Journal of the Indian Math. Soc.
52 (1987) 91–126

AN INTRODUCTION TO SELBERG'S TRACE FORMULA

M. RAM MURTY

[Received June 22, 1987]

[*Dedicated to the memory of Srinivasa Ramanujan*]

1. THE ANALOGY

Since the dawn of history, humankind has been fascinated with prime numbers. They followed no pattern of generation and enigmatic problems could be posed in simple ways. The problem of the distribution of prime numbers was one such problem that was seminal in the creation of mathematics.

It was Gauss who first observed (at the age of 17) a statistical pattern in the distribution of prime numbers. If we let $\pi(x)$ denote the numbers of prime numbers upto x, then Gauss conjectured that

$$\pi(x) \sim \int_2^x \frac{dt}{\log t}.$$

The proof had to wait for the creation of the field of complex analysis. Indeed, in his epoch making paper of 1860, Riemann [7] outlined a method of attacking the conjecture of Gauss. His central idea was the introduction of the analytic function, now called Riemann's zeta function, that is defined by

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

for $Re(s) > 1$. The connection with prime numbers comes via the identity:

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1},$$

where the product is over prime numbers. This identity is equivalent to the unique factorization of the natural numbers into prime numbers. (This product representation of $\zeta(s)$ was first noticed much earlier by Euler, who had worked with $\zeta(s)$ as a funcion of a real variable.) In his paper of 1860, Riemann made the following assertions:

(i) $\zeta(s)$ has an analytic continuation to the entire complex plane except for a simple pole at $s = 1$ and satisfies the functional equation:

$$\pi^{-s/2}\,\Gamma(s/2)\,\zeta(s) = \pi^{(1-s)/2}\,\Gamma((1-s)/2)\,\zeta(1-s).$$

(ii) Apart from trivial zeros at $s = -2, -4, \ldots$ all of the zeroes of $\zeta(s)$ lie in the critical strip $0 \leqslant Re(s) \leqslant 1$. Moreover, if $N(T)$ denotes the number of zeroes ρ satisfying $| Im\, \rho | \leqslant T$, then

$$N(T) = \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + O(\log T)$$

(iii) If we define,

$$\Lambda(n) = \begin{cases} \log p, & \text{if } n = p^{\alpha}, \\ 0, & \text{otherwise} \end{cases}$$

and set

$$\psi(x) = \sum_{n \leqslant x} \Lambda(n),$$

then

$$\psi(x) = x - \sum_{\rho} \frac{x^{\rho}}{\rho} - \frac{1}{2} \log(1 - x^{-2}) - \frac{\zeta'}{\zeta}(0)$$

where the sum is over the non-trivial zeroes of the Riemann zeta function.

(iv) All the non-trivial zeroes ρ of $\zeta(s)$ satisfy $Re\,\rho = \frac{1}{2}$.

This program of Riemann motivated the phenomenal development of complex analysis. By 1900, the first three assertions were rigorously proved and the foundations of complex analysis were laid. The conjecture of Gauss in the equivalent form

$$\psi(x) \sim x$$

turned out to be a consequence of the fact that the ζ function did not vanish on the line $Re\,(s) = 1$. This was proved independently by Hadamard and de la Vallée Poussin. But assertion (iv) which is the famous Riemann hypothesis, still remains a mystery.

The explicit formula (iii) connects prime numbers with zeroes of the

Riemann zeta function. More generally, Weil [9] noticed that the methods used to prove (iii) can be utilised to show the following.

Let

$$\rho = \tfrac{1}{2} + i\gamma, \quad \gamma \in \mathbf{C}$$

be a typical non-trivial zero of $\zeta(s)$, and let $h(r)$ be any analytic function on $| Im(r) | \leqslant \frac{1}{2} + \delta$ such that

(*) $\qquad h(-r) = h(r)$, *and* $| h(r) | \leqslant A(1 + | r |)^{-2-\delta}$ $(A > 0, \delta > 0)$.

Let

$$g(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h(r)\, e^{-iru}\, dr$$

be its Fourier transform. Then, we have:

Weil's explicit formula

Let h satisfy (*). Then,

$$\sum_{\gamma} h(\gamma) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h(r) \frac{\Gamma'}{\Gamma} \left(\frac{1}{4} + \frac{1}{2}\, ir\right) dr + h(i/2)$$

$$+ h(-i/2) - g(0) \log \pi - 2 \sum_{n=1}^{\infty} \frac{\Lambda(n)}{\sqrt{n}} g(\log n). \qquad (1)$$

The sum is over all non-trivial zeroes of $\zeta(s)$. This formula has an important connection with the Riemann hypothesis. Suppose that

$$g(u) = \int_{-\infty}^{\infty} g_0(t + u)\, \overline{g_0(t)} dt \quad (g_0 \textit{ even}).$$

Then, observe that $h(r) = h_0(r)\overline{h_0(\bar{r})}$ for $| Im(r) | \leqslant \frac{1}{2} + \delta$. Weil showed that the Riemann hypothesis is equivalent to showing that (1) is positive for all such g.

In 1950, Selberg [8] noticed a remarkable similarity between this formula and a similar one, which he derived connecting the eigenvalues of the Laplacian and the lengths of closed geodesics on Riemannian manifolds.

Indeed, let F be a compact Riemann surface of genus $g \geqslant 2$. The universal covering space of F' is conformally equivalent to the upper half plane

$$H = \{x + iy \colon y > 0\}.$$

We can express F in the form $\Gamma \backslash H$ where Γ is a strictly hyperbolic Fuchsian group. The group $\Gamma \subseteq PSL_2(\mathbf{R})$ will have a compact fundamental polygon $\mathcal{F}$. Recall that $\Gamma \simeq \pi_1(F)$, the fundamental group of F. There is the hyperbolic (Poincaré) metric on H:

$$ds = \frac{|dz|}{y},$$

whose corresponding area element is

$$d\mu(z) = \frac{dxdy}{y^2},\ z = x + iy.$$

The hyperbolic distance is given by

$$d(z_1, z_2) = arg\ cosh\left[1 + \frac{|z_1 - z_2|^2}{2y_1y_2}\right].$$

It is easily checked that this distance is invariant under the action of $SL_2(\mathbf{R})$. We consider

$$L^2(\Gamma \backslash H) = \{f \colon H \to \mathbf{C} \colon \iint_{\mathcal{F}} |f(z)|^2\, d\mu(z) < \infty, f(\gamma z) = f(z)$$

$$\forall \gamma \in \Gamma, z \in H\}.$$

Such functions are called *automorphic*. We can introduce an inner product on $L^2(\Gamma \backslash H)$ as follows:

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

The Laplace operator is the non-Euclidean Laplacian:

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

We write $\lambda_0 \leqslant \lambda_1 < \dots$ for the sequence of eigenvalues and $\phi_0, \phi_1, \dots$ for the corresponding eigenfunctions.

Let σ be an element of $PSL_2(\mathbf{R})$, which is not the identity. Then, the Jordan canonical form of σ is one of the following:

$$(i) \quad \begin{pmatrix} \omega & \\ & \bar{\omega} \end{pmatrix}, \quad \omega \in \mathbf{C}, |\omega| = 1$$

$$(ii) \quad \begin{pmatrix} t & \\ & t^{-1} \end{pmatrix}, \quad t \in \mathbf{R}$$

$$(iii) \quad \begin{pmatrix} 1 & x \\ & 1 \end{pmatrix}, \quad x \in \mathbf{R}$$

σ is called *elliptic, hyperbolic parabolic* according as the Jordan canonical form of σ is of type (i), (ii), (iii) respectively.

Now let us consider the hyperbolic conjugacy classes of Γ. We define the *norm* of P as t^2 and write $N(P) = t^2$. Moreover, the centraliser of P in Γ, denoted $C_\Gamma(P)$, must reduce to a cyclic group since Γ acts discontinuously on the upper half plane. Indeed, it is easily verified that $C_\Gamma(P)$ is generated by a primitive hyperbolic element P_0. Thus, if P is a hyperbolic element, $P = P^k$ for some integer k, and P_0 is a unique primitive hyperbolic element. We define $\Lambda(P) = \log N(P_0)$. Selberg proved the following:

Trace formula:

Let $h(r)$ denote any analytic function on $|Im(r)| \leqslant \frac{1}{2} + \delta$ such that

$$h(-r) = h(r) \text{ and } |h(r)| \leqslant A(1 + |r|)^{-2-\delta}, (A > 0, \delta > 0).$$

Then

$$\sum_{n=0}^{\infty} h(r_n) = \frac{vol(\mathcal{F})}{4\pi} \int_{-\infty}^{\infty} rh(r) \tanh \pi r dr + \sum_P \frac{\Lambda(P)}{N(P)^{1/2} - N(P)^{-1/2}} g(\log N(P))$$

where the sum over P is taken over all hyperbolic conjugacy classes and

$$s_n = \tfrac{1}{2} + ir_n, \lambda_n = s_n(1 - s_n).$$

Selberg found a zeta function which he denoted by $Z(s)$, and which corresponds to this formula in the same way as Weil's formula corres ponds to the Riemann zeta function. By means of his trace formula, h showed that the non-trivial zeroes of $Z(s)$ are located on the line $Re\, s = \frac{1}{2}$ It is this analogy that we would like to pursue in this short survey in th hope of gaining a better understanding of the elusive Riemann hypothesi For the sake of clarity, we suppress tedious calculations and proofs o technical lemmas. The interested reader may find all of these details i Hejhal [3, 4].

Acknowledgements. This survey is based on a seminar of informa lectures given by the author at Carleton University, Ottawa, Canad during May, 1987. The author would like to thank Professors R. Bedard J. Dixon, and M. Rahman for their participation and enthusiasm.

2. EIGENVALUE PROBLEMS

Let $D \subseteq \mathbf{C}$ be a smoothly bounded domain with connectivity $p < \infty$ Stretching a membrane across D and clamping it along the boundary o D, the oscillations $\psi(z, t)$ satisfy the wave equation

$$\frac{\partial^2\psi}{\partial t^2} = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2}, \; z = x + iy$$

with the boundary condition

$$\psi(z, t) = 0, \; z \in \partial D.$$

Separation of variables leads to

$$\psi(z, t) = u(z) \exp (i\sqrt{\lambda}t)$$

where

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \lambda u = 0$$

with the boundary condition

$$u(z) = 0, \; z \in \partial D.$$

These boundary conditions force $\lambda \geqslant 0$.

This is the classical Dirichlet problem and it is well known (see [1]) that the eigenvalues form an increasing sequence :

$$0 < \lambda_1 \leqslant \lambda_2 \leqslant \cdots$$

with corresponding eigenfunctions ϕ_n suitably normalised so that

$$\iint_D |\phi_n(z)|^2 \, dxdy = 1.$$

Green's theorem and the boundary condition immediately imply that these eigenfunctions are orthonormal. Moreover, these eigenfunctions form an orthonormal basis for $L^2(D)$. Any twice differentiable and compactly supported function f on D has a "Fourier" expansion of the form

$$f(z) = \sum_{n=1}^{\infty} c_n \phi_n(z).$$

In general, not much is known about the eigenvalues λ_n and we must be content with asymptotic information of the following kind.

Let $N(x)$ denote number of eigenvalues $\lambda_n \leqslant x$. Then, H. Weyl proved in 1912 that

$$N(x) \sim \frac{area\ (D)}{4\pi}\ x$$

as $x \to \infty$. In 1920, Courant showed that

$$N(x) = \frac{area\ (D)}{4\pi}\ x + O(x^{1/2} \log x).$$

The remainder function defined by

$$R(x) = N(x) - \frac{area(D)}{4\pi}\ x$$

satisfies

$$\int_0^{\infty} e^{-tx} dR(x) \sim \frac{-L}{8\sqrt{\pi t}}$$

as $t \to 0$. This would suggest that as $x \to \infty$,

$$R(x) \sim -\frac{L}{4\pi}\sqrt{x}$$

by analogy with classical Tauberian theorems But unfortunately, $dR(x)$ is a signed measure and its behaviour is a mystery. There are several conjectures, suggested mainly from problems in number theory, that are still unresolved.

Conjectures

(i) (Polya) $R(x) \leqslant 0$ for all $x \geqslant 0$.

(ii) (I.M. Vinogradov)

$$R(x) = -\frac{L}{4\pi}\sqrt{x} + O(x^{1/3}).$$

If we consider the special case when D is the unit square, then separation of variables leads to

$$u(z) = (\sin \pi m x)(\sin \pi n y)$$

with corresponding eigenvalue $\lambda = \pi^2(m^2 + n^2)$. Hence, $N(x)$ is related to the classical circle problem of counting the number of lattice points in a circle.

There are similar results and questions for eigenvalue problems on general Riemannian manifolds. We restrict ourselves to the case of a compact Riemann surface $\mathcal{F}$. It is assumed that $\mathcal{F}$ carries a C^∞ metric:

$$ds^2 = \sum_{j,k} g_{jk} dx_j dx_k.$$

We now form the Laplace-Beltrami operator:

$$\Delta u = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x_j}\left(\sqrt{g} g^{jk} \frac{\partial u}{\partial x_k}\right).$$

Here (g^{jk}) is the matrix inverse to (g_{jk}) and g denotes the absolute value of $det\,(g_{jk})$. This operator is strongly elliptic on $\mathcal{F}$ and commutes with the group of isometries. We consider the following eigenvalue problem:

$$\Delta u + \lambda u = 0$$

on $\mathcal{F}$. Since ∂F is empty, there are no boundary conditions. The eigenvalues and eigenfunctions have the usual properties. One can establish:

THEOREM:

$$\sum_{n=0}^{\infty} e^{-\lambda_n t} = \frac{vol\,(\mathcal{F})}{4\pi t} + \sum_{n=0}^{N} b_n t^n + O(t^{N+1}), \qquad (i)$$

$$N(x) = \frac{vol\,(\mathcal{F})}{4\pi}\, x + O(x^{1/2}). \tag{ii}$$

It is not difficult to see that (ii) can be deduced from (i) by a standard Tauberian theorem.

3. POISSON'S SUMMATION FORMULA

As a precursor to the trace formula, let us consider the discrete subgroup $\mathbf{Z}$ of the additive group of real numbers $\mathbf{R}$. A fundamental domain for the action of $\mathbf{Z}$ on $\mathbf{R}$, given by translation, can be taken as the interval $(0, 1]$. We can identify this with the circle S^1 via the map

$$p : \mathbf{R} \to S^1$$

given by

$$p(t) = e^{2\pi i t}.$$

This induces an isomorphism

$$\mathbf{Z} \backslash \mathbf{R} \simeq S^1.$$

We consider functions

$$f : \mathbf{R} \to \mathbf{C}$$

which satisfy

$$f(x + m) = f(x) \quad \forall m \in \mathbf{Z}$$

and such that

$$\int_0^1 | f(x) |^2 \, dx < \infty.$$

This space of functions we denote by $L^2(\mathbf{Z} \backslash \mathbf{R})$. For two functions f, g in this space, we can define an inner product by

$$(f, g) = \int_0^1 f(x)\overline{g(x)}dx.$$

The differential operator.

$$\Delta = \frac{d^2}{dx^2}$$

is the Laplacian on $\mathbf{Z}\backslash\mathbf{R}$. A function $f \in L^2(\mathbf{Z}\backslash\mathbf{R})$ is an eigenfunction for Δ if

$$\Delta f = \lambda f.$$

The solutions are of the form

$$f(x) = \exp(\sqrt{\lambda} x).$$

The condition $f(x+1) = f(x)$ forces $\exp(\sqrt{\lambda}) = 1$ so that

$$\lambda_n = -4\pi^2 n^2, \quad n = 0, 1, 2, \ldots$$

and the corresponding eigenfunctions are

$$e_n(x) = e^{2\pi i n x}.$$

The classical theorem of Fourier analysis says that every $f \in L^2(\mathbf{Z}\backslash\mathbf{R})$ has a Fourier expansion

$$f(x) = \sum_{n=-\infty}^{\infty} (f, e_n)\, e^{2\pi i n x}.$$

Now let us consider the Schwartz space $\mathcal{S}$ of functions $f : \mathbf{R} \to \mathbf{C}$ such that for every pair of integers $m, n \geqslant 0$, the function

$$\left| x^m \frac{d^n f}{dx^n} \right|$$

is bounded on $\mathbf{R}$. It is easy to see that if $f \in \mathcal{S}$, then the Fourier transform

$$\hat{f}(x) = \int_{-\infty}^{\infty} f(t)\, e^{-2\pi i t x}\, dt$$

is also in $\mathcal{S}$. Moreover,

$$\hat{\hat{f}}(x) = f(-x).$$

The most striking application of Fourier analysis to number theory is in the proof of the functional equation of the Riemann zeta function via:

Poissons summation formula:

Let $f \in \mathcal{S}$. Then,

$$\sum_{n=-\infty}^{\infty} f(n) = \sum_{n=-\infty}^{\infty} \hat{f}(n).$$

The proof is easy. Consider

$$g(x) = \sum_{n=-\infty}^{\infty} f(x+n).$$

Then, $g(x+1) = g(x)$ and so g has a Fourier Expansion:

$$g(x) = \sum_{n=-\infty}^{\infty} (g, e_n)\, e^{2\pi inx}.$$

We have

$$(g, e_n) = \int_0^1 g(x)\, e^{-2\pi inx}\, dx$$

$$= \int_{-\infty}^{\infty} f(x)\, e^{-2\pi inx}\, dx$$

$$= \hat{f}(n)$$

so that

$$g(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n)\, e^{2\pi inx}.$$

Setting $x = 0$ gives the Poisson summation formula.

Now let $F \in \mathcal{S}$, be an even function. If we define for fixed y,

$$f(x) = F(xy),$$

then it is easy to see that

$$\hat{f}(x) = y^{-1} \hat{F}(x/y).$$

Indeed,

$$\hat{f}(x) = \int_{-\infty}^{\infty} f(t)\, e^{-2\pi ixt}\, dt$$

$$= \int_{-\infty}^{\infty} F(ty)\, e^{-2\pi ixt}\, dt$$

$$= y^{-1} \int_{-\infty}^{\infty} F(w)\, e^{-2\pi ixw/y}\, dw$$

$$= y^{-1} \hat{F}(x/y)$$

The Poisson formula yields

$$\sum_{n=-\infty}^{\infty} F(ny) = y^{-1} \sum_{n=-\infty}^{\infty} \hat{F}(n/y).$$

As F is even, this translates as

$$F(0) + 2 \sum_{n=1}^{\infty} F(ny) = y^{-1} \hat{F}(0) + 2y^{-1} \sum_{n=1}^{\infty} \hat{F}(n/y).$$

Alternately,

$$\sum_{n=1}^{\infty} F(n/y) = y \sum_{n=1}^{\infty} \hat{F}(ny) + \frac{1}{2} \hat{F}(0)y - \frac{1}{2} F(0).$$

Now let us consider for $Re\ s > 1$,

$$\zeta(s) \int_0^{\infty} F(y)y^s \frac{dy}{y} = \int_0^{\infty} \sum_{n=1}^{\infty} F(ny)y^s \frac{dy}{y}$$

$$= \int_0^1 \left(\sum_{n=1}^{\infty} F(ny) \right) y^s \frac{dy}{y} + \int_1^{\infty} \left(\sum_{n=1}^{\infty} F(ny) \right) y^s \frac{dy}{y}$$

$$= \int_1^{\infty} \sum_{n=1}^{\infty} (F(ny)y^s + F(n/y)y^{-s}) \frac{dy}{y}$$

after a change of variable. By the above form of Poisson's formula, we obtain that this is

$$= \int_1^{\infty} \sum_{n=1}^{\infty} (F(ny)y^s + \hat{F}(ny)y^{1-s}) \frac{dy}{y} - \frac{1}{2}\left(\frac{F(0)}{s} + \frac{\hat{F}(0)}{1-s}\right).$$

As $f \in \mathcal{S}$, this integral converges for all values of s and therefore defines an analytic function of s. If we set $F(x)=e^{-\pi x^2}$, then, it is easily seen that

$$\int_0^{\infty} F(y)y^s \frac{dy}{y} = \pi^{-s/2}\,\Gamma(s/2).$$

Moreover, since $\hat{F} = F$, we obtain the functional equation

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \pi^{(1-s)/2}\Gamma((1-s)/2)\zeta(1-s).$$

Perhaps the symmetry of this equation led Riemann to make his famous hypothesis. Yet, this is unlikely. It is easy to produce Dirichlet series with functional equations of a similar type but which violate the analogue of the Riemann hypothesis. Many of these examples arise from the theory of Eisenstein series or Dirichlet series which are Mellin transforms of modular forms. These counterexamples, however, do not possess an Euler roduct and therefore, this fact must enter into any proof of the Riemann hypothesis.

There are other numerous applications of Poisson's summation formula. In our context, the analogue of the formula for the torus leads to an explicit error term for the classical circle problem alluded to in the previous section.

4. WEIL'S EXPLICIT FORMULA

Let $h(r)$ be an analytic function satisfying the properties

$$(*) \qquad h(r) = h(-r), \ | h(r) | \leqslant A(1 + | r |)^{-2-\delta}.$$

We will prove:

Proposition: (Weil)

$$\sum_{\gamma} h(\gamma) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h(r) \frac{\Gamma'}{\Gamma}\left(\frac{1}{4} + \frac{1}{2}\, ir\right) dr + 2h(i/2)$$

$$- g(0) \log \pi - 2 \sum_{n=1}^{\infty} \frac{\Lambda(n)}{\sqrt{n}}\, g(\log n)$$

where the first sum is over non-trivial zeroes $\frac{1}{2} + i\gamma$, $\gamma \in \mathbf{C}$, of $\zeta(s)$. Moreover,

$$g(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h(r)\, e^{-iru}\, dr.$$

Proof. To prove this formula, it is convenient to define

$$H(s) = \int_{-\infty}^{\infty} g(u)\, e^{(s-\frac{1}{2})u}\, du.$$

Then,

$$H\left(\frac{1}{2} + iw\right) = h(w)$$

by the Fourier inversion formula. Since h is even, $H(w) = H(1-w)$. As H is essentially the Laplace transform of g, it is useful to recall the Mellin inversion formula:

$$g(u) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} H(w)\, e^{(w-\frac{1}{2})u}\, dw.$$

In order to prove the explicit formula, consider the integral

$$\frac{1}{2\pi i} \int_{\partial R} H(s) \frac{\zeta'}{\zeta}(s)\, ds$$

where

$$R = [-\delta, 1+\delta] \times [-T, T]$$

where T is suitably chosen as an element of a sequence of T_m such that $T_m \to \infty$ and $\zeta(s) \neq 0$ for $Im(s) = T_m$.

The singularities of the integrand are the non-trivial zeroes of $\zeta(s)$. By Cauchy's theorem, the integral is

$$\sum_{|Im(i\gamma)| < T} h(\gamma).$$

It can also be arranged that for these T_m,

$$\frac{\zeta'}{\zeta}(s) = O(\log^2(|t|+1))$$

on the horizontal and vertical lines of the contour. The growth condition on h is sufficient to ensure that the contribution from the horizontal integrals tends to 0 as $T_m \to \infty$. Now by the functional equation,

$$-\frac{\zeta}{\zeta}(1-s) = -\log\pi + \frac{1}{2}\frac{\Gamma'}{\Gamma}(s/2) + \frac{1}{2}\frac{\Gamma'}{\Gamma}((1-s)/2) + \frac{\zeta'}{\zeta}(s)$$

so that one of the vertical integrals is transformed into the other vertical integral:

$$-\frac{1}{2\pi i} \int_{-\delta+i\infty}^{-\delta+i\infty} H(s) \frac{\zeta'}{\zeta}(s)\, ds = -\int_{1+\delta-i\infty}^{1+\delta+i\infty} H(1-w) \frac{\zeta'}{\zeta}(1-w)\, dw$$

which is

$$= \int_{+\delta-i\infty}^{1+\delta+i\infty} H(w)\left\{-\log\pi + \frac{1}{2}\frac{\Gamma'}{\Gamma}(w/2) + \frac{1}{2}\frac{\Gamma'}{\Gamma}((1-w)/2) + \frac{\zeta'}{\zeta}(w)\right\} dw.$$

The part of the integral involving the zeta function gives for each vertical integral:

$$-\sum_{n=1}^{\infty} \frac{\Lambda(n)}{\sqrt{n}} \int_{1+\delta-i\infty}^{1+\delta+i\infty} H(w)\, e^{(w-\frac{1}{2})\log n}\, dw = -\sum_{n=1}^{\infty} \frac{\Lambda(n)}{\sqrt{n}}\, g(\log n),$$

by the Mellin inversion formula. We also obtain the constant term

$$-\, g(0) \log \pi.$$

To evaluate the integrals involving the gamma function, we move the line of integration to $Re(w) = \frac{1}{2}$. This is valid since

$$\frac{\Gamma'}{\Gamma}(w) = O(\log (|w| + 2))$$

in the half plane under consideration. This completes the proof of the proposition.

The most interesting consequence of the explicit formula is the following equivalent formulation of the Riemann hypothesis.

COROLLARY. *Suppose that*

$$g(u) = \int_{-\infty}^{\infty} g_0(t+u)\, \overline{g_0(t)}\, dt$$

for some g_0. *Then the Riemann hypothesis is true if and only if*

$$\sum_{\gamma} \hat{g}(\gamma) \geqslant 0$$

for all such g. *(Here, the sum is over all the non-trivial zeroes*

$$\tfrac{1}{2} + i\gamma,\ \gamma \in \mathbf{C} \text{ of } \zeta(s).)$$

Proof. We can write

$$g(u) = g_0(u) * \overline{g_0(-u)}.$$

If we set

$$g_1(u) = \overline{g_0(-u)},$$

then, a simple calculation shows that

$$\hat{g}_1(r) = \int_{-\infty}^{\infty} \overline{g_0(-u)}\, e^{-iru}\, du$$

$$= \overline{\int_{-\infty}^{\infty} g_0(-u)\, e^{-i\bar{r}u}\, du}$$

$$= \overline{h_0(\bar{r})}$$

where $h_0 = \hat{g}_0$. Thus, by the theory of the Fourier transform,

$$h(r) = h_0(r)\, \overline{h_0(\bar{r})}.$$

If all the zeroes ρ of $\zeta(s)$ have $Re(\rho) = \frac{1}{2}$, then γ is real and

$$h(\gamma) = | h_0(\gamma) |^2 \geqslant 0$$

and hence

$$\sum_{\gamma} h(\gamma) \geqslant 0.$$

Conversely, suppose that the sum is non-negative for all such functions and there is a zero $\rho_0 = \beta_0 + i\gamma_0$ with $\beta_0 \neq \frac{1}{2}$. We will utilise the fact that the Fourier transform of a function of the form

$$P(x)\, e^{-Ax^2},$$

where P is a polynomial and $A > 0$, is again a function of the same form. We will set

$$h_0(x) = P(x)\, e^{-Ax^2}$$

for a suitable P and A, and show that for such a choice,

$$\sum_{\gamma} h(\gamma) < 0.$$

This will establish the corollary. Indeed, set

$$z = i(s - \tfrac{1}{2} - i\gamma_0)$$

and let $\eta = \eta_\rho = i(\rho - \frac{1}{2} - i\gamma_0)$. Then, $\eta_0 = i(\beta_0 - \frac{1}{2})$ corresponds to $\rho = \rho_0$. Let $Q(x)$ be the polynomial having for zeroes all the η's satisfying $|Re(\eta)| \leqslant 2$ except for η_0, $\overline{\eta_0}$. (Note that if ρ is a zero of $\zeta(s)$, then $\bar{\rho}$ and $1 - \rho$, $1 - \bar{\rho}$ are zeroes of $\zeta(s)$ by virtue of the functional equation and the fact that $\zeta(s)$ is real for real values of s.) Thus, the zeroes of $Q(x)$ are either real or occur in conjugate pairs. Hence, Q has real coefficients. We will set

$$P(x) = xQ(x)Q(-x),$$

and define

$$h_0(s) = P(z)\, e^{-Az^2}.$$

Then,

$$h(\gamma) = P(\eta)^2\, e^{-2A\eta^2}$$

and therefore,

$$\sum_{\gamma} h(\gamma) = \sum_{\eta} P(\eta)^2\, e^{-2A\eta^2}.$$

We decompose this sum as

$$\sum_{|Re(\eta)| \leqslant 2} + \sum_{|Re(\eta)| > 2}.$$

Let m be the order of the zero ρ_0. As Q vanishes at η satisfying $|Re(\eta)| \leqslant 2$, except at η_0, the first sum is

$$2m\eta_0^2\, |\, Q(\eta_0)\, |^4 \exp(-2A\eta_0^2) = -2m\left(\beta_0 - \frac{1}{2}\right)^2 Q(\eta_0)\, |^4\, e^{2A(\beta_0 - \frac{1}{2})^2}$$

and the second sum is

$$\sum_{|Re(\eta)| > 2} P(\eta)^2\, e^{-2A\eta^2} \leqslant e^{-2A} \sum_{|Re(\eta)| > 2} |\, P(\eta)^2\, e^{-\eta^2}\, |$$

as a simple calculation reveals. Choosing A large forces the total sum to be strictly negative. That is,

$$\sum_{\gamma} h(\gamma) < 0,$$

contrary to hypothesis. This completes the proof of the corollary.

As was done by Weil, the explicit formula can be stated in the more general context of L functions satisfying suitable functional equations and growth conditions. This leads to a variety of applications. For example, Serre utilised the number field analogue of the above explicit formula to obtain lower bounds for discriminants of these number fields.

5. TRACE FORMULA FOR A COMPACT RIEMANN SURFACE

Let F be a compact Riemann surface of genus $g \geqslant 2$. The universal covering space of F is conformally equivalent to the upper half plane

Moreover, F can be identified with $\Gamma\backslash H$ for some strictly hyperbolic Fuchsian group. The group $\Gamma \subseteq PSL_2(\mathbf{R})$ will have a compact fundamental domain $\mathcal{F}$. In addition, $\Gamma \simeq \pi_1(F)$, where π_1 denotes the fundamental group.

There is a metric on H given by

$$ds^2 = \frac{dx^2 + dy^2}{y^2}$$

with the corresponding area element

$$d\mu(z) = \frac{dx\, dy}{y^2}, \; z = x + iy.$$

The distance can be calculated explicitly :

$$d(z_1, z_2) = arg\ cosh \left(1 + \frac{|z_1 - z_2|^2}{2y_1 y_2}\right).$$

This metric is invariant under the action of $SL_2(\mathbf{R})$.

We consider

$$L^2(\Gamma\backslash H) = \{f : \int_{\mathcal{F}} |f(z)|^2\, d\mu(z) < \infty, f(\gamma z) = f(z),\ \forall \gamma \in \Gamma, z \in H\}.$$

These functions are called *automorphic functions*. The inner product on this space is given by

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

The Laplace operator is

$$D = y^2 \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right).$$

This is a self adjoint differential operator with respect to this inner product, and therefore generates a discrete set of real eigenvalues

$$0 = \lambda_0 < \lambda_1 \leqslant \lambda_2 \leqslant \ldots$$

$$D\phi_n + \lambda_n \phi_n = 0.$$

These functions were first studied by Delsarte and Maass.

We begin by mimicing the Poisson summation formula. Consider

$$g(z) = \sum_{\gamma \in \Gamma} f(\gamma z), f \in C_c^\infty(H).$$

We consider functions which are radially symmetric:

$$f(z) = \Phi\left(\frac{|z - z_0|^2}{yy_0}\right)$$

for some point $z_0 \in H$.

We have the eigenfunction expansion:

$$g(z) = \sum_{n=0}^{\infty} c_n \phi_n(z)$$

and

$$c_n = \int_{\mathcal{F}} g(z)\, \phi_n(z)\, d\mu(z)$$

$$= \sum_{\gamma \in \Gamma} \int_{\mathcal{F}} f(\gamma z)\, \phi_n(z)\, d\mu(z)$$

$$= \int_H f(z)\, \phi_n(z)\, d\mu(z).$$

We must therefore calculate:

$$\int_H \Phi\left(\frac{|z - z_0|^2}{yy_0}\right) \phi_n(z)\, d\mu(z).$$

To do this we utilise:

Main Lemma: *Suppose that* $\Phi \in C_c^\infty(\mathbf{R})$ *and* $\phi \in C^\infty(H)$ *and* $D\phi + \lambda\phi = 0$. *Then,*

$$\int_H \Phi\left(\frac{|z - z_0|^2}{yy_0}\right) \phi(z)\, d\mu(z) = H(\lambda)\, \phi(z_0)$$

where $H(\lambda)$ *depends solely on* Φ *and* λ.

Remark: H is called the *Harish-Chandra transform* of Φ.

The proof can be found in Kubota [6, p. 8] or in Hejhal [3, p. 453].

We now set:

$$k(z, w) = \Phi\left(\frac{|z-w|^2}{Im(z)\, Im(w)}\right).$$

The function $k(z, w)$ is called a *point pair invariant* because

$$k(\sigma z, \sigma w) = k(z, w), \ \forall \sigma \in SL_2(\mathbf{R}).$$

Thus, in the above,

$$c_n = H(\lambda_n)\, \phi_n(z_0).$$

Hence,

$$g(z) = \sum_{n=0}^{\infty} c_n\, \phi_n(z)$$

$$= \sum_{n=0}^{\infty} H(\lambda_n)\, \phi_n(z)\, \phi_n(z_0).$$

Therefore,

$$K(z, w) = \sum_{\gamma \in \Gamma} k(\gamma z, w)$$

$$= \sum_{n=0}^{\infty} H(\lambda_n)\, \phi_n(z)\, \phi_n(w)$$

By integrating over the fundamental domain, we can get rid of th
eigenfunctions:

$$\int_{\mathcal{F}} K(z, z)\, d\mu(z) = \sum_{n=0}^{\infty} H(\lambda_n).$$

This can indeed be interpreted as a trace. Consider the integral operato

$$Lf(z) = \int_{H} k(z, w)\, f(w)\, d\mu(w).$$

Then, by the main lemma,

$$L\phi_n = H(\lambda_n)\, \phi_n$$

so that

$$Tr(L) = \sum_{n=0}^{\infty} H(\lambda_n).$$

At first, it looks as if we do not know much about H. But since y^s is a
eigenfunction of the Laplacian, with eigenvalue $s(1-s)$, we get

$$\int_H k(i, w)\,(Im(w))^s\, d\mu(w) = H(s(1-s))$$

for any $s \in C$. We set

$$s = \frac{1}{2} + ir,\ \lambda = s(1-s) = \frac{1}{4} + r^2,\ r \in C.$$

We define the *Abel transform*

$$Q(x) = \int_x^\infty \frac{\Phi(t)}{\sqrt{t-x}}\, dt$$

and

$$g(u) = Q(e^u + e^{-u} - 2).$$

We have the inversion formula for Q:

$$\Phi(x) = -\frac{1}{\pi} \int_x^\infty \frac{dQ(t)}{\sqrt{t-x}},$$

which will be needed later. Then,

$$H(\lambda) = h(r) = \int_{-\infty}^{\infty} g(u)\, e^{iru}\, du.$$

To see this, we note that

$$H(\lambda) = \int_H \Phi\left(\frac{|z-i|^2}{y}\right) y^s \frac{dx\,dy}{y^2}$$

$$= \int_0^\infty \int_{-\infty}^{\infty} \Phi\left(\frac{x^2 + (y-1)^2}{y}\right) y^s \frac{dx\,dy}{y^2}$$

$$= \int_0^\infty \int_{-\infty}^{\infty} \Phi\left(\frac{x^2}{y} + y + y^{-1} - 2\right) y^s \frac{dx\,dy}{y^2},$$

We put

$$y + y^{-1} - 2 = w$$

$$w + \frac{x^2}{y} = t,$$

then

$$\frac{dw}{dx} + \frac{2x}{y} = \frac{dt}{dx}.$$

As w is a function of y, $\dfrac{dw}{dx} = 0$ and so

$$dx = \frac{y\,dt}{2x}.$$

Hence,

$$\int_{-\infty}^{\infty} \Phi\left(w + \frac{x^2}{y}\right) dx = 2 \int_{0}^{\infty} \Phi\left(w + \frac{x^2}{y}\right) dx$$

$$= \sqrt{y} \int_{w}^{\infty} \Phi(t) \frac{dt}{\sqrt{t-w}}.$$

Therefore,

$$h(r) = \int_{0}^{\infty} Q(y + y^{-1} - 2)\, y^{s+\frac{1}{2}} \frac{dy}{y^2}.$$

If we set $y = e^u$, then we obtain

$$H(\lambda) = \int_{-\infty}^{\infty} g(u)\, e^{u(s-\frac{1}{2})}\, du.$$

But as $s = \frac{1}{2} + ir$, this is $h(r)$ as desired. We have therefore succeede in expressing h as a Fourier transform of an appropriate function.

Now consider the sum

$$\sum_{n=0}^{\infty} h(r_n) = \int_{\mathcal{F}} K(z, z)\, d\mu\,(z).$$

The right hand side can now be expanded as a sum over the conjugac classes in Γ

$$\int_{\mathcal{F}} K(z, z)\, d\mu\,(z) = \sum_{\gamma \in \Gamma} \int_{\mathcal{F}} k(\gamma z, z)\, d\mu\,(z)$$

$$= \sum_{\{P\}} \sum_{\gamma \in \{P\}} \int_{\mathcal{F}} k(\gamma z, z)\, d\mu\,(z)$$

$$= \sum_{\{P\}} \sum_{\eta \in C_\Gamma(P)\backslash\Gamma} \int_{\mathcal{F}} k(\eta^{-1} P\eta z, z)\, d\mu\,(z)$$

$$= \sum_{\{P\}} \sum_{\eta \in C_\Gamma(P)\backslash\Gamma} \int_{\mathcal{F}} k(P\eta z.\ \eta z)\, d\mu\,(z)$$

because k is a point pair invariant. Changing variables, we obtain

$$= \sum_{\{P\}} \sum_{\eta \in C_\Gamma(P)\backslash\Gamma} \int_{\eta\mathcal{F}} k(Pw, w)\, d\mu(w)$$

$$= \sum_{\{P\}} \int_{C_\Gamma(P)\backslash H} k(Pw, w)\, d\mu(w).$$

Clearly the sum is over distinct conjugacy classes. In order to calculate these integrals, we note that Γ has no elliptic or parabolic elements since Γ is compact. If $P = I$, then $k(Pw, w) = \Phi(0)$ and the contibution is

$$\Phi(0)\ vol\,(\mathcal{F}).$$

This can be evaluated explicitly in the following way. By the inversion formula

$$\Phi(x) = -\frac{1}{\pi} \int_x^\infty \frac{dQ(t)}{\sqrt{t-x}}$$

so that

$$\Phi(0) = -\frac{1}{\pi} \int_0^\infty \frac{g'(u)\, du}{e^{u/2} - e^{-u/2}}$$

$$= \frac{1}{2\pi^2} \int_{-\infty}^\infty \int_0^\infty rh(r) \frac{\sin ru}{e^{u/2} - e^{-u/2}} du\, dr.$$

From the formula for Fourier transforms,

$$\int_0^\infty \operatorname{cosech} \frac{1}{\pi} u \sin ru\, du = \pi \tanh \pi r,$$

so that

$$\Phi(0) = \frac{1}{4\pi} \int_{-\infty}^\infty rh(r) \tanh \pi r dr.$$

We may now suppose that P is hyperbolic. As noted earlier, by making a suitable conjugation, we may suppose that

$$P(z) = \lambda z, \quad 1 < \lambda < \infty.$$

As Γ is discrete, $C_\Gamma(P)$ is a cyclic group generated by P_0 (say),

$$P_0(z) = \lambda_0 z, \quad P = P_0^k,$$

for some $k \geqslant 1$. We have defined the *norm* of P as $N(P) = \lambda$. Note also that

$$|Tr(P)| = N(P)^{1/2} + N(P)^{-1/2}.$$

The transformation P is called primitive if $C_\Gamma(P) = \langle P \rangle$. Since Γ is strictly hyperbolic, the classes $\{P_0^k\}$ are mutually disjoint for $k \in \mathbf{Z}$. Thus,

$$\int_{C_\Gamma(P)\backslash H} k(Pw, w)\, d\mu(w) = \int_{C_\Gamma(P_0)\backslash H} k(Pw, w)\, d\mu(w)$$

It is known that $C_\Gamma(P)$ is a cyclic group generated by P_0. It is then easy to see that a fundamental domain of $C_\Gamma(P_0)$ is given by the region

$$\{w : 1 \leqslant Im(w) < \lambda_0\}.$$

Hence,

$$\int_{C_\Gamma(P_0)\backslash H} k(\lambda w, w)\, d\mu(w) = \int_1^{\lambda_0} \int_{-\infty}^{\infty} \Phi\left(\frac{(\lambda-1)^2(x^2+y^2)}{\lambda y^2}\right) \frac{dx\, dy}{y^2}$$

$$= 2 \int_1^{\lambda_0} \int_0^{\infty} \Phi\left(\frac{(\lambda-1)^2}{\lambda}(1+t^2)\right) \frac{dt\, dy}{y}$$

$$= 2(\log \lambda_0) \int_0^{\infty} \Phi\left(\frac{(\lambda-1)^2}{\lambda}(1+t^2)\right) dt$$

after a change of variable. We set

$$u = \frac{(\lambda-1)^2}{\lambda}(1+t^2)$$

so that

$$t = \sqrt{\frac{\lambda u}{(\lambda-1)^2} - 1}$$

and the integral becomes

$$\frac{\lambda}{(\lambda-1)^2}\int_{(\lambda-1)^2/\lambda}^{\infty}\Phi(u)\frac{du}{2\sqrt{\frac{(\lambda-1)^2u}{\lambda}-1}}=\frac{\sqrt{\lambda}}{2(\lambda-1)}\int_{(\lambda-1)^2/\lambda}^{\infty}\frac{\Phi(u)\,du}{\sqrt{u-\frac{(\lambda-1)^2}{\lambda}}}$$

This is easily seen to be

$$\frac{\lambda^{1/2}}{2(\lambda-1)}Q(\lambda+\lambda^{-1}-2)=\frac{g(\lambda)}{2(\lambda^{1/2}-\lambda^{-1/2})}.$$

Hence, the total contribution is

$$\frac{\Lambda(P)}{N(P)^{1/2}-N(P)^{-1/2}}\,g(\log N(P)),$$

This establishes the trace formula in the compact case

$$\sum_{n=0}^{\infty}h(r_n)=\frac{\mathrm{vol}(\mathcal{F})}{4\pi}\int_{-\infty}^{\infty}r\,h(r)\tanh\pi r\,dr+$$

$$+\sum_{\{P\}}\frac{\Lambda(P)}{N(P)^{1/2}-N(P)^{-1/2}}\,g(\log N(P)).$$

6. SELBERG'S ZETA FUNCTION

We begin by noting the trace formula allows us to deduce asymptotic information about the distribution of the eigenvalues by making appropriate choice of the test function. Indeed, let

$$h(r)=\exp\left[\left(\tfrac{1}{4}+r^2\right)T\right]$$

then the corresponding Fourier transform is

$$g(u)=\frac{e^{-T/4}}{\sqrt{4\pi T}}\exp\left(-u^2/4T\right).$$

The trace formula becomes

$$\sum_{n=0}^{\infty}\exp\left(-\lambda_nT\right)=\frac{\mathrm{vol}\,\mathcal{F}}{4\pi}\int_{-\infty}^{\infty}r\exp\left[\left(-r^2+\tfrac{1}{4}\right)T\right]\tanh\pi r\,dr+$$

$$+\frac{e^{-T/4}}{\sqrt{4\pi T}}\sum_{\{P\}}\frac{\Lambda(P)}{N(P)^{1/2}-N(P)^{-1/2}}\exp\left(-(\log N(P))^2/4T.\right)$$

It is easy to see that as $T \to 0^+$, the sum arising from the hyperbolic classes contributes $o(1)$. Therefore,

$$\sum_{n=0}^{\infty} \exp(-\lambda_n T) = \frac{\operatorname{vol}(\mathcal{F})}{4\pi} \int_{-\infty}^{\infty} r \exp(-(r^2 + \tfrac{1}{4})T) \tanh \pi r \, dr + o(1)$$

as $T \to 0^+$. Now by a classical Tauberian theorem, we obtain the result of Weyl cited in section 2:

$$N(x) \sim \frac{\operatorname{vol}(\mathcal{F})}{4\pi} x$$

as $x \to \infty$.

To obtain information about the distribution of $N(P)$, we let $T \to \infty$. Let us set

$$\pi_0(x) = \#\{P_0 : N(P_0) \leqslant x\}.$$

Then, by partial integration,

$$1 + O(e^{-cT}) = \frac{e^{-T/4}}{\sqrt{4\pi T}} \int_1^{\infty} \frac{\log x}{\sqrt{x}} \exp(-(\log x)^2/4T) \, d\pi_0(x),$$

so that multiplying by e^{-sT} and integrating with respect to T in the range $[1, \infty]$, gives

$$\frac{1}{s} + O(1) = \int_0^{\infty} u \exp(-(1+s)u) \, d\pi_0(e^u).$$

Again, by the Tauberian theorem, we deduce that

$$\sum_{\{P\}} \frac{\Lambda(P)}{N(P)^{s+1}} = \frac{1}{s} + O(1)$$

and hence

$$\sum_{N(P) \leqslant x} \frac{\Lambda(P)}{N(P)} \sim \log x.$$

This is highly reminescent of prime number theory. It is therefore natural to define:

$$\frac{Z'}{Z}(s) = \sum_{\{P\}} \frac{\Lambda(P)}{N(P)^s}$$

for $Re(s) > 1$, which gives

$$Z(s) = \prod_{\{P_0\}} \prod_{k=0}^{\infty} (1 - N(P_0)^{-s-k}).$$

This is called the *Selberg zeta function* and the above product is the analogue of the Euler product for the Riemann zeta function. It is different from the classical analogue in two striking aspects. Firstly, there is an "extra" infinite product over k and secondly, the factor in the product does not appear to the exponent -1. $Z(s)$ is an analytic function for $Re\,(s) > 1$. If we choose,

$$g(u) = \frac{e^{-\alpha|u|}}{2\alpha} - \frac{e^{-\beta|u|}}{2\beta},$$

then

$$h(r) = \frac{1}{r^2+\alpha^2} - \frac{1}{r^2+\beta^2}$$

is the corresponding Fourier transform. Let $\alpha = s - \frac{1}{2}$ and $\beta \geqslant 2$ be fixed. This leads to

$$\sum_{n=0}^{\infty}\left[\frac{1}{r_n^2+(s-\frac{1}{2})^2} - \frac{1}{r_n^2+\beta^2}\right] = \frac{\mathrm{vol}(\mathcal{F})}{4\pi}\int_{-\infty}^{\infty} r\left[\frac{1}{r^2+(s-\frac{1}{2})^2} - \right.$$

$$\left. - \frac{1}{r^2+\beta^2}\right]\tanh \pi r dr + \frac{1}{2s-1}\frac{Z'}{Z}(s) - \frac{1}{2\beta}\frac{Z'}{Z}(\tfrac{1}{2}+\beta).$$

By complex analysis, the integral in the above formula can be evaluated as

$$2\sum_{k=0}^{\infty}\left[\frac{1}{\beta+\frac{1}{2}+k} - \frac{1}{s+k}\right].$$

Therefore

$$\frac{1}{2s-1}\frac{Z'}{Z}(s) = \frac{1}{2\beta}\frac{Z'}{Z}(\tfrac{1}{2}+\beta) + \sum_{n=0}^{\infty}\left[\frac{1}{r_n^2+(s-\frac{1}{2})^2} - \frac{1}{r_n^2+\beta^2}\right] +$$

$$+\frac{\mathrm{vol}\,(\mathcal{F})}{2\pi}\sum_{k=0}^{\infty}\left[\frac{1}{\beta+\frac{1}{2}+k} + \frac{1}{s+k}\right].$$

Observing that the right hand side is meromorphic for all $s \in \mathbb{C}$, we obtain a meromorphic continuation of $Z(s)$. One can prove

THEOREM. (Selberg)

(i) $Z(s)$ is an entire function.

(ii) $Z(s)$ has trivial zeroes at $s = -k$, $k \geqslant 1$, with multiplicity $(2g-2)(2k+1)$, and at $s = 0$ of order $2g-1$ at $s = 1$, a simple zero.

(iii) the non-trivial zeroes of $Z(s)$ correspond to the eigenvalues of the Laplacian in the following way: $\lambda_n = \frac{1}{4} + r_n^2$ is an eigenvalue if and only if $\frac{1}{2} + ir_n$ is a non-trivial zero of $Z(s)$.

(iv) $Z(s)$ satisfies a functional equation of the form:

$$Z(s) = \chi(s)\, Z(1 - s)$$

where

$$\chi(s) = \exp\left\{\mathrm{vol}\,(\mathcal{F}) \int_0^{s-\frac{1}{2}} v \tan v\pi \, dv\right\}.$$

$Z(s)$ has a greater growth rate than $(s - 1)\,\zeta(s)$. It is an entire function of order 2 and $(s - 1)\,\zeta(s)$ has order 1. In fact,

$$\#\,\{n : 0 \leqslant r_n \leqslant T\} \sim \frac{\mathrm{vol}\,(\mathcal{F})}{4\pi}\, T^2$$

in contrast to

$$N(T) \sim \frac{T}{2\pi} \log \frac{T}{2\pi}$$

for the Riemann zeta function. By analogy, let us set

$$\psi_Z(x) = \sum_{N(P) \leqslant x} \Lambda(P).$$

Then, Selberg [8] derived an explicit formula for $\psi_Z(x)$ and showed as a consequence

$$\psi_Z(x) = x + O\,(x^{3/4}).$$

Recently, Iwaniec [5] improved the exponent 3/4 to $\frac{35}{48} + \epsilon$. It is conjectured, in analogy with the classical distribution of prime numbers, that

$$\psi_Z(x) = x + O(x^{1/2+\epsilon})$$

but this is intractible though we know the Riemann hypothesis for $Z(s)$.

7. TRACE FORMULA FOR A NON-COMPACT RIEMANN SURFACE

We shall consider now Fuchsian groups of the first kind, that is, vol $(\Gamma \backslash H) < \infty$. The fundamental domain, $\mathcal{F}$, is now not necessarily compact, but will have at most a finite number of cusps. Geometrically, one can visualise this as a compact Riemann surface with a finite number of points removed. The simplest example is the case when $\Gamma = PSL_2(\mathbf{Z})$ and $\Gamma\backslash H$ is homeomorphic to the punctured sphere.

As before, let us consider $L^2(\Gamma\backslash H)$ and would like to decompose this space as a sum of eigenspaces. For clarity of exposition, let us confine ourselves to the case when $\Gamma\backslash H$ has only one cusp, which we can take, without loss of generality to be $i\infty$. Moreover, the stabiliser of this cusp, denoted Γ_∞, is an infinite cyclic group, generated by

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Therefore, any function $f \in L^2(\Gamma\backslash H)$ can be expanded as a Fourier series

$$f(z) = \sum_{-\infty}^{\infty} W_m(y) \exp(2\pi imx).$$

If f is to be an eigenfunction of the Laplacian, then

$$Df + \lambda f = 0$$

leads to the ordinary differential equation

$$y^2 W_m''(y) - (\lambda - 4\pi^2 m^2 y^2)\, W_m(y) = 0 \tag{1}$$

for each $m \in \mathbf{Z}$. This is related to the classical Bessel's equation:

$$y^2 u'' + yu' - (y^2 + r^2)\, u = 0. \tag{2}$$

Indeed, if in (2), we set $W(y) = y^{1/2} u\,(2\pi \mid m \mid y)$ then the equation becomes transformed into (1) with $\lambda = \frac{1}{4} + r^2$. If we define,

$$K_r(y) = \frac{1}{2} \int_0^\infty \exp\left(y\,(u + u^{-1})/2\right) u^{r-1}\, du$$

then K_r satisfies (2). Moreover, as $y \to \infty$,

$$K_r(y) \sim \sqrt{\frac{\pi}{2y}}\, e^{-y}$$

and so this is the unique solution of (2) which is regular at infinity. This leads to the following expansion for the eigenfunction f:

$$f(z) = c_0\, y^s + \bar{c}_0 y^{1-s} + y^{1/2} \sum_{m \neq 0} c_m K_{s-\frac{1}{2}}(2\pi \mid m \mid y) \exp(2\pi imx)$$

where $\lambda = s(1-s)$ is the corresponding eigenvalue. This series converges absolutely for $y > 0$. Using this formula, in conjunction with Green's theorem, we easily deduce that $\lambda \geqslant 0$.

An eigenfunction is called a *cusp form* if $c_0 = \tilde{c}_0 = 0$. This is equivalent to

$$\int_0^1 f(x+iy)\, dx = 0$$

in the case of the one cusp at $i\infty$. (In the general case, we say that f is a cusp form if it vanishes at all the cusps.)

Now let $\eta \in C_c^\infty(0, \infty)$ and define $\psi(x+iy) = \eta(y)$. Then define the *Poincaré series*

$$\theta_\psi(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \psi(\gamma z).$$

This belongs to $L^2(\Gamma \backslash H)$. As η varies over the space $C_c^\infty(0, \infty)$, the functions θ_ψ will generate a subspace of $L^2(\Gamma \backslash H)$ which we will denote $\Theta(\Gamma \backslash H)$.

Now consider the Mellin transform

$$L_\eta(s) = \int_0^\infty \frac{\eta(y)\, dy}{y^{s+1}}.$$

Then, $L_\eta(s)$ is an entire function which is rapidly decreasing in vertical strips. By the Mellin inversion formula,

$$\eta(y) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} L_\eta(s)\, y^s\, ds$$

for $\sigma \in \mathbf{R}$. Therefore,

$$\theta_\psi(z) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} L_\eta(s)\, E(z, s)\, ds$$

where

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

is the *Eisenstein series*. It is easy to see that

$$E(\gamma z, s) = E(z, s) \;\forall\; \gamma \in \Gamma$$

and that

$$DE(z, s) = s(s-1)\, E(z, s).$$

Moreover, $E(z, s)$ has a Fourier expansion of the form:

$$E(z, s) = y^s + \phi(s)\, y^{1-s} + y^{1/2} \sum_{m \neq 0} c_m(s)\, K_{s-\frac{1}{2}}(2\pi \mid m \mid y) \exp(2\pi i m x)$$

where

$$\phi(s) = \pi^{1/2} \frac{\Gamma(s-\frac{1}{2})}{\Gamma(s)} \sum_{c>0} \frac{N(c)}{c^{2s}},$$

$$N(c) = \# \left\{ 0 \leqslant d < c : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma, \textit{for some } a, b \right\}$$

in the special case under consideration.

One can prove that $E(z, s)$ has a meromorphic continuation for all $s \in \mathbf{C}$, and satisfies a functional equation

$$E(z, s) = \phi(s)\, E(z, 1-s).$$

The poles of $E(z, s)$ are independent of z and are finite in number, say, $s_1, \ldots, s_R$, corresponding to the (simple) poles of $\phi(s)$ in the half plane $Re(s) \geqslant \frac{1}{2}$.

In the general case of more than one cusp, we have an Eisenstein series defined by

$$E_{\mathcal{A}}(z, s) = \sum_{\gamma \in \Gamma_{\mathcal{A}} \backslash \Gamma} (Im(\gamma z))^s$$

where $\Gamma_{\mathcal{A}}$ denotes the stabiliser of the cusp $\mathcal{A}$ in Γ. About each cusp $\mathcal{B}$, we can write down a Fourier expansion of the form

$$E_{\mathcal{A}}(z, s) = y^s + \phi_{\mathcal{A}\mathcal{B}}(s)\, y^{1-s} + \ldots$$

Then, each Eisenstein series has a meromorphic continuation for all $s \in \mathbf{C}$ and the column vector

$$(E_{\mathcal{A}}(z, s))^{tr} = \mathcal{E}(z, s)$$

satisfies the functional equation

$$\mathcal{E}(z, s) = \Phi(s)\, \mathcal{E}(z, 1-s)$$

where

$$\Phi(s) = (\phi_{\mathcal{A}\mathcal{B}}(s))$$

is called the scattering matrix.

Returning to our special case, we see by the residue theorem that

$$\theta_\psi(z) = \sum_{j=0}^{R} L_\eta(s_j)\,\phi_j(z) + \frac{1}{2\pi i}\int_{\frac{1}{2}-i\infty}^{\frac{1}{2}+i\infty} L_\eta(s)\,E(z, s)\,ds$$

where

$$\phi_j(z) = \mathrm{Res}_{s=s_j} E(z, s).$$

Moreover, $\phi_j(z)$ is an eigenfunction of D, and the self adjointness of D implies that $\phi_j(z)$ is orthogonal to $E(z, \frac{1}{2} + ir)$ for $j = 1, 2, \ldots, R$. In a sense, every element of $\ominus(\Gamma\backslash H)$ is a finite linear combination of the ϕ_j's and a continuous sum over $E(z, \frac{1}{2} + ir)$.

What is the orthogonal complement of $\ominus(\Gamma\backslash H)$? It consists of f such that for every $\theta_\psi \in \ominus(\Gamma\backslash H)$,

$$0 = \int_{\Gamma\backslash H} \theta_\psi(z)\,\overline{f(z)}\,\frac{dx\,dy}{y^2}$$

$$= \int_0^\infty\int_0^1 \psi(z)\,\overline{f(z)}\,\frac{dx\,dy}{y^2},$$

by unfolding the integral in the usual way. This implies that

$$\int_0^1 f(x+iy)\,dx = 0$$

and hence f is a cusp form. It is then clear that the subspace generated by all cusp forms, denoted $L_0^2(\Gamma\backslash H)$, is contained in $\ominus(\Gamma\backslash H)^\perp$. In fact, one can obtain the decomposition

$$L^2(\Gamma\backslash H) = L_0^2(\Gamma\backslash H) \oplus \ominus(\Gamma\backslash H).$$

This leads to the eigenfunction expansion of $f \in L^2(\Gamma\backslash H)$ as

$$f(z) = \sum_{n=0}^{\infty} (f, f_n)\,f_n(z) + \frac{1}{4\pi}\int_{-\infty}^{\infty} (f, E(\cdot, \tfrac{1}{2} + ir))\,E(z, \tfrac{1}{2} + ir)\,dr.$$

If we define L by

$$L\,f(z) = \int_H k(z, w)\,f(w)\,d\mu(w)$$

for $f \in L_0^2(\Gamma \backslash H)$, then

$$Tr(L \mid L_0^2(\Gamma \backslash H)) = \sum_{n=0}^{\infty} h(r_n).$$

The automorphic kernel function can expanded as

$$K(z, w) = \sum_{n=0}^{\infty} h(r_n) f_n(z) f_n(w) + \frac{1}{4\pi} \int_{-\infty}^{\infty} h(r) E(w, \tfrac{1}{2} - ir) E(z, \tfrac{1}{2} + ir)\, dr$$

for each $w \in \mathcal{F}$. If we set

$$\tilde{K}(z, w) = \sum_{n=0}^{\infty} h(r_n) f_n(z) f_n(w)$$

then

$$\int_{\mathcal{F}} \tilde{K}(z, z)\, d\mu(z) = \sum_{n=0}^{\infty} h(r_n).$$

Now, by considering appropriately truncated fundamental domains and expanding $\tilde{K}(z, z)$ and integrating term by term as before, but very carefully, one derives the following trace formula.

Trace formula (Selberg)

$$\sum_{n=0}^{\infty} h(r_n) = \frac{\operatorname{vol}(\mathcal{F})}{4\pi} \int_{-\infty}^{\infty} r h(r) \tanh(\pi r)\, dr +$$

$$+ \sum_{\{P\}} \frac{\Lambda(P)}{N(P)^{1/2} - N(P)^{-1/2}} g(\log N(P)) +$$

$$+ \sum_{\{T\}} \frac{1}{2m \sin\theta} \int_{-\infty}^{\infty} \frac{e^{-2\theta r}}{1 + e^{-2\pi r}} h(r)\, dr +$$

$$+ \frac{1}{4\pi} \int_{-\infty}^{\infty} h(r) \left\{ \frac{\phi'}{\phi}(\tfrac{1}{2} + ir) - 2 \frac{\Gamma'}{\Gamma}(1 + ir) \right\} dr +$$

$$+ \tfrac{1}{4}(1 - \phi(\tfrac{1}{2}))\, h(0) - g(0) \log 2,$$

where the sum $\{P\}$ is over hyperbolic conjugacy classes, $\{T\}$ is over the elliptic conjugacy classes. In the latter sum, there are only finitely many terms, m denotes the order of $C_\Gamma(T)$ and $\theta = \cos^{-1}(Tr(T)/2)$. The last four terms correspond to the cusp at ∞, the single parabolic term in this case.

In the general case, there will be additional terms arising from the other (finitely many) parabolic classes. Indeed, if we let the matrix integral:

$$\mathcal{M} = \frac{1}{4\pi}\int_{-\infty}^{\infty} h(r)\,\Phi'\left(\tfrac{1}{2}+ir\right)\Phi\left(\tfrac{1}{2}+ir\right)^{-1} dr +$$

$$\tfrac{1}{4}\left(1-\Phi(\tfrac{1}{2})\right)h(0) - \left[g(0)\log 2 + \frac{1}{2\pi}\int_{-\infty}^{\infty} h(r)\,\frac{\Gamma'}{\Gamma}(1+ir)\,dr\right] I$$

then the parabolic contribution is $Tr(\mathcal{M})$. For example, if $\Gamma = PSL_2(\mathbf{Z})$, we have only one cusp at $i\infty$, and

$$\phi(s) = \pi^{1/2}\,\frac{\Gamma\left(s-\frac{1}{2}\right)}{\Gamma(s)}\,\frac{\zeta(2s-1)}{\zeta(2s)}$$

and

$$Tr(\mathcal{M}) = g(0)\log\frac{\pi}{2} - \frac{1}{2\pi}\int_{-\infty}^{\infty} h(r)\left\{\frac{\Gamma'}{\Gamma}\left(\tfrac{1}{2}+ir\right) + \frac{\Gamma'}{\Gamma}(1+ir)\right\} dr +$$

$$+\, 2\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n}\,g(\log n).$$

The classical von Mangoldt function appears because the logarithmic derivative of ϕ involves the zeta function. The details can be found in Hejhal [4].

8. CONCLUDING REMARKS

The idea of interpreting the zeroes of $\zeta(s)$ as eigenvalues of some Hermitian operator was first pointed out independently by Polya and Hilbert. The analogy deserves further exploration.

We have already gained much insight by pursuing analogies. Perhaps, the most compelling evidence for the Riemann hypothesis arises from the function field analogue (proved by Weil) and the Weil conjectures (proved by Deligne). On the other hand, it is not quite clear if this is such a sharp analogy. One of the most striking consequences of the Langlands program, for instance, is the interpretation of the Ramanujan-Petersson conjecture for modular forms of half-integral weight, as an analogue of the Lindelöf hypothesis (via a theorem of Waldspurger). The Lindelöf hypothesis is the assertion that

$$\zeta\left(\tfrac{1}{2} + it\right) \ll t^{\epsilon},$$

for every $\epsilon > 0$. This was shown to be a consequence of the Riemann hypothesis by Littlewood. It is a milder hypothesis. To gain some understanding of the intractibility of this hypothesis, let us cite the following *mean value hypothesis*

$$\frac{1}{T}\int_{-T}^{T} \left| \zeta\left(\tfrac{1}{2} + it\right) \right|^{2k} dt \ll T^{\epsilon}$$

for every $k \geqslant 1$. At present, this is only known for $k = 1$ and $k = 2$. The mean value hypothesis would be a consequence of the Lindelöf hypothesis.

When Riemann wrote his eight page paper in 1860, he may have been unaware of the tremendous excitement that it would generate in mathematics. Neither, did he single out his famous hypothesis for further study in that paper or give it a special position. After proving the functional equation of $\zeta(s)$, he simply stated a series of unproved propositons which included the celebrated unsolved hypothesis. In a fragmentary note found among his posthumous papers, Riemann wrote that these properties "follow from an expression for the function $\zeta(s)$ which I have not yet simplified enough to publish." In his book, "The Psychology of Invention in the Mathematical Field," Hadamard refers to this passage of Riemann. After a century, he says, "we still have not the slightest idea of what the expression could be ... In general, Riemann's intuition is highly geometrical, but this is not the case for his memoir on prime numbers, the one in which that intuition is the most powerful and mysterious."

It would seem that mathematics is born out of highly intuitive minds confronted by narrow margins and expressions that never simplified.

References

1. Courant, R. and D. Hilbert, *Methods of Mathematical Physics*, vol. 1, Wiley Interscience, New York, 1953.

2. GELFAND, I.M., M.I. GRAEV AND I.I. PIATETSKII-SHAPIRO, *Rspresentation Theor and Automorphic Functions*, W.B. Saunders Company, Philadelphia, 1969.
3. HEJHAL, D.A. The Selberg trace formula and the Riemann zeta function, *Duk Math. Journal*, 43 (1976) 441–482.
4. D.A. HEJHAL, *The Selberg trace formula for* PSL_2 (R), volume 1, Springe Lecture Notes, 548, Springer Verlag, New York.
5. IWANIEC, H. Prime geodesic theorem, *J. Reine Angew. Math.*, 349 (1984) 136 159.
6. KUBOTA T. *Elementary Theory of Eisenstein Series*, Halsted Press, New York 1973.
7. RIEMANN, B. *Gessammelte Math. Werke*, Zweite Auflage, B.G. Teubner Leipzig, 1892.
8. SELBERG, A. Harmonic analysis and discontinuous groups in weakly symmetri Riemannian spaces with applications to Dirichlet series, *J. Indian Math. Soc.* 20 (1956) 47–87.
9. A. WEIL, Sur les formules explicites de la théorie des nombres premiers, *Comm Sèm. Math. Lund.*, Tome Supplementaire, (1952) 252–265.

Department of Mathematics,
McGill University,
Montréal, Canada.