to $\mathbb{R} \times \mathbb{R}/\mathbb{Z}$ equipped with the metric

$$dr^2 + \ell^2 \cdot \cosh^2(r) dx^2. \quad (2.1)$$

A *funnel* is the subspace $[0, \infty) \times \mathbb{R}/\mathbb{Z}$ equipped with this metric. In particular, the boundary, ∂F , of F is a geodesic of length ℓ .

Let Γ_∞ the abelian discrete group generated by the parabolic isometry $z \mapsto z + 1$. For us the *horn* is the quotient

$$H = \Gamma_\infty \backslash \mathbb{H}.$$

It is well-known and easy to see that the horn is isometric to $\mathbb{R} \times \mathbb{R}/\mathbb{Z}$ equipped with the metric

$$dr^2 + e^{-2r} dx^2.$$

A *cuspidal end* is a subspace of the form $[a, \infty) \times \mathbb{R}/\mathbb{Z}$ equipped with this metric. A level curve of the coordinate r is called a *horocycle*. In particular, a cuspidal end is bounded by a horocycle of $r = a$. Note that in the decomposition of (1.1) the parameter a can be chosen to be 0 independent of the surface. This follows from a collar type lemma for cusps to be found, for example, in [5].

We shall regularize integrals over X by introducing a ‘defining function’ ρ for the boundary at infinity, namely, choose $\rho \in C^\infty(X)$ so that in each of the cusp and funnel ends $\rho = e^{-r}$ with respect to the geodesic coordinate systems introduced above. With this choice we can define a regularized integral: Suppose $f \in C^\infty(X)$ is *polyhomogeneous* in ρ , meaning that it has an asymptotic expansion as $\rho \rightarrow 0$ in powers of ρ and $\log \rho$, with smooth coefficients. (Actually, in our applications below the expansions involve only powers of ρ , with no $\log \rho$ terms.) We then define

$$\int_X^0 f dg = \text{FP}_{\varepsilon \downarrow 0} \int_{\{\rho \geq \varepsilon\}} f dg,$$

where FP denotes the finite part, meaning that the limit is taken after the divergent terms in the asymptotic expansion as $\varepsilon \rightarrow 0$ are subtracted off. The 0-volume of a region is correspondingly defined as the 0-integral of dg . These regularizations would depend on the choice of ρ in general. Our definition exploits the canonical choice of ρ made possible for hyperbolic surfaces by the standard models for the ends.

Lemma 2.1. *With the 0-integral as defined above,*

$$0\text{-vol}(X) = -2\pi \cdot \chi,$$

where χ is the Euler characteristic of X .

Proof. The area of the region $\{\rho \geq \varepsilon\} \subset F$ is $\ell(\varepsilon^{-1} - \varepsilon)/2$, so $0\text{-vol}(F) = 0$. Hence $0\text{-vol}(X) = \text{vol}(X \setminus Y)$ and the result follows from the Gauss–Bonnet theorem, since the funnels have geodesic boundary. \square

For a smoothing operator T defined on $X \times X$, with continuous kernel $K_T(x, y)$ with respect to Riemannian measure on X , the 0-trace of T is defined by

$$0\text{-tr } T = \int_X^0 K_T(z, z) dg(z).$$

For this to be well-defined, the function $K_T(z, z)$ must be polyhomogeneous as a function of ρ , which will be the case for the operators we consider.

There are two regularized traces formed from the resolvent that will be our main tools in this paper. Let $G_X(s; z, z')$ denote the Green’s function (the integral kernel of the resolvent operator $R_X(s) = (\Delta_X - s(1 - s))^{-1}$). One way to cancel the kernel’s singularity on the diagonal so as to produce a trace is to subtract off the Green’s function for the model hyperbolic space, which has the same diagonal singularity. This technique was used prominently by Patterson [25]. By lifting G_X to $\mathbb{H} \times \mathbb{H}$, we can define

$$\varphi_X(s; z) = (2s - 1)[G_X(s; z, w) - G_{\mathbb{H}}(s; z, w)]_{w=z},$$

as a function on \mathbb{H} (meromorphic in s). By the invariance properties of the model hyperbolic Green’s function it is easy to see that $\varphi_X(s)$ descends to define a smooth function on X . This function is known to be polyhomogeneous in ρ by the general theory of [19], [18]. Or, more directly, one can deduce this from Guillopé’s parametrix construction in §1 of [7]. This construction shows in particular that in the funnel ends,

$$\varphi_X(s)|_{F_j} \in \rho^{2s} C^\infty(\bar{F}_j), \tag{2.2}$$

where \bar{F}_j is the compactification of F_j to the cylinder $[0, 1]_\rho \times (\mathbb{R}/\mathbb{Z})_x$. In the cusps,

$$\varphi_X(s)|_{C_j} \in \rho^{2s-2} C^\infty([0, 1]_\rho) + \rho^{-1} C^\infty([0, 1]_\rho) + O(\rho^\infty). \tag{2.3}$$

(Coefficients of the asymptotic expansions in the cusps do not depend on the x variable.)

It follows from (2.2) and (2.3) that, as $\varepsilon \rightarrow 0$,

$$\int_{\{\rho > \varepsilon\}} \varphi_X(s; z) dg(z) \sim a(s) + b_0(s) \log \varepsilon + \sum_{k=1}^\infty b_k(s) \varepsilon^k + \sum_{l=0}^\infty c_l(s) \varepsilon^{2s-1+l},$$

and all of the coefficients are meromorphic by the meromorphy of the resolvents. For $s \neq 1/2 - \mathbb{N}_0/2$ (where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$) we define the function

$$\Phi_X(s) = \int_X^0 \varphi_X(s; z) dg(z) = a(s), \tag{2.4}$$

which extends meromorphically to the whole plane because $a(s)$ is meromorphic. (The formula (2.4) does not apply at the points $s \in 1/2 - \mathbb{N}_0/2$ because the coefficient c_l would also contribute to the 0-integral for $s = (1 - l)/2$.)

Another function of interest is the renormalized trace of the spectral density:

$$\Upsilon_X(s) = (2s - 1) 0\text{-tr}[R_X(s) - R_X(1 - s)]. \tag{2.5}$$

This 0-trace is well defined because the singularities of $G_X(s, z, w)$ and $G_X(1 - s, z, w)$ on the diagonal cancel (and again because these kernels are polyhomogeneous in ρ). By examining the boundary asymptotics as above, we can see that $\Upsilon_X(s)$ also extends meromorphically to the whole plane, although the function is only given by the formula (2.5) for $s \notin \frac{1}{2}\mathbb{Z}$.

As one would expect, $\Phi_X(s)$ and $\Upsilon_X(s)$ are closely related.

Proposition 2.2. *With Φ and Υ as defined above,*

$$\Phi_X(s) + \Phi_X(1 - s) = \Upsilon_X(s) + (2s - 1)\pi \cdot \chi(X) \cot(\pi s). \tag{2.6}$$

Proof. For $s \notin \frac{1}{2}\mathbb{Z}$, we have

$$\begin{aligned} \Phi_X(s) + \Phi_X(1 - s) &= (2s - 1) \int_X^0 [G_X(s; z, w) - G_X(1 - s; z, w) \\ &\quad - G_{\mathbb{H}}(s; z, w) + G_{\mathbb{H}}(1 - s; z, w)]_{w=z} dg(z). \end{aligned}$$

The difference $G_{\mathbb{H}}(s; z, w) - G_{\mathbb{H}}(1 - s; z, w)$ is continuous, and its restriction to the diagonal is necessarily a constant since these kernels depend only on $d(z, w)$. This constant is shown to be $(1/2) \cot \pi s$ in (A.3), and the result then follows from Lemma 2.1. □

The Selberg zeta function for the cylinder $M = \Gamma_\ell \backslash \mathbb{H}$ is

$$Z_M(s) = \prod_{k \geq 0} (1 - e^{-(s+k)\ell})^2.$$

In Proposition 3.3 of [25] Patterson computed

$$\Phi_M(s) = \frac{Z'_M}{Z_M}(s). \tag{2.7}$$

where $\Phi_M(s)$ is the function defined by (2.4).

Now let Δ_F be the Laplacian on F with Dirichlet boundary conditions on the single closed geodesic bounding F . We shall define

$$Z_F(s) = e^{-s\ell/4} \prod_{k \geq 0} (1 - e^{-(s+2k+1)\ell})^2, \tag{2.8}$$

where ℓ is the length of the geodesic boundary of F .

Technically, $\Phi_X(s)$ was defined by (2.4) only for X a complete manifold. To extend this definition to the manifold with boundary, F , we simply apply the same definition using $G_F(s)$ the Green's function for Δ_F and taking a 0-integral over F (with regularization needed only at the open end). Together with the definition (2.8), this allows us to make a statement analogous to (2.7):

Proposition 2.3. *For $\text{Re}(s) > 1/2$,*

$$\Phi_F(s) = \frac{Z'_F(s)}{Z_F(s)}. \tag{2.9}$$

The proof will be deferred to Appendix A.

Since $\chi(F) = 0$, by Proposition 2.2 we have

$$\Upsilon_F(s) = \frac{d}{ds} \log \left(\frac{Z_F(s)}{Z_F(1-s)} \right). \tag{2.10}$$

By Fourier analysis in the x variable (in the coordinates (2.1)), one can write Δ_F as an infinite direct sum of ordinary differential operators on the half-line; see [8] for details. This leads to the following explicit computation of the resonances of Δ_F as the set

$$\left\{ \zeta_{k,m} = -(2k+1) + \frac{2\pi im}{\ell} \text{ (with multiplicity 2) } \mid k \in \mathbb{N}_0, m \in \mathbb{Z} \right\}, \tag{2.11}$$

where $\ell = \ell(\partial F)$. We can form a convergent Hadamard product whose zero set is (2.11):

$$P_F(s) = \prod_{k=0}^{\infty} \prod_{m=-\infty}^{\infty} (1 - s/\zeta_{k,m})^2 e^{(2s/\zeta_{k,m} + (s/\zeta_{k,m})^2)}, \tag{2.12}$$

defining an entire function of order two. Note that (2.11) is also the zero set of Z_F defined in (2.8), and hence

$$Z_F(s) = e^{p(s)} \cdot P_F(s) \tag{2.13}$$

for a polynomial p of order at most two whose coefficients depend only on ℓ .

The calculation of $\Phi_H(s)$ for the horn H is another special function calculation which we will defer to Appendix A. The result is:

Proposition 2.4. For $H = \Gamma_\infty \backslash \mathbb{H}$ we have

$$\Phi_H(s) = -\log 2 - \Psi(s + 1/2) + \frac{1}{2s - 1}, \tag{2.14}$$

where $\Psi(z)$ is the digamma function $\Gamma'(z)/\Gamma(z)$.

Because $\chi(H) = 0$, Propositions 2.2 and 2.4 immediately yield:

Corollary 2.5. For $H = \Gamma_\infty \backslash \mathbb{H}$ we have

$$\Upsilon_H(s) = -\frac{d}{ds} \log \left[4^s \frac{\Gamma(s - 1/2)}{\Gamma(1/2 - s)} \right].$$

3. Scattering theory

As above, X denotes a geometrically finite hyperbolic surface. We assume that X has infinite area ($n_F \geq 1$), which implies that the positive Laplacian, denoted Δ_X , has at most finitely many eigenvalues in the interval $[0, 1/4]$ and purely continuous spectrum in $[1/4, \infty)$ (no embedded eigenvalues). The resolvent $R_X(s) = (\Delta_X - s(1 - s))^{-1}$, initially defined as a meromorphic $L^2(X)$ -operator-valued function in the half-plane $\text{Re}(s) > 1/2$, is known to extend, as an operator from $C_0^\infty(X)$ to $C^\infty(X)$, to a meromorphic function in \mathbb{C} (see [19], [9] if $n_C = 0$ and [10] if $n_C \neq 0$). A pole of $R_X(s)$ is called a *resonance*. We let \mathcal{R}_X denote the set of resonances, counted with multiplicity m_ζ given for $\zeta \neq 1/2$ by the rank of the residue of $R_X(s)$:

$$m_\zeta = \text{rank} \int_{\gamma_\zeta} (2s - 1)R_X(s) ds,$$

where γ_ζ is a simple closed contour surrounding ζ and no other resonance of Δ_X . (When $\zeta(1 - \zeta)$ is an eigenvalue, m_ζ is equal to the multiplicity.)

The point $\zeta = 1/2$ is exceptional because $R_X(s)$ has possibly a second order singularity there. For $\gamma_{1/2}$ a contour surrounding $1/2$ and no other resonance, let

$$a_{1/2} = \text{rank} \int_{\gamma_{1/2}} (2s - 1)R_X(s) ds,$$

and

$$b_{1/2} = \text{rank} \int_{\gamma_{1/2}} R_X(s) ds.$$

Lemma 4.1 of [10] shows that $a_{1/2}$ is the multiplicity of $1/4$ as an L^2 -eigenvalue of Δ_X (which is zero if $n_C = 0$). To get the proper zero of the zeta function we set

$$m_{1/2} = 2a_{1/2} + b_{1/2}.$$

See the proof of Proposition 4.5 for the justification of this.

Let us recall some of the basics of scattering theory in our context; see [10] for details and complete proofs. Note that in [10] the basic objects of scattering theory – the resolvent, generalized eigenfunctions, and scattering operator – are defined as sections of certain bundles over X ; we in effect trivialize these bundles by our explicit choice of defining function for the boundary of X (see the remarks at the beginning of §2).

First we define the scattering operator $S_X(s)$ as a map from $C^\infty(\partial_\infty X) = \mathbb{C}^{n_C} \oplus (\bigoplus_{j=1}^{n_F} C^\infty(S^1))$ to itself. The kernel of $S_X(s)$ could be defined purely algebraically (as an average over Γ where $X = \Gamma \backslash \mathbb{H}$). But a more intuitive definition, following the general philosophy explained for example in [21], is based on generalized eigenfunctions:

Proposition 3.1. *Given $f = \{z_1, \dots, z_{n_C}, f_1, \dots, f_{n_F}\} \in C^\infty(\partial_\infty X)$, $s \notin \mathbb{Z}/2$, there is a unique smooth solution u of the eigenvalue problem*

$$(\Delta_X - s(1-s))u = 0$$

with

$$u \sim \rho^s \tilde{f} + \rho^{1-s} \tilde{g},$$

where $\tilde{f}, \tilde{g} \in C^\infty(X)$ are polyhomogeneous of degree zero in ρ , and $\lim_{\rho \rightarrow 0} \tilde{f} = f$.

See [13] for a proof which covers our case $n_C = 0$; the generalization to $n_C > 0$ is straightforward using the parametrix construction from [7]. From this proposition we obtain a linear map $S_X(s): f \mapsto g = \lim_{\rho \rightarrow 0} \tilde{g}$. It follows from the definition that $S_X(s)S_X(1-s)$ is the identity operator.

In what follows, we write i or i' for a cusp index $1 \leq i, i' \leq n_C$ and j or j' for a funnel index $1 \leq j, j' \leq n_F$. We denote by $S_X^{i i'}(s)$ the component of $S_X(s)$ mapping $\partial_\infty C_{i'}$ to $\partial_\infty C_i$, etc. From the detailed description of $S_X(s)$ in [10] we extract:

Proposition 3.2. *The map $S_X(s)$ extends to a meromorphic operator-valued function with the following properties:*

- (1) $\{S_X^{i i'}(s) \mid 1 \leq i, i' \leq n_C\}$ is a matrix-valued meromorphic function of s ,
- (2) $S_X^{j j'}(s)$ is a smoothing operator if $j \neq j'$, and $S_X^{j j}(s) - S_{F_j}(s)$ extends to a meromorphic family of smoothing operators on $C^\infty(S^1)$, and
- (3) $S_X^{i j}(s)$ and $S_X^{j i}(s)$ have integral kernels in $C^\infty(S^1)$ which are meromorphic functions of s .

The poles of the scattering operator can be deduced from the resonance set \mathcal{R}_X , as shown in §2 of [10]. (They coincide precisely except for infinite rank poles of $S_X(s)$)

at positive integer points, and possibly finitely many points related to the discrete spectrum.)

In [8], Guillopé and Zworski proved that number of resonances (including multiplicities) in the ball $|\zeta| < r$ grows at most quadratically in r :

$$\sum_{\zeta \in \mathcal{R}_X, |\zeta| < r} m_\zeta = O(r^2).$$

Hence the Hadamard product

$$P_X(s) = s^{m_0} \prod_{\zeta \in \mathcal{R}_X} (1 - s/\zeta)^{m_\zeta} e^{m_\zeta(s/\zeta + s^2/2\zeta^2)} \tag{3.1}$$

converges.

We recall from [10] that the relative scattering determinant $\tau_X(s) = \det(S_{\text{rel}}(s))$, defined in §1, is determined up to finitely many parameters by the lengths ℓ_j of geodesics bounding the F_j together with the resonance set \mathcal{R}_X . To state this result more precisely, recall the definition (2.12) of $P_F(s)$ and define

$$P_Y(s) = \prod_{j=1}^{n_F} P_{F_j}(s).$$

Proposition 3.3 (Proposition 3.4 of [10]). *The function τ_X extends to a meromorphic function in the complex plane having the form*

$$\tau_X(s) = e^{h(s)} \frac{P_X(1-s)}{P_X(s)} \frac{P_Y(s)}{P_Y(1-s)},$$

where $h(s)$ is a polynomial of degree at most four.

(We will find that $h(s)$ has degree at most two at the end of §4.)

Let $R_X(s) = (\Delta_X - s(1-s))^{-1}$ and let $R_Y(s) = (\Delta_Y - s(1-s))^{-1}$, where Y is the union of the funnels as before with Dirichlet conditions on ∂Y . Note that the operators $R_X(s) - R_X(1-s)$ and $R_Y(s) - R_Y(1-s)$ actually have smooth kernels although with insufficient decay at infinity to be trace-class. There is a natural identification of Y with the corresponding submanifold of X and we denote by $\mathbf{1}_Y$ the obvious induced mapping from $C^\infty(X)$ onto $C^\infty(Y)$. We set

$$Q_X(s) = R_X(s) - \mathbf{1}_Y^t \circ R_Y(s) \circ \mathbf{1}_Y. \tag{3.2}$$

Proposition 3.4 (Equation (4.14) of [10]²). *For $\text{Re}(s) = \frac{1}{2}$, $s \neq 1/2$,*

$$(2s - 1) \cdot 0\text{-tr} [Q_X(s) - Q_X(1-s)] = -\frac{\tau'_X}{\tau_X}(s). \tag{3.3}$$

²We believe there to be a typo in formula (4.14) of [10]: the terms on the right-hand side of their equation should carry negative signs.

In the original formula in [10], the trace in (3.3) is interpreted as a distribution on the line $\operatorname{Re}(s) = 1/2$ and the right-hand side includes a delta-function singularity at $s = 1/2$. The singularity occurs for reasons discussed in §2; the same issue is behind our restriction of formula (2.5) to $s \notin \frac{1}{2}\mathbb{Z}$.

4. Selberg's zeta function

Let $X = \Gamma \backslash \mathbb{H}$ be a geometrically finite hyperbolic surface. Recall that Selberg's zeta function, $Z_X(s)$, is defined for $\operatorname{Re}(s) > 1$ by

$$Z_X(s) = \prod_{\{\gamma\}} \prod_{k=0}^{\infty} (1 - e^{-(s+k)\ell(\gamma)}). \quad (4.1)$$

where the outer product goes over conjugacy classes of primitive hyperbolic elements of Γ , and $\ell(\gamma)$ is the length of the corresponding closed geodesic. Following [27], we define

$$Z_{\infty}^{\chi}(s) = \left[\frac{(2\pi)^s \Gamma_2(s)^2}{\Gamma(s)} \right]^{-\chi}.$$

where $\Gamma_2(s)$ is Barnes' double Gamma function and $\chi = \chi(X)$ is the Euler characteristic of X . The function $Z_{\infty}^{\chi}(s)$ has no zeros, and poles of order $(2k+1)(-\chi)$ at $s = -k$, $k \in \mathbb{N}_0$.

The purpose to this section is to prove the following result.

Theorem 4.1. *The Selberg zeta function has a factorization*

$$Z_X(s) \cdot Z_{\infty}^{\chi}(s) = e^{q(s)} \cdot \Gamma(s - 1/2)^{n_C} \cdot P_X(s), \quad (4.2)$$

where $q(s)$ is a polynomial of degree ≤ 2 and $P_X(s)$ is the Hadamard product over resonances (3.1).

Theorem 1.1 is an immediate consequence of Theorem 4.1. But before taking up the proof of Theorem 4.1, we will show how Corollary 1.2 follows.

Proof of Corollary 1.2. Suppose that the resonance set is fixed, which determines $P_X(s)$. We claim that $P_X(s)$ determines χ (hence Z_{∞}^{χ}), n_C , and $q(s)$, and therefore fixes $Z_X(s)$, from which the length spectrum may be deduced by a standard argument. To see this we take the log of (4.2) and analyze the asymptotics as $\operatorname{Re}(s) \rightarrow \infty$. On the left-hand side, $\log Z_X(s)$ decays exponentially, while

$$\frac{1}{\chi} \log Z_{\infty}^{\chi}(s) \sim \frac{2 + \log 2\pi}{2} - 2\zeta'(-1) - \left(\frac{1}{2}s(s-1) - \frac{1}{6} \right) \log s(s-1) + \frac{3}{2}s(s-1).$$

By Stirling's formula,

$$\log \Gamma(s - 1/2) \sim \frac{1 + \log 2\pi}{2} + s \log(s - 1/2) - s.$$

This implies that $\log P_X(s)$ has an asymptotic expansion as $\text{Re}(s) \rightarrow \infty$. The value of χ may be read off from the $s^2 \log s$ coefficient. Then, after subtracting off the $\log Z_\infty^\chi(s)$ term, the $s \log s$ coefficient determines n_C . Once the n_C terms are removed, what remains in the asymptotic expansion is precisely $q(s)$, so this is also determined by $P_X(s)$.

Now assume that the length spectrum is known, giving $Z_X(s)$. The order of the zero of $Z_X(s)$ at $s = -k, k \in \mathbb{N}_0$, is $m_{-k} + (2k + 1)(-\chi)$. This gives us for any k a bound

$$0 \geq \chi \geq \frac{-1}{2k + 1} \text{ord}_{s=-k} Z_X(s),$$

implying that only finitely many values of χ are possible. And hence finitely many choices of n_C , since $n_C \leq 2 - \chi$. If n_C and χ were fixed, then $Z_X(s)$ would determine the divisor of $P_X(s)$ and the resonance set could be read off directly. \square

Remark 4.2. Corollary 1.2 could also have been deduced from the wave trace formula proved by Guillopé and Zworski in [11]. The combination of Theorem 1 and Equation (3.1) of that paper yields the following trace formula. As distributions on \mathbb{R}_+ ,

$$\begin{aligned} \sum_{\zeta \in \mathcal{R}_X} m_\zeta e^{(\zeta - 1/2)t} &= \sum_{\{\gamma\}} \sum_{k=1}^\infty \frac{\ell(\gamma)}{\sinh(k\ell(\gamma)/2)} \delta(t - k\ell(\gamma)) \\ &\quad + \chi \frac{\cosh(t/2)}{2 \sinh^2(t/2)} + \frac{n_C}{2} \coth(t/2), \end{aligned}$$

where the $\{\gamma\}$ denotes a list of conjugacy classes of primitive hyperbolic elements of Γ , as in the definition of the zeta function. Note that

$$\frac{\cosh(t/2)}{2 \sinh^2(t/2)} = \sum_{k=0}^\infty (2k + 1) e^{(-k - 1/2)t},$$

for $t > 0$. The problem of overlap between resonances and topological zeros would not be avoided by this alternate route.

Proposition 4.3. *The logarithmic derivative of the zeta function is given by*

$$\frac{Z'_X}{Z_X}(s) = \Phi_X(s) - n_C \Phi_H(s). \tag{4.3}$$

Proof. Since both sides of (4.3) are known to be meromorphic, it suffices to consider $\text{Re}(s) > 1$.

For $\text{Re}(s) > 1$, the Green’s function $G_X(s; z, w)$ can be written as a convergent sum

$$G_X(s; z, w) = G_{\mathbb{H}}(s; z, w) + \sum_{\gamma \in \Pi_h} \sum_{g \in \Gamma / \langle \gamma \rangle} \sum_{k \neq 0} G_{\mathbb{H}}(s; z, g^{-1} \gamma^k g w) + \sum_{\gamma \in \Pi_p} \sum_{g \in \Gamma / \langle \gamma \rangle} \sum_{k \neq 0} G_{\mathbb{H}}(s; z, g^{-1} \gamma^k g w),$$

where Π_h and Π_p are lists of representatives of conjugacy classes of maximal hyperbolic and parabolic subgroups of Γ , respectively. It follows that

$$\Phi_X(s) = (2s - 1) \int_X [G_X(s; z, w) - G_{\mathbb{H}}(s; z, w)]_{z=w} dg(z)$$

into corresponding sums over Π_h and Π_p . For any $\gamma \in \Gamma$, the standard trace formula technique for summing over a conjugacy class gives:

$$\begin{aligned} \sum_{g \in \Gamma / \langle \gamma \rangle} \sum_{k \neq 0} \int_X G_{\mathbb{H}}(s; z, g^{-1} \gamma^k g z) dg(z) &= \int_{\langle \gamma \rangle \backslash \mathbb{H}} [G_{\langle \gamma \rangle \backslash \mathbb{H}}(s; z, w) - G_{\mathbb{H}}(s; z, w)]_{z=w} dg(z) \\ &= \Phi_{\langle \gamma \rangle \backslash \mathbb{H}}. \end{aligned}$$

If γ is hyperbolic, then by Patterson’s formula (2.7) this is equal to $(\log Z_{\langle \gamma \rangle \backslash \mathbb{H}})'(s)$. And then since, by definition,

$$Z_X(s) = \prod_{\gamma \in \Pi_h} Z_{\langle \gamma \rangle \backslash \mathbb{H}}(s),$$

summing these hyperbolic terms over Π_h gives the logarithmic derivative of $Z_X(s)$.

For any γ parabolic, the quotient $\langle \gamma \rangle \backslash \mathbb{H}$ is isometric to the model horn H from §2, so the contribution from each cusp is $\Phi_H(s)$. \square

Since Y is a direct sum of the funnels F_j , we set

$$Z_Y(s) = \prod_{j=1}^m Z_{F_j}(s) = \prod_{j=1}^m \left[e^{-s \ell_j / 4} \prod_{k \geq 0} (1 - e^{-(s+2k+1)\ell_j})^2 \right],$$

Proposition 4.4.³ For some constant c ,

$$\tau_X(s) = e^{c - (n_C \log 4)s} \frac{Z_X(1-s)}{Z_X(s)} \frac{Z_{\infty}^X(1-s)}{Z_{\infty}^X(s)} \frac{Z_Y(s)}{Z_Y(1-s)} \left(\frac{\Gamma(s-1/2)}{\Gamma(1/2-s)} \right)^{n_C}.$$

³In the case $n_C = 0$, such a formula was noted in [10], equation (5.3).

Proof. From Proposition 3.4 and (2.5) we find that for $s \notin \frac{1}{2}\mathbb{Z}$

$$-\frac{\tau'_X(s)}{\tau_X(s)} = \Upsilon_X(s) - (2s - 1) 0\text{-tr}_X[\mathbf{1}_Y^t \circ (R_Y(s) - R_Y(1 - s)) \circ \mathbf{1}_Y].$$

Since R_Y breaks up into a direct sum of R_{F_j} on the individual funnels, the second term on the right-hand side can be reduced to

$$(2s - 1) 0\text{-tr}_X[\mathbf{1}_Y^t \circ (R_Y(s) - R_Y(1 - s)) \mathbf{1}_Y] = \sum_{j=1}^{n_F} \Upsilon_{F_j}(s),$$

by the definition of $\Upsilon_{F_j}(s)$. From (2.10) we find

$$\sum_{j=1}^{n_F} \Upsilon_{F_j}(s) = \frac{d}{ds} \log \left(\frac{Z_Y(s)}{Z_Y(1 - s)} \right). \tag{4.4}$$

Given the computation of $\Upsilon_H(s)$ in Corollary 2.5, to prove the proposition it will suffice to show that

$$\Upsilon_X(s) = \frac{d}{ds} \log \left(\frac{Z_X(s)}{Z_X(1 - s)} \frac{Z_\infty^\chi(s)}{Z_\infty^\chi(1 - s)} \right) + n_C \Upsilon_H(s). \tag{4.5}$$

To deduce (4.5), one can use the product expansion of $\Gamma_2(s)$ to derive the identity

$$\frac{d}{ds} \log Z_\infty^\chi(s) = \chi \cdot (2s - 1) \cdot (\Psi(s) - 1). \tag{4.6}$$

where $\Psi(s) = \Gamma'(s)/\Gamma(s)$. From the functional equation $\Gamma(s) \cdot \Gamma(1 - s) = \pi \csc(\pi s)$ it then follows that

$$\frac{d}{ds} \log \left(\frac{Z_\infty^\chi(s)}{Z_\infty^\chi(1 - s)} \right) = -\chi \cdot (2s - 1) \cdot \pi \cdot \cot(\pi s). \tag{4.7}$$

Equation (4.5) follows by combining Equation (4.7), Proposition 2.2, and Proposition 4.3. □

Proposition 4.5. *The meromorphic function $Z'_X(s)/Z_X(s)$ is analytic in $\text{Re}(s) > 1/2$ except for first-order poles at points ζ where $\zeta(1 - \zeta)$ is an eigenvalue (with residues equal to the eigenvalue multiplicities). Also, $Z'_X(s)/Z_X(s)$ has no poles on $\text{Re}(s) = 1/2$ except possibly at $s = 1/2$. The residue of $Z'_X(s)/Z_X(s)$ at $s = 1/2$ is $m_{1/2} - n_C$.*

Proof. Except at $s = 1/2$ the argument is essentially the same as in Theorem 6.2 of [26]. Recall the integrand $\varphi_X(s; w)$ appearing in the definition (2.4) of $\Phi_X(s)$.

The poles of $\varphi_X(s)$ for $\text{Re}(s) \geq 1/2$, $s \neq 1/2$, and its behavior near those poles can be deduced from Lemmas 2.4 of [10]. In particular, in $\text{Re}(s) > 1/2$, $\varphi_X(s)$ is meromorphic with poles only at points $\zeta \in (1/2, 1)$ for which $\zeta(1 - \zeta) \in \sigma(\Delta_X)$. Near such a point ζ it has the form

$$\varphi_X(s) = \frac{1}{s - \zeta} \sum_{k=1}^{m_\zeta} \psi_k^2 + \tilde{\varphi}(s),$$

where $\{\psi_k\}$ form an orthonormal basis of the eigenspace for $\zeta(1 - \zeta)$, and $\tilde{\varphi}(s)$ is holomorphic near $s = \zeta$. Since $\tilde{\varphi}(s)$ will be integrable for $\text{Re}(s) > 1/2$, from Proposition 4.3 we have near $s = \zeta$:

$$\frac{Z'_X}{Z_X}(s) = \frac{m_\zeta}{s - \zeta} + h(s).$$

where h is holomorphic.

For $\text{Re}(s) = 1/2$, $s \neq 1/2$, the function $\varphi_X(s)$ is holomorphic (there are no embedded eigenvalues as $n_F > 0$), but it is not integrable. Nonetheless, from the behavior of $\varphi_X(s)$ at infinity (see (2.2) and (2.3)) it follows that the 0-integral of $\varphi_X(s)$ vanishes for this case.

This leaves finally the point $s = 1/2$. By Lemma 4.1 of [10] the structure of $R_X(s)$ near $s = 1/2$ is

$$R_X(s) = \frac{A}{(2s - 1)^2} + \frac{B}{(2s - 1)} + C(s), \tag{4.8}$$

where $C(s)$ is analytic near $s = 1/2$. Our definitions in §3 set $a_{1/2} = \text{rank } A$, $b_{1/2} = \text{rank } B$, and $m_{1/2} = 2a_{1/2} + b_{1/2}$. Substituting (4.8) into the definition of φ_X shows that

$$\varphi_X(s; z) = \frac{a(z)}{2s - 1} + b(s; z),$$

where $b(s; z)$ is analytic near $s = 1/2$.

Suppose u is an eigenvalue of Δ_X with eigenvalue $1/4$. Applying $R_X(s)$ to

$$(\Delta_X - s(1 - s))u = (s - 1/2)^2 u$$

shows that $u = \frac{1}{4}Au$. Hence $A/4$ is the projector onto the $1/4$ -eigenspace of Δ_X . The contribution of the a term to the residue of $\Phi_X(s)$ at $s = 1/2$ is

$$\frac{1}{2} \int_X a \, dg = \frac{1}{2} \text{tr } A = 2a_{1/2}.$$

Although $b(s; z)$ is analytic near $s = 1/2$, its 0-integral does contribute to the pole in $\Phi_X(s)$ because of its boundary asymptotics. In a funnel F_j , $b(s; \cdot)|_{F_j} \in \rho^{2s} C^\infty(\bar{F}_j)$

and the contribution can be deduced exactly as in Theorem 6.2 of [26]. Namely, the 0-integral of $b|_{F_j}$ contributes a residue of $\frac{1}{2} \text{tr}[S_X^{jj}(1/2) + I]$. (For comparison of formulas, note that the normalized scattering matrix $\mathfrak{S}_X(s)$ used in [26] is related to our definition by $\mathfrak{S}_X(1/2) = -S_X(1/2)$.)

The contribution of a cusp is slightly different. Consider cusp C_i and introduce a cutoff η which is supported in C_i and equal to 1 in a neighborhood of $\rho = 0$. The contribution from cusp C_j is

$$\text{Res}_{s=1/2} \int_{C_j}^0 \eta(z) \varphi_X(s; z) dg(z)$$

(which will independent of η .) Identifying C_j with the cusp end of H , and using the parametrix construction in [7] we can identify

$$\varphi_X(s; z)|_{C_i} = \varphi_H(s; z) + \rho^{2s-2} k(s; z),$$

where, under our convention for $S_X(s)$, $\lim_{\rho \rightarrow 0} k(s, z) = S_X^{ii}(s)$. The contribution to the residue from the $k(s; z)$ term is then easily computed to be $\frac{1}{2} S_X^{ii}(1/2)$. Lemma A.1 shows that

$$\text{Res}_{s=1/2} \int_{C_j}^0 \eta(z) \varphi_H(s; z) dg(z) = 0,$$

so the $\varphi_H(s; z)$ term does not contribute to the residue.

The full contribution of the b term to the residue of $\Phi_X(s)$ at $s = 1/2$ is thus $(1/2) \text{tr}[S_X^{FF}(s) + I] + (1/2) \text{tr}[S_X^{CC}(s)]$. By adding and subtracting $\frac{1}{2} \text{tr} I^{CC} = n_C/2$ we can write this as $\text{tr}[S_X(1/2) + I] - n_C/2$. By Lemma 4.3 of [10] (with our convention for $S_X(s)$),

$$\frac{1}{2} \text{tr}[S_X(1/2) + I] = b_{1/2}.$$

Thus the contribution of the $b(s; z)$ term to the residue is $b_{1/2} - n_C/2$.

Finally, combining the a and b contributions shows that

$$\text{Res}_{s=1/2} \Phi_X(s) = 2a_{1/2} + b_{1/2} - n_C/2.$$

And the stated result follows because $\text{Res}_{s=1/2} \Phi_H(s) = 1/2$. □

Proof of Theorem 4.1. Combining Propositions 3.3 and 3 and Equation (2.13) yields

$$\frac{Z_X(s) Z_\infty^X(s)}{Z_X(1-s) Z_\infty^X(1-s)} = e^{h_1(s)} \frac{P_X(s)}{P_X(1-s)} \cdot \left(\frac{\Gamma(s-1/2)}{\Gamma(1/2-s)} \right)^{n_C} \tag{4.9}$$

for a polynomial h_1 of order at most four.

Now let

$$G(s) = \frac{Z_X(s)Z_\infty^\chi(s)}{P_X(s)\Gamma(s - 1/2)^{nc}}.$$

Proposition 4.5 says that $G(s)$ has no zeroes or poles in $\text{Re}(s) \geq 1/2$. And (4.9) says $G(s) = e^{h_1(s)}G(1 - s)$, implying that G has no zeroes or poles in $\text{Re}(s) \leq 1/2$ either. Hence G is the exponential of an entire function $q(s)$.

It remains to show that $q(s)$ is a polynomial with degree equal to 2. From the convergent Euler product (4.1) and elementary estimates on the counting function for lengths of closed geodesics, we can easily show that $Z_X(s)$ is bounded in the half-plane $\text{Re}(s) \geq 2$. Since $P_X(s)$ and $1/Z_\infty^\chi(s)$ are entire of order 2 and $1/\Gamma(s - 1/2)$ is order 1, we conclude that $|q(s)| \leq C_\kappa |s|^{2+\kappa}$ for $\kappa > 0$ and $\text{Re}(s) \geq 2$. The functional relation 4.9 and the fact that h_1 is order 4 then imply a bound $|q(s)| \leq C|s|^4$ for $\text{Re}(s) \leq -1$.

To control $q(s)$ within the strip $S_\varepsilon = \{-1 - \varepsilon < \text{Re}(s) < 2 + \varepsilon\}$, we first consider the function

$$F_X(s) = 0\text{-tr}[Q_X(s) - Q_X(s_0)]$$

for s_0 fixed and not a pole of $Q_X(s)$, where Q_X was defined in (3.2). Introduce, as in §5 of [1], disks D_j such that $\mathcal{R}_X \subset \cup_j D_j$ and $d(s, \mathcal{R}_X) \geq C\langle s \rangle^{-2-\delta}$ for all $s \in \mathbb{C} \setminus (\cup_j D_j)$. Using the parametrix construction of [10] and arguing exactly as in Appendix B of [1], we obtain the bound

$$|F_X(s)| \leq C(\eta) \exp(|s|^{2+\eta}), \tag{4.10}$$

for $s \in S_\varepsilon \setminus (\cup_j D_j)$.

We can use this bound on $|F_X(s)|$ to control $q(s)$, as follows. For convenience, let us define

$$\mathcal{L}_s = \frac{1}{2s - 1} \frac{d}{ds}. \tag{4.11}$$

Then note that

$$\mathcal{L}_s F_X(s) = -0\text{-tr}[R_X(s)^2 - \mathbf{1}_Y^t \circ R_Y(s)^2 \circ \mathbf{1}_Y]. \tag{4.12}$$

If we differentiate the definition of $\Phi_X(s)$, then the traces can be separated to obtain

$$\mathcal{L}_s \left(\frac{\Phi_X(s)}{2s - 1} \right) = -0\text{-tr} R_X(s)^2 + 0\text{-tr} R_{\mathbb{H}}(s)^2.$$

The first term on the right-hand side makes sense because $R_X(s)^2$ has a continuous kernel. The second term on the right makes sense (as a 0-integral over X) because the restriction of the integral kernel of $R_{\mathbb{H}}(s)^2$ to the diagonal is a constant. In (A.2) this constant is computed to be $(2\pi)^{-1} \mathcal{L}_s \Psi(s)$. Since $0\text{-vol}(X) = -2\pi \chi$, we obtain

$0\text{-tr } R_{\mathbb{H}}(s)^2 = -\chi \mathcal{L}_s \Psi(s)$, and by (4.6) this is in turn equal to $-\mathcal{L}_s^2 \log Z_{\infty}^{\chi}(s)$. We thus have

$$\mathcal{L}_s \left(\frac{\Phi_X(s)}{2s - 1} \right) = -0\text{-tr } R_X(s)^2 - \mathcal{L}_s^2 \log Z_{\infty}^{\chi}(s).$$

On the other hand, by differentiating (4.3) we could also write

$$\mathcal{L}_s \left(\frac{\Phi_X(s)}{2s - 1} \right) = \mathcal{L}_s^2 \log Z_X(s) - n_C \mathcal{L}_s \left(\frac{\Phi_H(s)}{2s - 1} \right),$$

so that

$$-0\text{-tr } R_X(s)^2 = \mathcal{L}_s^2 \log [Z_X(s) Z_{\infty}^{\chi}(s)] - n_C \mathcal{L}_s \left(\frac{\Phi_H(s)}{2s - 1} \right). \tag{4.13}$$

Applying similar arguments for Y gives

$$-0\text{-tr}[\mathbf{1}_Y^t \circ R_Y(s)^2 \circ \mathbf{1}_Y] = \mathcal{L}_s^2 \log Z_Y(s). \tag{4.14}$$

Comparing (4.12), (4.13), and (4.14), we see that $\mathcal{L}_s^2 q(s)$ can be expressed as $\mathcal{L}_s F_X(s)$ plus a sum of other quantities which are known explicitly.

Using the bound (4.10) and the maximum modulus principle (for the disks D_j which had been removed) we thus obtain a bound

$$|q(s)| \leq C(\eta) \exp(|s|^{2+\eta}) \quad \text{for } s \in S_{\varepsilon}.$$

The Phragmén–Lindelöf theorem then allows us to extend the polynomial bound on $q(s)$ into the strip S_{ε} and conclude that the $q(s)$ is a polynomial of degree at most four.

To improve the order bound, we examine the log of (4.2) as $\text{Re}(s) \rightarrow \infty$. In this limit, $\log Z_X(s)$ decays exponentially, while $\log Z_{\infty}^{\chi}(s) = O(s^2 \log s)$ and $n_C \log \Gamma(s - 1/2) = O(s \log s)$. Since P_X is entire of order 2, it follows that $q(s)$ has degree no more than two. \square

With Theorem 4.1 proven, we can revisit Proposition 3 and deduce the following:

Corollary 4.6. *The relative scattering determinant $\tau_X(s)$ is the ratio of entire functions of order two (i.e. the polynomial $h(s)$ in Proposition 3.3 has degree at most two).*

5. Determinant of the Laplacian

For X an infinite area geometrically finite hyperbolic surface, we can define (up to two constants of integration as in [1]) a ‘determinant’ of the Laplacian $D_X(s)$ formally equal to $\det(\Delta - s(1 - s))$ by integrating

$$\mathcal{L}_s^2 \log D_X(s) = -0\text{-tr } R_X(s)^2, \tag{5.1}$$

where \mathcal{L}_s was introduced above in (4.11). Comparing (5.1) to (4.13), we see that formulas of the previous section could be recast as a computation of this determinant.

To complete this computation, note that by Proposition 2.4,

$$\frac{\Phi_H(s)}{2s-1} = -\mathcal{L}_s \log [2^s (s-1/2)^{1/2} \Gamma(s-1/2)].$$

So (4.13) becomes

$$-0\text{-tr } R_X(s)^2 = \mathcal{L}_s^2 \log \left[\frac{Z_X(s) Z_\infty^\chi(s)}{2^{sn_C} (s-1/2)^{n_C/2} \Gamma(s-1/2)^{n_C}} \right]. \quad (5.2)$$

In view of Theorem 4.1, this proves:

Theorem 5.1. *If $D_X(s)$ is a function satisfying (5.1), then $(s-1/2)^{n_C/2} D_X(s)$ is an entire function with zeros given by \mathcal{R}_X (with multiplicities), which can be expressed as*

$$\begin{aligned} (s-1/2)^{n_C/2} D_X(s) &= e^{[Fs(1-s)+E+sn_C \log 2]} \frac{Z_X(s) Z_\infty^\chi(s)}{\Gamma(s-1/2)^{n_C}} \\ &= e^{q_1(s)} P_X(s), \end{aligned}$$

where F, E are the constants of integration left unspecified by (5.1) and $q_1(s)$ is a polynomial of degree ≤ 2 .

This is consistent with previous calculations of the determinant in [1], [6], [27]. Another way to interpret (5.2) is as an analytic expression for the second derivative of $\log P_X(s)$:

$$\mathcal{L}_s^2 \log P_X(s) = -0\text{-tr } R_X(s)^2 + \frac{A}{(2s-1)^3} - \frac{4n_C}{(2s-1)^4},$$

for some constant A .

6. Isospectral finiteness for geometrically finite hyperbolic surfaces

The purpose of this section is to prove Theorem 1.3. First, we make the notion of *length spectrum counted with multiplicities* more precise. Given a complete hyperbolic surface X , let $N_X(L)$ denote the number of closed geodesics on X with length less than L . The function N_X is the length spectrum counting function. We say that two surfaces X and Y are (*length*) *isospectral (including multiplicities)* if and only if $N_X = N_Y$ as functions.

By Corollary 1.2 the resonance set determines the length spectrum, and the length spectrum fixes the Euler characteristic up to finitely many possibilities. The Euler characteristic in turn determines the homeomorphism class up to finitely many possibilities. So to prove Theorem 1.3 it suffices to prove that if X is a geometrically finite hyperbolic surface of infinite area, then there are only finitely many hyperbolic surfaces Y , homeomorphic to X , such that Y is (length) isospectral to X . Moreover, as there are only finitely many ends, we may assume that the number of cusps and number of funnels are constant.

We restate and prove this claim in the (hyperbolic) language of Teichmüller theory [5], [29]. Let S be a fixed smooth surface of finite Euler characteristic with $\partial S = \emptyset$. Here g will denote a hyperbolic metric which is (Riemannian) complete and has exactly n_C cusps and n_F funnels. The set of all such metrics g on S has a natural topology coming from the C^k convergence of tensors on S [29]. The *moduli space*, $\mathcal{M}(S)$, is the quotient of this space by the group of smooth self-diffeomorphisms of S acting via pull-back. In other words, $\mathcal{M}(S)$ is the set of isometry classes $[g]$ of complete hyperbolic metrics g on S .

The *Teichmüller space* $\mathcal{T}(S)$ of S is the quotient of the space of hyperbolic metrics by the diffeomorphisms isotopic to the identity. It is well-known that $\mathcal{T}(S)$ is naturally homeomorphic to a ball in some Euclidean space, and that the natural projection $\pi: \mathcal{T}(S) \rightarrow \mathcal{M}(S)$ is an orbifold covering map.

In this language, Theorem 1.3 is reduced by virtue of Corollary 1.2 to the following:

Theorem 6.1. *Let $\mathcal{A} \subset \mathcal{M}(S)$ be an isospectral set of isometry classes. Then \mathcal{A} is a finite set.*

To prove this theorem⁴, we first note that it suffices to prove that the isospectral set \mathcal{A} is precompact. Indeed, this would imply that the lift of \mathcal{A} to a fundamental domain in $\mathcal{T}(S)$ is precompact, and the lengths of closed geodesics locally separate points in $\mathcal{T}(S)$. (For example, Fenchel–Nielsen coordinates on $\mathcal{T}(S)$ are determined by lengths of closed geodesics [5].) Hence, to prove the theorem we need only discard the possibility that there exists a *divergent* sequence, $[g_n] \subset \mathcal{A}$, a sequence that leaves every compact set in $\mathcal{M}(S)$.

Remark 6.2. The injectivity radius of (S, g) is determined by the length spectrum. Indeed, since the curvature is negative, the injectivity radius equals half of the length of the shortest curve, in other words, half of the infimum of the length spectrum. In the case that the number of funnels $n_F = 0$, Mumford's lemma [22] would give the desired precompactness of (a lift of) \mathcal{A} . If, however, $n_F > 0$, then Mumford's lemma does not apply (directly) to $\mathcal{T}(S)$.

⁴During the preparation of this work, we received a preprint from In Kang Kim [15] with an isofiniteness result (for fixed homeomorphism class). We do not fully understand his argument.

Recall that a topologically finite, complete, hyperbolic surface (S, g) is determined by its convex core. The *convex core* can be described as either the convex hull of the set of closed g -geodesics on S or as the hyperbolic surface (S, g) with its funnels removed.

A given cusp end of (S, g) can be compactified by adding one point. A closed curve $\alpha \subset S$ will be called *non-cuspidal* if and only if α is not null-homotopic after compactifying any single cusp. In other words, given a deck group representation $\rho: \pi_1(S) \rightarrow \mathrm{SL}(2, \mathbb{R})$ for (S, g) , a closed curve α is non-cuspidal if and only if $|\mathrm{tr}(\rho([\alpha]))| > 2$. In particular, α is non-cuspidal if and only if there exists a (unique) g -geodesic that is homotopic to α .

Lemma 6.3 (Geometric limits). *Let S be a (connected) differentiable surface with $\chi(S) < 0$. Let $[g_n] \in \mathcal{M}(S)$ be a divergent sequence with injectivity radius uniformly bounded from below by a positive constant. Then there exist*

- (1) *a subsequence of metric representatives, g_n ,*
- (2) *a topologically finite, complete, hyperbolic surface (R, h) ,*
- (3) *a precompact neighborhood U of the convex core of (R, h) , and*
- (4) *a smooth embedding $f: U \rightarrow S$,*

such that

- (A) *each metric $f^*(g_n)$ extends to a complete hyperbolic metric h_n on R ,*
- (B) *$h_n \rightarrow h$ in $C^k(U)$ for each $k \in \mathbb{N}$,*
- (C) *for each n , the convex core of (R, h_n) lies inside U ,*
- (D) *given $L > 0$, there exists $M > 0$ such that if $m > M$ and $\alpha \subset S$ is a non-cuspidal closed curve that is not homotopic to a curve in $f(U)$, then the g_m -length of α is larger than L , and*
- (E) *$\chi(S) < \chi(R)$.*

The surface (R, h) is a geometric limit of (S, g_n) .

We postpone the proof of Lemma 6.3 to the end of this section.

Proof of Theorem 6.1. By the discussion immediately after the statement of Theorem 6.1, it suffices to show that there cannot exist a divergent isospectral sequence. Suppose to the contrary that such a sequence exists. By Remark 6.2, the injectivity radius of this sequence is bounded below.

Note that if a connected, complete, hyperbolic surface (S, g) satisfies $\chi(S) \geq 0$, then S is either a cylinder or the disk. In the latter case, there is only one isometry class. In the former, either S is a horn in which case there is one isometry class, or S is a cylinder in which case there is exactly one primitive closed geodesic whose

length – the infimum of the length spectrum – determines the isometry class of the metric. In particular, any isospectral sequence of metrics on the cylinder cannot be divergent.

Thus we assume that $\chi(S) < 0$, and hence apply Lemma 6.3 to obtain a geometric limit (R, h) . Let f, U , and h_n be as in Lemma 6.3.

Let β (resp. β_n) be the unique h -geodesic representative (resp. h_n -geodesic representative) of the homotopy class of a non-cuspidal closed curve $\beta \subset R$. We claim that, by passing to a further subsequence if necessary, we may assume that β_n converges to β . Indeed, each geodesic is a solution to an ordinary differential equation whose coefficients depend continuously on the metric. By part (C) of Lemma 6.3, the convex core of each (R, h_n) lies in the precompact set U , and hence there exists a subsequence such that the initial conditions converge. Thus, the claim follows from part (B) of Lemma 6.3 and the continuity of solutions to ordinary differential equations with respect to coefficients and initial data.

There exists a finite number of homotopy classes of simple closed curves, γ_i , on R such that the h_n -lengths (resp. h -lengths) of the corresponding geodesics determine the metric h_n (resp. h) [5]. By the above, the h_n -lengths of the h_n -representatives of the γ_i converge to the h -length of the h -geodesic representing γ_i . Since, by assumption, the length spectrum is constant, the h_n -length of each γ_i is constant for sufficiently large n . Therefore, by passing to a further subsequence if necessary, we may assume that (R, h_n) is isometric to (R, h) .

Let $N_n^0(x)$ (resp. $N_n^\infty(x)$) denote the number of closed g_n -geodesics having length less than x that are homotopic (resp. not homotopic) to a curve in $f(U)$. Then we have

$$N_{(S, g_n)}(x) = N_n^0(x) + N_n^\infty(x) = N_{(R, h_n)}(x) + N_n^\infty(x).$$

By assumption, the left hand side does not depend on n , and since (R, h_n) is isometric to (R, h) for all n , we have that $N_{(R, h_n)}$ is independent of n . Thus, to obtain a contradiction, it suffices to show that N_n^∞ does depend on n .

Let $\alpha \subset S$ be a non-cuspidal g_1 -geodesic that is not homotopic to a curve in $f(U)$. That such an α exists follows from, for example, homological considerations. Since the first homology $H^1(S, \mathbb{Z}) \cong \mathbb{Z}^{2-\chi(S)}$, we can take α to be a geodesic representative of a homology class whose intersection number with the class of some component of γ is not zero. If the class of α were homotopic to a loop in $f(U)$, then this intersection number would be zero since $f(U) \cap \gamma = \emptyset$. We apply part (D) of Lemma 6.3 to the curve α , choosing L to be equal to twice the g_1 -length of α . Thus, there exists M – necessarily larger than 1 – such that the g_M -length of any non-cuspidal curve that is not homotopic to a curve in $f(U)$ is larger than L . In particular, the g_M -length of any closed g_M -geodesic that is not homotopic to a curve in $f(U)$ is greater than L .

It follows that for x satisfying $\frac{L}{2} < x < L$, we have $N_M^\infty(x) = 0$ and $N_1^\infty(x) > 0$. Thus, the functions N_n^∞ depend on n , thus giving the desired contradiction. \square

Proof of Lemma 6.3. Let K, K' be two isometric copies of the convex core of (S, g_n) , and let D be the surface obtained by gluing K and K' along their respective boundaries. The number of funnels and cusps of (S, g_n) does not depend on n , hence the diffeomorphism class of D does not depend on n . Since $\partial K, \partial K'$, are geodesic, the metric g_n extends to a complete hyperbolic metric on D having finite area. Abusing notation slightly, we will refer to this extended metric as g_n .

Note that the divergence of $[g_n]$ implies the divergence of $[g_n]$ on the double D . Let $\langle g_n \rangle \in \mathcal{T}(D)$ such that $\pi(\langle g_n \rangle) = [g_n]$.

Mumford's lemma [22] (cofinite case) implies that there exists a sequence, γ_n , of multicurves⁵ on D whose $\langle g_n \rangle$ -length, $\ell_n(\gamma_n)$, tends to zero as n tends to infinity.

By Lemma 6.4 below, each component of γ_n meets ∂K orthogonally. It follows that the group, $\text{Mod}_{\partial K}(D)$, of mapping classes that preserve (the homotopy class of) ∂K acts on the collection $\{\gamma_n\}$ with finitely many orbits. Hence, by taking a subsequence and applying elements of $\text{Mod}_{\partial K}(D)$, we may assume that $\langle g_n \rangle \subset \mathcal{T}(S)$ was chosen so that sequence γ_n is equal to a constant γ .

Standard arguments (using, for example, pants decompositions) show that there exist representatives g_n of $\langle g_n \rangle$ such that ∂K is g_n -geodesic for all n , such that the g_n -geodesic representative (still denoted γ) of the multicurve γ is independent of n , and such that g_n converges uniformly on compact subsets of $D \setminus \gamma$ to a complete (finite area) hyperbolic metric g on $D \setminus \gamma$.

Let $K^* = K \setminus (K \cap \gamma)$. Then (K^*, g_n) has piecewise geodesic boundary with interior angles equal to $\frac{\pi}{2}$ by Lemma 6.4. It follows that K^* is geodesically convex and, moreover, the g_n -holonomy (in the sense of hyperbolic structures) of any closed curve in K^* cannot be elliptic. Thus, the hyperbolic surface (K^*, g_n) extends to a complete hyperbolic surface (R, h_n) with R homeomorphic to the interior of K^* .

Similarly, we obtain a metric h on R such that (R, h) is a complete hyperbolic surface that is an extension of (K^*, g) . Since the metrics g_n converge to g , it follows, by analytic continuation, that h_n converges to h uniformly on compact subsets of R .

Note that since (K^*, g_n) is geodesically convex, the convex core of (R, h_n) is bounded by simple closed geodesics that lie within K^* . By Lemma 6.5 below, there exists, independent of n , a collar neighborhood C of γ such that no simple closed g_n -geodesic in $K \setminus \gamma$ intersects C . Thus, the convex core of each (R, h_n) lies within the compact set $U = K^* \setminus C$. Let $f: U \rightarrow S$ be the associated inclusion. Then R, U, f, h , and h_n satisfy (A), (B), and (C). See Figure 2.

By the collar lemma [5], given $L > 0$, there exists n such that if a g_n -geodesic on D intersects γ , then the g_n -length must be greater than L . Part (D) of the lemma follows.

Let k be the number of components of the multicurve γ . Note that ∂K^* has $4 \cdot k$ corners each with an exterior angle equal to $\pi/2$ by Lemma 6.4. Therefore, it follows

⁵Recall that a *multicurve* is a free homotopy class of a finite disjoint union of mutually non-homotopic simple closed curves. A well-known argument shows that a multicurve has at most $3g - 3$ components.

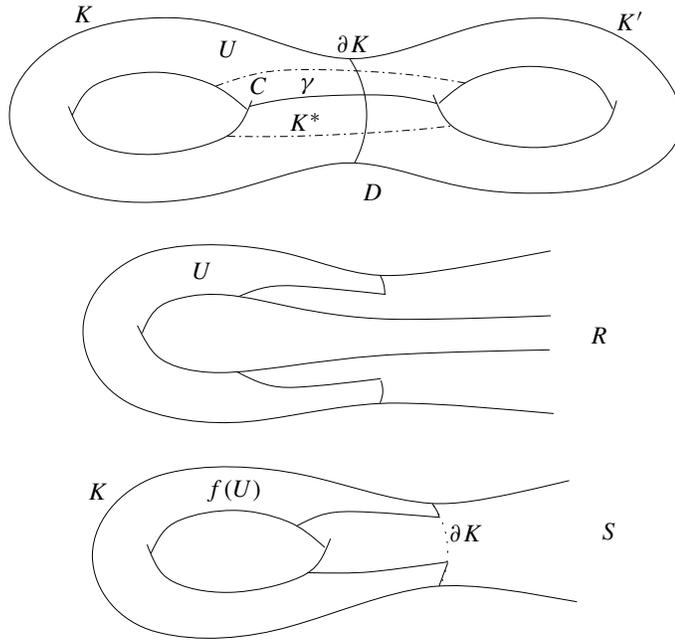


Figure 2. The set up for geometric limits.

from the Gauss–Bonnet theorem that

$$\chi(K^*) = \chi(K) + k. \tag{6.1}$$

Since $\chi(R) = \chi(K^*)$ and $\chi(S) = \chi(K)$ and $k > 0$, we have part (E). □

Lemma 6.4. *There exists a constant ℓ^* such that if γ is the g_n -geodesic representative of a multicurve on D of g_n -length $\ell(\gamma) < \ell^*$, then each component of γ intersects ∂K orthogonally. Moreover, ℓ^* , depends only on a lower bound on the injectivity radius of (D, g_n) .*

Proof. If $\ell(\gamma)$ is less than the injectivity radius of (D, g_n) , then no component of γ can lie entirely inside one of the two copies of K . Hence each component intersects ∂K transversally. To see that this intersection is orthogonal, let $\phi: D \rightarrow D$ be the reflection isometry associated to the doubling. Then $\phi(\gamma)$ is another geodesic multicurve on D that intersects γ at the fixed point set of ϕ , ∂K . Since ϕ is an isometry, $\ell(\phi(\gamma)) = \ell(\gamma)$. Recall that by the collar lemma, there exists ℓ' (independent of the complete hyperbolic surface) such that if β_1 and β_2 are intersecting, simple closed geodesics each with length less than ℓ' , then $\beta_1 = \beta_2$.

Let ℓ^* be the minimum of ℓ' and the injectivity radius. □

Lemma 6.5 (Point of No Return). *Let γ be a simple closed geodesic of length ℓ on a complete hyperbolic surface X . If α is a simple closed geodesic that does not intersect γ , then its distance d from γ is at least $\log(\coth(\ell/4))$.*

Proof. Consider the universal covering of X by the upper half-plane $\{(x, y) \in \mathbb{R}^2 \mid y > 0\}$ such that the lift of the geodesic γ is the y -axis: $\tilde{\gamma} = \{(0, y) \mid y > 0\}$. Let τ be the hyperbolic isometry $\tau(x, y) = (e^\ell x, e^\ell y)$ stabilizing $\tilde{\gamma}$. Each tubular neighborhood of γ is the quotient by $\langle \tau \rangle$ of a (Euclidean) sector that is bisected by $\tilde{\gamma}$.

Since both γ and α are closed, $d > 0$. Take $\tilde{\alpha}$ to be a lift of α such that the distance between $\tilde{\gamma}$ and $\tilde{\alpha}$ is d . (Note that $\tilde{\alpha}$ cannot be vertical.)

Let Σ be the smallest sector bisected by $\tilde{\gamma}$ that intersects $\tilde{\alpha}$. (See Figure 3). Then $\tilde{\alpha}$ meets $\partial\Sigma$ tangentially at one point (x, y) and $\Sigma/\langle \tau \rangle$ is a d -tubular neighborhood of γ . Conjugating by a suitable hyperbolic element having axis $\tilde{\gamma}$ and a reflection

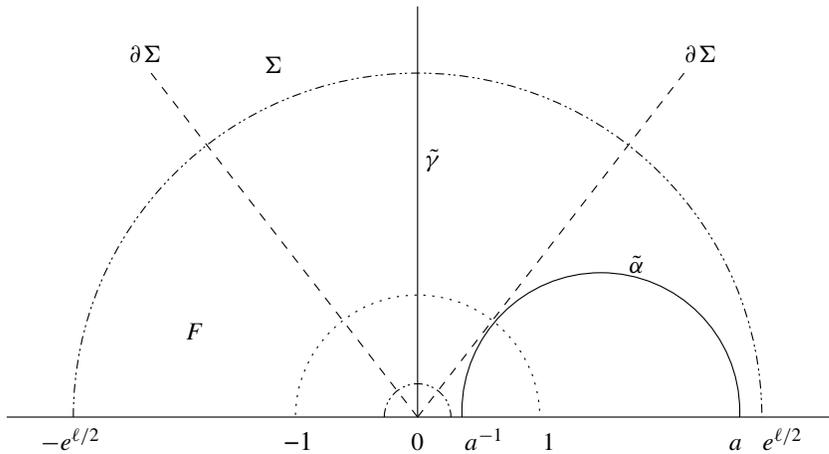


Figure 3. Universal cover of X in Lemma 6.5.

about $\tilde{\gamma}$ if necessary, we may assume without loss of generality, that $x > 0$ and $x^2 + y^2 = 1$. Thus, $\tilde{\alpha}$ is a semi-circle with endpoints at $(1/a, 0)$ and $(a, 0)$ where $a > 1$.

Consider the following fundamental domain of τ :

$$F = \{(r \cos(\theta), r \sin(\theta)) \mid 0 < \theta < \pi; e^{-\ell/2} \leq r \leq e^{\ell/2}\}.$$

We claim that $\tilde{\alpha}$ does not intersect either of the boundary components of F : $\partial F = \{r = e^{\ell/2}\} \cup \{r = e^{-\ell/2}\}$. To see this note that the geodesic $\tilde{\alpha}$ meets the (Euclidean)

unit circle $\{r = 1\}$ orthogonally and hence $\tilde{\alpha}$ is preserved by the reflection $r \mapsto 1/r$. In addition, this reflection exchanges the two boundary components of F . It follows that if $(x', y') \in \tilde{\alpha} \cap \partial F$, then either $\tau(x', y')$ or $\tau^{-1}(x', y')$ also belongs to $\tilde{\alpha}$. The corresponding self-intersection of $\tilde{\alpha}$ is transverse, a contradiction to the simplicity of α .

Therefore, we must have $e^{-\ell/2} \leq a^{-1} < 1 < a \leq e^{\ell/2}$. A straightforward calculation shows that the hyperbolic width of Σ is at least $2 \log(\coth(\ell/4))$. \square

A. Special function calculations

To calculate $\Phi_F(s)$ and $\Phi_H(s)$ explicitly, we derive and use an explicit representation of the integral kernel $G_{\mathbb{H}}(s; z, z')$ of the resolvent operator $(\Delta_{\mathbb{H}} - s(1 - s))^{-1}$. In particular, if we denote

$$\sigma(z, z') = \cosh^2(d(z, z')/2) = \frac{(x - x')^2 + (y + y')^2}{4yy'}$$

then $G_{\mathbb{H}}(s, z, z') = g_s(\sigma)$ where g_s can be expressed explicitly in terms of a hypergeometric function:

$$g_s(\sigma) = \frac{1}{4\pi} \frac{\Gamma(s)^2}{\Gamma(2s)} \sigma^{-s} F(s, s; 2s, \sigma^{-1}),$$

where

$$F(a, b; c; z) = 1 + \frac{ab}{c}z + \frac{a(a + 1)b(b + 1)}{2!c(c + 1)}z^2 + \dots$$

(see, for example, formula (6.6.1) of [12]). Euler’s representation of the hypergeometric function,

$$F(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c - b)} \int_0^1 \frac{t^{b-1}(1 - t)^{c-b-1}}{(1 - tz)^a} dt,$$

gives us the useful representation:

$$G_{\mathbb{H}}(s, z, z') = g_s(\sigma) = \frac{1}{4\pi} \int_0^1 \frac{t^{s-1}(1 - t)^{s-1}}{(\sigma - t)^s} dt, \tag{A.1}$$

for $\sigma > 1, \text{Re}(s) > 0$.

As an application, we can evaluate

$$\lim_{\sigma \rightarrow 1} \frac{d}{ds} g_s(\sigma) = \frac{1}{4\pi} \int_0^1 \frac{t^{s-1}}{1 - t} \log t dt = -\frac{1}{2\pi} \Psi'(s).$$

Since $(d/ds)R_{\mathbb{H}}(s) = -(2s - 1)R_{\mathbb{H}}(s)^2$ this implies that

$$(R_{\mathbb{H}}(s)^2)(s, w, w) = \frac{1}{2\pi}(2s - 1)^{-1}\Psi'(s). \tag{A.2}$$

By combining this with the same expression evaluated at $1 - s$ and using the identity $\Gamma(s)\Gamma(1 - s) = \pi \csc(\pi s)$, we can also derive

$$\frac{d}{ds}[G_{\mathbb{H}}(s, z, w) - G_{\mathbb{H}}(1 - s, z, w)]_{z=w} = -\frac{\pi}{2} \csc^2(\pi s).$$

This can be integrated to give

$$[G_{\mathbb{H}}(s, z, w) - G_{\mathbb{H}}(1 - s, z, w)]_{z=w} = \frac{1}{2} \cot(\pi s). \tag{A.3}$$

A.1. Funnel. As above, F is the model funnel (hyperbolic half-cylinder) and Δ_F denotes the Laplacian with Dirichlet boundary conditions. F is half of the cylinder $M = \Gamma_\ell \backslash \mathbb{H}$. The resolvent kernel for Δ_F is obtained from that of Δ_M by the method of images:

$$G_F(s; z, w) = G_M(s; z, w) - G_M(s; z, \tau w), \tag{A.4}$$

where τ denotes the involution $(x, y) \mapsto (-x, y)$ and this is really a statement about the lifts of the kernels to $\mathbb{H} \times \mathbb{H}$.

Proof of Proposition 2.3. Using (A.4) the integrand in the definition on $\Phi_F(s)$ can be written

$$\begin{aligned} [G_F(s; z, w) - G_{\mathbb{H}}(s; z, w)]_{w=z} &= \sum_{m \neq 0} g_s \left(\frac{(x - e^{m\ell}x)^2 + (y + e^{m\ell}y)^2}{4e^{m\ell}y^2} \right) \\ &\quad - \sum_{m \in \mathbb{Z}} g_s \left(\frac{(x + e^{m\ell}x)^2 + (y + e^{m\ell}y)^2}{4e^{m\ell}y^2} \right). \end{aligned} \tag{A.5}$$

With an inverse Mellin transform, Patterson [25] computes that

$$(2s - 1) \int_{-\infty}^{\infty} g_s \left((1 + t^2) \frac{(1 + \kappa)^2}{4\kappa} \right) dt = \frac{\max(\kappa, \kappa^{-1})^{1/2-s}}{\kappa^{1/2} + \kappa^{-1/2}}. \tag{A.6}$$

We will apply this by choosing fundamental domain $\mathcal{F} = \mathbb{R}_+ \times [1, e^\ell]$. When (A.5) is integrated over \mathcal{F} , the $m = 0$ term in the second sum contributes

$$\begin{aligned} -(2s - 1) \int_1^{e^\ell} \int_0^\infty g_s \left(1 + \frac{x^2}{y^2} \right) \frac{dx dy}{y^2} &= -\frac{2s - 1}{2} \int_1^{e^\ell} \int_{-\infty}^\infty g_s(1 + t^2) \frac{dt dy}{y} \\ &= -\frac{\ell}{4}. \end{aligned}$$

For the $m \neq 0$ the appropriate substitution is

$$t = \frac{(1 \mp e^{m\ell}) x}{(1 + e^{m\ell}) y},$$

(with \mp depending on whether the involution τ is inserted or not). Integrating an $m \neq 0$ from (A.5) over \mathcal{F} then gives

$$\begin{aligned} (2s - 1) \int_{\mathcal{F}} g_s \left(\frac{(x \mp e^{m\ell} x)^2 + (y + e^{m\ell} y)^2}{4e^{m\ell} y^2} \right) \frac{dx dy}{y^2} \\ = \frac{\ell}{2} \cdot \frac{(1 + e^{m\ell})}{(1 \mp e^{m\ell})} (2s - 1) \int_{-\infty}^{\infty} g_s \left((1 + t^2) \frac{(1 + \lambda^2)^2}{4e^{m\ell}} \right) dt \\ = \frac{\ell}{2} \cdot \frac{(1 + e^{m\ell})}{(1 \mp e^{m\ell})} \frac{e^{(1/2-s)|m|\ell}}{e^{m\ell/2} + e^{-m\ell/2}} \\ = \frac{\ell}{2} \cdot \frac{e^{-s|m|\ell}}{1 \mp e^{-|m|\ell}}. \end{aligned}$$

Thus the full contribution from the $m \neq 0$ terms is

$$\begin{aligned} \frac{\ell}{2} \sum_{m \neq 0} \left(\frac{e^{-s|m|\ell}}{1 - e^{-|m|\ell}} - \frac{e^{-s|m|\ell}}{1 + e^{-|m|\ell}} \right) &= 2\ell \sum_{m \geq 1} \left(\frac{e^{-(s+1)m\ell}}{1 - e^{-2m\ell}} \right) \\ &= 2 \frac{d}{ds} \sum_{k \geq 1} \log(1 - e^{-(s+2k+1)\ell}). \end{aligned}$$

Thus, taking into account all of the terms in (A.5) gives

$$\begin{aligned} (2s - 1) \int_{\mathcal{F}} [G_F(s; z, w) - G_{\mathbb{H}}(s; z, w)]_{w=z} dg(z) \\ = -\frac{\ell}{4} + \frac{d}{ds} \sum_{k \geq 1} \log(1 - e^{-(s+2k+1)\ell})^2. \end{aligned} \quad \square$$

A.2. Horn. The resolvent and generalized eigenfunctions for the horn $H = \Gamma_{\infty} \backslash \mathbb{H}$ may be obtained by summing $G_{\mathbb{H}}$ over Γ_{∞} . This makes it possible to compute $\Phi_H(s)$ fairly directly.

Proof of Proposition 2.4. For the proof it suffices to assume $\text{Re}(s) > 1/2$. The integrand in (2.14) can be written

$$\begin{aligned} [G_H(s; z, w) - G_{\mathbb{H}}(s; z, w)]_{w=z} &= \sum_{k \neq 0} G_{\mathbb{H}}(s, z, z + k) \\ &= 2 \sum_{k=1}^{\infty} g_s \left(1 + \frac{k^2}{4y^2} \right). \end{aligned}$$

We will use $[0, 1] \times \mathbb{R}_+$ as our fundamental domain for H . The integral in the definition of $\Phi_H(s)$ is well-behaved at the funnel end, but needs regularization at the cusp end:

$$\Phi_H(s) = 2(2s - 1) \text{FP}_{\varepsilon \downarrow 0} \int_0^{1/\varepsilon} \int_0^1 \sum_{k=1}^{\infty} g_s \left(1 + \frac{k^2}{4y^2} \right) \frac{dx dy}{y^2}.$$

Substituting $u = \frac{k}{2y}$, this becomes

$$\begin{aligned} \Phi_H(s) &= 4(2s - 1) \text{FP}_{\varepsilon \downarrow 0} \sum_{k=1}^{\infty} \int_{k\varepsilon/2}^{\infty} g_s(1 + u^2) \frac{1}{k} du \\ &= 4(2s - 1) \text{FP}_{\varepsilon \downarrow 0} \int_0^{\infty} g_s(1 + u^2) \left[\sum_{1 \leq k \leq 2u/\varepsilon} \frac{1}{k} \right] du. \end{aligned}$$

In order to extract the finite part, we use

$$\sum_{1 \leq k \leq 2u/\varepsilon} \frac{1}{k} = \log \frac{2u}{\varepsilon} + \gamma + O(\sqrt{\varepsilon/2u}),$$

where γ is Euler's constant. The $O(\sqrt{\varepsilon/2u})$ contribution will vanish in the $\varepsilon \rightarrow 0$ limit, and $\log \varepsilon$ is dropped in taking the finite part, leaving

$$\begin{aligned} \Phi_H(s) &= (\log 2 + \gamma) \cdot 4(2s - 1) \int_0^{\infty} g_s(1 + u^2) du \\ &\quad + 4(2s - 1) \int_0^{\infty} g_s(1 + u^2) \log u du. \end{aligned} \tag{A.7}$$

The formula (A.6) gives:

$$4(2s - 1) \int_0^{\infty} g_s(1 + u^2) du = 1. \tag{A.8}$$

For the other integral in (A.7) we use (A.1) to write

$$4(2s - 1) \int_0^{\infty} g_s(1 + u^2) \log u du = \frac{2s - 1}{\pi} \int_0^{\infty} \int_0^1 \frac{t^{s-1}(1-t)^{s-1}}{(1-t+u^2)^s} \log u du dt.$$

Substituting $u = \sqrt{(1-t)(1-w)/w}$ yields

$$\begin{aligned} &4(2s - 1) \int_0^1 g_s(1 + u^2) \log u du \\ &= \frac{2s - 1}{4\pi} \int_0^1 \int_0^1 t^{s-1}(1-t)^{-1/2} w^{s-3/2} (1-w)^{-1/2} \log[(1-t)(1-w)/w] dw dt. \end{aligned}$$

These are straightforward Beta-type integrals, whose evaluation gives:

$$4(2s - 1) \int_0^1 g_s(1 + u^2) \log u \, du = -\gamma - \log 4 - \Psi(s + 1/2) + \frac{1}{2s - 1}.$$

Using this last equation together with (A.8) in (A.7) gives the stated result. \square

The preceding calculation shows that $\text{Res}_{s=1/2} \Phi_H(s) = 1/2$. This residue can be identified as originating in the funnel end of H , with no contribution from the cusp. To make this statement precise, we introduce a cutoff $\eta \in C^\infty(\mathbb{R}_+)$ which is equal to 0 in a neighborhood of 0 and equal to 1 in a neighborhood of ∞ . Note that the integrand used to define $\Phi_H(s)$ is explicitly given by

$$\varphi_H(s; y) = 2(2s - 1) \sum_{k=1}^\infty g_s \left(1 + \frac{k^2}{4y^2} \right).$$

Lemma A.1. *The contribution from the cusp end to the 0-trace in $\text{Res}_{s=1/2} \Phi_H(s) = 1/2$ is zero, in the sense that:*

$$\text{Res}_{s=1/2} \left[\text{FP}_{\varepsilon \rightarrow 0} \int_0^{1/\varepsilon} \eta(y) \varphi_H(s; y) \frac{dy}{y^2} \right] = 0.$$

Proof. It is equivalent to show that the 0-trace comes exclusively from the funnel end, i.e. that

$$\text{Res}_{s=1/2} \left[\text{FP}_{\varepsilon \rightarrow 0} \int_\varepsilon^\infty (1 - \eta(y)) \varphi_H(s; y) \frac{dy}{y^2} \right] = \frac{1}{2}.$$

Using the hypergeometric series representation of $g_s(\sigma)$, we can deduce that $(1 - \eta(y))\varphi_H(s; y)$ is of the form $y^{2s} f_s(y)$, with $f_s \in C_c^\infty(\mathbb{R}_+)$ and

$$f_s(0) = \frac{(2s - 1) \Gamma(s)^2}{2\pi \Gamma(2s)} 4^s \zeta(2s).$$

Integration by parts shows that, for $s \neq 1/2$,

$$\text{FP}_{\varepsilon \rightarrow 0} \int_\varepsilon^\infty y^{2s-2} f_s(y) \, dy = -\frac{1}{(2s - 1)} \int_0^\infty y^{2s-1} f'_s(y) \, dy.$$

Thus

$$\text{Res}_{s=1/2} \left[\text{FP}_{\varepsilon \rightarrow 0} \int_0^{1/\varepsilon} (1 - \eta(y)) \varphi_H(s; y) \frac{dy}{y^2} \right] = \frac{1}{2} f_{1/2}(0).$$

The result then follows from $f_{1/2}(0) = 1$, which holds because the factor $(2s - 1)$ exactly cancels the pole of $\zeta(2s)$ at $s = 1/2$. \square

References

- [1] D. Borthwick, C. Judge, P. A. Perry, Determinants of Laplacians and isopolar metrics on surfaces of infinite area. *Duke Math. J.* **118** (2003), 61–102. [Zbl 1040.58013](#) [MR 1978883](#)
- [2] D. Borthwick, P. Perry, Scattering poles for asymptotically hyperbolic manifolds. *Trans. Amer. Math. Soc.* **354** (2002), 1215–1231. [Zbl 1009.58021](#) [MR 1867379](#)
- [3] R. Brooks, O. Davidovich, Isoscattering on surfaces. *J. Geom. Anal.* **13** (2003), 39–53. [Zbl 1040.58010](#) [MR 1967035](#)
- [4] U. Bunke, M. Olbrich, Group cohomology and the singularities of the Selberg zeta function associated to a Kleinian group. *Ann. of Math. (2)* **149** (1999), 627–689. [Zbl 0969.11019](#) [MR 1689342](#)
- [5] P. Buser, *Geometry and Spectra of Compact Riemann Surfaces*. Progr. Math. 106, Birkhäuser, Boston 1992. [Zbl 0770.53001](#) [MR 1183224](#)
- [6] I. Efrat, Determinants of Laplacians on surfaces of finite volume. *Comm. Math. Phys.* **119** (1988), 443–451; Erratum, *Comm. Math. Phys.* **138** (1991), 607. [Zbl 0661.10038](#); [Zbl 0722.11029](#) [MR 0969211](#); [MR 1110458](#)
- [7] L. Guillopé, Fonctions zeta de Selberg et surfaces de géométrie finie. In *Zeta functions in geometry*, Adv. Stud. Pure Math. 21, Kinokuniya, Tokyo 1992, 33–70. [Zbl 0794.58044](#) [MR 1210782](#)
- [8] L. Guillopé, M. Zworski, Upper bounds on the number of resonances for non-compact Riemann surfaces. *J. Funct. Anal.* **129** (1995), 364–389. [Zbl 0841.58063](#) [MR 1327183](#)
- [9] L. Guillopé, M. Zworski, Polynomial bounds on the number of resonances for some complete spaces of constant negative curvature near infinity. *Asymptotic Anal.* **11** (1995), 1–22. [Zbl 0859.58028](#) [MR 1344252](#)
- [10] L. Guillopé, M. Zworski, Scattering asymptotics for Riemann surfaces. *Ann. of Math. (2)* **145** (1997), 597–660. [Zbl 0898.58054](#) [MR 1454705](#)
- [11] L. Guillopé, M. Zworski, The wave trace for Riemann surfaces. *Geom. Funct. Anal.* **9** (1999), 1156–1168. [Zbl 0947.58022](#) [MR 1736931](#)
- [12] D. A. Hejhal, *The Selberg Trace Formula for PSL(2, \mathbb{R})* (Vol. II). Lecture Notes in Math. 1001, Springer-Verlag, Berlin 1983. [Zbl 0543.10020](#) [MR 0711197](#)
- [13] M. Joshi, A. Sá Barreto, Inverse scattering on asymptotically hyperbolic manifolds. *Acta Math.* **184** (2000), 41–86. [Zbl 01541222](#) [MR 1756569](#)
- [14] H. Khuri, Heights on the moduli space of Riemann surfaces with circle boundaries. *Duke Math. J.* **64** (1991), 555–570. [Zbl 0755.30037](#) [MR 1141285](#)
- [15] I. Kim, On spectral finiteness of hyperbolic 3-manifolds. Preprint.
- [16] P. Lax, R. S. Phillips, The asymptotic distribution of lattice points in Euclidean and non-Euclidean spaces. *J. Funct. Anal.* **46** (1982), 280–350. [Zbl 0497.30036](#) [MR 0661875](#)
- [17] P. Lax, R. S. Phillips, Translation representation for automorphic solutions of the wave equation in non-Euclidean spaces, I–III. *Comm. Pure Appl. Math.* **37** (1984), 303–328; **37** (1984), 779–813; **38** (1985), 179–207. [Zbl 0544.10024](#); [Zbl 0549.10019](#); [Zbl 0578.10033](#) [MR 07399235](#); [MR 0762873](#); [MR 0780072](#)

- [18] R. R. Mazzeo, Elliptic theory of differential edge operators I. *Comm. Partial Differential Equations* **16** (1991), 1615–1664. [Zbl 0745.58045](#) [MR 1133743](#)
- [19] R. Mazzeo, R. Melrose, Meromorphic extension of the resolvent on complete spaces with asymptotically constant negative curvature. *J. Funct. Anal.* **75** (1987), 260–310. [Zbl 0636.58034](#) [MR 0916753](#)
- [20] H. P. McKean, Selberg's trace formula as applied to a compact Riemann surface. *Comm. Pure Appl. Math.* **25** (1972), 225–246. [MR 473166](#)
- [21] R. B. Melrose, *Geometric Scattering Theory*. Stanford Lectures, Cambridge University Press, New York 1995. [Zbl 0849.58071](#) [MR 1350074](#)
- [22] D. Mumford, A remark on Mahler's compactness theorem. *Proc. Amer. Math. Soc.* **28** (1971), 289–294. [Zbl 0215.23202](#) [MR 0276410](#)
- [23] W. Müller, Spectral geometry and scattering theory for certain complete surfaces of finite volume. *Invent. Math.* **109** (1992), 265–305. [Zbl 0772.58063](#) [MR 1172692](#)
- [24] M. Olbrich, Cohomology of convex cocompact groups and invariant distributions on limit sets. Preprint (arXiv:math.DG/0207301). [Zbl 1006.22012](#)
- [25] S. J. Patterson, The Selberg zeta-function of a Kleinian group. In *Number Theory, Trace Formulas, and Discrete Groups: Symposium in Honor of Atle Selberg* (Oslo, Norway, July 14–21, 1987), Academic Press, New York 1989, 409–441. [Zbl 0668.10036](#) [MR 0993330](#)
- [26] S. J. Patterson, P. A. Perry, The Divisor of Selberg's Zeta Function for Kleinian Groups. *Duke Math. J.* **106** (2001), 321–390. [Zbl 1012.11083](#) [MR 1813434](#)
- [27] P. Sarnak, Determinants of Laplacians. *Comm. Math. Phys.* **110** (1987), 113–120. [Zbl 0618.10023](#) [MR 0885573](#)
- [28] J. Sjöstrand, M. Zworski, Lower bounds on the number of scattering poles. *Comm. Partial Differential Equations* **18** (1993), 847–857. [Zbl 0784.35070](#) [MR 1218521](#)
- [29] A. J. Tromba, *Teichmüller theory in Riemannian geometry*. Lectures in Mathematics ETH Zürich, Birkhäuser, Basel 1992. [Zbl 0785.53001](#) [MR 1164870](#)

Received October 23, 2003

David Borthwick, Department of Mathematics and Computer Science, Emory University, Atlanta, Georgia 30322, U.S.A.

E-mail: davidb@mathcs.emory.edu

Chris Judge, Department of Mathematics, Indiana University, Bloomington, Indiana 47401, U.S.A.

E-mail: cjudge@indiana.edu

Peter A. Perry, Department of Mathematics, University of Kentucky, Lexington, Kentucky 40506–0027, U.S.A.

E-mail: perry@ms.uky.edu