SPECTRAL THEORY OF THE RIEMANN ZETA-FUNCTION

Y. MOTOHASHI

Yoichi Motohashi

Nihon University

Spectral Theory of the
Riemann Zeta-Function

Preface

EVER since Riemann's use of the theta transformation formula in one of his proofs of the functional equation for the zeta-function, number-theorists have been fascinated by various interactions between the zeta-function and automorphic forms. These experiences, however, have remained episodic like rare glimpses of crests, for most of them ensued from apparently spontaneous relations of the zeta-function with a variety of Eisenstein series. Nevertheless such glimpses are highly suggestive of a grand view over and far beyond the Eisenstein ridge, and bring forth the notion of a *kamuy-mintar* where the entire collection of automorphic forms contribute to the formation of the zeta-function.

My aim in the present monograph is to try to substantiate this belief by demonstrating that the zeta-function has indeed a structure tightly supported by all automorphic forms. The story begins with an unabridged treatment of the spectral resolution of the non-Euclidean Laplacian, and continues to a theory of trace formulas. The fundamental means thus readied are subsequently mustered up for the quest to find an explicit formula for the fourth power moment of the zeta-values. Then the zeta-function emerges as a magnificent peak embracing infinitely many gems called automorphic L-functions representing the spectrum.

My best thanks are due to my friends A. Ivić and M. Jutila for their unfailing encouragement, and to D. Tranah, P. Jackson, and all of the personnel of the Cambridge University Press engaged in this project for sharing their professional vigor. I must also thank my family for the comfort and the music that have been sustaining my scientific life.

Tokyo
January, 1997

Y. M.

Convention and assumed background

Once introduced most symbols will remain effective throughout the sequel. Some of them are naturally standard. Thus $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are sets of all integers, rationals, reals, and complex numbers, respectively. For example the group composed of all $n \times n$ integral matrices with determinant equal to 1 is denoted by $SL(n, \mathbb{Z})$. The arithmetic functions $\sigma_a(n)$ and $d_k(n)$ stand, respectively, for the sum of the ath powers of divisors of n and for the number of ways of expressing n as a product of k integral factors. In particular, $d(n) = d_2(n)$ is the divisor function. The Bessel functions are denoted by I_v, J_v, K_v as usual. We use the term K-Bessel function to indicate K_v without the specification of the order v; and the same convention applies to other Bessel functions as well. The symbol Γ is for the gamma function, and Γ is for the full modular group introduced in Section 1.1. The dependency of implied constants on others will not always be explained, since it is more or less clear from the context.

Some knowledge of integrals involving basic transcendental functions is certainly helpful. For this purpose Lebedev's book [38] is quite handy. But there are occasions when Titchmarsh [69], Watson [74], and Whittaker and Watson [75] give more precise information, though proofs of most integral formulas and relevant estimates are given or at least briefly indicated either in the text or in the respective notes. In addition to these standard books, Sonine's article [68], Vilenkin's book [71], and the table [15] by Gradshteyn and Ryzhik are recommended.

Readers are supposed to have ample knowledge of the zeta-function such as that developed in Titchmarsh [70] as well as in Ivić [19]. In fact this monograph is, in part, a continuation of their books. Thus, for

instance, bounds like

$$(\log t)^{-1} \ll |\zeta(1 + it)| \ll \log t, \qquad \zeta(\tfrac{1}{2} + it) \ll t^{\frac{1}{6}} \log t,$$

$$\int_0^T |\zeta(\tfrac{1}{2} + iu)|^4 du \ll T(\log T)^4 \quad (t, T \geq 2)$$

are used freely under the term *classical estimates*. On the other hand no experience in the theory of automorphic functions is assumed. The first three chapters can be taken for an introduction to the subject.

The references are limited to the essentials. Suggestions for further readings may be found in the notes and the articles quoted there.

Contents

1

Non-Euclidean harmonics

Our story of the Riemann zeta-function is to be unfolded on a stage filled with non-Euclidean harmonics. Accordingly we need first to tune our principal instrument. We are going to prove in this initial chapter a spectral resolution of the non-Euclidean Laplacian

$$\Delta = -y^2((\partial/\partial x)^2 + (\partial/\partial y)^2)$$

with minimum prerequisites. The entire theory originated in a seminal work of H. Maass, which was later developed by W. Roelcke, A. Selberg, and many others. Our account is an elementary approach to their theory in the case of the full modular group. Despite this specialization it will not be hard to see that our argument extends to general arithmetic situations.

1.1 Basic concepts

To begin with, we shall equip the upper half plane

$$\mathcal{H} = \{z = x + iy \ : \ -\infty < x < \infty, \ y > 0\}$$

with the non-Euclidean differentiable structure. For this purpose we introduce the group $\mathbb{T}(\mathcal{H})$ consisting of all real fractional linear transformations

$$\gamma : z \mapsto \frac{az + b}{lz + h} \qquad (ah - bl = 1; \, a, b, l, h \in \mathbb{R}). \qquad (1.1.1)$$

The γ's map $\mathcal{H}$ onto itself conformally. To see this it is enough to note that γ has the inverse map $z \mapsto (hz - b)/(-lz + a)$, and that

$$\operatorname{Im}\gamma(z) = \frac{y}{|lz + h|^2}, \qquad \frac{d}{dz}\gamma(z) = \frac{1}{(lz + h)^2}. \qquad (1.1.2)$$

1

The elements of $\mathbb{T}(\mathcal{H})$ are also rigid motions acting on $\mathcal{H}$ in the sense that $\mathcal{H}$ carries the non-Euclidean metric

$$y^{-1}|dz| = y^{-1}((dx)^2 + (dy)^2)^{\frac{1}{2}}$$

which is invariant with respect to any $\gamma \in \mathbb{T}(\mathcal{H})$. This is a simple consequence of the relations in (1.1.2), since they imply

$$(\mathrm{Im}\,\gamma(z))^{-1}\Big|\frac{d}{dz}\gamma(z)\Big| = (\mathrm{Im}\,z)^{-1}.$$

The invariance of the metric is inherited by Δ as being the negative of the corresponding Laplace–Beltrami operator. It can also be checked by direct computation: Putting $f(x,y) = F(u,v)$ with $\gamma(x+iy) = u+iv$ and invoking the Cauchy–Riemann equation for the function $\gamma(z)$, we have

$$\Delta f(x,y) = -y^2(u_x^2 + v_x^2)(F_{uu} + F_{vv})$$
$$= -y^2\Big|\frac{d}{dz}\gamma(z)\Big|^2(F_{uu} + F_{vv}) = -v^2(F_{uu} + F_{vv}),$$

which amounts to $\Delta \cdot \gamma = \gamma \cdot \Delta$, i.e., the invariance of Δ. We have also the invariance of the non-Euclidean area element

$$d\mu(z) = y^{-2}dxdy$$

induced by the metric. This can be confirmed by computing the Jacobian of the map γ :

$$\Big|\frac{\partial(u,v)}{\partial(x,y)}\Big| = \Big|\frac{d}{dz}\gamma(z)\Big|^2 = (v/y)^2.$$

Further, we have the invariance of the non-Euclidean outer-normal derivative

$$y\frac{\partial}{\partial n} = y\Big\{\frac{dy}{|dz|}\frac{\partial}{\partial x} - \frac{dx}{|dz|}\frac{\partial}{\partial y}\Big\}$$

taken along any piecewise smooth curve in $\mathcal{H}$: In fact we have, for f, F as above,

$$y\frac{\partial f}{\partial n} = y\Big\{\frac{dy}{|dz|}(F_u u_x + F_v v_x) - \frac{dx}{|dz|}(F_u u_y + F_v v_y)\Big\}$$
$$= y\Big\{\frac{dy}{|dz|}(F_u v_y - F_v u_y) - \frac{dx}{|dz|}(-F_u v_x + F_v u_x)\Big\}$$
$$= y\Big\{\frac{dv}{|dz|}F_u - \frac{du}{|dz|}F_v\Big\} = v\frac{\partial F}{\partial n}.$$

This will be used in conjunction with Green's formula, which is a basic tool in the discussion below.

We next define the full modular group Γ as the subgroup of $\mathbb{T}(\mathcal{H})$ that is composed of those maps with $a, b, l, h \in \mathbb{Z}$ in (1.1.1). This signifies in particular that we do not regard Γ as a matrix group. Thus, if an element of Γ is pulled back to $SL(2, \mathbb{Z})$ in an obvious way, then we get two image matrices with corresponding entries having opposite signs; that is,

$$\Gamma \cong SL(2, \mathbb{Z})/\{\pm 1\}.$$

In any event readers should bear in mind that we are always dealing with transformations of $\mathcal{H}$.

The most basic fact about the motions caused by the elements of Γ is that they are discontinuous. This means that the action of Γ on $\mathcal{H}$ is comparable to, e.g., that of the group generated by two independent linear translations acting on the Euclidean plane, which is equivalent to tessellating $\mathbb{C}$ with congruent parallelograms. To make the situation with Γ explicit, we introduce the fundamental domain of Γ

$$\mathcal{F} = \left\{ z \in \mathcal{H} : |z| \geq 1, |x| \leq \tfrac{1}{2} \right\}, \tag{1.1.3}$$

and also the notation

$$z \equiv z' \bmod \Gamma$$

that indicates the existence of a $\gamma \in \Gamma$ such that $\gamma(z) = z'$. Then we have

Lemma 1.1 *The family of domains $\{\gamma(\mathcal{F}), \gamma \in \Gamma\}$ induces a tessellation of $\mathcal{H}$.*

Proof We fix an arbitrary $z \in \mathcal{H}$, and consider $\max[\operatorname{Im} \gamma(z)]$ as γ given in (1.1.1) varies in Γ. This should exist. For the first relation in (1.1.2) implies that $\operatorname{Im} \gamma(z)$ takes its maximum when $|lz + h|$ takes its minimum; and the latter can readily be seen to exist by observing that $lz + h$, $\gamma \in \Gamma$, are among the lattice points generated by 1 and z. We assume that $z_0 = x_0 + iy_0$ has the maximum imaginary part in this context; naturally we may assume also that $|x_0| \leq \tfrac{1}{2}$. Then we note that $-1/z_0 \equiv z_0 \bmod \Gamma$, and thus $\operatorname{Im}(-1/z_0) = y_0|z_0|^{-2} \leq y_0$. Hence we have $|z_0| \geq 1$, namely $z_0 \in \mathcal{F}$. This means that the tiles $\gamma(\mathcal{F})$, $\gamma \in \Gamma$, cover $\mathcal{H}$. We shall next show that these tiles have common points only on their boundaries. This is clearly equivalent to the assertion that if $z_1 = x_1 + iy_1$ and $\gamma(z_1) = x_2 + iy_2$ with a non-trivial $\gamma \in \Gamma$ are in $\mathcal{F}$ then both are on $\partial \mathcal{F}$, the boundary of $\mathcal{F}$. To see this let γ be as in (1.1.1) with integral coefficients. Obviously we may suppose also that $l \geq 0$ as well as $y_2 \geq y_1$. Comparing the imaginary

parts of z_1 and $\gamma(z_1)$ we have $|lz_1 + h| \leq 1$, which implies $ly_1 \leq 1$. On noting that $y_1 \geq \frac{1}{2}\sqrt{3}$, we find that l is equal either to 0 or to 1. If $l = 0$ then $ah = 1$; and $\gamma(z) = z \pm b$. Here $b = 0$ is excluded because of an obvious reason. Thus $b = \pm 1$, and $z_1, \gamma(z_1) \in \partial\mathcal{F}$. On the other hand, if $l = 1$ then $|z_1 + h| \leq 1$. Thus $|h| \leq 1$. If $h = 0$ then $|z_1| = 1$ and $\gamma(z) = a - 1/z$; and if $h = \pm 1$ then $z_1 = \frac{1}{2}(\mp 1 + i\sqrt{3})$ and $\gamma(z) = a - 1/(z \pm 1)$. We readily get $|a| \leq 1$, and hence $z_1, \gamma(z_1) \in \partial\mathcal{F}$ again. This ends the proof.

We note that tiles $\gamma(\mathcal{F})$, $\gamma \in \Gamma$, are generally different in shape for our Euclidean eyes, but if they were corrected with the metric $y^{-1}|dz|$ they would look just like each other. We remark also that the left and the right vertical edges of $\mathcal{F}$ are obviously equivalent to each other mod Γ, and that the circular part of $\partial\mathcal{F}$ is mapped onto itself by $z \mapsto -1/z$; thus the left half of the arc is equivalent to the right half. The identification of the equivalent boundary elements of $\mathcal{F}$ yields a punctured Riemann surface. The puncture corresponds to the point at infinity, and will be called the cusp of Γ in the sequel. The Riemann surface thus obtained is designated as the *manifold* $\mathcal{F}$, which carries the metric $y^{-1}|dz|$ with an obvious localization. Without this specification the symbol $\mathcal{F}$ stands for the fundamental domain of Γ. In passing we stress that the possible overlapping of $\gamma(\mathcal{F})$'s on their boundaries will not raise any pathological situations in our later discussion.

Turning to the analytical aspect, we introduce the concept of automorphy: A function f defined on $\mathcal{H}$ is said to be Γ-automorphic if $f(\gamma(z)) = f(z)$ for all $\gamma \in \Gamma$. This is the same as to have f defined originally on the manifold $\mathcal{F}$ and to view it as a function over $\mathcal{H}$ in an obvious way. In this context the invariance of Δ means precisely that Δ can be regarded as a differential operator acting on the manifold $\mathcal{F}$.

A very important example of Γ-automorphic functions is the Poincaré series: We put, for a non-negative integer m and a complex number s,

$$P_m(z,s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} (\operatorname{Im} \gamma(z))^s e(m\gamma(z)) \qquad (z \in \mathcal{H}, \operatorname{Re} s > 1), \qquad (1.1.4)$$

where $e(z) = \exp(2\pi i z)$, Γ_∞ is the stabilizer in Γ of the cusp, i.e., the cyclic subgroup generated by the translation $z \mapsto z + 1$, and γ runs over a representative set of the left cosets of Γ_∞ in Γ. The summands are independent of the choice of the representatives, and the sum converges absolutely, as can be seen from the expression (1.1.5) below. Hence $P_m(z,s)$ is Γ-automorphic. We note that the relation $\eta\gamma^{-1} \in \Gamma_\infty$, where $\gamma, \eta \in \Gamma, \gamma(z) = (az + b)/(lz + h), \eta(z) = (a'z + b')/(l'z + h')$ with $l, l' \geq 0$,

is equivalent either to $l = l'$, $h = h'$ or to $\gamma, \eta \in \Gamma_\infty$. Also we observe that for $l > 0$

$$\gamma(z) = \frac{a}{l} - \frac{1}{l(lz + h)}, \quad ah \equiv 1 \bmod l.$$

Thus we have

$$P_m(z,s) = y^s e(mz) + y^s \sum_{l=1}^{\infty} \sum_{\substack{h=-\infty \\ (h,l)=1}}^{\infty} |lz + h|^{-2s} e\left(\frac{mh^\bullet}{l}\right) \exp\left(-\frac{2\pi mi}{l(lz + h)}\right)$$

(1.1.5)

with $hh^\bullet \equiv 1 \bmod l$. We classify the summands according to $h \bmod l$, so that we have

$$P_m(z,s) = y^s e(mz) + y^s \sum_{l=1}^{\infty} l^{-2s} \sum_{\substack{h=1 \\ (h,l)=1}}^{l} e(mh^\bullet/l)$$

$$\times \sum_{n=-\infty}^{\infty} |z + h/l + n|^{-2s} \exp\left(-\frac{2\pi mi}{l^2(z + h/l + n)}\right).$$

Applying Poisson's sum-formula to the last sum, we get, for $\mathrm{Re}\, s > 1$,

$$P_m(z,s) = y^s e(mz) + y^{1-s} \sum_{n=-\infty}^{\infty} e(nx) \sum_{l=1}^{\infty} l^{-2s} S(m,n;l)$$

$$\times \int_{-\infty}^{\infty} \exp\left(-2\pi ny\xi i - \frac{2\pi m}{l^2 y(1 - \xi i)}\right)(1 + \xi^2)^{-s} d\xi, \quad (1.1.6)$$

where

$$S(m,n;l) = \sum_{\substack{h=1 \\ (h,l)=1}}^{l} e\big((hm + h^\bullet n)/l\big) \quad (1.1.7)$$

is the Kloosterman sum. Here we have performed an exchange of the order of summation, which is legitimate because the convergence is absolute, as can be seen by shifting the path to $\mathrm{Im}\, \xi = -\frac{1}{2}\mathrm{sgn}(n)$ in the integral. The formula (1.1.6) will be used in the next chapter.

In particular we put

$$E(z,s) = P_0(z,s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} (\mathrm{Im}\, \gamma(z))^s, \quad (1.1.8)$$

and call it the Eisenstein series attached to Γ. This function will appear at various important stages of our discussion. Its principal properties are collected in

Lemma 1.2 *For any* $z \in \mathcal{H}$ *the function* $E(z,s)$ *is meromorphic in s over the whole of* $\mathbb{C}$, *and we have the expansion*

$$E(z,s) = y^s + \varphi_\Gamma(s)y^{1-s}$$
$$+ \frac{2\pi^s\sqrt{y}}{\Gamma(s)\zeta(2s)}\sum_{n\neq 0}|n|^{s-\frac{1}{2}}\sigma_{1-2s}(|n|)K_{s-\frac{1}{2}}(2\pi|n|y)e(nx), \qquad (1.1.9)$$

where

$$\varphi_\Gamma(s) = \sqrt{\pi}\frac{\Gamma(s-\frac{1}{2})\zeta(2s-1)}{\Gamma(s)\zeta(2s)} = \frac{\pi^{2s-1}\Gamma(1-s)\zeta(2(1-s))}{\Gamma(s)\zeta(2s)}. \qquad (1.1.10)$$

Hence $E(z,s)$ *is regular for* $\mathrm{Re}\, s \geq \frac{1}{2}$ *save for the simple pole at* $s = 1$ *with residue* $3/\pi$, *and satisfies the functional equation*

$$E(z,s) = \varphi_\Gamma(s)E(z,1-s) \qquad (1.1.11)$$

as well as the differential equation

$$\Delta E(z,s) = s(1-s)E(z,s). \qquad (1.1.12)$$

Proof We invoke first the functional equation

$$\zeta(s) = 2^s\pi^{s-1}\sin(\tfrac{1}{2}s\pi)\Gamma(1-s)\zeta(1-s)$$
$$= \pi^{s-\frac{1}{2}}\frac{\Gamma(\frac{1}{2}(1-s))}{\Gamma(\frac{1}{2}s)}\zeta(1-s) \qquad (1.1.13)$$

and the representation

$$\sigma_\xi(n) = \zeta(1-\xi)\sum_{l=1}^{\infty}l^{\xi-1}c_l(n) \quad (n > 0,\ \mathrm{Re}\,\xi < 0), \qquad (1.1.14)$$

where

$$c_l(n) = S(n,0;l) = \sum_{\substack{h=1\\(h,l)=1}}^{l}e(nh/l) \qquad (1.1.15)$$

is the Ramanujan sum. When $\mathrm{Re}\, s > 1$ the expansion (1.1.9) is a consequence of the relations (1.1.6) and (1.1.14) with (1.1.13), since we have

$$\int_{-\infty}^{\infty}\frac{1}{(1+\xi^2)^{\nu+\frac{1}{2}}}d\xi = \sqrt{\pi}\frac{\Gamma(\nu)}{\Gamma(\nu+\frac{1}{2})} \qquad (\mathrm{Re}\,\nu > 0)$$

and

$$\int_{-\infty}^{\infty}\frac{\cos(y\xi)}{(1+\xi^2)^{\nu+\frac{1}{2}}}d\xi = 2\sqrt{\pi}\frac{(y/2)^\nu}{\Gamma(\nu+\frac{1}{2})}K_\nu(y) \qquad (y > 0;\ \mathrm{Re}\,\nu > -\tfrac{1}{2}).$$
$$(1.1.16)$$

We then note the other representation

$$K_\nu(y) = \frac{1}{2}\int_0^\infty \xi^{\nu-1}\exp\left(-\tfrac{1}{2}y(\xi+\xi^{-1})\right)d\xi, \qquad (1.1.17)$$

which shows that $K_\nu(y)$ is entire in ν, $K_\nu(y) = K_{-\nu}(y)$ for $\mathrm{Re}\, y > 0$, and moreover

$$K_\nu(y) = (1+o(1))\left(\frac{\pi}{2y}\right)^{\frac{1}{2}}e^{-y} \qquad (y\to+\infty) \qquad (1.1.18)$$

for any fixed ν. In fact this asymptotic formula can readily be proved by putting $\xi = 1+r$ in (1.1.17) and observing that the main part of the integral comes from the short interval $|r| \le y^{-2/5}$. Hence $E(z,s)$ exists as a meromorphic function of s over $\mathbb{C}$. The assertion (1.1.11) follows from (1.1.13) and these properties of $K_\nu(y)$. As to (1.1.12) it is a consequence of the definition (1.1.8) and

$$\Delta[(\mathrm{Im}\,\gamma(z))^s] = s(1-s)(\mathrm{Im}\,\gamma(z))^s \qquad (s\in\mathbb{C},\, \gamma\in\mathbb{T}(\mathcal{H})), \qquad (1.1.19)$$

which is due to the invariance of Δ. An alternative proof is to use the expansion (1.1.9) and the fact that $\sqrt{y}K_{s-\frac{1}{2}}(2\pi|n|y)$ is a solution of the differential equation

$$[D_{s,n}g](y) \equiv -y^2g''(y) + ((2\pi ny)^2 + s(s-1))g(y) = 0 \quad (y>0) \quad (1.1.20)$$

(see the next lemma). This ends the proof of the lemma.

It is appropriate to make here a little digression about the nature of the differential operator $D_{s,n}$: It is a result of the application of the separation of variables to the operator $\Delta + s(s-1)$. In fact we have formally

$$(\Delta + s(s-1))\left\{\sum_n a_n(y)e(nx)\right\} = \sum_n [D_{s,n}a_n](y)e(nx). \qquad (1.1.21)$$

This relation and the following assertion will be used in our later discussion.

Lemma 1.3 *The differential equation (1.1.20) with $n > 0$ has linearly independent solutions $\sqrt{y}K_{s-\frac{1}{2}}(2\pi ny)$ and $\sqrt{y}I_{s-\frac{1}{2}}(2\pi ny)$. Thus the resolvent kernel of the differential operator $y^{-2}D_{s,n}$, $n > 0$, is equal to*

$$g_{s,n}(y,v) = \begin{cases} \sqrt{vy}I_{s-\frac{1}{2}}(2\pi nv)K_{s-\frac{1}{2}}(2\pi ny) & \text{if } v \le y, \\[2mm] \sqrt{vy}I_{s-\frac{1}{2}}(2\pi ny)K_{s-\frac{1}{2}}(2\pi nv) & \text{if } v \ge y. \end{cases} \qquad (1.1.22)$$

Proof That the above two functions are solutions of (1.1.20), and that the Wronskian of these is equal to 1, can be checked by using the recurrence relations

$$I_{v-1}(z) + I_{v+1}(z) = 2I_v'(z), \quad I_{v-1}(z) - I_{v+1}(z) = 2vz^{-1}I_v(z),$$
$$K_{v-1}(z) + K_{v+1}(z) = -2K_v'(z), \quad K_{v-1}(z) - K_{v+1}(z) = -2vz^{-1}K_v(z),$$
$$(1.1.23)$$

and also the definitions

$$I_v(z) = \sum_{k=0}^{\infty} \frac{(z/2)^{v+2k}}{\Gamma(k+1)\Gamma(v+k+1)}, \qquad (1.1.24)$$

$$K_v(z) = \frac{\pi}{2\sin(\pi v)}\{I_{-v}(z) - I_v(z)\}, \qquad (1.1.25)$$

where $z^v = \exp(v \log z)$ with $|\arg z| < \pi$. The rest of the proof is a standard application of the general theory of ordinary differential equations. The excluded case $n = 0$ is easy, and left for readers. This ends the proof.

Proceeding to our main issue, we let $L^2(\mathcal{F}, d\mu)$ stand for the set of all Γ-automorphic f's such that

$$\|f\|^2 = \int_{\mathcal{F}} |f(z)|^2 d\mu(z) < +\infty.$$

It should be observed that $\mathbb{C} \subset L^2(\mathcal{F}, d\mu)$, since we have

$$\int_{\mathcal{F}} d\mu(z) = \frac{\pi}{3}.$$

The set $L^2(\mathcal{F}, d\mu)$ is a Hilbert space equipped with the Petersson inner-product

$$\langle f_1, f_2 \rangle = \int_{\mathcal{F}} f_1(z)\overline{f_2(z)}d\mu(z). \qquad (1.1.26)$$

We are going to diagonalize the operator Δ in $L^2(\mathcal{F}, d\mu)$; that is, we shall try to find a set of Γ-automorphic functions which spans $L^2(\mathcal{F}, d\mu)$ and in which the operator Δ is well-defined and reduces, in an informal sense, to a scalar multiplication at each element. To this end we introduce the linear set defined on the manifold $\mathcal{F}$:

$$B^{\infty}(\mathcal{F}) = \{f \in C^{\infty}(\mathcal{F}) :$$

each partial derivative of $f(z)$ is of rapid decay$\}$.   (1.1.27)

Here, that a Γ-automorphic function $g(z)$ is of rapid decay means that

$g(z) = O(y^{-M})$ for any $M > 0$ as $z \in \mathcal{F}$ tends to the cusp, where the implied constant may depend on M. This is dense in $L^2(\mathcal{F}, d\mu)$, for it contains all C^∞-functions having compact supports on the manifold $\mathcal{F}$. Then Δ is a symmetric operator in the sense that

$$\langle \Delta f_1, f_2 \rangle = \langle f_1, \Delta f_2 \rangle \qquad (f_1, f_2 \in B^\infty(\mathcal{F})). \tag{1.1.28}$$

In fact we have, by Green's formula,

$$\langle \Delta f_1, f_2 \rangle = \int_{\mathcal{F}} y^2 \nabla f_1(z) \cdot \overline{\nabla f_2(z)} d\mu(z), \tag{1.1.29}$$

where ∇ denotes the ordinary gradient. For the manifold $\mathcal{F}$ has no boundary, and the integrand is of rapid decay around the cusp. More precisely, we apply Green's formula to

$$\langle \Delta f_1, f_2 \rangle = \lim_{Y \to \infty} \int_{\mathcal{F}_Y} \Delta f_1(z) \overline{f_2(z)} d\mu(z), \tag{1.1.30}$$

where $\mathcal{F}_Y = \mathcal{F} \cap \{z : \operatorname{Im} z \leq Y\}$. We have

$$
\begin{aligned}
\langle \Delta f_1, f_2 \rangle &= \lim_{Y \to \infty} \int_{\mathcal{F}_Y} \nabla f_1(z) \cdot \overline{\nabla f_2(z)} \, dx dy \\
&\quad - \lim_{Y \to \infty} \int_{\partial \mathcal{F}_Y} y \frac{\partial f_1}{\partial n}(z) \overline{f_2(z)} \frac{|dz|}{y} \\
&= \int_{\mathcal{F}} \nabla f_1(z) \cdot \overline{\nabla f_2(z)} \, dx dy \\
&\quad - \lim_{Y \to \infty} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\partial f_1}{\partial y}(x + iY) \overline{f_2(x + iY)} dx, \tag{1.1.31}
\end{aligned}
$$

which gives (1.1.29). Here we have used the invariance of $(\partial f_1 / \partial n) \overline{f_2} |dz|$ and the consequential cancellation of integrals on the boundary elements of $\mathcal{F}$ which are equivalent mod Γ to each other. This cancellation is due to the invariance of the non-Euclidean outer-normal derivative, and to the reverse of orientation in the corresponding boundary elements. Intuitively it is the same as what happens when the process of folding and pasting is applied to $\mathcal{F}$ to transform it into a Riemann surface.

The formula (1.1.29) implies, in particular,

$$\langle \Delta f, f \rangle > 0 \tag{1.1.32}$$

for any non-constant $f \in B^\infty(\mathcal{F})$. This and (1.1.28) mean that Δ is a semi-bounded symmetric operator which has a dense domain in $L^2(\mathcal{F}, d\mu)$. Thus we could appeal to the general theory on the self-adjoint extension of such an operator with the effect of a shorter presentation. We