The Riemann zeta function and Hamburger's theorem

S.J.Patterson
Mathematisches Institut
Georg-August-Universität Göttingen

This talk was partly historical and partly philosophical in its conception. The intention was to read between the lines of some of our forebearers. Although most of the material is well-known it seemed worth-while to to discuss it along with some of its implications.

Hamburger's theorem: [3] Let $f_j(s) = \sum a_j(n)n^{-s} \quad (j = 1, 2)$ be Dirichlet series convergent in $\mathrm{Re}(s) > 1$ with analytic continuations to $\mathbb{C}$ as functions of finite order with only finitely many singularities and, if with $G(s) = \pi^{-s/2}, \Gamma(s/2)$ we have $G(s)f_1(s) = G(1-s)f_2(1-s)$ then $f_1 = f_2 = \zeta$.

To get some sort of feel for this I shall first sketch various proofs and discuss them.

We consider a more general situation and follow essentially [5].

$$f_1(s) = \sum_{n=1}^{\infty} a_1(n)\lambda_n^{-s} \quad \text{conv. in } \mathrm{Re}(s) > \sigma_1,$$

$$f_2(s) = \sum_{n=1}^{\infty} a_2(n)\mu_n^{-s} \quad \text{conv. in } \mathrm{Re}(s) > \sigma_2.$$

We assume that f_1, f_2 have analytic continuations to $\mathbb{C}$ as meromorphic functions of finite order and only finitely many singularities. We assume also that

$$G(s)f_1(s) = G(1-s)f_2(s).$$

By Cauchy's theorem we get

$$\frac{1}{2\pi i}\int_{(\sigma)} f_1(s)\frac{X^s}{s(s+1)}ds = \sum \mathrm{Res}\left(f_1(s)\frac{X^s}{s(s+1)}\right) + \frac{1}{2\pi i}\int_{(\sigma')} f_2(s)\frac{G(s)}{G(1-s)}\frac{X^{1-s}ds}{(s-1)(s-2)}$$

where $\sigma > \sigma_1$, $\sigma_2 < \sigma' < 3/2$. The two sides can be evaluated and we obtain

$$\sum_{\lambda_n \le X} a_1(n)(1 - \frac{\lambda_n}{X}) = \sum \mathrm{Res}\left(\frac{f_1(s)X^s}{s(s+1)}\right) - \sum_m a_2(m)\frac{\cos(2\pi\mu_m X) - 1}{(2\pi\mu_m X)^2}X.$$

Now let $R(X) = X \cdot \sum \mathrm{Res}\left(\frac{f_1(s)X^s}{s(s+1)}\right)$; this is a smooth function. *If* we assume that $\{\mu_m : m > 1\} \subset \mathbb{N}$ then $\sum_{\lambda_n \le X} a_1(n)(X - \lambda_n) - R(X)$ is periodic, whence we deduce that $\sum_{X < \lambda_n \le X+1} a_n(n) = R(X+1) - R(X)$. The left-hand side is piecewise constant; the right-hand side is smooth. ¿From this we deduce that $R(X+1) = R(X) + k$ for some constant k, and that $\sum_{X<\lambda_n \le X+1} a_n(n) = k$. This latter statement means that with perhaps the exception of some "small" terms (i.e. with λ_n small) $f_1(s)$ is of the form $\sum d_\alpha \zeta(s,\alpha)$ where $\zeta(s,\alpha) = \sum\limits_{n)1}^{\infty} \frac{1}{(n+\alpha)^\sigma}$ (the Hurwitz zeta-function) and $\sum d_\alpha = k$. We are mow in the realm of known functions and we deduce that $\alpha : 0 \leq \alpha < 1$, $d_\alpha = d_{1-\alpha} (\alpha \not\equiv 0)$. Note that the arithmetic assumption was on $\{\mu_m \mid m \in \mathbb{N}\}$. We now deduce Hamburger's theorem by demanding additionally that $\{\lambda_n : n \in \mathbb{N}\} \subset \mathbb{N}$. If, for example, we make the weaker assumption that $\{\lambda_n : n \in \mathbb{N}\} \subset \frac{1}{D}\mathbb{N}$ then f_1 is a linear combination of $D^5 L(s,\chi)$ with $f_\chi | D$ and χ even. The assumption $\sigma_2 < 3/2$ is quite strong. This argument can be modified in various ways. The argument can become trickier if instead of $\mathbb{N}$ or $\mathbb{Z}$ a multidimensional lattice is involved. Then one has to "twist" with, for example, harmonic polynomials. In this way F. Sato [12] showed a "Hamburger theorem" for Epstein zeta functions.

Let us look at another variant, this time following the ideas of [6]. We consider

$$\frac{1}{2\pi i} \int\limits_{\sigma} f_1(s)\Gamma(s/2)\pi^{-s/2} y^{-s/2} ds = 2 \sum a_1(n) e^{-\pi\lambda_n^2 Y}.$$

The same type of argument as before shows that

$$\begin{aligned} \sum a_1(n) e^{\pi i \lambda_n^2 z} &= \sum \operatorname{Res}(f_1(s)\Gamma(\frac{s}{2})\pi^{-s/2}(z/i)^{-s/2}) \\ &+ (\frac{i}{z})^{\frac{1}{2}} \sum a_2(n) e^{-\pi i \mu_n^2/z} \quad (\operatorname{Im}(z) > 0). \end{aligned}$$

Let us write $\tilde{f}_1(z), \tilde{f}_2(z)$ for $\sum\limits_{n=1}^{\infty} a_1(n) e^{\pi i \lambda_n^2 z}$ and $\sum\limits_{n=1}^{\infty} a_2(n) e^{\pi i \mu_n^2 z}$. Let us suppose now that we know that $\lambda_n^2 \in \mathbb{N}$, $\mu_n^2 \in \mathbb{N}$. Then we have $\tilde{f}_1(z+2) = \tilde{f}_1(z)$; $\tilde{f}_2(z+1) = \tilde{f}_2(z)$. Since the group generated by $\left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ is free on these two generators we cannot make any further deductions directly. If we assume that there are at most simple poles at 0 and 1 of Gf_1 (and so of Gf_2), with residues p_0 and p_1 then we get

$$-p_0 + \sum a_1(n) e^{\pi i \lambda_n^2 z} = (\frac{i}{z})^{\frac{1}{2}} (p_1 + \sum a_2(n) e^{-\pi i \mu_n^2/z}).$$

Let us write $\overset{\approx}{f}_1$ and $\overset{\approx}{f}_2$ for the functions on the left and right hand sides here. Then we would get that $\overset{\approx}{f}_1\ (z+2) = \overset{\approx}{f}_1\ (z)$; $\overset{\approx}{f}_2\ (z+2) -\ \overset{\approx}{f}\ (z)$; $\overset{\approx}{f}_1\ (z) = (\frac{1}{z})^{\frac{1}{2}}\ \overset{\approx}{f}_2\ (\frac{1}{z})$. The group $\mathfrak{G}(2) = \langle \left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right), \left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right) \rangle$ has two cusps, one at ∞ and one at 1 (it is of index 3 in the modular group). The classical theta function $\theta(z) = \sum e^{\pi i n^2 z}$ has a transformation under $\mathfrak{G}(2)$ and so $(\overset{\approx}{f}_1 \pm \overset{\approx}{f}_1)/\theta$ is a function (with character) on $\mathfrak{G}(2)$ and, as the only singularity is at the cusp 1, we can classify the possibilities through a polynomial in a Hauptmodul; the particular polynomial is identified by the behaviour of $\sum a_1(n)(-1)^{\lambda_n^2} \lambda_n^{-s}$ and $\sum a_2(n)(-1)^{\mu_n^2} \mu_n^{-s}$. It is possible to use these ideas to construct many more solutions of the functional equation for the Riemann zeta function.

We let (as an example) $\mathfrak{G}(\lambda) = \langle \left(\begin{smallmatrix} 1 & \lambda \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \rangle$, $\lambda = 2\cos\frac{\pi}{q}$ $(q \geq 3)$; this is a triangle group of the form $(2, q, \infty)$. We can construct forms of weight $\frac{1}{2}$ on $\mathbb{H}$ under $\mathfrak{G}(\lambda_q)$; it has a Fourier expansion of the form

$$\sum_{m=0}^{\infty} c_m \cdot e^{\frac{2\pi i}{\lambda_q}(m+\frac{q-1}{8q})z}$$

and so the Dirichlet series is of the form

$$\sum_{m=0}^{\infty} c_m (8q \cdot m + q - 1)^{-\frac{s}{2}} \cdot (8q\lambda_q)^{s/2}, \quad (c_0 \neq 0)$$

and so we obtain other families of solutions with

$$\{\lambda_n\} = \{\mu_n\} = \left\{ \left(\frac{8q \cdot m + q - 1}{8q \cdot \lambda_q} \right)^{1/2} : m \geq 0 \right\}.$$

This gives a wider class of solutions and these show the relevance of the assumptions on $\{\lambda_n\}, \{\mu_n\}$. Apparently very little is known about the c_m is these cases.

Having discussed this particular approach suggests a more systematic approach valid for more general situations. Here the general idea is that one has a deduction of the following type:

$$\left.\begin{array}{c} \text{Dirichlet series} \\ \text{Euler product} \end{array}\right\} \quad \to \text{Automorphic form.}$$

On the left-hand side one needs various "twists". In the first case the classic example is Weil's theorem [14]. It is worth noting that in this case one can deal with fairly complicated sets of poles and, in fact one can use group theoretic arguments to restrict the possibilities of poles (cf. theory of cubic metaplectic forms, [9]). The second possibility was explored first by Jacquet & Langlands and by Piatetski-Shapiro cf. [7], [10] but see also [15]. there are many variants although in general one avoids poles. There are often difficulties at a finite number of "bad" factors. Nevertheless over the past 25 years or so a range of tricks have been developed. this method has been used mainly to investigate possible "correspondences" between either Galois L-series or automorphic L-series on one group with automorphic forms on a group of type $GL(r)$. This is somewhat different emphasis than the problem underlying Hamburger's theorem.

There are two questions open. The first is that having found an automorphic form we have to identify it. This is not all that easy; even to prove. that spaces of automorphic forms are finite dimensional is sophisticated. One method — the most popular nowadays — is to use methods of integral operators and functional analysis. The other is due to Siegel (and apparently inspired by Hardy's work on the Ramanujan τ-function and the circle method) is much more effective and it can, at least in some cases, be make the basis of explicit computations - see [11].

We should however discuss one other rather odd feature. The ζ-function of Riemann is associated with the classical θ-function $\sum q^{n^2}$ $(q = e^{\pi i z})$, i.e. an automorphic form on $\widetilde{GL}_2(\mathbb{Q}_\mathbb{A})$, the 2-fold metaplectic cover of $GL_2(\mathbb{Q}_\mathbb{A})$. It "should" be associated with $GL_1(\mathbb{Q}_\mathbb{A})$ and we see that the Shimura correspondence yields a correspondence between 1-dimensional representations of $GL_2(\mathbb{Q}_\mathbb{A})$ and "exceptional" representations of $\widetilde{GL}_2(\mathbb{Q}_\mathbb{A})$. If we consider $K/\mathbb{Q}$ quadratic then we obtain

a) $K/\mathbb{Q}$ real (Maaß, [8]; see also [13])

$$f(x,v) = c \cdot v^{\frac{1}{2}} + \sum_{\substack{\alpha \in \mathfrak{O} \\ \text{mod.units}}} v^{\frac{1}{2}} K_0(4\pi N(\alpha)v) e(2\pi N(\alpha)x);\ c \equiv \log \varepsilon^2$$

if

$$\theta(z_1, z_2, z_{12}) = \sum_{\alpha \in 0} \exp(2\pi i (i_1(\alpha)^2 z_1 + i_2(\alpha)^2 z_2 + i_1(\alpha) i_2(\alpha) z_{12})) \text{ “Siegelsche } \theta - f^{\underline{n}}\text{”}$$

then

$$f(x,v) = v^{\frac{1}{2}} \cdot \int\limits_{\mathbb{R}_+^y / \langle \varepsilon^2 \rangle} \theta(iyv, iy^{-1}v, x) d^y y$$

b) $K/\mathbb{Q}$ imaginary (Hecke, [6])

$$f(z,v) = \sum_{\alpha \in 0} e^{2\pi i N(\alpha) z}$$

a theta function of weight 1.

As was first observed by Hecke these can be analyzed in terms of non-holomorphic (Maaß-) and holomorphic Eisenstein series, essentially giving another way of looking at the fundamental theorems of class-field theory. Although these are very suggestive concerning the relationships between automorphic forms on various groups andtheir metaplectic covers this has not, as yet, proved to be real significance in understanding the arithmetic aspects of the zeta function. Roughly speaking, this theory is at the level of the Poisson Summation Formula; the arithmetic theory is at the level of the explicit formula of prime number theory.

Now let us come back to the Riemann zeta function. In 1921 in a lecture to the Mathematical Society G.H. Hardy [4] expressed the opinion "You must make me a present of hypothesis R [the Riemann Hypothesis for $L(s,\chi)$, all primitive Dirichlet character s χ] [in order to attack the Goldbach conjecture]. I presume that the hypothesis of Riemann will some day be proved within a week from then, and the proofs will be substantially the same. There is nothing whatever to suggest that, in these respects, one L-function behaves unlike another" Hardy was expressing the conventional wisdom of his day, and an opinion that is, I think, generally accepted today. I think that one should reappraise it. The ideas about the zeta function which are perhaps the most widely accepted are of the general form:

— the explicit formulae of prime number theory should be interpreted in terms of a topological vector space
— the eigenvalues should be interpreted in terms of an operator (preferably a differential operator)
— the Riemann Hypothesis should be equivalent to the self-adjointness of this differential operator with respect to an inner product (positive).

This idea — at least according to the mythology – goes back to Hilbert and the original devolopment of this theory. In the case of the characteristic p analogue a closely analogous approach works and it suggests more precise versions of the strategy above. One thing that one learns from this is the following: the structure of the space and operator are crucial — it does not suffice simply to use "explicit formulae". Examples like:

If

$$N(i) = \sum_{d|i} \mu(\frac{i}{d})\{1 + q^d - (\sqrt{-3})^d(1 + 3^d)(1 + (-1)^d)\}$$

then

$$\prod_{j=1}^{\infty}(1 - X^j)^{-N(j)} = \frac{1 - 30X^2 + 81X^4}{(1 - X)(1 - 9X)}$$

and $N(i) \geq 0$. This looks like a zeta function over $\mathbb{F}_q$ but the roots are the inverses of $\pm\sqrt{3}$ and $\pm 3\sqrt{3}$. Generally speaking it seems as if an Euler product structure is very common and that it does not really mean all that much. One would therefore expect that if one had a proof which was merely based on the explicit formulae then it would apply *mutatis mutandi* to the function above, and this would be a contradiction. This situation appears to be incompatible with Hamburger's theorem, since the explicit formulae determine the zeta function. Although this may appear to be the case it is not such a contradiction as appears at first sight.

Let us first observe that the Hasse-Weil argument proves a statement of the form

$$\forall\ k\colon \mathrm{char}(k) > 0,\ k \text{ an } \mathbb{A} \text{ field the } RH \text{ holds for } \zeta_k.$$

This is a *very* strong statement. On the other hand the functional equation implies

$$\forall\ k\colon \mathrm{char}(k) > 0,\ \text{ one can decide the } RH \text{ for } k.$$

In the case of characteristic 0 essentially the same method is used to show that one can localize the zeros as closely as one likes in $\{s\colon 0 < \mathrm{Re}(s) < 1,\ |\mathrm{Im}(s)| < T\}$ and to verify that the Riemann Hypothesis exactly if one has provably simple zeros and approximately if one has apparently multiple zeros. (It is conjectured that the Riemann zeta function has simple zero — and I never expect to actually see a multiple one — but this does *not* hold for ζ_k with $\mathrm{char}(k) > 0$, for example for certain supersingular elliptic curves;

if $\mathrm{char}(k) = 0$, $k/\mathbb{Q}$ Galois then $\zeta_k(s) = \prod L(s,\chi)^{\deg(\chi)}$ and so the simplicity does not hold here either.)

What we see is that the Hasse-Weil proof is of a much more general assertion and one where the Hamburger theorem does not hold. What is also clear is that the Hasse-Weil approach is extreme by ambitions. Comparing the difficulties of the proofs of the two assertions we conclude (at least I do) that a much more approachable goal would be to seek a proof of the second statement. I should also add that Matyasevich has investigated the logical complexity of the classical Riemann Hypothesis. It is at the level of a Diophantine statement. On the other hand the general statement above — and its generalizations — are extremely difficult even to formulate in a usable form, for example one which would allows one to investigate the non-standard models of Robinson-Roquette.

References

[1] J.W.S.Cassels, Footnote to a note of Davenport and Heilbronn, J. London Math.Soc.36(1961)177-184

[2] H.Davenport, H.Heilbronn, On the zeros of certain Dirichlet series, J. London Math.Soc.11(1936)181-185,307-312

[3] H.Hamburger, Über die Riemannsche Funktionalgleichung der ζ-Funktion, Math.Z. 10(1921)240-254

[4] G. Hardy, Goldbach's Theorem, Coll. Papers I, 545-560 = Mat Tidssk.B (1922), 1-16.

[5] E.Hecke, Über die Lösungen der Riemannschen Funktionalgleichung, Math.Z. 16(1923)301-307 = Ges.Werke, 19

[6] E.Hecke, Über die Bestimmung Dirichletscher Reihen durch ihre Funktionalgleichung, Math.Ann.112(1936)664-699

[7] H.Jacquet, R.P.Langlands, Automorphic forms on GL(2), Springer Lecture Notes 114(1970)

[8] H.Maaß Über eine neue Art von nichtanalytischen automorphen Funktionen und ihre Bestimmung durch Funktionalgleichungen, Math.Ann.121(1949)141-183

[9] S.J.Patterson, A cubic analogue of the theta series, J.reine angew.Math.296(1977)125-161

[10] I.I.Piatetski-Shapiro, On the Weil-Jacquet-Langlands theorem, Proc. Summer School on Group Representations, Bolyai János Math. Soc. Budapest, Halsted, 1971, 583-595

[11] W.Roelcke, Über die Wellengleichung bei Grenzkreisgruppen erster Art, Sitzber. Akad.Heidelberg, Math.Naturw.Kl.1953/54

[12] F.Sato, The Hamburgr theorem for the Epstein zeta function, Algebraic Analysis, Papers ded. to M.Sato on the occasion of his sixtieth birthday, Vol.2, 1989, 789-807

[13] M.-F.Vignéras, Séries thêta der formes indefinies, Modular Forms in one variable, VI, Proceedings, Bonn, 1976, Springer Lecture Notes 627, 1977, 227-239

[14] A.Weil, Über die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen, Math.Ann.168(1967)149-156 = Coll.Papers, [1967a]

[15] A.Weil, Automorphic forms and Dirichlet series, Springer Lecture Notes 189, 1971

Mathematisches Institut
Bunsenstr. 3 - 5
D–37073 Göttingen
Germany