Automorphic forms, L-functions and number theory (March 12–16)

Three Introductory lectures

E. Kowalski

Université Bordeaux I - A2X, 351, cours de la Libération, 33405 Talence Cedex, France

E-mail address: emmanuel.kowalski@math.u-bordeaux.fr

CHAPTER 1

Elementary theory of L-functions, I

1.1. Introduction

In this first lecture we will define and describe, in a roughly chronological order from the time of Euler to that of Hecke, some interesting classes of holomorphic functions with strange links to many aspects of number theory. Later lectures will explain how some, at least, of the mysterious aspects are understood today. But it should be emphasized that there are still many points which are not fully explained, even in a very sketchy, philosophical way.

As my background is in analytic number theory, I will particularly try to mention some of the more peculiar features of the theory of L-functions (and of automorphic forms) which arise from this point of view. I will also give indications at the places where lecturers coming after me will bring new perspectives.

The next lecture will develop the points presented here, and in particular will sketch proofs of some of the most important ones, especially when such a proof yields new insights into the theory.

The mathematicians whose name are most important for us now are: Euler, Gauss, Dirichlet, Riemann, Dedekind, Kronecker, Hecke, Artin, Hasse.

1.2. The Riemann zeta function

The first L-function has been given Riemann's name. This fact is convenient for us: it seems to call for some explanation, since no one denies that other mathematicians, most notably Euler, considered this function before, and these explanations are the best entrance to our subject.

The function in question is defined by the series

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

For integers $s \geqslant 1$, this was studied even before Euler, and for even $s \geqslant 2$, it is well-known that Euler first found an exact formula (see (2.14)). However the starting point for the theory of L-functions is Euler's discovery that the existence and uniqueness of factorization of an integer as a product of prime powers implies that

$$(1.1) \qquad\qquad \zeta(s) = \prod_p \frac{1}{1 - \dfrac{1}{p^s}},$$

a product over all prime numbers. From the divergence of the harmonic series, Euler deduced from this a new proof of Euclid's theorem that there are infinitely many primes, and with

some care even obtained the more precise formulation that

$$\sum_{p \leqslant X} \frac{1}{p} = \log \log X + O(1)$$

as $X \to +\infty$.[1] Thus (1.1) clearly contained new information about the distribution of prime numbers.

Riemann's insight [**Rie**] was to remark that the function $\zeta(s)$ thus defined, if s is taken to be a complex variable ($s = \sigma + it$ in his notation), is holomorphic in its region of convergence. This justifies looking for its (maximal) analytic continuation, but even that had been done before (see [**We**]). It is the combination of the Euler product expansion (1.1) and the analytic continuation given by the functional equation described below, which is the cause for all this rejoicing among mankind, as it reveals the strange "duality" between the complex zeros of $\zeta(s)$ and prime numbers.

To be more specific, Riemann stated that $\zeta(s)$ has a meromorphic continuation to the whole complex plane, the only singularity being a simple pole with residue 1 at $s = 1$, and that moreover this analytic continuation satisfied the following property, aptly named the *functional equation*: the function

(1.2) $$\Lambda(s) = \pi^{-s/2} \Gamma(s/2) \zeta(s)$$

is meromorphic except for simples poles at $s = 0$ and $s = 1$ and satisfies for all $s \in \mathbf{C}$ the relation

(1.3) $$\Lambda(1 - s) = \Lambda(s).$$

From the simple poles of $\zeta(s)$ at $s = 1$ and of $\Gamma(s/2)$ at $s = 0$ one deduces in particular that $\zeta(0) = -\pi^{-1/2} \Gamma(1/2)/2 = -1/2$. Moreover, the other poles at $s = -2n$, $n \geqslant 1$ integer, of $\Gamma(s/2)$ show that (1.3) implies that $\zeta(-2n) = 0$, for $n \geqslant 1$: those zeros are called the *trivial zeros* of $\zeta(s)$.

This, and in fact all of (1.3) for integers $s \geqslant 1$, was already known to Euler in the language of divergent series!

The presence of the "gamma factor" in the functional equation was not well understood: indeed (1.3) was often written in completely different ways (see e.g. [**Ti**, Ch. 2]). The more transparent and conceptual proofs of (1.3), and its various generalizations, by the Poisson summation formula (see the next lecture) and theta functions gave the gamma factor a clear *de facto* standing, but it is only by the time of Tate's thesis [**Ta**] that its role was made clear as the Euler factor at the archimedean place of $\mathbf{Q}$. In general I should mention that an Euler product is a product over primes of the form

(1.4) $$\prod_p L_p(s)$$

---

[1] To avoid any controversy, here are the definitions of Landau's $O(\cdots)$ and Vinogradov's $\ll$ symbols: $f = O(g)$ as $x \to x_0$ means that there exists some (unspecified) neighborhood U of x_0, and a constant $C \geqslant 0$ such that $|f(x)| \leqslant Cg(x)$ for $x \in U$; this is equivalent to $f \ll g$ for $x \in U$ where now U is specified beforehand, and in this latter case one can speak of the "implicit constant" C in $\ll$. However we sometimes also speak of estimates involving $O(\cdots)$ being "uniform" in some parameters: this means that the U and C above can be chosen to be independent of those parameters.

and that $L_p(s)$ is called the Euler factor at p.[2]

The interplay of the zeta function with primes was revealed by Riemann in the form of the so-called "explicit formula". His version (and that proved later by van Mangoldt and others) obscured somewhat the essential point in purely analytic difficulties; going around them by means of smooth test functions, one can state it in the following straightforward manner:

PROPOSITION 1.2.1. *Let* φ : $]0, +\infty[\to \mathbf{C}$ *be a* C^∞ *function with compact support, let*

$$\hat{\varphi}(s) = \int_0^{+\infty} \varphi(x)x^{s-1}dx,$$

be its Mellin transform, which is entire and decays rapidly in vertical strips. Let

$$\psi(x) = \frac{1}{x}\varphi(x^{-1}).$$

Then we have

$$(1.5) \quad \sum_p \sum_{k \geqslant 1} (\log p)(\varphi(p^k) + \psi(p^k)) = \int_0^{+\infty} \varphi(x)dx$$

$$- \sum_{\substack{\zeta(\rho)=0 \\ 0<\mathrm{Re}(\rho)<1}} \hat{\varphi}(\rho) + \frac{1}{2i\pi} \int_{(-1/2)} \left(\frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{s}{2}\right) - \frac{1}{2}\frac{\Gamma'}{\Gamma}\left(\frac{1-s}{2}\right)\right)\hat{\varphi}(s)ds.$$

A more general case of this will be sketched in the next lecture (Proposition 2.3.5).

In this general formula, which expresses sums over prime numbers in terms of sums over zeros of $\zeta(s)$, the first term on the right-hand side is really the contribution of the pole at $s = 1$ (very often, giving the main term in some asymptotic formula). We implicitly make use of the fact that $\zeta(s) \neq 0$ for $\mathrm{Re}(s) > 1$. It is very important to realize that this is by no means obvious from the series representation: it follows immediately from the Euler product expansion (an absolutely convergent infinite product is non-zero). This must not be underestimated: it is for instance the key analytical input in Deligne's first proof of the Riemann Hypothesis for varieties over finite fields [**De**, 3.8]. The functional equation then shows that $\zeta(s) \neq 0$ for $\mathrm{Re}(s) < 0$, except for the trivial zeros already identified.

Riemann immediately expresses that it is "likely" that all the (non-trivial) zeros of $\zeta(s)$ satisfy $\mathrm{Re}(s) = 1/2$. This is indeed the best possible situation if one is interested in prime numbers: if one takes for φ test functions which converge (in the sense of distributions, say) to the characteristic function of $[0, X]$, one finds rather easily that one has an estimate for the error term in the prime number theorem of the form

$$(1.6) \qquad \psi(X) = \sum_{p^k \leqslant X} (\log p) = X + O(X^{\beta^*}) \text{ as } X \to +\infty$$

if and only if $\beta^* > \sup\{\mathrm{Re}(\rho)\}$. Since the functional equation implies that $\sup\{\mathrm{Re}(\rho)\} \geqslant 1/2$, the Riemann Hypothesis is seen to be simply the statement that primes are distributed "in the best possible way".

---

[2] Be careful that when an Euler product is defined over prime ideals $\mathfrak{p}$ in a number field $\neq \mathbf{Q}$, as will be described below, the data of the Euler factors $L_{\mathfrak{p}}(s)$ is *stronger* than the product, even if the latter defines a holomorphic function.

Following the standard terminology, the strip $0 \leqslant \operatorname{Re}(s) \leqslant 1$, which must contain the non-trivial zeros is called the *critical strip* and the line $\operatorname{Re}(s) = 1/2$ is called the *critical line*. This will apply to all the L-functions in the first two lectures, but for automorphic L-functions, the critical strip may (depending also on normalization) be translated to the right by some amount.

1.3. Dirichlet L-functions

Although it is fundamental, the case of the Riemann zeta function does not exhibit some very important features of the general theory. Those, notably the notions of "conductor" and of "primitivity", and the link with class-field theory and algebraic number theory more generally, appear first in the case of Dirichlet L-functions.

Dirichlet defined those functions [**Di**] to prove his famous theorem:

THEOREM 1.3.1. *Let $q \geqslant 1$ and $a \geqslant 1$ such that $(a, q) = 1$. Then there are infinity many primes $p \equiv a \,(\mathrm{mod}\, q)$ and more precisely*

$$\sum_{\substack{p \leqslant X \\ p \equiv a \,(\mathrm{mod}\, q)}} \frac{1}{p} = \frac{1}{\varphi(q)} \log\log X + O(1)$$

as $X \to +\infty$.

He proved this result by detecting invertible congruence classes modulo q by means of harmonic analysis in $(\mathbf{Z}/q\mathbf{Z})^\times$, which lead him to Dirichlet characters: an arithmetic function $\chi : \mathbf{Z} \to \mathbf{C}$ is a Dirichlet character modulo $q \geqslant 1$ if there exists a group homomorphism

$$\tilde{\chi} : (\mathbf{Z}/q\mathbf{Z})^\times \to \mathbf{C}^\times$$

such that $\chi(x) = \tilde{\chi}(x \,(\mathrm{mod}\, q))$ for $(x, q) = 1$ and $\chi(x) = 0$ if $(x, q) \neq 1$. To such a character, extending Euler's definition, one associates the L-function

$$L(\chi, s) = \sum_{n \geqslant 1} \chi(n) n^{-s} = \prod_p (1 - \chi(p) p^{-s})^{-1}$$

the last equation, the Euler product, being a consequence of unique factorization and of the complete multiplicativity of χ:

$$\chi(mn) = \chi(m)\chi(n) \text{ for all } m \geqslant 1, \ n \geqslant 1.$$

The orthogonality relations for characters of a finite group imply immediately the relation

$$\sum_{\substack{p \leqslant X \\ p \equiv a \,(\mathrm{mod}\, q)}} \frac{1}{p} = \frac{1}{\varphi(q)} \sum_{p \leqslant X} \frac{1}{p} + \frac{1}{\varphi(q)} \sum_{\chi \neq 1} \overline{\chi}(a) \sum_{p \leqslant X} \frac{\chi(p)}{p}$$

and Dirichlet's Theorem is easily seen to be equivalent with the assertion that if $\chi \neq 1$ is a Dirichlet character, then $L(\chi, 1) \neq 0$. For Dirichlet (coming before Riemann) this is not in the sense of analytic continuation, but rather a statement about the sum of the (conditionally) convergent series which "is" $L(\chi, 1)$.[3] For the non-trivial character χ_4 modulo

---

[3] For $\chi \neq 1$, the partial sums are bounded so the series converges.

4, for instance, this series is simply

$$L(\chi_4, 1) = 1 - \frac{1}{3} + \frac{1}{5} - \cdots = \frac{\pi}{4}.$$

However, since χ is a periodic function of $n \geqslant 1$, it is not difficult to extend the proof of the analytic continuation of $\zeta(s)$ and its functional equation based on the Poisson summation formula and theta functions (explained in the second lecture) to establish analogue statements for Dirichlet characters. This almost forces us to introduce the notion of primitivity: a Dirichlet character χ modulo q is called *primitive*, and q is called its *conductor*, if there does not exist $\tilde{q} \mid q$, $\tilde{q} < q$ and a character $\tilde{\chi}$ of $(\mathbf{Z}/\tilde{q}\mathbf{Z})^\times$ such that

$$\chi(n) = \tilde{\chi}(n \,(\mathrm{mod}\,\tilde{q})) \text{ for } (n, \tilde{q}) = 1.$$

If q is prime, then any non-trivial character is primitive modulo q. Any Dirichlet character is *induced* in the way described above by a unique primitive character $\tilde{\chi}$, and one has the relation

(1.7)
$$L(\chi, s) = L(\tilde{\chi}, s) \prod_{p \mid q/\tilde{q}} (1 - \chi(p)p^{-s})$$

which shows that the analytic properties of Dirichlet L-functions can be reduced immediately to that associated to primitive characters, the second factor being a finite product (which is an entire function).

For χ primitive, the case $q = 1$ corresponds to the zeta function and is special; if $q > 1$, the character χ is non-trivial and one shows that $L(\chi, s)$ has an extension to $\mathbf{C}$ as an entire function. The functional equation requires more disctinctions: if $\chi(-1) = 1$, χ is called *even*, and one defines

$$\Lambda(\chi, s) = \pi^{-s/2} \Gamma(s/2) L(\chi, s),$$

whereas for $\chi(-1) = -1$ (χ is *odd*), one defines

$$\Lambda(\chi, s) = \pi^{-(s+1)/2} \Gamma((s+1)/2) L(\chi, s).$$

Moreover, let $\tau(\chi)$ be the Gauss sum attached to χ,

(1.8)
$$\tau(\chi) = \sum_{x \,(\mathrm{mod}\,q)} \chi(x) e(x/q)$$

where $e(z) = e^{2i\pi z}$. Then the functional equation is

(1.9)
$$\Lambda(\chi, s) = \varepsilon(\chi) q^{1/2-s} \Lambda(\overline{\chi}, 1 - s)$$

where

(1.10)
$$\varepsilon(\chi) = \begin{cases} \dfrac{\tau(\chi)}{\sqrt{q}} & \text{if } \chi(-1) = 1 \\ -i\dfrac{\tau(\chi)}{\sqrt{q}} & \text{if } \chi(-1) = -1. \end{cases}$$

Thus a number of features appear which are ubiquitous in the more general context:

- The functional equation (1.9) involves certain global invariants of the Dirichlet character. First, its conductor q; notice that even when starting with an imprimitive character, the functional equation will only be possible for the associated ("inducing") primitive character, and thus the *a priori* unknown conductor $\tilde{q}$ will appear in this way. This remark, in other contexts, is very fruitful (see Section 1.4).
- The second invariant is the *argument* of the functional equation $\varepsilon(\chi)$, also called the *root number*. It is a complex number of absolute value 1, as a simple calculation (or a second application of the functional equation) reveals (in the case of primitive character), and is related to the Gauss sum $\tau(\chi)$. Notice that the latter is also a kind of finite field analogue of the gamma function (multiplicative Fourier transform of an additive character). In general, the argument of the functional equation is a very delicate invariant.
- The functional equation relates $L(\chi, s)$ with the value of at $1-s$ of the *dual* character $\overline{\chi}$.

An important special case is that of real-valued characters, i.e. characters of order 2 (or quadratic characters). In such a case, we have $\overline{\chi} = \chi$ and the argument $\varepsilon(\chi)$ becomes a *sign* ± 1. Actually, Gauss had computed the exact value of $\tau(\chi)$ (Dirichlet gave a simple analytic proof) which implies that $\varepsilon(\chi) = +1$ for any primitive quadratic character χ, i.e. $\tau(\chi) = \sqrt{q}$ is $\chi(-1) = 1$ and $\tau(\chi) = i\sqrt{q}$ if $\chi(-1) = -1$.

We now come back to Dirichlet's proof of Theorem 1.3.1. It is enough to show that $L(\chi, 1) \neq 0$ for $\chi \neq 1$, and because of (1.7) one may assume that χ is primitive. Dirichlet easily showed that if χ is not real, then $L(\chi, 1) \neq 0$ (if not, 1 would also be zero of $L(\overline{\chi}, s) \neq L(\chi, s)$ and this is easy to exclude). The difficult case, as it has remained to this day, is that of a quadratic χ. Dirichlet's proof that $L(\chi, 1) \neq 0$ is a direct application of his analytic class number formula (see (1.15) below) and thus introduces another recurrent and all-important theme, that of linking L-functions defined by analytic means (here the Dirichlet L-function) to others coming from algebraic number theory, or algebraic geometry (here the Dedekind zeta function of a quadratic field).[4] We explain this in the next section in the more general context of number fields, and will simply conclude by saying that this proof of non-vanishing of $L(\chi, 1)$ is spectacular, but somewhat misleading. One can indeed give very natural elementary proofs of the fact that

$$L(\chi, 1) \gg \frac{1}{\sqrt{q}},$$

for $q > 1$, which is the more precise outcome of the class number formula (at least in the imaginary case.) And more importantly, all progress concerning the problem (going back to Gauss) of understanding the order of magnitude of the class number of quadratic fields has come from the opposite direction, by using the class number formula to relate it to the special value of the L-function and estimating (more often, failing to estimate...) the latter quantity by analytic means (see e.g. Landau [**L**], Siegel [**Si**], Goldfeld [**Go**]).

---

[4] Today one would say of "motivic" origin.

1.4. Dedekind zeta functions

Let K be a number field, so that $K/\mathbf{Q}$ is a finite field extension, of degree $d = [K : \mathbf{Q}]$. We denote $\mathcal{O}$, or $\mathcal{O}_K$, the ring of integers in K. Dedeking defined the zeta function of K by the series

$$\zeta_K(s) = \sum_{\mathfrak{a}} \frac{1}{(N\mathfrak{a})^s}$$

(over non-zero integral ideals $\mathfrak{a} \subset \mathcal{O}$) which is easily seen to be absolutely convergent for $\mathrm{Re}(s) > 1$. So, for $K = \mathbf{Q}$, one recovers $\zeta(s)$. Since in $\mathcal{O}$ there is unique factorization of ideals into products of prime ideals, the same argument as for $\zeta(s)$ proves that there is an Euler product expansion

$$\zeta_K(s) = \prod_{\mathfrak{p}} (1 - (N\mathfrak{p})^{-s})^{-1},$$

and indeed $\zeta_K(s)$ carries much the same information about the distribution of prime ideals in K as $\zeta(s)$ does about ordinary prime numbers.

However, there is more to it: if we write $\zeta_K(s)$ as an ordinary Dirichlet series with some coefficients $r_K(n)$,

(1.11) $$\zeta_K(s) = \sum_{n \geqslant 1} r_K(n) n^{-s},$$

we immediately see that knowing $\zeta_K(s)$ implies knowing how prime numbers split in $\mathcal{O}$: indeed $r_K(p^k)$ is the number of ideals with norm $= p^k$, hence $r_K(p) = d$ if and only if p splits completely in K, $r_K(p) = 1$ means that p is totally ramified, etc...

This information is one of the most relevant to the study of the arithmetic of K, and this brings a new urgency to the exploration of the properties of $\zeta_K(s)$, which concentrates really on the zeta function itself instead of its logarithmic derivative (which is the key to the analytic distribution of prime ideals) and considers $\zeta(s)$ almost as a "trivial case".

Because of the definition by a series expansion, it is possible again to extend the method used for $\zeta(s)$ (and for $L(\chi, s)$) to derive analogous analytic properties of $\zeta_K(s)$. This is not completely straightforward however, and contains or uses most of the "basic" notions and results of algebraic number theory, which we now recall (see e.g. [**La**], [**CF**]):

- The field K has $d = r_1 + 2r_2$ embeddings $\sigma : K \hookrightarrow \mathbf{C}$, r_1 real embeddings and r_2 pairs of complex embeddings. To each is associated an absolute value $| \cdot |_\sigma$, $|z|_\sigma = |\sigma(z)|$ is σ is real and $|z|_\sigma = |\sigma(z)|^2$ if σ is complex. Those absolute values are also called the *archimedean places* of K; note σ and $\bar{\sigma}$ give rise to the same place. We use usually v to denote a place. We let $d_v = 1$ is v is real and $d_v = 2$ otherwise. Note that the norm of an element $z \in K$ is given by

$$Nz = \prod_v |z|_v \text{ (product over the places).}$$

- The ring $\mathcal{O}$ is of rank $d = [K : \mathbf{Q}]$. Let (ω_i) be a $\mathbf{Z}$-basis, $1 \leqslant i \leqslant d$. The discriminant D of K is

$$D = \det(\omega_i^\sigma)_{i,\sigma}^2.$$

For every prime ideal $\mathfrak{p}$ in $\mathcal{O}$, the quotient $\mathbf{F}_\mathfrak{p} = \mathcal{O}_K/\mathfrak{p}$ is a finite field and the *(absolute) norm* of $\mathfrak{p}$ is the cardinality of $\mathbf{F}_\mathfrak{p}$.

- The multiplicative group $U = \mathcal{O}^{\times}$ of units of $\mathcal{O}$ is a finitely generated abelian group of rank $r_1 + r_2 - 1$, the free part of which admits an injection into $\mathbf{R}^d$ through the "logarithmic" map
$$\ell : u \mapsto (\log |\sigma(u)|)_{\sigma}$$
the image of which is a lattice in the subspace $V = \{(x_{\sigma}) \mid \sum x_{\sigma} = 0 \text{ and } x_{\sigma} = x_{\bar{\sigma}}\}$. The volume $R > 0$ of a fundamental domain for $\ell(U) \subset V$ (with respect to the Lebesgue measure on V) is called the *regulator* of K.
- One denotes by $w = w_K$ the order of the torsion subgroup of U (the number of roots of unity in K).
- The group of ideal classes in K, denoted $\mathrm{Pic}(\mathcal{O})$ or $H(K)$, is a finite abelian group of order denoted $h(K) = h(\mathcal{O})$, called the *class number* of K. The ring $\mathcal{O}$ is principal (i.e. every ideal is principal) if and only if $h(K) = 1$.

If E/K is a finite Galois extension of number fields with Galois group G, for every prime ideal $\mathfrak{p}$ in K and prime ideal $\mathfrak{P} \mid \mathfrak{p}$ in E above $\mathfrak{p}$, one defines the *decomposition group* of $\mathfrak{P}$
$$G_{\mathfrak{P}} = \{\sigma \in G \mid \sigma\mathfrak{P} = \mathfrak{P}\}$$
and the *inertia subgroup*
$$I_{\mathfrak{P}} = \{\sigma \in G_{\mathfrak{P}} \mid \sigma(x) \equiv x \,(\mathrm{mod}\,\mathfrak{P}) \text{ for all } x \in \mathcal{O}_E\} \lhd G_{\mathfrak{P}},$$
and $\mathfrak{P}$ is unramified if $I_{\mathfrak{P}} = 1$.

The *Frobenius conjugacy class* $\sigma_{\mathfrak{P}}$, also denoted $[\mathfrak{P}, E/K]$, is the conjugacy class in $G_{\mathfrak{P}}/I_{\mathfrak{P}}$ (i.e. in the decomposition group if $\mathfrak{P}$ is unramified) of the elements σ such that
$$\sigma(x) \equiv x^{N\mathfrak{p}} \,(\mathrm{mod}\,\mathfrak{P}) \text{ for all } x \in \mathcal{O}_E,$$
i.e. the "reduction modulo $\mathfrak{P}$" of $\sigma_{\mathfrak{P}}$ is the canonical generator of the finite field extension $\mathbf{F}_{\mathfrak{P}}/\mathbf{F}_{\mathfrak{p}}$. The conjugacy class of σ *in* G only depends on $\mathfrak{p}$ and is denoted $\sigma_{\mathfrak{p}}$.

If G is an abelian group, the Frobenius conjugacy class is reduced to a single element, the Frobenius automorphism, which only depends on $\mathfrak{p}$ and is denoted $\sigma_{\mathfrak{p}}$.

The precise statement about analytic continuation and functional equation for $\zeta_K(s)$ takes the following form: let
$$\Lambda_K(s) = \pi^{-r_1 s/2}(2\pi)^{-r_2 s}\Gamma(s/2)^{r_1}\Gamma(s)^{r_2}\zeta_K(s).$$

Then $\Lambda_K(s)$ admits analytic continuation to a meromorphic function on $\mathbf{C}$ with simple poles at $s = 1$ and at $s = 0$. Thus $\zeta_K(s)$ is meromorphic with a simple pole at $s = 1$. Moreover the residue of $\zeta_K(s)$ at $s = 1$ is given by the *analytic class number formula*

$$(1.12) \qquad\qquad \mathrm{Res}_{s=1}\,\zeta_K(s) = \frac{2^{r_1}(2\pi)^{r_2}h(K)R}{w\sqrt{|D|}}.$$

The functional equation satisfied by $\Lambda_K(s)$ is

$$(1.13) \qquad\qquad \Lambda_K(s) = |D|^{1/2-s}\Lambda_K(1-s).$$

Compared to the previous functional equations, one notices that the discriminant appears as the conductor did for Dirichlet L-functions, and that the corresponding root number is always $+1$.

The formula (1.12) is a newcomer in these lectures, the prototypical example of a large class of formulae, proved or conjectured to hold, expressing the *special values* of L-functions

in terms of global invariants of the "motivic" objects they are related to. The best known generalization is the conjecture of Birch and Swinnerton-Dyer but there are many others.

The notion of a "special value" can be put in a rigorous conceptual framework (Deligne, etc...) explaining which values should have those properties for a given L-function. One should maybe mention that analytic number theorists often have reasons to consider as "special" points which escape this algebraic context: a typical example is the deformation theory of Phillips and Sarnak [**PS**] where the point in question is (probably) transcendental.

Let me now explain how this general formula relates to Dirichlet's proof of Theorem 1.3.1. Let χ be a non-trivial primitive quadratic character of conductor q. The essence of the argument is that there is a relation

$$(1.14) \qquad \zeta(s)L(\chi, s) = \zeta_K(s)$$

between the Dirichlet L-function and the Dedekind zeta function of the quadratic field $K = \mathbf{Q}(\sqrt{\chi(-1)q})$ (so K is imaginary if $\chi(-1) = -1$ and real if $\chi(-1) = 1$). There is a lot of mathematics behind this seemingly simple statement, namely it is a reformulation of the quadratic reciprocity law of Gauss, and as such it has had considerable influence on the development of the whole theory of L-functions.

Before discussing this further, observe that since $\zeta(s)$ and $\zeta_K(s)$ on both sides of (1.14) have a simple pole at $s = 1$, it follows that $L(\chi, 1) \neq 0$, and more precisely that

$$(1.15) \qquad L(\chi, 1) = \begin{cases} \dfrac{2\pi h(K)}{w\sqrt{q}} & \text{if } K \text{ is imaginary} \\ \dfrac{2h(K)\log \varepsilon}{\sqrt{q}} & \text{if } K \text{ is real, } \varepsilon > 1 \text{ being the fundamental unit of } K \end{cases}$$

which implies $L(\chi, 1) > 0$ (even $L(\chi, 1) \gg q^{-1/2}$ if K is imaginary); see also the discussion in Section 2.3.

We come back to (1.14). A prime number p is either ramified, inert or split in the quadratic field K. Expliciting this in the Euler product expression of $\zeta_K(s)$, we have

$$(1.16) \qquad \zeta_K(s) = \prod_{p \text{ split}} (1 - p^{-s})^{-2} \prod_{p \text{ inert}} (1 - p^{-2s})^{-1} \prod_{p \text{ ramified}} (1 - p^{-s})^{-1}$$

and comparing with (1.14) it follows that the latter is equivalent with the following characterization of the splitting of primes:

$$(1.17) \qquad \begin{cases} p \text{ is ramified if and only if } \chi(p) = 0, \text{ if and only if } p \mid q \\ p \text{ is split in } K \text{ if and only if } \chi(p) = 1 \\ p \text{ is inert in } K \text{ if and only if } \chi(p) = -1. \end{cases}$$

The point is that this characterization is in terms of the Dirichlet character χ which is a *finite amount of data* related to $\mathbf{Q}$ and not to K (it is after all only a particular periodic arithmetic function of period q).

Even more concretely, basic algebraic number theory shows that another characterization, involving Legendre symbols, exists:

$$
(1.18) \qquad
\begin{cases}
p \text{ is ramified if and only if } p \mid q \\[2mm]
p \text{ is split in } K \text{ if and only if } \left(\dfrac{D}{p}\right) = 1 \\[2mm]
p \text{ is inert in } K \text{ if and only if } \left(\dfrac{D}{p}\right) = -1
\end{cases}
$$

(where D is the discriminant of K), but there is no reason *a priori* to expect that the map

$$
p \mapsto \left(\frac{D}{p}\right)
$$

(defined only on the rather chaotic set of prime numbers!) has any particularly good property. Now the primitive quadratic characters are very easy to characterize using the structure of $(\mathbf{Z}/q\mathbf{Z})^{\times}$. In particular, if q is squarefree and odd, there is a unique primitive character modulo q, given by

$$
\chi(n) = \prod_{p \mid q} \left(\frac{n}{p}\right)
$$

(this is clearly a quadratic character modulo q, and since for any $p \mid q$ there are quadratic residues and non-residues modulo $p > 2$, χ can not be induced from q/p).

In this case, one has $\chi(-1) \equiv q \,(\mathrm{mod}\,4)$, so the quadratic field is $K = \mathbf{Q}(\sqrt{q})$ if $q \equiv 1 \,(\mathrm{mod}\,4)$ and $K = \mathbf{Q}(\sqrt{-q})$ if $q \equiv 3 \,(\mathrm{mod}\,4)$, with discriminant $D = \chi(q)q$. Take q to be a prime > 2: then $\chi(-1) = (-1)^{(q-1)/2}$ so comparison of (1.17) and (1.18) gives (after some easy checking)

$$
\left(\frac{p}{q}\right) = \left(\frac{D}{p}\right) = \left(\frac{(-1)^{(q-1)/2}}{p}\right)\left(\frac{q}{p}\right) = (-1)^{(p-1)(q-1)/4}\left(\frac{q}{p}\right),
$$

the original form of the quadratic reciprocity law. Similarly for quadratic characters with even conductor. In fact (1.14) for all quadratic fields is equivalent with quadratic reciprocity.

This explains why it may seem natural and desirable to seek generalizations of the factorization (1.14) for other L-functions, specifically at first for Dedekind zeta functions, and why one would see them as some form of "reciprocity law", if the factors appearing are (like $L(\chi, s)$) defined in terms of objects of the base field $\mathbf{Q}$, since this would amount to describing the splitting of primes in K solely in terms of data belonging to $\mathbf{Q}$.

Of course, it is also natural to expect similar situations to exist for relative extensions E/K of number fields.

The general (still very much conjectural) form of this idea belongs to the theory of Artin L-functions and to some of the conjectures of Langlands, and will be covered later. There is however one important special case where a (mostly) satisfactory theory exists and with which I will finish this first lecture. It is the case when one has an *abelian* extension L/K and the results are known as *class-field theory*.

1.5. Hecke L-functions, class-field theory

Let E/K be an extension of number fields and assume that it is a Galois extension with *abelian* Galois group G. It follows in particular that for every prime ideal $\mathfrak{p}$ of K unramified in E, the Frobenius element $\sigma_{\mathfrak{p}} \in G$ is well-defined.

A formal computation, although non-trivial (a special case of the invariance by induction of Artin L-functions), shows that there is a factorization of $\zeta_E(s)$ as

$$(1.19) \qquad \zeta_E(s) = \prod_{\rho \in \hat{G}} L(\rho, s)$$

where $\hat{G}$ is the (dual) character group of G and

$$L(\rho, s) = \prod_{\mathfrak{p} \text{ in } K} (1 - \rho(\sigma_{\mathfrak{p}})(N\mathfrak{p})^{-s})^{-1},$$

where $\rho(\sigma_{\mathfrak{p}})$ refers to the image of the Frobenius element by the Galois representation (either 0 or 1-dimensional) induced on $\mathbf{C}^{I_{\mathfrak{p}}}$, to take care of the ramification. Of course, if $\mathfrak{p}$ is unramified in E, this is literally $\rho(\sigma_{\mathfrak{p}})$. (For instance, if $\mathfrak{p}$ is totally split in E with $N\mathfrak{p} = q$, the $\mathfrak{p}$-factor on the left of (1.19) is made of $d = [E : K]$ factors $(1 - q^{-s})^{-1}$, one for each prime above $\mathfrak{p}$; on the right there are $|\hat{G}| = |G| = d$ characters ρ, and for each $\rho(\sigma_{\mathfrak{p}}) = \rho(1) = 1$, hence the $\mathfrak{p}$-factor is indeed the same).

This new Euler product is, for trivial reasons, still absolutely convergent for $\mathrm{Re}(s) > 1$ since $|\rho(\sigma)| = 1$ for any $\sigma \in G$, but its analytic continuation, for individual ρ, is by no means obvious: one can see that a simple definition as a series instead of a product does not appear immediately. There is no clear reason that the coefficients of this Dirichlet series should have any good property (compare with the discussion of the quadratic reciprocity law above).[5] In particular, at first sight, computing the coefficient of $(N\mathfrak{a})^{-s}$ seems to require knowing the factorization of $\mathfrak{a}$ in E.

EXAMPLE 1.5.1. For a quadratic field $K/\mathbf{Q}$, there are two characters, the trivial one with L-function equal to $\zeta(s)$ and a quadratic character χ_2 with

$$L(\chi_2, s) = \prod_{p \text{ split}} (1 - p^{-s})^{-1} \prod_{p \text{ inert}} (1 + p^{-s})^{-1}$$

since $G \simeq \{\pm 1\}$ and $\chi_2(\sigma_p) = +1$ (resp. -1) if and only if p is split (resp. if p is inert).

Of course one has $L(\chi_2, s) = L(\chi, s)$ in the factorization (1.14) involving the Dirichlet character (compare (1.18)).

The L-function part of class-field theory can be thought of as the identification of the L-functions of Galois characters in terms of generalizations of the quadratic Dirichlet character appearing in (1.14). For a general base field K, those were defined by Hecke. However, in the case of $\mathbf{Q}$ it suffices to consider the original Dirichlet characters (not necessarily quadratic), which we state, without striving for the utmost precision:

THEOREM 1.5.2 (Kronecker-Weber theorem). *Let $K/\mathbf{Q}$ be an abelian extension with Galois group G, let $\rho : G \to \mathbf{C}^{\times}$ be a Galois character and $L(\rho, s)$ its L-function as above.*

---

[5] For instance, are the partial sums of the coefficients bounded, as for Dirichlet characters?

There exists a unique primitive Dirichlet character χ modulo q for some $q \geqslant 1$ such that

$$L(\rho, s) = L(\chi, s).$$

In particular, the analytic continuation, functional equation and other properties of $L(\chi, s)$ are thus inherited by all those L-functions. This gives as a consequence "reciprocity laws" describing the splitting of prime numbers in abelian extensions of $\mathbf{Q}$, and also the following corollary (the more usual form of the Kronecker-Weber Theorem):

COROLLARY 1.5.3. *Let $K/\mathbf{Q}$ be an abelian extension. Then K is contained in some cyclotomic field, i.e. there exists some $m \geqslant 1$ such that $K \subset \mathbf{Q}(\boldsymbol{\mu}_m)$ where $\boldsymbol{\mu}_m$ is the group of m-th roots of unity in $\mathbf{C}$.*

PROOF OF THE COROLLARY. Let

(1.20)
$$\zeta_K(s) = \prod_{\chi} L(\chi, s)$$

be the factorization obtained from (1.19) and Theorem 1.5.2 in terms of (some) Dirichlet characters. Let m be the l.c.m of the conductors q_χ of the characters occuring. Then in fact we have $K \subset \mathbf{Q}(\boldsymbol{\mu}_m)$.

To prove this, we recall a simple general fact about number fields (other proofs are possible): we have $E \subset K$ for Galois extensions $E/\mathbf{Q}$ and $K/\mathbf{Q}$ if and only if all but finitely many of the prime numbers p which are totally split in K are totally split in E. Let thus p be a prime number totally split in $\mathbf{Q}(\boldsymbol{\mu}_m)$. By elementary theory of cyclotomic fields [**CF**, Ch. 3] this means that $p \equiv 1 \, (\mathrm{mod}\, m)$. Thus for any character χ modulo m, we have $\chi(p) = 1$, and in particular for all characters occuring in (1.20) we have $\chi(p) = 1$, i.e. (by the Kronecker-Weber Theorem) $\rho(\sigma_p) = 1$ for all $\rho \in \hat{G}$. This means that $\sigma_p = 1 \in G$, hence that p is totally split in K. $\qquad\square$

For more about the relationships between Dirichlet characters and cyclotomic fields, see e.g. [**Wa**, Ch. 3].

REMARK 1.5.4. The Kronecker-Weber Theorem, as stated here, bears a striking resemblance to the L-function form of the modularity conjecture for elliptic curves (explained in de Shalit's lectures). One can prove Theorem 1.5.2 by following the general principles of Wiles's argument [**Tu**] (deformation of Galois representations, and computation of numerical invariants in a commutative algebra criterion for isomorphism between two rings).

We come to the definition of Hecke L-functions of a number field K. As one may easily imagine, they are L-functions of the form

(1.21)
$$L(\chi, s) = \sum_{\mathfrak{a}} \chi(\mathfrak{a})(N\mathfrak{a})^{-s} = \prod_{\mathfrak{p}} (1 - \chi(\mathfrak{p})(N\mathfrak{p})^{-s})^{-1}$$

for some "arithmetic" function χ defined on integral ideals in $\mathfrak{a}$ in K. The proper definition, in the language of ideals, requires some care however; this becomes much clearer in the idèle-theoretic description. See e.g. [**Co**] for further details in the "classical" language. The difficulties arise because of the various archimedean places and the class number being possibly > 1.

Let $\mathfrak{m}$ be a non-zero integral ideal of K, which will play the role of modulus. One defines the subgroups $I_\mathfrak{m}$ (resp. $P_\mathfrak{m}$) of the group I of fractional ideals in K (resp. of the subgroup of principal ideals) by

$$I_\mathfrak{m} = \{\mathfrak{a} \in I \mid (\mathfrak{a}, \mathfrak{m}) = 1\}$$
$$P_\mathfrak{m} = \{\mathfrak{a} = (\alpha) \in I_\mathfrak{m} \cap P \mid \alpha \equiv 1 \,(\mathrm{mod}\,\mathfrak{m})\}.$$

The finite abelian group $H_\mathfrak{m} = I_\mathfrak{m}/P_\mathfrak{m}$ is called the ray-class group modulo $\mathfrak{m}$. Let also $U_\mathfrak{m}$ be the group of units in $P_\mathfrak{m}$.

Let ξ_∞ be a (unitary) character

$$\xi_\infty \,:\, K^\times/\mathbf{Q}^\times \to \mathbf{C}^\times$$

(these can be easily described using the various places at infinity, see the example below) such that $U_\mathfrak{m} \subset \ker \xi_\infty$. Hence ξ_∞ induces a homomorphism

$$\xi_\infty \,:\, P_\mathfrak{m} \to \mathbf{C}^\times.$$

DEFINITION. A Hecke character of weight ξ_∞ for the modulus $\mathfrak{m}$ is a homomorphism

$$\chi \,:\, I_\mathfrak{m} \to \mathbf{C}^\times$$

(unitary) such that

$$\chi((\alpha)) = \xi_\infty(\alpha)$$

if $\mathfrak{a} = (\alpha) \in P_\mathfrak{m}$. The character χ is extended to I by putting $\chi(\mathfrak{a}) = 0$ if $(\mathfrak{a}, \mathfrak{m}) \neq 1$.

One can define primitive Hecke characters in much the same way as for Dirichlet characters, with the same basic properties: any character is induced by a unique primitive one, and their L-functions are the same up to a finite Euler product, as in (1.7).

For a Hecke character χ the L-function $L(\chi, s)$ has the Euler product expansion (1.21) by multiplicativity and converges absolutely for $\mathrm{Re}(s) > 1$.

EXAMPLE 1.5.5. If we take $\mathfrak{m} = 1$ and $\xi_\infty = 1$, then the corresponding Hecke characters are just the *ideal class-group characters*, i.e. the (finitely many) characters $H(K) \to \mathbf{C}^\times$ of the ideal class-group of K. More generally, if $\mathfrak{m}$ is arbitrary but $\xi_\infty = 1$, the resulting Hecke characters are called ray-class characters modulo $\mathfrak{m}$.

For the trivial Hecke character χ_0, one has $L(\chi_0, s) = \zeta_K(s)$.

EXAMPLE 1.5.6. Let $K = \mathbf{Q}$. Then the only possibility is $\xi_\infty = 1$, $\mathfrak{m} = (m)$ for some unique $m \geqslant 1$, Hecke characters modulo $\mathfrak{m}$ are the same thing as Dirichlet characters modulo m. Primitivity also corresponds.

This extends very similarly for any field with class number 1.

EXAMPLE 1.5.7. Let K be a quadratic field and σ the non-trivial element in the Galois group of K. Then $K^\times/\mathbf{Q}^\times \simeq \{a \in K^\times \mid Na = 1\}$ by $a \mapsto a/a^\sigma$. Assume first that K is imaginary. Then one checks easily that ξ_∞ must be of the form

$$\xi_\infty(a) = \left(\frac{a}{|a|}\right)^u$$

(i.e. $\arg(a)^u$) for some integer $u \in \mathbf{Z}$, to which the weight can be identified.

If, on the other hand, K is real, then one finds that

$$\xi_\infty(a) = \left(\frac{a}{|a|}\right)^{u_1}\left(\frac{a^\sigma}{|a^\sigma|}\right)^{u_2}$$

with u_1, $u_2 \in \{0,1\}$.

Hecke managed to prove the fundamental analytic properties of his L-series, in complete analogy with the case of Dirichlet L-functions. The next lecture will sketch the proof, which becomes more transparent in the language of adèles: this translation is the subject of Tate's famous thesis [**Ta**].

THEOREM 1.5.8 (Hecke). *Let $\chi \neq 1$ be a primitive, non-trivial Hecke character of K. Then $L(\chi,s)$ admits analytic continuation as en entire function and there is a gamma factor $\Gamma(\xi_\infty,s)$, depending only on the weight, which is a product of gamma functions, such that*

$$\Lambda(\chi,s) = \Gamma(\xi_\infty,s)L(\chi,s)$$

satisfies

(1.22) $$\Lambda(\chi,s) = \varepsilon(\chi)(|D|N\mathfrak{m})^{1/2-s}\Lambda(\overline{\chi},1-s)$$

for some complex number $\varepsilon(\chi)$ of absolute value 1. Here D is the (absolute) discriminant of $K/\mathbf{Q}$ and $N\mathfrak{m}$ the norm from K to $\mathbf{Q}$.

We are a little vague: one can indeed write a formula for $\varepsilon(\chi)$ (in terms of a Gauss sum for χ) as well as for $\Gamma(\xi_\infty,s)$. The point is that they appear "naturally" in the course of the proof. Writing the gamma factor "as a function of the weight" is actually reminiscent of adelic arguments. Of course, one should compare this to (1.9) and (1.13).

Now one can state the analogue of Theorem 1.5.2 for abelian extensions of a number field K, which is extremely similar and encompasses much of class-field theory:

THEOREM 1.5.9 (Artin). *Let K be a number field, E/K a finite abelian extension with Galois group G, let $\rho : G \to \mathbf{C}^\times$ be a Galois character and $L(\rho,s)$ the associated L-function. Then there exists a unique primitive Hecke character χ of K, of modulus $\mathfrak{m}$ say, such that*

$$L(\rho,s) = L(\chi,s).$$

Actually, this identity holds locally, i.e. the $\mathfrak{p}$-component of the Euler products are equal for all prime ideal $\mathfrak{p}$ of K, namely[6]

$$\rho(\sigma_\mathfrak{p}) = \chi(\mathfrak{p}) \text{ for all } \mathfrak{p} \text{ coprime to } \mathfrak{m}.$$

REMARK 1.5.10. Actually, as in the more general case of Artin L-functions, one can define beforehand, in terms simply of ρ, a *conductor* $\mathfrak{f}(\rho)$ and a weight $\xi_\infty(\rho)$. The proof of the theorem shows that $\mathfrak{f} = \mathfrak{m}$ and the weight of χ is ξ_∞. This added precision is very important in applications, if only because it makes this statement theoretically verifiable for any given ρ by brute-force search (this is similar to Weil's stipulation that the level of the modular form associated to an elliptic curve over $\mathbf{Q}$ should be its conductor).

---

[6] As often remarked by Serre, this last statement is stronger than the first one if $K \neq \mathbf{Q}$; compare the footnote after (1.4).

REMARK 1.5.11. Comparing the functional equations on both sides of (1.19) after applying Theorem 1.5.9, one obtains a relation between the conductors of the Galois (or Hecke) characters related to E/K and the discriminant of E. In the simple case $K = \mathbf{Q}$ this takes the form

$$|D| = \prod_\chi q_\chi$$

where D is the discriminant of $E/\mathbf{Q}$, χ are the Dirichlet characters in (1.20) and q_χ their conductors. For $E = \mathbf{Q}(\boldsymbol{\mu}_\ell)$ for $\ell > 2$ prime, the factorization is

$$\zeta_E(s) = \prod_{\chi \,(\mathrm{mod}\,\ell)} L(\chi, s),$$

there are $\ell - 1$ Dirichlet characters, of which one has conductor 1 and $\ell - 2$ have conductor ℓ, hence we recover the well-known discriminant $|D| = \ell^{\ell-2}$.

Similarly, since the root number for $\zeta_E(s)$ is 1, one deduces a relation between the root numbers $\varepsilon(\chi)$. One can also fruitfully compare the residues at $s = 1$ on both sides...

Just as in the case of Dirichlet characters, the Hecke characters of a number field K have (independently of class-field theory) many applications to the equidistribution of ideals in various classes. Actually, those split naturally into two kinds: one concerns the distribution of prime ideals $\mathfrak{p}$ in K, and is based on the same methods using the logarithmic derivatives of Hecke L-functions. The other concerns the distribution (in $\mathbf{N}$) of the *norms* of integral ideals in K, i.e. the (average) properties of the arithmetic function $r_K(n)$ in (1.11). Here one uses $\zeta_K(s)$ and the Hecke L-functions themselves, and it is somewhat simpler since their singularities are completely known. One deduces for instance that

$$|\{\mathfrak{a} \subset \mathcal{O} \mid N\mathfrak{a} \leqslant X\}| = (\mathrm{Res}_{s=1}\,\zeta_K(s))X + O(X^{1-1/d})$$

as $X \to +\infty$, and that the ideals are equidistributed in ideal classes. For example, taking a quadratic field K, one gets the asymptotic formula for the number of integers $n \leqslant X$ represented by a quadratic form (with multiplicity). Standard methods of analytic number theory can also be used to deduce an asymptotic formula for the number of integers which are norm of an ideal in K (excluding multiplicity).

EXAMPLE 1.5.12. Take $K = \mathbf{Q}(i)$. Then $\mathcal{O} = \mathbf{Z}[i]$ has class number one and the number of ideals with norm $\leqslant X$ is equal to the number of lattice points in $\mathbf{C}$ inside a disc of radius $\sqrt{R}$ (the Gauss circle problem). One gets in this case trivially

$$\sum_{n^2+m^2 \leqslant X} 1 = \pi X + O(\sqrt{X}) \text{ as } X \to +\infty$$

($r(n)$ is the number of representations of n as sum of two squares). More subtle arguments yield the same formula with $X^{1/3}$ as error term. It is conjectured that any exponent $> 1/4$ will do, which is best possible. In suitably smoothed form, even better results are known, for instance the automorphy of theta functions (see Lecture III) easily implies that if φ is a smooth function with compact support in $[0, +\infty[$ we have

$$\sum_{n,m} \varphi\Big(\frac{n^2 + m^2}{X}\Big) = \pi X \int_0^{+\infty} \varphi(x)dx + c_0\varphi(0) + O(X^{-1/2})$$

as $X \to +\infty$, where c_0 is some constant. (Use Mellin transform and move the line of integration to $\mathrm{Re}(s) = -1/2$.)

1.6. Function fields

I will only say a very few words about the "geometric" analogue (or "function field case") of the various theories described before, mostly by lack of proper competence. More detailed explanations will come from other lectures.

The analogy between the arithmetic of $\mathbf{Z}$ and that of the ring of polynomials over a field $k[X]$ is a very old one. If one takes for k a finite field $\mathbf{F}_q$ with $q = p^d$ elements, the analogy deepens. For instance, if P is an irreducible polynomial, the residue field $\mathbf{F}_q[X]/(P)$ is a finite field, as happens with number fields.

This leads to a theory which is very similar to that described previously, in many respects, although the geometric intuition and some other intrinsic characteristics tend to make it simpler (e.g. one can differentiate polynomials, but not integers.) It provides both a different set of problems, and a way to get evidence for conjectures which remain intractable over number fields, for instance concerning the zeros of L-functions (for example the work of Katz and Sarnak [**KS**] on the fine distribution of zeros, or the recent proof by Lafforgue of the global Langlands correspondance for $GL(n)$ over function fields).

We will describe here briefly the case of curves. Let $C/\mathbf{F}_q$ be an algebraic curve (are there any others?) over $\mathbf{F}_q$, say given as the set of zeros in $\mathbf{P}^2$ of an homogeneous polynomial $F \in \mathbf{F}_q[X, Y, Z]$. Let $D/\mathbf{F}_q$ be "the" smooth projective model of C. In the language of Hasse and Artin, this would be identified with the *function field* $K = \mathbf{F}_q(C)$ of C, which is a finite extension of the field $\mathbf{F}_q(X)$ of rational functions in one variable (geometrically, the latter corresponds to $\mathbf{P}^1$ and the extension to a ramified covering $C \to \mathbf{P}^1$).

The non-archimedean valuations of K correspond in one-to-one fashion with the set of closed points of C (in scheme theoretic language), in other words with the *orbits* of the action of $\mathrm{Gal}(\bar{\mathbf{F}}_q/\mathbf{F}_q)$ on the points of C defined over the algebraic closure of $\mathbf{F}_q$. If C is affine, they also correspond to the maximal ideals in the ring of functions regular on C. Define the *zeta function* of C by the Euler product

$$Z(C; s) = \prod_x (1 - (Nx)^{-s})^{-1},$$

over the set of closed points, with Nx equal to the cardinality of the residue field at x. This is clearly an analogue of the Dedekind zeta function for the field K, and it turns out to satisfy much the same properties.

Since the norm of x is always of the form q^a for some $a \geqslant 1$, the degree $\deg(x)$ of x, it is convenient to put $T = q^{-s}$ and consider $Z(C)$ as a formal power series in T

$$Z(C) = \prod_x (1 - T^{\deg(x)})^{-1}.$$

This turns out to have another expression, peculiar to the geometric case, which gives a direct diophantine interpretation of this zeta function:

LEMMA 1.6.1. *We have*

$$Z(C) = \exp\Big(\sum_{n \geqslant 1} \frac{|C(\mathbf{F}_{q^n})|}{n} T^n\Big),$$

as formal power series.

SKETCH OF PROOF. On applies the operator $T d \log$ to both sides. On the right the result is

$$\sum_{n \geqslant 1} |C(\mathbf{F}_{q^n})| T^n,$$

while on the right after expanding a geometric series one gets

$$\sum_{x} \deg(x) \sum_{k \geqslant 1} X^{k \deg(x)} = \sum_{n} N_n T^n,$$

with

$$N_n = \sum_{d|n} d|\{x \mid \deg(x) = d\}|.$$

The degree d is the cardinality of the Galois orbit of a closed point x, so we see that $N_n = |C(\mathbf{F}_{q^n})|$, as desired. $\qquad \square$

The basic properties of the zeta function are as follows:

THEOREM 1.6.2. *Suppose C is smooth and projective. Then the zeta function has analytic continuation and functional equation. More precisely (!), $Z(C)$ is a rational function of T of the form*

$$Z(C) = \frac{P_{2g}}{(1-T)(1-qT)}$$

where $P_{2g} \in \mathbf{Z}[T]$ is a monic polynomial of even degree $2g$, g being equal to the genus of C. It satisfies

$$Z(C) = \varepsilon(C) q^{-\chi/2} T^{-\chi} Z(1/(qT))$$

where $\varepsilon(C) = \pm 1$ and $\chi = 2 - 2g$ is the Euler-Poincaré characteristic of C.

The polynomial P_{2g} can be expressed over $\mathbf{C}$ as

$$P_{2g} = \prod_{i=1}^{2g} (1 - \alpha_i T)$$

with $|\alpha_i| = \sqrt{q}$ for all i.

The first part was proved by Schmidt in general; it is based on the Riemann-Roch theorem. The second follows also from this argument. The last part is the Riemann Hypothesis for $Z(C)$: it was first proved by A. Weil. Weil also made conjectures generalizing those statements to higher dimensional varieties, and they were proved in even stronger and much more powerful form through the work of Grothendieck's school and particularly, of course, of P. Deligne. It seems particularly significant that these results are based on expressions for the polynomial P_{2g} as characteristic polynomials of a certain operator (the Frobenius operator) acting on some cohomology group: a (natural) analogue of such a phenomenon is eagerly sought in the number field case.

REMARK 1.6.3. In terms of s, the functional equation becomes

$$Z(C; s) = \varepsilon(C) q^{\chi(s-1/2)} Z(C; 1-s),$$

19

and the Riemann Hypothesis says that the zeros of $Z(C; s)$ satisfy
$$q^{-\sigma} = |T| = |\alpha_i|^{-1} = q^{-1/2},$$
i.e. $\mathrm{Re}(s) = 1/2$. So the analogy is indeed very clear.

Elementary theory of L-functions, II

2.1. Introduction

This second lecture is partly a development, with some sketches of proofs, of the main points of the first lecture, and partly a survey of the important topic of the zeros of L-functions.

2.2. The functional equation of Hecke L-functions

We will describe the proof of the functional equation for Dedekind zeta functions and briefly mention the changes needed for the general case of Hecke L-functions. See [**La**, XIII-3] or Hecke's original paper for this classical approach. It is well motivated historically (based on one of Riemann's proofs), and has clear connections with modular forms via theta functions, but is not entirely satisfactory in some respects (for the appearance of the gamma factors for instance). However, its expliciteness makes it still quite valuable (see the formula (2.6) below).

Let K be a number field of degree d over $\mathbf{Q}$ and

$$(2.1) \qquad\qquad \zeta_K(s) = \sum_{\mathfrak{a}} (N\mathfrak{a})^{-s}$$

its Dedekind zeta function.

The fundamental tool from harmonic analysis to obtain the analytic continuation is the *Poisson summation formula*:

PROPOSITION 2.2.1. *Let $f : \mathbf{R}^d \to \mathbf{C}$ be a Schwartz function on (euclidean) $\mathbf{R}^d$, let*

$$\hat{f}(\xi) = \int_{\mathbf{R}^d} f(x)e(-<x,\xi>)dx$$

be its Fourier transform. Then we have for all $a \in \mathbf{R}^d$

$$\sum_{m \in \mathbf{Z}^d} f(m+a) = \sum_{\mu \in \mathbf{Z}^d} \hat{f}(\mu)e(<a,\mu>),$$

both series being absolutely convergent, uniformly for a in compact sets.

SKETCH OF PROOF. For the proof, one simply sees that the left-hand side is, by averaging, a function on $\mathbf{R}^d$ invariant under the translation action of $\mathbf{Z}^d$, and the right-hand side is simply its expansion into Fourier series. Compare with the discussion of Poincaré and Eisenstein series in Lecture 3. □

It is clear that such a result is relevant; one may ask why not apply it directly to $\zeta(s)$ with $f(x) = x^{-s}$, $x > 0$, but of course this function is not in Schwartz space.

We split the series (2.1) into ideal classes in order to obtain summation sets easily parameterized by d-tuples of integers: we have

$$(2.2) \qquad \zeta_K(s) = \sum_a \zeta(s; a)$$

where

$$\zeta(s; a) = \sum_{[\mathfrak{a}]=a} (N\mathfrak{a})^{-s}$$

for any ideal class a. Using the same notation for the invariants of K as in Section 1.4, one has

PROPOSITION 2.2.2. *Let a be any ideal class in K, let*

$$\Lambda(s; a) = \pi^{-r_1 s/2}(2\pi)^{-r_2 s}\Gamma(s/2)^{r_1}\Gamma(s)^{r_2}\zeta(s; a).$$

Then $\Lambda(s; a)$ admits analytic continuation to a meromorphic function on $\mathbf{C}$ with simple poles at $s = 1$ and at $s = 0$ and it satisfies the functional equation

$$\Lambda(s; a) = |D|^{1/2-s}\Lambda(1 - s; (a\mathfrak{d})^{-1}),$$

where $\mathfrak{d}$ is the ideal-class of the different of $K/\mathbf{Q}$.

The partial zeta function $\zeta(s; a)$ is meromorphic with a simple pole at $s = 1$ with residue equal to

$$\mathrm{Res}_{s=1} \zeta(s; a) = \frac{2^{r_1}(2\pi)^{r_2} R}{w\sqrt{|D|}}.$$

Summing over $a \in H(K)$, this proposition implies the analytic continuation and functional equation of $\zeta_K(s)$ as stated in Section 1.4.

If one knows beforehand that $\zeta_K(s)$ has at most a simple pole at $s = 1$, Proposition 2.2.2 actually gives a proof of the finiteness of the class number since the series for $\zeta_K(s)$, hence the expansion (2.2), are absolutely convergent for $\mathrm{Re}(s) > 1$, and since all partial zeta functions have the same residue at $s = 1$.

REMARK 2.2.3. The analytic continuation is crucially based on the series expression (2.1) to which the Poisson summation formula can be applied, and has nothing to do with the Euler product, as the proof through partial zeta functions (which have no Euler product if $h > 1$) shows. However it is the Euler product which has the deepest arithmetic content, in particular through its local-global interpretation and its consequences on the location of the zeros of L-functions.

Assume for simplicity that $a = 1$ is the trivial ideal class, thus represented by $\mathcal{O} = \mathcal{O}_K$. The ideals in the class thus correspond bijectively with non-zero integers $z \in \mathcal{O}$ up to multiplication by units $u \in U$.

Let $z \in \mathcal{O}$, $z \neq 0$. For each embedding σ of K, we have by definition of the gamma function

$$\pi^{-s/2}\Gamma(s/2)|z|_\sigma^{-s} = \int_0^{+\infty} \exp(-\pi y|z|_\sigma^2)y^{s/2}\frac{dy}{y},$$

$$(2\pi)^{-s}\Gamma(s)|z|_\sigma^{-s} = \int_0^{+\infty} \exp(-2\pi y|z|_\sigma^2)y^s\frac{dy}{y}.$$

We apply the former for real embeddings and the latter for (pairs of) complex ones, and sum over $z \in \mathcal{O}$ modulo units for $\mathrm{Re}(s) > 1$. We get a formula

$$(2.3) \qquad \Lambda(s; \mathcal{O}) = \int \Theta_1(y; \mathcal{O}) ||y||^{s/2} \frac{dy}{y}$$

where the integral is over $(\mathbf{R}^+)^{r_1+r_2}$, the coordinates corresponding to the archimedean places of K, with

$$||y|| = \prod_v y_v^{d_v}, \quad \frac{dy}{y} = \prod_v \frac{dy_v}{y_v}$$

and the kernel is almost a theta function:

$$\Theta_1(y; \mathcal{O}) = \sum_{\substack{z \in \mathcal{O}/U \\ z \neq 0}} \exp(-\pi \sum_v y_v d_v |z|_v^2).$$

The sum is absolutely convergent, uniformly for $\mathrm{Re}(s) > 1 + \delta$ for any $\delta > 0$, since this holds for $\zeta(s, \mathcal{O})$.

We now rearrange this integral formula, integrating over $t = ||y||$ first: we let

$$G_1 = \{y \mid ||y|| = 1\};$$

observe also that U acts on G_1 (we actually let u act by the "obvious" action of u^2) and the quotient G_1/U is compact by Dirichlet's Unit Theorem. We rewrite

$$\int \sum_{\mathcal{O}/U} \cdots = \int_0^{+\infty} \int_{G_1/U} \sum_{u \in U} \sum_{\mathcal{O}/U} \cdots$$

$$= \int_0^{+\infty} \int_{G_1/U} \sum_{u \in \mathcal{O}} \cdots$$

so we get after some rearranging (and taking care of roots of unity and other details):

$$\Lambda(s; \mathcal{O}) = \frac{1}{w} \int_0^{+\infty} \int_{G_1/U} (\Theta(t^{1/d} x; \mathcal{O}) - 1) t^{s/2} d\mu(x) \frac{dt}{t}$$

($d\mu(x)$ being the appropriate measure on G_1/U) where $\Theta(x; \mathcal{O})$ is a Hecke theta function (for $a = \mathcal{O}$):

$$(2.4) \qquad \Theta(x; \mathcal{O}) = \sum_{z \in \mathcal{O}} \exp(-\pi {<} |z|^2, x {>})$$

for $x \in (\mathbf{R}^+)^{r_1+r_2}$, with the somewhat awkward shorthand notation

$$<x, y> = \sum_v d_v x_v y_v \text{ and } |z|^2 = (|z|_v^2)_v.$$

In this integral expression, there is no problem at $+\infty$ for any $s \in \mathbf{C}$ because the theta function minus its constant term 1 decays very rapidly, but there might be at 0 for s small. So we split the integral over t in two parts and for t small we need to analyze the behavior of the theta function: that is where the Poisson summation formula will be useful.

LEMMA 2.2.4. *With notation as above, the theta function satisfies*

(2.5)
$$\Theta(x; \mathcal{O}) = \frac{1}{\sqrt{|D| \prod x_i}} \Theta(|D|^{-2/d} x^{-1}; \mathfrak{d}^{-1}).$$

SKETCH OF PROOF. Choosing a basis (ω_i) of $\mathcal{O}$ and letting ℓ_i be the "component of ω_i" linear form, the value of the theta function at x is of the shape

$$\sum_{n \in \mathbf{Z}^d} f(n)$$

for the Schwartz function $f(n) = \exp(-\pi Q(n))$, Q being a positive definite quadratic form on $\mathbf{R}^d$ (depending on x) defined by

$$Q(n_1, \ldots, n_d) = \sum_v d_v x_v |z|_v^2 = \sum_{\sigma : K \hookrightarrow \mathbf{C}} x_\sigma \left| \sum_{j=1}^d n_j \omega_j^\sigma \right|^2.$$

The Fourier transform of such a function is well-known (by diagonalization):

$$\hat{f}(\xi) = \frac{1}{|\det(Q)|} \exp(-\pi Q'(\xi)),$$

where $Q'(x)$ is the dual quadratic form (its matrix is the inverse of that of Q). Computing the determinant, the result follows (the discriminant arises as a determinant). $\square$

Let

$$m = \int_{G_1/U} d\mu(x) < +\infty$$

(since G_1/U is compact). We split the integral over t at $t = \alpha$, and change t into β/t in the first integral, after separating the part involving -1 which is explicitly evaluated:

$$w\Lambda(s; \mathcal{O}) = \int_\alpha^{+\infty} \int_{G_1/U} (\Theta(t^{1/d} x; \mathcal{O}) - 1) t^{s/2} d\mu(x) \frac{dt}{t}$$

$$+ \int_0^\alpha \int_{G_1/U} \Theta(t^{1/d} x; \mathcal{O}) t^{s/2} d\mu(x) \frac{dt}{t} - \frac{2m}{s} \alpha^{s/2}$$

$$= \int_\alpha^{+\infty} \int_{G_1/U} (\Theta(t^{1/d} x; \mathcal{O}) - 1) t^{s/2} d\mu(x) \frac{dt}{t}$$

$$+ \beta^{s/2} \int_{\beta/\alpha}^{+\infty} \int_{G_1/U} \Theta(\beta^{1/d} t^{-1/d} x; \mathcal{O}) t^{-s/2} d\mu(x) \frac{dt}{t} - \frac{2m}{s} \alpha^{s/2}$$

$$= \int_\alpha^{+\infty} \int_{G_1/U} (\Theta(t^{1/d} x; \mathcal{O}) - 1) t^{s/2} d\mu(x) \frac{dt}{t}$$

$$+ |D|^{1/2} \beta^{(s-1)/2} \int_{\beta/\alpha}^{+\infty} \int_{G_1/U} \Theta(|D|^{-2/d} \beta^{-1/d} t^{1/d} x; \mathfrak{d}^{-1}) t^{(1-s)/2} d\mu(x) \frac{dt}{t} - \frac{2m}{s} \alpha^{s/2}$$

24

$$= \int_\alpha^{+\infty} \int_{G_1/U} (\Theta(t^{1/d}x;\mathcal{O}) - 1)t^{s/2}d\mu(x)\frac{dt}{t}$$

$$+ |D|^{1/2}\beta^{(s-1)/2} \int_{\beta/\alpha}^{+\infty} \int_{G_1/U} (\Theta(|D|^{-2/d}\beta^{-1/d}t^{1/d}x;\mathfrak{d}^{-1}) - 1)t^{(1-s)/2}d\mu(x)\frac{dt}{t}$$

$$- \frac{2m}{s}\alpha^{s/2} - |D|^{-1/2}\beta^{(s-1)/2}\frac{2m}{1-s}\left(\frac{\beta}{\alpha}\right)^{(1-s)/2}$$

One uses the fact that $d\mu$ is invariant under $x \mapsto x^{-1}$ in this computation and that $||x|| = 1$ for $x \in G_1$, when applying the transformation formula for the theta function.

In the final formula, derived under the assumption that $\mathrm{Re}(s) > 1$ so that every manipulation was justified, both integrals are now entire functions of $s \in \mathbf{C}$ if $\alpha > 0$, $\beta > 0$. Hence it follows that $\Lambda(s;\mathcal{O})$ is meromorphic with simple poles at $s = 1$ and $s = 0$.

Moreover, taking $\beta = |D|^{-2}$, $\alpha = 1/|D|$, one gets

$$(2.6) \quad \Lambda(s;\mathcal{O}) = \frac{1}{w}\left\{|D|^{1/2-s} \int_{\beta/\alpha}^{+\infty} \int_{G_1/U} (\Theta(t^{1/d}x;\mathfrak{d}^{-1}) - 1)t^{(1-s)/2}d\mu(x)\frac{dt}{t}\right.$$

$$\left. + \int_\alpha^{+\infty} \int_{G_1/U} (\Theta(t^{1/d}x;\mathcal{O}) - 1)t^{s/2}d\mu(x)\frac{dt}{t} - \frac{2m}{s}\alpha^{s/2} - |D|^{1/2-s}\frac{2m}{1-s}\left(\frac{\beta}{\alpha}\right)^{(1-s)/2}\right\}.$$

and since $\beta/\alpha = \alpha$, the functional equation follows immediately.

Computing the residues explicitly requires the computation of m, which is done by describing more explicitly a fundamental domain in G_1 for the action of U.

EXAMPLE 2.2.5. For the Riemann zeta function we get

$$\pi^{-s/2}\Gamma(s/2)\zeta(s) = \frac{1}{2}\int_0^{+\infty} (\theta(y) - 1)y^{s/2}\frac{dy}{y},$$

with

$$\theta(y) = \sum_{n\in\mathbf{Z}} e^{-\pi n^2 y}.$$

The functional equation for θ is

$$\theta(y) = y^{-1/2}\theta(y^{-1}).$$

REMARK 2.2.6. We have used Dirichlet's Unit Theorem for K to say that G_1/U is compact: those two statements are in fact obviously equivalent, and equivalent to the statement that the measure m of G_1/U is finite. Now since the first integral expression (2.3) is a priori valid for $\mathrm{Re}(s) > 1$, it is not difficult to perform the previous computation without the information that $m < +\infty$, and *deduce* this from (2.6). (I learned this proof from Bill Duke).

2.3. Zeros of L-functions, explicit formula

For all the L-functions considered in Lecture I, the Generalized Riemann Hypothesis (abbreviated GRH) is expected to hold as for the Riemann zeta function: all non-trivial zeros of a Hecke L-function[1] $L(\chi, s)$ of a number field K, i.e. those in the critical strip

---

[1] All the other L-functions of Lecture I are special cases of Hecke L-functions.

$0 < \mathrm{Re}(s) < 1$, should be on the critical line $\mathrm{Re}(s) = 1/2$. Although this is still completely open, much has been proved about the zeros so that in many circumstances one can manage to prove unconditionnally results which were at first only established on the assumption of GRH.

The first basic results give a rough idea of the distribution of the zeros. Let

$$N(\chi; T_1, T_2) = |\{\rho = \beta + i\gamma \mid L(\chi, \rho) = 0,\ 0 < \beta < 1 \text{ and } T_1 \leqslant \gamma \leqslant T_2\}|$$
$$N(\chi; T) = N(\chi; -|T|, |T|).$$

Also let Λ be the van Mangoldt function for K, i.e. for any non-zero integral ideal $\mathfrak{a} \subset \mathcal{O}$ we have

$$\begin{cases} \Lambda(\mathfrak{a}) = \log N\mathfrak{p} & \text{if } \mathfrak{a} = \mathfrak{p}^k \text{ for some } k \geqslant 1 \\ \Lambda(\mathfrak{a}) = 0 & \text{otherwise,} \end{cases}$$

the point being that the Euler product (1.21) for $L(\chi, s)$ implies that the logarithmic derivative of $L(\chi, s)$ has an absolutely convergent Dirichlet series expansion for $\mathrm{Re}(s) > 1$ given by

$$(2.7) \qquad\qquad -\frac{L'}{L}(\chi, s) = \sum_{\mathfrak{a}} \chi(\mathfrak{a})\Lambda(\mathfrak{a})(N\mathfrak{a})^{-s}.$$

PROPOSITION 2.3.1. *Let K be a number field of degree $d = [K : \mathbf{Q}]$, χ a primitive Hecke character of K of modulus $\mathfrak{m}$. Then we have*

$$(2.8) \qquad\qquad N(\chi; T) = \frac{T}{2\pi} \log\Big(\frac{T^d |D| N\mathfrak{m}}{(2\pi e)^d}\Big) + O(\log(T^d |D| N\mathfrak{m}))$$

for $T \geqslant 2$, the estimate being uniform in all parameters. For any $\varepsilon > 0$, the series

$$(2.9) \qquad\qquad \sum_{\rho} \frac{1}{|\rho|^{1+\varepsilon}}$$

is absolutely convergent.

The proof of this asymptotic formula is straightforward in principle: one applies Cauchy's Theorem to the logarithmic derivative of the function $\Lambda(\chi, s)$ (which has the same zeros as $L(\chi, s)$) in the rectangle $[-1/2, 3/2] \times [-T, T]$. The major contribution comes from the gamma factor for χ. To control the part coming from $L(\chi, s)$, one uses the expansion of $\Lambda(\chi, s)$ in Weierstraß product, due to Hadamard: if $\chi \neq 1$, then[2] there exist a and $b \in \mathbf{C}$, depending on χ, such that

$$\Lambda(\chi, s) = e^{a+bs} \prod_{\rho} \Big(1 - \frac{s}{\rho}\Big) e^{s/\rho},$$

the product being over all non-trivial zeros of $L(\chi, s)$ (this is the general theory of entire functions of finite order, of which $\Lambda(\chi, s)$ is an example; (2.9) is actually also part of this theory). This product is used to show that $-L'(\chi, s)/L(\chi, s)$ is, with good precision, the sum of the terms arising from the zeros close to s, namely those with $|s - \rho| \leqslant 1$ (we know *a fortiori* from (2.8) that there are a fair number of them, if s is in the critical strip). The

---

[2] One has to multiply by $s(1 - s)$ if $\chi = 1$ to get rid of the poles.

constant a is not a problem, but b requires some ingenuity to be dealt with, and not too much is known about it as a function of χ.

The next important information, which is the key to the equidistribution theorems for prime ideals (the Chebotarev Density Theorem), is that there is no zero on the line $\mathrm{Re}(s) = 1$ (the edge of the critical strip). This was first proved by Hadamard and de la Vallée Poussin (independently) in 1896, and their method still remains essentially the only known to prove a zero-free region for L-functions. It is also here that the so-called Landau-Siegel zeros[3] first appear.

THEOREM 2.3.2. *There exists an absolute constant $c_1 > 0$ with the property that for any K and χ as above, the Hecke L-function $L(\chi, s)$ has no zero in the region $s = \sigma + it$ where*

$$\sigma > 1 - \frac{c_1}{\log(T^d |D| N\mathfrak{m})},$$

except *possibly for a single simple real zero β_1. The latter can only exist if χ is a quadratic (real) character, and it is then called the* Landau-Siegel zero *for χ (and c_1).*[4]

In what follows c_1 is fixed such that the Theorem holds, and we will let $\delta^*(\mathfrak{m}) = 1$ if there is a Landau-Siegel zero for a Hecke character of K modulo $\mathfrak{m}$, and $\delta^*(\mathfrak{m}) = 0$ otherwise. One can show easily that the corresponding real character modulo $\mathfrak{m}$ is unique.

EXAMPLE 2.3.3. The classical case is when $K = \mathbf{Q}$ and then the possible existence of β_1 for a primitive odd quadratic character χ modulo q is directly related to the possibility that the class number of the corresponding imaginary quadratic field $K = \mathbf{Q}(\sqrt{-q})$ be very small. This follows easily from the class number formula (1.15). Improving the "trivial bound" $h(K) \gg q^{-1/2}$ is extremely hard, although the Generalized Riemann Hypothesis implies that in fact

$$L(\chi, 1) \gg \frac{1}{\log q}, \text{ hence } h(K) \gg \frac{\sqrt{q}}{\log q}.$$

Siegel [**Si**] has proved that for any $\varepsilon > 0$ one has

(2.10) $$L(\chi, 1) \gg_\varepsilon q^{-\varepsilon}$$

for $q > 1$, but this results is *non-effective*, i.e. for given $\varepsilon > 0$ there is no (known) way of really computing a constant $C(\varepsilon)$ such that the promised inequality

$$L(\chi, 1) \geqslant C(\varepsilon) q^{-\varepsilon}$$

holds for $q > 1$. This ineffectivity is similar to that present in the Thue-Siegel theorem: the proof of Siegel's estimate is based on assuming that there exists χ with very small $L(\chi, 1)$ and using this hypothetical character to prove a lower bound for the others. As it is not expected that such a bad character exists, one understands why it seems so difficult to make Siegel's theorem effective.

Goldfeld's theorem [**Go**], based on known cases of the Birch and Swinnerton-Dyer conjecture for elliptic curves, is the only effective improvement on the trivial bound currently

---

[3] Iwaniec and Sarnak have proposed this name for what was usually called Siegel zeros, on the strength of Landau's contribution [**L**].

[4] This is well-defined only after fixing a possible value of c_1.

known:

$$L(\chi, 1) \gg \frac{(\log q)}{\sqrt{q}(\log \log q)}$$

for $q \geqslant 3$, with an effective implied constant (which has indeed been computed by Oesterlé). Only a logarithm factor is gained...

Roughly speaking, the hypothetical character used by Siegel has small $L(\chi, 1)$ if and only if it has a real zero close to $s = 1$, and the ineffectivity comes here because such a zero would contradict the Riemann Hypothesis and is not believed to exist. Its good effect would be to "repulse" other zeros, thereby providing a lower bound for them. In Goldfeld's argument,[5] a *real* lever is used, namely an L-function which does have a real zero, which necessarily must be the central critical point $s = 1/2$. Being far from $s = 1$ makes its repulsing effect much weaker, with two consequences: first, that it must be a zero of order $\geqslant 3$ for the argument to go through, and secondly, that the resulting lower bound is much weaker than Siegel's...

REMARK 2.3.4. The methods which yield Theorem 2.3.2 can not be expected to yield much better: indeed, a *lower bound* for $L(\chi, s)$ in the region described is actually produced (explaining also the example above), and one can show for instance that there is no "simple" lower-bound for $L(\chi, s)$ in (say) a strip of positive width $1 - \delta < \mathrm{Re}(s) \leqslant 1$. However the exponential sum methods of Vinogradov (see e.g. [**Ti**]) can be used to enlarge a little bit the zero-free region, which is sometimes significant in applications.

From the zero-free region for the Dedekind zeta function, the original method of Hadamard and de la Vallée Poussin (or the explicit formula, see below) derives the analogue of the Prime Number Theorem. Fix a modulus $\mathfrak{m}$ and a ray-class $a \in H_\mathfrak{m}$ (see Section 1.5), and let

$$\pi_K(X; \mathfrak{m}, a) = |\{\mathfrak{p} \mid \text{ the class of } \mathfrak{p} \text{ is } a \text{ and } N\mathfrak{p} \leqslant X\}|.$$

Then we have

$$(2.11) \qquad \pi_K(X; \mathfrak{m}, a) = \frac{1}{|H_\mathfrak{m}|} \mathrm{li}(X) + \frac{1}{|H_\mathfrak{m}|} \delta^*(\mathfrak{m}) X^{\beta_1} + O(X \exp(-c_2 \sqrt{\log(X|D|N\mathfrak{m})})),$$

as $X \to +\infty$. Here the estimate is uniform in all parameters, and c_2 is absolute and effective. Since $\beta_1 < 1$, for a fixed $\mathfrak{m}$ this gives the asymptotic behavior of $\pi_K(X; \mathfrak{m}, a)$, but in most applications uniformity in $\mathfrak{m}$ is the key issue.

For example, when $K = \mathbf{Q}$ (the classical case), Siegel's estimate (2.10) immediately implies the Siegel-Walfisz theorem: for any $A > 0$, we have

$$\pi(X; q, a) \sim \frac{1}{\varphi(q)} \mathrm{li}(X)$$

uniformly for all $q < (\log X)^A$ and all a modulo q. If there are Landau-Siegel zeros, this is non-effective. The Generalized Riemann Hypothesis implies the corresponding statement uniformly for $q < \sqrt{X}(\log X)^{-2}$. In many applications to analytic number theory,[6] it is indispensable to have such uniformity, although often only on average. The best known

---

[5] J. Friedlander had similar ideas.

[6] For instance, the *Titchmarsh divisor problem* (first solved by Linnik) of estimating asymptotically as $X \to +\infty$ the sum

$$\sum_{p \leqslant X} d(p - 1)$$

results – proved using the spectral theory of automorphic forms (see Lecture III) – for primes in arithmetic progressions go *beyond* what is immediately provable from GRH (Bombieri-Friedlander-Iwaniec); they are commonly used in applications.

The same results and difficulties occur in the Chebotarev density theorem for Artin L-functions, often exacerbated because the degree of interesting families of fields is larger than that of cyclotomic fields, making even the form of the prime ideal theorem based on GRH insufficient for applications.[7]

Considering the fact that so little progress has been made over almost a century on the Generalized Riemann Hypothesis for individual L-functions, it is hard to avoid thinking that some deep structure lies undiscovered: this is all the more tempting when compared with the case of function fields (briefly mentioned in 1.6) with the rich geometric and cohomological formalism. In recent years, a lot of work has been devoted to trying to probe evidence and clues to such a structure by analyzing (often on GRH) the finer vertical distribution of zeros of L-functions: this led to the remarkable discovery of links with the distribution of eigenvalues of random matrices of large rank . For $\zeta(s)$ this was first attempted by Montgomery (leading to the pair-correlation conjecture), and has been much generalized in particular by Rudnick and Sarnak [**RS**]. Some of the conjectures emerging have been proved in the function field case by Katz and Sarnak [**KS**]. As a spectacular vindication of the interest of such studies, Soundararajan and Conrey [**CS**] were able to prove, after using heuristics based on random matrices, that for a positive proportion of real Dirichlet characters, the L-function $L(\chi, s)$ has no Landau-Siegel zero!

We now come to the explicit formula. The version below obviously reduces to Proposition 1.2.1 when applied to $\zeta(s)$.

PROPOSITION 2.3.5. *Let $\varphi :]0, +\infty[\to \mathbf{C}$ be a C^∞ function with compact support, let*

$$\hat{\varphi}(s) = \int_0^{+\infty} \varphi(x) x^{s-1} dx,$$

be its Mellin transform, which is entire and decays rapidly in vertical strips. Let

$$\psi(x) = \frac{1}{x}\varphi(x^{-1}).$$

Then we have

$$(2.12) \quad \sum_{\mathfrak{a}} \Lambda(\mathfrak{a})\big(\chi(\mathfrak{a})\varphi(N\mathfrak{a}) + \overline{\chi}(\mathfrak{a})\psi(N\mathfrak{a})\big) = (\log|D|N\mathfrak{m})\varphi(1) + \delta(\chi)\int_0^{+\infty}\varphi(x)dx$$

$$- \sum_{\substack{L(\chi,\rho)=0 \\ 0<\mathrm{Re}(\rho)<1}} \hat{\varphi}(\rho) + \frac{1}{2i\pi}\int_{(-1/2)}\Big(\frac{\Gamma'}{\Gamma}(\xi_\infty, s) - \frac{\Gamma'}{\Gamma}(\xi_\infty, 1-s)\Big)\hat{\varphi}(s)ds + \mathrm{Res}_{s=0}\Big(-\frac{L'}{L}(\chi, s)\hat{\varphi}(s)\Big),$$

which can be rewritten as

$$2\sum_{d\leqslant\sqrt{X-1}}(\pi(X;d,1) - \pi(d^2+1;d,1)) + O(\sqrt{X}).$$

---

[7] For instance, if $E/\mathbf{Q}$ is an elliptic curve without CM, its d-torsion $\mathbf{Q}(E[d])$ has degree roughly d^4 by Serre's Theorem, and GRH only yields an asymptotic formula uniformly for $d < X^{1/8}$.

29

where $\delta(\chi) = 1$ if χ is trivial and is $= 0$ otherwise, and the sum over zeros includes multiplicity.

Although seemingly complicated, the last two terms are in most circumstances very easy to deal with (using the explicit form of the gamma factor and Stirling's formula for instance).

SKETCH OF PROOF. By Mellin inversion and (2.7) one has

$$\sum_{\mathfrak{a}} \chi(\mathfrak{a})\Lambda(\mathfrak{a})\varphi(N\mathfrak{a}) = \frac{1}{2i\pi} \int\limits_{(3/2)} -\frac{L'}{L}(\chi, s)\hat{\varphi}(s)ds.$$

One moves the line of integration to $\mathrm{Re}(s) = -1/2$ (this requires some simple estimates on the growth of the L-function in vertical strips, which is easy to get). The poles which occur are at $s = 1$ (if $\chi = 1$), possibly at $s = 0$ (depending on the shape of the gamma factor) – with contribution corresponding to the second and last term, respectively, of the right-hand side of (2.12) – and at the non-trivial zeros ρ of $L(\chi, s)$. The latter are simple with residue $-k\hat{\varphi}(\rho)$ where k is the multiplicity of ρ, thus their contribution is the third term in (2.12).

On the line $\mathrm{Re}(s) = -1/2$, one uses the functional equation (1.22) to obtain the relation

$$-\frac{L'}{L}(\chi, s) = (\log|D|N\mathfrak{m}) + \left(\frac{\Gamma'}{\Gamma}(\xi_\infty, s) - \frac{\Gamma'}{\Gamma}(\xi_\infty, 1 - s)\right) - \frac{L'}{L}(\overline{\chi}, 1 - s).$$

The middle term yields the fourth term in the right-hand side of (2.12). Then one writes

$$\frac{1}{2i\pi} \int\limits_{(-1/2)} -\frac{L'}{L}(\overline{\chi}, 1 - s)\hat{\varphi}(s)ds = -\frac{1}{2i\pi} \int\limits_{(3/2)} -\frac{L'}{L}(\overline{\chi}, w)\hat{\varphi}(1 - w)dw$$

and one can again apply (2.7) since w is back in the region of absolute convergence. The result follows after expanding in Dirichlet series and observing (and using also in the term involving $\log|D|N\mathfrak{m}$) that

$$\frac{1}{2i\pi} \int\limits_{(3/2)} \hat{\varphi}(1 - w)x^{-w}dw = \psi(x)$$

by the simple fact that $\hat{\varphi}(1 - w) = \hat{\psi}(w)$. $\qquad\square$

2.4. The order of magnitude of L-functions in the critical strip

There are many applications were the order of magnitude of L-functions in the critical strip is very important. Of course, it is necessary to have at least some simple estimate for all the usual analytic manipulations (contour shifts, etc...) to succeed, but there are much deeper reasons. For instance, the class number formula involves interesting invariants of number fields and relates them to the value (or residue) of L-functions at $s = 1$. As for the case of imaginary quadratic fields already discussed, much of the information gained about these invariants has come from working with the L-functions. And of course, there are close links with the distribution of the zeros of the L-functions.

This first example also illustrates that "order of magnitude" does not refer only to the size as the complex argument s varies, but may refer instead to the size in terms of the

conductor of the L-function: indeed, in many arithmetic applications, this is the really interesting problem and often one doesn't need strong bounds in terms of $|s|$.

From the proof of the analytic continuation of the L-functions, one does not get much information, but enough to bootstrap the further investigations (and it is crucial to get it): namely from the formula (2.4), or its analogues, it follows that $\Lambda(\chi, s)$ is an entire[8] function of order at most 1, i.e. for $s \in \mathbf{C}$ we have

$$\Lambda(\chi, s) \ll \exp(|s|^{1+\varepsilon})$$

for any $\varepsilon > 0$, the implied constant depending on χ and ε. Since the inverse of the gamma factor is known to be entire of order $= 1$, it follows that $L(\chi, s)$ is also.

This is actually very far from the truth. Of course, for $\mathrm{Re}(s) > 1$, where the series converges absolutely, it follows that

$$L(\chi, s) \ll 1, \text{ uniformly for } \mathrm{Re}(s) > 1 + \delta, \ \delta > 0,$$

for any χ, the implied constant depending only on K. For $\mathrm{Re}(s) < 0$, on the other hand, this implies, thanks to the functional equation, Stirling's formula and the shape of the gamma factor, that

$$L(\chi, s) \ll |s|^{(1-\sigma)/2}, \text{ uniformly for } \sigma = \mathrm{Re}(s) < -\delta, \ \delta > 0.$$

The well-known Phragmen-Lindelöf principle of complex analysis implies that a function of order 1 polynomially bounded in terms of $|s|$ on the boundary of a vertical strip is actually polynomially bounded inside the strip, and that the "rate of growth" is a convex function of the real part. This is in general best possible but for L-functions one expects actually a very different answer:

CONJECTURE 2.4.1. *Let K be a number field, χ a Hecke character to modulus $\mathfrak{m}$ with weight ξ_∞. Then we have*

$$L(\chi, s) \ll (|D|N\mathfrak{m}(1 + |s|))^\varepsilon$$

for any $\varepsilon > 0$, for $1/2 \leqslant \mathrm{Re}(s) < 1$, with an implied constant depending only on ε and on ξ_∞. If $\chi = 1$, multiply on the left by $(s - 1)$.

This is called the Lindelöf Conjecture. One can quickly give two motivations: for the first, it is implied by GRH (see e.g. [**Ti**, 14.2]). The second reason is more interesting and is based on a principle in analytic number theory sometimes referred to (rather incorrectly) as the *approximate functional equation*, a variant of which is the next proposition:

PROPOSITION 2.4.2. *Let K and χ be as above and let $s = \sigma + it$ with $0 < \sigma < 1$. Let $G(w)$ be any function real-valued on $\mathbf{R}$ and holomorphic in the strip $-4 < \mathrm{Re}(s) < 4$, with rapid decay as $|\mathrm{Im}(w)| \to +\infty$, and such that*

$$\begin{cases} G(-w) = G(w) \\ G(0) = 1 \\ G(w)\Gamma(\xi_\infty, w) \text{ is holomorphic for } w = 0, \ -1, \ -2, \ -3. \end{cases}$$

Let X, $Y > 0$ be real numbers such that $XY = |D|N\mathfrak{m}$. Then we have

$$(2.13) \qquad L(\chi, s) = \sum_{\mathfrak{a}} \frac{\chi(\mathfrak{a})}{(N\mathfrak{a})^s} V_s\left(\frac{N\mathfrak{a}}{X}\right) + \varepsilon(\chi)(|D|N\mathfrak{m})^{1/2-s} \sum_{\mathfrak{a}} \frac{\overline{\chi}(\mathfrak{a})}{(N\mathfrak{a})^{1-s}} W_s\left(\frac{N\mathfrak{a}}{Y}\right)$$

---

[8] Multiply by $s(1 - s)$ if $\chi = 1$

where V_s and W_s are the functions defined for $y > 0$ by

$$V_s(y) = \frac{1}{2i\pi} \int_{(3)} G(w) y^w \frac{dw}{w}$$

$$W_s(y) = \frac{1}{2i\pi} \int_{(3)} \frac{\Gamma(\xi_\infty, 1 - s + w)}{\Gamma(\xi_\infty, s + w)} G(w) y^w \frac{dw}{w}$$

As the (simple) proof will show, the choice of $G(w)$ is rather arbitrary and can be modulated in many ways. It is obviously very easy to write down a concrete choice of G if need be. (Note that G may depend on s; however, very often the point s is fixed in the applications and it's the character which varies).

The point of this formula is that V_s and W_s can be easily estimated using contour shifts and the Stirling formula. This shows that they behave much as cutoff functions, i.e. they decay very rapidly for y large. More precisely, the decay rate in terms of s is such that in (2.13), one sees easily that the first sum is essentially a sum over ideals of norm $N\mathfrak{a} \leqslant X$, while the second is over ideals of norm $N\mathfrak{a} \leqslant Y(1 + |s|)^{d/2}$ (this means the tails are very small). So (2.13) gives an expression of the L-function *inside the critical strip* by very rapidly convergent series. Such a formula is the basis of most (analytic) work on L-functions outside the region of absolute convergence.

In any case, choosing $X = Y = \sqrt{|D|N\mathfrak{m}}$, estimating trivially (i.e. term by term) both sums (recall that $|\varepsilon(\chi)| = 1$), one immediately gets

COROLLARY 2.4.3. *We have for $s = \sigma + it$ with $0 < \mathrm{Re}(s) < 1$*

$$L(\chi, s) \ll_\varepsilon \left((1 + |t|)^d |D| N\mathfrak{m}\right)^{\ell(\sigma) + \varepsilon}$$

for any $\varepsilon > 0$, the implied constant depending only on ε and on ξ_∞, where $\ell(\sigma)$ is the affine function on $[0, 1]$ such that $\ell(1) = 0$ and $\ell(0) = 1/2$. If $\chi = 1$, multiply on the left by $(s - 1)$.

In particular (the most interesting case) we have

$$L(\chi, s) \ll_\varepsilon \left((1 + |t|)^d |D| N\mathfrak{m}\right)^{1/4 + \varepsilon}$$

on the critical line $\mathrm{Re}(s) = 1/2$.

The bound of the corollary is called the *convexity bound*. It can also be derived by the Phragmen-Lindelöf principle. However, from the point of view of (2.13), one sees that it amounts indeed to estimating individually each term in the sums. However those terms, for $N\mathfrak{a}$ "small", are oscillating non-trivially, this coming both from $\chi(\mathfrak{a})$ (if $\chi \neq 1$) and from $(N\mathfrak{a})^{-s} = (N\mathfrak{a})^{-\sigma - it}$. Such oscillations might be expected to yield some cancellation in the sum, hence a better estimate for $L(\chi, s)$. Indeed the well-known *square-rooting philosophy* of oscillatory sums[9], if it applies here, would immediately yield the Lindelöf conjecture (by partial summation).

All this discussion extends quite easily to the more general automorphic L-functions discussed in the other lectures. A significant phenomemon, philosophically badly understood,

---

[9] If

$$S = \sum_{1 \leqslant n \leqslant N} e(\theta_n)$$

is a truly oscillating (unbiased) exponential sum, then $|S| \ll \sqrt{N}$, "up to small amounts".

is that many significant problems can be completely solved (qualitatively) by proving, for some family of L-function, *any* estimate which improves on the convexity bound of the corollary (i.e. replacing the exponent $\ell(\sigma) + \varepsilon$ by $\ell(\sigma) - \delta$ for *some* $\delta > 0$, the size of which is irrelevant to the conclusion). Even apart from applications, the sketch above shows that any such estimate is tantamount to showing that there is some non-trivial cancellation in the two sums in (2.13) and this is obviously a deep arithmetical result.

Such *convexity breaking estimates* go back to Weyl, who proved, for $\zeta(s)$, that

$$\zeta(1/2 + it) \ll_\varepsilon (1 + |t|)^{1/6+\varepsilon}$$

for $t \in \mathbf{R}$, the constant depending only on $\varepsilon > 0$. Nowadays, many more cases are known, for Dirichlet characters both in terms of s and in terms of the conductor, the latter being the most interesting (and most difficult) case, first proved by Burgess and only recently improved by Conrey and Iwaniec (for quadratic characters, using crucially automorphic L-functions).

Remarkable applications of these estimates for modular forms are due to Iwaniec (equidistribution of integral points on spheres), Sarnak and others. Sarnak [**Sa**] in particular has developped the underlying philosophy and shown how it relates to "arithmetic quantum chaos", i.e. the study of the repartition of eigenfunctions of the Laplace operator on negatively curved manifolds as the eigenvalue gets large.

We finish with the proof of the Proposition.

PROOF. Consider the integral over the line $\mathrm{Re}(s) = 3$:

$$I = \frac{1}{2i\pi} \int\limits_{(3)} L(\chi, s + w) G(w) \frac{dw}{w}.$$

By expanding into Dirichlet series the L-function, it is equal to

$$\sum_{\mathfrak{a}} \frac{\chi(\mathfrak{a})}{(N\mathfrak{a})^s} V_s\left(\frac{N\mathfrak{a}}{X}\right)$$

i.e. the first sum in (2.13).

On the other hand, shifting the contour to the line $\mathrm{Re}(s) = -3$ (using the fact that $G(w)$ has been chosen to cancel any possible pole of the gamma factor), picking up the single simple pole at $s = 0$, we obtain

$$I = G(0)L(\chi, s) + \frac{1}{2i\pi} \int\limits_{(-3)} L(\chi, s + w) G(w) \frac{dw}{w}.$$

Now apply the functional equation of $L(\chi, s)$: the last integral (say J) becomes

$$J = \varepsilon(\chi)(|D|N\mathfrak{m})^{1/2-s} \frac{1}{2i\pi} \int\limits_{(-3)} \frac{\Gamma(\xi_\infty, 1 - s - w)}{\Gamma(\xi_\infty, s + w)} L(\overline{\chi}, 1 - s - w) G(w) \left(\frac{X}{|D|N\mathfrak{m}}\right)^w \frac{dw}{w}$$

$$= -\varepsilon(\chi)(|D|N\mathfrak{m})^{1/2-s} \frac{1}{2i\pi} \int\limits_{(3)} \frac{\Gamma(\xi_\infty, 1 - s - w)}{\Gamma(\xi_\infty, s + w)} L(\overline{\chi}, 1 - s + w) G(w) \left(\frac{X}{|D|N\mathfrak{m}}\right)^{-w} \frac{dw}{w}$$

33

We are back in the region of absolute convergence, and expanding again into Dirichlet series, it follows that

$$J = -\varepsilon(\chi)(|D|N\mathfrak{m})^{1/2-s} \sum_{\mathfrak{a}} \frac{\overline{\chi}(\mathfrak{a})}{(N\mathfrak{a})^{1-s}} W_s\left(\frac{N\mathfrak{a}}{Y}\right).$$

Hence the result since $G(0) = 1$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

2.5. Odds and ends, including poles

One of the features, partly known and partly conjectural, of the theory of L-functions, is that it seems that (when unitarily normalized, i.e. the critical line is translated to $\mathrm{Re}(s) = 1/2$) the point $s = 1$ is the only possible pole for an (automorphic) L-function, and even more, that such a pole is always accounted for by the simple pole of the Riemann zeta function, in the sense that the L-function $L(f, s)$ has a factorization

$$L(f, s) = \zeta(s)^k L(f_1, s),$$

where $L(f_1, s)$ is another L-function which is entire. This is indeed the case for the Dedekind zeta function as (1.19) shows, and is also true in a number of other cases (for the Rankin-Selberg convolution L-functions $L(f \otimes \bar{f}, s)$ on $GL(2)$, for instance, where the "quotient" f_1 is the symmetric square L-function, as will be explained in other lectures). It is known also for Artin L-functions, using Brauer's Theorem and the non-vanishing of Hecke L-functions at $s = 1$. The order of the pole at $s = 1$ of $L(\rho, s)$ is then the multiplicity of the trivial representation in ρ. This formulation is recurrent in many conjectures of arithmetic-geometry, such as the Tate conjectures about the dimension of spaces of k-cycles on varieties over finite fields.

A pole occurs typically (only?) when the coefficients of the L-function (as a Dirichlet series) are positive. Analytically, this conjecture can be paraphrased as saying that "the harmonic series is the only really divergent one".

The following cute observation (mentioned to me by Serre) gives an interesting illustration: recall Euler's formula for the values of the zeta function at negative integers

(2.14) $\qquad\qquad\qquad \zeta(0) = -1/2, \text{ and } \zeta(1 - k) = -b_k/k \text{ for } k \geqslant 2$

where b_k is the k-th Bernoulli number (see e.g. [**Wa**]). Now theorems of von Staudt and J.C. Adams [**IR**, Ch. 15, Th. 3 and Pr. 15.2.4] prove that the denominator of $-b_k/k$ contains all the primes p such that $p - 1 \mid 2k$, *and only them*, so that for a p in the denominator we have

$$x^{k-1} \equiv x^{-1} \,(\mathrm{mod}\, p).$$

Hence even the "poles" modulo primes of $\zeta(1-k)$ are still "explained" by the divergence of the harmonic series! In the other direction, maybe it is not so surprising that the numerator of Bernoulli numbers should remain so mysterious...

We finally only mention that the congruence properties of the values of the zeta function at negative integers were the motivation for the discovery by Leopoldt of the p-adic zeta function, later much generalized by others to p-adic L-functions of various kinds. See for instance [**Wa**, Ch. 5] for an introduction.

Classical Automorphic Forms

3.1. Introduction

With automorphic forms and their associated L-functions, one enters into new territory; the catch-phrase here is that we will describe the $GL(2)$ analogue of the $GL(1)$ theory of Dirichlet characters (i.e. over $\mathbf{Q}$). The corresponding work for more general groups and general base field K will be described in later lectures.

For more complete accounts of this beautiful theory, see for instance the books of Iwaniec ([**I1**], [**I2**]), Miyake [**Mi**] or Shimura [**Sh**], chapter VII of Serre's book [**Se1**] or chapter I of Bump's book [**Bu**].

3.2. Three motivations for automorphic forms

First we will give three reasons for introducing automorphic forms. Two are obviously related to problems mentioned in the first two lectures, and the third one has a very natural algebro-geometric content. One of the amazing features of automorphic forms is that one could find many other apparently disjoint motivating examples; those included here make no mention of such important topics as the links with Galois representations, congruences, partitions, etc...

3.2.1. Theta functions. The proof of the functional equation of the Riemann zeta function (by specialization of that in Lecture II) makes use of the basic theta function

$$(3.1) \qquad \theta(z) = \sum_{n \in \mathbf{Z}} e(n^2 z / 2)$$

for $z = iy$, $y > 0$. In fact, this defines a function holomorphic on the ubiquitous Poincaré upper half-plane $\mathbf{H} = \{z \in \mathbf{C} \mid y > 0\}$. The Poisson summation formula (with the fact that $x \mapsto e^{-\pi x^2}$ is self-dual for the Fourier transform) implies by analytic continuation the functional equation

$$\theta\left(-\frac{1}{z}\right) = (-iz)^{1/2} \theta(z),$$

where $(-iz)^{1/2}$ is given by the branch of this function on $\mathbf{H}$ which sends iy to $\sqrt{y}$.

Obviously $\theta(z)$ is also 1-periodic so there is some transformation formula for $f(\gamma z)$ where $z \mapsto \gamma z$ is any mapping in the group generated by $z \mapsto z + 1$ and $z \mapsto -1/z$. This group $\bar{\Gamma}$ turns out to be isomorphic to $PSL(2, \mathbf{Z})$ acting on $\mathbf{H}$ by

$$(3.2) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az + b}{cz + d}.$$

Similarly, it follows that for any $k \geqslant 1$, $\theta(2z)^k$ satisfies similar transformation formulae. Since

$$\theta(2z)^k = \sum_{n \geqslant 0} r_k(n)e(nz)$$

with $r_k(n)$ the number of representations of n as sums of k squares of integers, it is not surprising that much classical interest was expanded around such functions. Identities such as

$$r_2(n) = 4 \sum_{d|n} \chi_4(d)$$

(3.3)
$$r_4(n) = 8(2 + (-1)^n) \sum_{\substack{d|n \\ d \text{ odd}}} d$$

$$r_6(n) = 16 \sum_{d|n} d^2 \chi_4\left(\frac{n}{d}\right) - 4 \sum_{d|n} d^2 \chi_4(d)$$

can be derived by (more or less) elementary or ad-hoc means, but the theory of modular forms can explain the general shape of such formulas, why there is none so "elementary" for k large, and yields asymptotic relations implying that the size of $r_k(n)$ is about n^{k-1} for $k \geqslant 5$. Much deeper results are due to Kloosterman for general quadratic forms in $k = 4$ variables and to Iwaniec for $k = 3$ variables (see [**I1**, 5-3,11-4]), yielding for instance equidistribution results for integral points on the sphere.

3.2.2. Counting solutions of determinant equations. In studying the order of magnitude of $\zeta(s)$ or of other L-functions on the critical strip, a natural analytic approach is through the moments

$$I_k(T) = \int_{-T}^{T} |\zeta(1/2 + it)|^k dt$$

since if we could prove that for all $k \geqslant 1$ we have

(3.4)
$$I_k(T) \ll_{k,\varepsilon} T^{1+\varepsilon} \text{ for } T \geqslant 1,$$

the Lindelöf Conjecture $\zeta(1/2 + it) \ll_\varepsilon (1 + |t|)^\varepsilon$ would easily follow. Note that even for a single $k \geqslant 1$, (3.4) is a confirmation of the Lindelöf Conjecture on average, stronger for larger values of k, and as such does have many applications in analytic number theory.

For $k = 2$ an asymptotic evaluation of $I_2(T)$ is rather simple and classical (see e.g. [**Ti**, 7.3]), but already the case $k = 4$ is challenging and (3.4) is unknown for any value of $k > 4$. For $k = 4$, standard techniques reduce (3.4) to estimating the fourth moment of exponential sums

$$\left(\sum_{n \leqslant X} n^{-it}\right)^4 \text{ for } X \sim \sqrt{T},$$

and one expands the sum as

$$S = \sum_{a,b,c,d \leqslant X} \left(\frac{bc}{ad}\right)^{it}.$$

It is natural to split the sum according to the value of the *determinant* $h = ad - bc$ so that
$$S = \sum_{|h| \leqslant X^2} \sum_{ad - bc = h} \left(1 - \frac{h}{ad}\right)^{it}.$$
The idea is the following: if $h = 0$, we obtain the equation $ad = bc$ which is easily solved and yields a "main term". Then for $h \neq 0$, but "small", the exponential $(1 - h/ad)^{it}$ is not oscillating very much and will also yield an important contribution, whereas for h "large", the oscillations should cause a large amount of cancellation. To account for the contribution of the small values of h, one needs to be able to count with precision the number of solutions of the determinant equation
$$ad - bc = h$$
with a, b, c, $d \leqslant X$. This can be done by "usual" harmonic analysis[1] but much better results are obtained if one remarks that $SL(2, \mathbf{Z})$ acts on the solutions (when there is no size condition) and if one tries to apply a form of harmonic analysis (i.e. a form of the Poisson summation formula) adapted to this action: here we go from the classical case of $\mathbf{Z}^n$ discrete in $\mathbf{R}^n$ (and the applications to lattice-point counting in euclidean space) to that of $SL(2, \mathbf{Z})$ discrete in $SL(2, \mathbf{R})$ with corresponding lattice point problems. Those, however, are of a completely different nature because $SL(2, \mathbf{R})$ is not abelian and the whole geometry is of hyperbolic – negative curvature – type (so for instance most of the area of a domain accumulates on its boundary). For more about those aspects, see [**I2**].

3.2.3. Invariants of elliptic curves. Consider the standard torus $\mathbf{T} = \mathbf{R}^2/\mathbf{Z}^2$; it is a 2-dimensional smooth Lie group. What compatible structures of complex Lie group can be put on $\mathbf{T}$? The answer to this question naturally leads to the study of the space of lattices in $\mathbf{R}^2 = \mathbf{C}$ up to unimodular transformation: indeed if $\Lambda \subset \mathbf{C}$ is a lattice, then the quotient $\mathbf{C}/\Lambda$ is diffeomorphic to $\mathbf{T}$ and inherits from $\mathbf{C}$ a compatible complex structure and a group structure. Moreover, all such structures must arise in this way.

Since all complex automorphisms of $\mathbf{C}$ (as a group) are linear $z \mapsto \alpha z$ one sees that two such *complex tori* $\mathbf{C}/\Lambda_1$ and $\mathbf{C}/\Lambda_2$ are isomorphic as Riemann surfaces if and only if $\Lambda = \alpha \Lambda'$ for some $\alpha \in \mathbf{C}^\times$. Now if $\Lambda = \omega_1 \mathbf{Z} \oplus \omega_2 \mathbf{Z}$ and $\Lambda' = \omega_1' \mathbf{Z} \oplus \omega_2' \mathbf{Z}$, then $\Lambda = \Lambda'$ if and only if the elements of the basis are related by a unimodular linear transformation
$$\omega_1' = a\omega_1 + b\omega_2$$
$$\omega_2' = c\omega_1 + d\omega_2$$
with
$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbf{Z}), \text{ ie } ad - bc = 1.$$
Up to homothety, $\Lambda \simeq \mathbf{Z} \oplus \tau \mathbf{Z}$ with $\tau = \omega_1/\omega_2$ (resp. $\Lambda' \simeq \mathbf{Z} \oplus \tau' \mathbf{Z}$ with $\tau' = \omega_1'/\omega_2'$) so the equation becomes
$$\tau' = \gamma \cdot \tau = \frac{a\tau + b}{c\tau + d}$$
and finally we can also still switch ω_1 and ω_2, i.e. replace τ by $1/\tau$: one and one only of the two choices is in $\mathbf{H}$.

---

[1] Actually, this quickly involves Kloosterman sums, and whichever way one estimates those, there are automorphic forms lurking in the background.

Thus the set of complex tori can be identified with the quotient of $\mathbf{H}$ by this action of $SL(2, \mathbf{Z})$, and various invariants of such complex tori appear as functions $f : \mathbf{H} \to \mathbf{C}$ invariant under this action. Traditionally, meromorphic functions on $\mathbf{C}$ with two $\mathbf{R}$-linearly independent periods ω_1 and ω_2 are called *elliptic functions* and many naturally arise in such a way that they give rise to functions $f(z; \omega_1, \omega_2)$ homogeneous of some weight k so that

$$f(\lambda z; \lambda \omega_1, \lambda \omega_2) = \lambda^{-k} f(z; \omega_1, \omega_2),$$

and thus

$$f(u; \omega_1, \omega_2) = \omega_2^{-k} g(u/\omega_2, \omega_1/\omega_2),$$

for some function g defined on $\mathbf{C} \times (\mathbf{H} \cup \overline{\mathbf{H}})$, which restricted to $\mathbf{C} \times \mathbf{H}$ is such that

$$g(z; \gamma\tau) = (c\tau + d)^k g(z; \tau) \text{ for } \gamma \in SL(2, \mathbf{Z}).$$

EXAMPLE 3.2.1. The Weierstraß $\wp$-function of the lattice Λ is

$$\wp(z; \omega_1, \omega_2) = \frac{1}{z^2} + \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right),$$

it is Λ-invariant (this is not so obvious: differentiate to see that $\wp'$ is definitely Λ-invariant, then integrate). Clearly

$$\wp(\lambda z; \lambda\omega_1, \lambda_2\omega_2) = \lambda^{-2}\wp(z; \omega_1, \omega_2).$$

Moreover

$$\wp(z; \Lambda) = z^{-2} + \sum_{m \geqslant 1} (m+1) z^m G_{m+2}(\Lambda)$$

where

$$G_k(\Lambda) = \sum_{\omega \neq 0} \frac{1}{\omega^k},$$

and $G_k(\lambda\Lambda) = \lambda^{-k} G_k(\Lambda)$ furnishes another example of those *modular functions*. This theme of elliptic functions is again very classical and much beloved of mathematicians of the 18th and 19th century.

3.3. Definitions and examples

We will now briefly survey the main definitions and results of what can be called the classical theory of automorphic forms and L-functions. This belongs historically to the early 20th century and the most important names for us are Eisenstein, Poincaré, Hecke, Ramanujan, Petersson and (later) Maass and Selberg. We will speak simultaneously of holomorphic modular forms and of non-holomorphic (Maass) forms, although the latter were introduced quite a bit later: this will strongly motivate the development of more group-theoretic methods to unify the work being done.

For simplicity we will mostly restrict our attention to congruence subgroups of $SL(2, \mathbf{Z})$, which is arithmetically the most important.

Recall that $\mathbf{H}$ is the Poincaré upper half-plane. It is a simply connected Riemann surface and a model of the hyperbolic plane (of constant negative curvature -1) when equipped with the riemannian metric

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

The group $GL(2, \mathbf{R})^+$ of invertible matrices with positive determinant acts on $\mathbf{H}$ by fractional linear transformation as in (3.2). The scalar matrices act trivially so often one restricts the attention to $SL(2, \mathbf{R})$ or to the quotient $PGL(2, \mathbf{R})^+ = PSL(2, \mathbf{R})$, which is the automorphism group of $\mathbf{H}$ as a Riemann surface, and the group of orientation preserving isometries of $\mathbf{H}$ as riemannian manifold; the full group of isometries is the semi-direct product

$$PSL(2, \mathbf{R}) \rtimes \mathbf{Z}/2\mathbf{Z},$$

where a representative of the orientation-reversing coset is the symmetry $z \mapsto -\bar{z}$ with respect to the imaginary axis.

Associated to the Riemann metric there is the volume element $d\mu(z) = y^{-2}dxdy$, and the Laplace operator $\Delta = -y^2(\partial_x^2 + \partial_y^2)$, both of which are invariant by isometries.

In analogy with the study of periodic functions on $\mathbf{R}$, or of elliptic functions, one considers subgroups[2] $\Gamma < SL(2, \mathbf{R})$ which are discrete, and then functions on $\mathbf{H}$ which are "periodic" with respect to Γ, i.e. functions on the quotient space $\Gamma \backslash \mathbf{H}$, or differential forms, etc...

A considerable difference is that those quotients are very diverse in shape and form. Topologically, almost all (orientable) surfaces arise as such quotients, and all Riemann surfaces of genus > 1 (it is known since Riemann that the set of all such surfaces, up to complex isomorphism, depends on $3g - 3$ complex "parameters" or "moduli").

For arithmetic purposes, the subgroup $SL(2, \mathbf{Z})$ springs to attention: it is indeed discrete in $SL(2, \mathbf{R})$ and therefore so is every subgroup of $SL(2, \mathbf{Z})$. In general a *congruence subgroup* $\Gamma < SL(2, \mathbf{Z})$ is one such that $\Gamma \supset \Gamma(q) = \{\gamma \in SL(2, \mathbf{Z}) \mid \gamma \equiv 1 \,(\mathrm{mod}\, q)\}$ for some q. Among congruence groups, the *Hecke congruence groups*

$$\Gamma_0(q) = \{\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbf{Z}) \mid q \mid c\}$$

for any integer $q \geqslant 1$ will be of particular interest. The number q is called the *level* of the group.

The index of $\Gamma_0(q)$ in $\Gamma_0(1) = SL(2, \mathbf{Z})$ is easily computed

$$[\Gamma_0(1) : \Gamma_0(q)] = q \prod_{p \mid q} \left(1 + \frac{1}{q}\right).$$

In the "abelian" case, $\mathbf{R}/\mathbf{Z}$ is compact; however this is not so of the quotient $SL(2, \mathbf{Z}) \backslash \mathbf{H}$ and this introduces the important notion of *cusps*.

We will describe geometrically the cusps as follows: one can visualize the quotient space $\Gamma \backslash \mathbf{H}$, for any $\Gamma < SL(2, \mathbf{R})$ discrete, by a *fundamental domain*, i.e. an open subset F of $\mathbf{H}$ such that
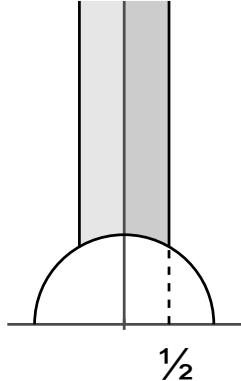
- No two points in F are Γ-equivalent.
- Any Γ-orbit Γz intersects the closure $\overline{F}$ of F.

In the case of $SL(2, \mathbf{Z})$, one can take the familiar triangle with one vertex at infinity

$$F_1 = \{z = x + iy \in \mathbf{H} \mid -1/2 < x < 1/2, \text{ and } |z| > 1\}$$

(see e.g. [**Se1**, VII-1.2]).

---

[2] The context will always clearly distinguish between a subgroup and the gamma function.

½

Then for $\Gamma < SL(2, \mathbf{Z})$ (the case which will occupy us), one can take

$$F_\Gamma = \bigcup_{g \in \Gamma \backslash \mathbf{H}} g \cdot F_1.$$

The Gauss-Bonnet theorem or direct integration shows that the hyperbolic volume of F_1 is finite (it is equal to $\pi/3$), but since $in \in F_1$ for $n > 1$, it is not compact. The situation is the same for $\Gamma_0(q)$, with volume multiplied by the index of $\Gamma_0(q)$ in $SL(2, \mathbf{Z}) = \Gamma_0(1)$.

The lack of compactness lies of course in the points where the fundamental domain reaches the "boundary" $\overline{\mathbf{R}} = \mathbf{R} \cup \infty$. Note that $GL(2, \mathbf{R})^+$ also acts on the boundary, by the same formula (3.2).

If the covolume $\mathrm{Vol}(\Gamma \backslash \mathbf{H}) = \mathrm{Vol}(F)$ of Γ is finite, there are only finitely many cusps because around each of them small disjoint neighborhoods with constant positive volume can be constructed. There are no cusps if and only if $\Gamma \backslash \mathbf{H}$ is compact. (We will not really discuss such groups here; however, some very important arithmetic examples exist, based on unit groups of quaternion algebras, and they occur in the Jacquet-Langlands correspondance that will be discussed in other lectures). In terms of group theory, any cusp $\mathfrak{a}$ for a group Γ is the unique fixed point of some $\gamma \in \Gamma$ (such elements are called *parabolic*). The geometric action of such a γ on $\mathbf{H}$ is by translation along the horocycles "around the cusp": for $\mathfrak{a} = \infty$,

(3.5) $$\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ acting by } z \mapsto z + 1,$$

and the horocycles are the horizontal lines $\mathrm{Im}(z) = y_0$.

By transitivity, for any cusp $\mathfrak{a}$, there exists a *scaling matrix* $\sigma_\mathfrak{a} \in SL(2, \mathbf{R})$ such that

$$\begin{cases} \sigma_\mathfrak{a} \infty = \mathfrak{a} \\ \sigma_\mathfrak{a}^{-1} \gamma_\mathfrak{a} \sigma_\mathfrak{a} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{cases}$$

which, by conjugacy, is used to reduce all notions about cusps to the case of $\mathfrak{a} = \infty$ which is simpler to visualize (having a distinguished cusp is a very convenient feature of the Poincaré upper half-plane, compared to other models of the hyperbolic plane). Here $\gamma \in \Gamma_\mathfrak{a} < \Gamma$ is a generator of the stabilizer of the cusp in Γ, which is an infinite cyclic group (when projected to PSL_2 at least).

DEFINITION. Let $f : \mathbf{H} \to \mathbf{C}$ be a 1-periodic smooth function on $\mathbf{H}$ ($f(z+1) = f(z)$),

$$f(z) = \sum_{n \in \mathbf{Z}} a_f(n; y) e(nx),$$

its Fourier expansion. One says that
   (i) f is of *moderate growth* at ∞ if there exists $A > 0$ such that

$$f(z) \ll |y|^A \text{ for } y \geqslant 1;$$

   (ii) f is *cuspidal* at ∞ if $a_0(y) = 0$.

Besides the natural right-action of $SL(2, \mathbf{R})$ on functions on $\mathbf{C}$ by $(f \,|\, \gamma)(z) = f(\gamma z)$, one defines for any integer k an action

$$f \mapsto f \,|_k \gamma \text{ where } (f \,|_k \gamma)(z) = (cz+d)^{-k} f(\gamma z).$$

One checks that indeed $f \,|_k \gamma \,|_k \gamma' = f \,|_k \gamma\gamma'$. Also, the matrix (3.5) still acts by $(f \,|_k \gamma)(z) = f(z+1)$.

We can finally define modular forms.

DEFINITION. Let $k \in \mathbf{Z}$ be an integer, $\lambda \in \mathbf{C}$, $q \geqslant 1$ an integer and χ a Dirichlet character modulo q. A smooth function $f : \mathbf{H} \to \mathbf{C}$ on $\mathbf{H}$ is called a *modular form* of weight k, level q, with nebentypus χ, resp. an automorphic function with eigenvalue λ, level q and nebentypus χ, if f is holomorphic on $\mathbf{H}$ and satisfies

$$f \,|_k \gamma = \chi(a) f, \text{ for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(q),$$

resp. f satisfies $\Delta f = \lambda f$ and

$$f \,|\, \gamma = \chi(a) f, \text{ for } \gamma \in \Gamma_0(q).$$

REMARK 3.3.1. Taking $\gamma = -1$, on gets the relation $f = f \,|_k \gamma = \chi(-1)(-1)^k f$, so there can exist non-zero holomorphic modular forms only if the character satisfies the consistency condition $\chi(-1) = (-1)^k$, which is tacitly assumed to be the case in what follows. Similarly for automorphic functions, we must have $\chi(-1) = 1$.

One can also define non-holomorphic forms of weight $k \neq 0$, using a modified differential operator. Both holomorphic and Maass forms can be most convincingly put into a single framework through the study of the representation theory of $GL(2, \mathbf{R})$ (or of the adèle group $GL(2, \mathbf{A})$ in the arithmetic case).

Using the definition above, one can impose more regularity conditions at the cusps:

DEFINITION. Let f be a modular form (resp. an automorphic function). One says that f is holomorphic at the cusps (resp. is a Maass form) if and only if it is of moderate growth at all cusps, i.e. for any $\mathfrak{a}$, the 1-periodic function $f_{\mathfrak{a}} = f \,|_k \sigma_{\mathfrak{a}}$ is of moderate growth at infinity. We let $M_k(q, \chi)$ (resp. $M^\lambda(q, \chi)$) denote the vector space of modular forms of weight k, level q and character χ (resp. Maass forms).

One says that f is a cusp form (resp. a Maass cusp form) if it is cuspidal at all cusps, and we denote $S_k(q, \chi) \subset M_k(q, \chi)$ (resp. $S^\lambda(q, \chi)$) the subspace of cusp forms (resp. Maass cusp forms).

REMARK 3.3.2. Other equivalent formulations can be given. In particular, for $\Gamma_0(q)$, an automorphic function f is a Maass cusp form if and only if f is non-constant (i.e. $\lambda \neq 0$) and in $L^2(\Gamma_0(q)\backslash\mathbf{H})$ (with respect to the hyperbolic measure). This amounts to saying that 0 is the only residual eigenvalue for $\Gamma_0(q)$ (see e.g. [**I2**, 11.2]).

For holomorphic forms, the Fourier expansion at infinity is of the form

$$(3.6) \qquad\qquad f(z) = \sum_{n \in \mathbf{Z}} a_f(n)e(nz) \text{ with } a_f(n) \in \mathbf{C},$$

and f being holomorphic means $a_n = 0$ for $n < 0$, since $|e(nz)| = \exp(-2\pi ny)$ is not polynomially bounded for $n < 0$.

REMARK 3.3.3. Let $Y_0(q) = \Gamma_0(q)\backslash\mathbf{H}$. This is a non-compact Riemann surface (there are some fixed points in $\mathbf{H}$ under the action, all $SL(2, \mathbf{Z})$-equivalent to either i, fixed by $z \mapsto -1/z$, or $\exp(i\pi/3)$, fixed by $z \mapsto (z - 1)/z$; however suitable coordinates at those points still provide a complex structure, see e.g. [**Sh**, Ch. 1] for the details).

Moreover, by adding the finitely many cusps, one can compactify $Y_0(q)$, getting a compact Riemann surface denoted $X_0(q)$. For instance, at ∞ one uses $q = e(z)$ as a coordinate chart (compare (3.6)). For $q = 1$, one gets $X_0(1) \simeq \mathbf{P}(1)$, through the j-invariant $j(z)$ (see e.g. [**Se1**, VII-3.3]).

By the general theory of Riemann surfaces, it follows that $X_0(q)$ is actually an algebraic curve. However much more is true: this algebraic curve is actually defined over $\mathbf{Q}$, and very good models over $\mathbf{Z}$ exist (so the "bad fibers" are quite well understood). All this is largely based on the interpretation of $X_0(q)$ as a *moduli space*: it "classifies" pairs (E, H), where E is an elliptic curve and H is a cyclic subgroup of order q on E. The image in $X_0(q)$ of the point $z \in \mathbf{H}$ then corresponds to the pair $(\mathbf{C}/(\mathbf{Z} \oplus z\mathbf{Z}), <1/q>)$. These are very important aspects that will be discussed in further lectures, and it is the key to many of the deep arithmetical properties of modular forms. This is related to our third motivating example for modular forms.

An easy example of those links is the important isomorphism

$$\begin{cases} S_2(q) \to \Gamma(X_0(q), \Omega_1) \\ f \mapsto f(z)dz \end{cases}$$

between the space of weight 2 cusp forms of level q (with trivial nebentypus) and the space of holomorphic 1-forms on the modular curve $X_0(q)$ (not $Y_0(q)$). This is fairly easy to prove: the weight 2 action is just what is needed to show that $f(z)dz$ is invariant under $\Gamma_0(q)$, hence descends to a 1-form on $X_0(q)$. In addition, $dq/q = 2i\pi dz$ for the local coordinate $q = e(z)$ at ∞, so the condition of vanishing at the cusp ∞ is equivalent with $f(z)dz$ being holomorphic at ∞, and similarly at the other cusps.

The last paragraph of this remark shows (since $X_0(q)$ is compact so the space of global differentials $\Gamma(X_0(q), \Omega_1)$ is finite dimensional) that $\dim_{\mathbf{C}} S_2(q) < +\infty$. Indeed, it is a general fact that $\dim M_k(q, \chi)$ (resp. $\dim M^\lambda(q, \chi)$) is finite.[3] For holomorphic forms, one can argue in an elementary way, using Cauchy's Theorem (see e.g. [**Se1**, VII-3]), but deeper results, including exact formulae for $\dim M_k(q, \chi)$, when $k > 1$, follow from the Riemann-Roch theorem as in [**Sh**, 2.6]: for instance $\dim S_2(q)$ is the genus of $X_0(q)$, by the above, which

---

[3] This holds in much greater generality.

can be computed explicitly, using the natural (ramified) covering $X_0(q) \to X_0(1) = \mathbf{P}(1)$ induced by the inclusion $\Gamma_0(q) < \Gamma_0(1)$.

With small fluctuations, the dimension increases as the level and the weight do: $\dim M_k(q)$ is of size (roughly) $qk/12$, so there are many modular forms as soon as the weight or the level is not too small (for instance, the genus of $X_0(q)$ is > 1 when $q > 49$, or if q is prime > 19).

The excluded case[4] $k = 1$ is indeed quite different: the problem of computing its dimension is of great arithmetic significance, and remains largely open (see [**Se2**] and [**Du**]). Spectrally, the weight $k = 1$ corresponds to an eigenvalue which is in the continuous spectrum so it is very hard to pick it up using the tools of harmonic analysis like the trace formula: any reasonable test function also detects the (large) contribution of eigenvalues close to it.

For the case of Maass forms, $\dim M^\lambda(q, \chi)$ is proved to be finite using the spectral theory of operators in Hilbert space. In contrast to the holomorphic case, this is just a qualitative statement and no formula for the dimension of this space is known. This is understandable, as the spectral theory also implies that $S^\lambda(q, \chi) = 0$ for all but (at most!) countably many eigenvalues $\lambda > 0$, tending to $+\infty$ (and M^λ/S^λ is easy to describe). The eigenvalues are completely mysterious and only few special examples are known ($\lambda = 1/4$ is a special case, also of deep arithmetic significance, a recurring sentence here, and a few others can be shown to have associated Maass forms using Hecke characters of real quadratic fields, as shown by Maass; see the example in [**Ge**]).

Selberg developed his celebrated trace formula at least in part to address this problem of the existence of Maass cusp forms. Using it, he was able to prove (see e.g. [**I2**, Ch. 11]) that *for congruence subgroups*, in particular for $\Gamma_0(q)$, the Weyl law holds:

$$|\{\lambda \mid 0 \leqslant \lambda \leqslant X \text{ and } S^\lambda(q) \neq 0\}| \sim \frac{\mathrm{Vol}(X_0(q))}{4\pi} X$$

as $X \to +\infty$, the eigenvalues being counted with multiplicity.[5]

On the other hand, it is now considered likely (contrary to the expectation when Selberg proved the result above) that for "generic" (non-arithmetic) groups Γ, this statement fails so badly that there are only finitely many (if any) eigenvalues λ for Maass cusp forms for Γ. This shift in common belief is due to the beautiful theory of Phillips and Sarnak of disappearance of cusp forms under deformation of the group [**PS**].

We conclude this section by giving some concrete examples of modular forms of various types.

EXAMPLE 3.3.4. The most natural way of constructing a function which is "periodic" under some action of a group G is by averaging (see the proof of the Poisson summation formula, Proposition 2.2.1) a function over the group. In fact, if the function is already invariant under a subgroup $G_1 < G$, one can average only over the cosets of G_1 in G. This leads to Poincaré and Eisenstein series.

---

[4] For $k \leqslant 0$, there are no modular forms.

[5] The problem of bounding from above in a precise way the multiplicity of Maass cusp forms is one of the most inscrutable open problems in analytic number theory.

Let $m \geqslant 0$ be an integer. Define the m-th holomorphic Poincaré series of weight k by

$$P_m(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(q)} \overline{\chi}(a)(cz+d)^{-k} e(m\gamma z).$$

For $k > 2$, this series converges absolutely uniformly on compact subsets to a holomorphic function on $\mathbf{H}$ which, for the reason just mentioned, is a modular form of level q. One can compute the expansion of P_m at the various cusps and see that $P_m \in M_k(q, \chi)$, and in fact that if $m \neq 0$, then $P_m \in S_k(q, \chi)$ (see below). For $m = 0$, this is called an *Eisensten series*. In the case of $q = 1$, we get the classical Eisenstein series

$$E_k(z) = \sum_{(a,b)=1} \frac{1}{(az+b)^k}.$$

Eisenstein series and Poincaré series can be defined in the non-holomorphic setting also, with a complex variable s instead of k:

(3.7) $$P_m(z; s) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma_0(q)} \overline{\chi}(a) \operatorname{Im}(\gamma z)^s e(m\gamma z), \text{ and } E(z; s) = P_0(z; s)$$

(notice that $\operatorname{Im}(\gamma z) = |cz + d|^{-2}$ for $\gamma \in SL(2, \mathbf{R})$). This converges for $\operatorname{Re}(s) > 1$, and defines then a real-analytic function on $\mathbf{H}$ which satisfies

$$P_m(\cdot; s) \,|\, \gamma = \overline{\chi}(d) P_m(\cdot; s) \text{ for } \gamma \in \Gamma_0(q),$$

but it is not an eigenvalue of the laplacian if $m \neq 0$: in fact

$$\Delta P_m(\cdot; s) = (s(1-s) - 4\pi^2 m^2) P_m(\cdot; s) + 4\pi m s P_m(\cdot; s+1).$$

The Eisenstein series $E(z; s)$ for $\operatorname{Re}(s) > 1$, on the other hand, is indeed an eigenfunction of Δ with eigenvalue $\lambda = s(1-s)$ and it is quite easy to check that it has polynomial growth at the cusps . As will be discussed in further lectures, non-holomorphic Eisenstein series turn out to play a very important role even for studying holomorphic forms, through their appearance in the Rankin-Selberg method for instance (see Section 3.6.2 for a very short introduction).

In general, it is often convenient to write *any* eigenvalue in this way (so two values of s exist for a given λ), and to further put $s = 1/2 + it$ with $t \in \mathbf{C}$. For cusp forms, $\lambda > 0$ translates into $s \in]0, 1[$ (if $\lambda \leqslant 1/4$) or $\lambda = 1/2 + it$ with $t \in \mathbf{R}$. In the first case, the eigenvalue is called *exceptional*. Selberg made the following conjecture:

CONJECTURE 3.3.5. *For any congruence subgroup Γ, in particular $\Gamma_0(q)$, there is no exceptional eigenvalue, i.e. the first non-zero eigenvalue for Δ acting on $L^2(\Gamma \backslash \mathbf{H})$ satisfies $\lambda_1 \geqslant 1/4$.*

Moreover, Selberg proved $\lambda_1 \geqslant 3/16$. This is an arithmetic statement because it is easy to construct non-congruence subgroups for which there are arbitrarily many exceptional eigenvalues. In problems such as those mentioned in Section 3.2.2, exceptional eigenvalues for $\Gamma_0(q)$ have an effect as the Landau-Siegel zero for Dirichlet characters: the uniformity in q (say in counting solutions to $ad - bc = h$ with $c \equiv 0 \,(\mathrm{mod}\, q)$ with $a^2 + b^2 + c^2 + d^2 \leqslant X$) is affected by the presence of "many" exceptional eigenvalues (the closer to 0, the worse the effect). Hence Selberg's theorem indicates that the situation is a little bit better controlled. This conjecture is also clearly understood today as the archimedean analogue

of the Ramanujan-Petersson conjecture (see Section 3.4.2 below), and indeed the significant improvements to the 3/16 bound proved by Luo, Rudnick and Sarnak [**LRS**] is based on this analogy.

The spectral analysis of the general Poincaré series is the essence of the Kuznetsov-Bruggeman formula which is of great importance in the applications of Maass forms to analytic number theory (see [**I2**, Ch. 9], [**CP**]).

Properly speaking, we have defined the Poincaré and Eisenstein series relative to the cusp ∞. One can define analogous functions for any cusp, and the computation of the Fourier expansions of Eisenstein series shows that $M_k(q, \chi) = S_k(q, \chi) \oplus E_k(q, \chi)$, where $E_k(q, \chi)$ is the vector space spanned by the Eisenstein series (at all cusps).

EXAMPLE 3.3.6. We now come to theta functions, with notation as in [**I1**, Ch. 10], in a fairly general situation. Let A be a symmetric positive definite $(r \times r)$-matrix with integral coefficients, with all diagonal coefficients *even*, and $N \geqslant 1$ any integer such that NA^{-1} has integer coefficients. We let
$$A[x] = {}^t x A x, \text{ for } x \in \mathbf{R}^r$$
be the associated quadratic form. *Assume the number of variables r is even.*[6] Let
$$\Theta(z; A) = \sum_{m \in \mathbf{Z}^r} e(A[m]z/2).$$

PROPOSITION 3.3.7. *We have* $\Theta(z; a) \in M_{r/2}(2N)$.

This is proved using the Poisson summation formula and generalizes, of course, the formula (2.5).

For instance, let $A_4 = \operatorname{diag}(2, 2, 2, 2)$. We can take $N = 2$, and we have $A_4[x]/2 = x_1^2 + x_2^2 + x_3^2 + x_4^2$, so $\Theta(z; A_4) = \theta(2z)^4$ where θ is the basic theta function (3.1). Hence $\Theta(\cdot; A_4) \in M_2(4)$. Now $X_0(4)$ has genus zero hence $S_2(4) = 0$, so as explained in the previous example $\Theta \in E_2(4)$. There are 3 cusps for $\Gamma_0(4)$, namely 0, 1/2 and ∞ and one can explicitly compute the Fourier expansions of the three corresponding Eisenstein series of weight 2, E_0, $E_{1/2}$ and E_∞ say. Checking the first few coefficients, one identifies the linear combination equal to $\Theta(z; A_4)$:
$$\Theta(z; A_4) = \alpha E_\infty(z) + \beta E_0(z),$$
for some explicit α, β, and (3.3) follows from this.

EXAMPLE 3.3.8. The Ramanujan Delta function is
$$\Delta(z) = e(z) \prod_{n \geqslant 1} (1 - e(nz))^{24}.$$

One has $\Delta \in S_{12}(1)$, and in fact the latter space is one dimensional. Using the Eisenstein series $E_k \in M_k(1)$, there is another expression
$$\Delta = (60E_4)^3 - 27(140E_6)^2.$$

In terms of elliptic curves, if $E \simeq \mathbf{C}/(\mathbf{Z} \oplus z\mathbf{Z})$, $\Delta(z)$ is (up to a non-zero constant) the discriminant of the curve E.

---

[6] For r odd, the theory leads to half-integral modular forms, which we will not describe for lack of space and time.

3.4. Fourier expansions and L-series

3.4.1. Definition. We have already mentioned that modular forms are in particular 1-periodic on $\mathbf{H}$ so that a Fourier expansion of the type

(3.8) $$f(z) = \sum_{n \geqslant 0} a_f(n)e(nz) \text{ (for a holomorphic form } f)$$

(3.9) $$f(x + iy) = \sum_{n \in \mathbf{Z}} a_f(n;y)e(nx) \text{ (for a non-holomorphic one)}$$

exists. Similar expansions, after conjugating by the scaling matrices $\sigma_{\mathfrak{a}}$, hold at every cusp of the group considered.

In the case of (3.9), if f also satisfies $\Delta f = \lambda f$, applying Δ on both sides yields (separation of variable!) an ordinary differential equation of order 2 satisfied by $a_f(n;y)$, namely

$$y^2 w'' + (\lambda - 4\pi^2 n^2 y^2)w = 0,$$

which is of Bessel type. If $n \neq 0$, two linearly independent solutions are

$$w_1(y) = \sqrt{y}K_{s-1/2}(2\pi|n|y)$$
$$w_2(y) = \sqrt{y}I_{s-1/2}(2\pi|n|y)$$

(writing $\lambda = s(1 - s)$ with $s \in \mathbf{C}$), where K_s and I_s denote standard Bessel functions (see e.g. [**I2**, App. B-4]). For $n = 0$, two solutions are y^s and y^{1-s} (or $y^{1/2}$ and $y^{1/2}\log y$ for $s = 1/2$).

The important fact is that the asymptotics of I_s and K_s at infinity are different: $w_1(y) \sim \frac{\pi}{2}e^{-2\pi|n|y}$ and $w_2(y) \sim \frac{1}{2\pi}e^{2\pi|n|y}$ (there is no typo, but the legacy of wildly inconsistent normalizations). For a Maass form, of moderate growth, the second solution can therefore not appear for $n \neq 0$, and we write the Fourier expansion (at ∞) of f in the following normalized way:

(3.10) $$f(z) = a_f(0)y^s + b_f(0)y^{1-s} + \sqrt{y}\sum_{n \neq 0} a_f(n)K_{s-1/2}(2\pi|n|y)e(nx),$$

with $a_f(n) \in \mathbf{C}$, $b_f(0) \in \mathbf{C}$. This represents a form cuspidal at ∞ if and only if $a_f(0) = b_f(0) = 0$.

In the case of theta functions, the Fourier coefficients at infinity are related to representations of integers by quadratic forms and are clearly of arithmetic significance. It is however not obvious at all that for more general forms there should be some interest in the Fourier coefficients, or that they should have special properties. It turns out however that this is the case, and this can be at least partly revealed through L-functions in a way closely connected to the use of theta functions in the study of Hecke L-functions.

3.4.2. Examples; order of magnitude. But before going in this direction, we mention some explicit computations of Fourier coefficients.

EXAMPLE 3.4.1. Assume $k \geqslant 3$. Let $m \geqslant 0$ be an integer. The n-th Fourier coefficient of the Poincaré series $P_m(z)$ (at ∞) is given for $n \geqslant 1$ by

(3.11) $$p(m,n) = \left(\frac{m}{n}\right)^{(k-1)/2}\left\{\delta(m,n) + 2\pi i^{-k}\sum_{q|c}c^{-1}S_{\overline{\chi}}(m,n;c)J_{k-1}\left(\frac{4\pi\sqrt{mn}}{c}\right)\right\}$$

(see e.g. [**I1**, 3.2] for a proof), where J_{k-1} is a Bessel function (again) and $S_{\overline{\chi}}(m, n; c)$ a *twisted Kloosterman sum*

$$S_{\overline{\chi}}(m, n; c) = \sum_{x \,(\mathrm{mod}\, c)}^{*} \overline{\chi}(x) e\left(\frac{mx + n\bar{x}}{c}\right),$$

the sum being over invertible elements and $\bar{x}$ the inverse of x modulo c. The 0-th Fourier coefficient turns out to vanish, as well as those at the other cusps, showing that P_m is a cusp form.

The Bessel function may look strange and unfamiliar, but one has

$$J_{k-1}(y) = O(y^{k-1}) \text{ as } y \to 0,$$

(from its power series expansion) showing the convergence of the series. Also, various integral expressions would reveal the fact that $J_{k-1}(y)$ is really an archimedean analogue of the Kloosterman sums (or the other way around). This is best explained, once more, in the adèlic language (see [**CP**]).

Using the Petersson inner product on $S_k(q, \chi)$

(3.12)
$$<f, g> = \int_{\Gamma_0(q) \backslash \mathbf{H}} f(z) \overline{g(z)} y^k d\mu(z)$$

(the integrand is $\Gamma_0(q)$-invariant), one shows that

$$<f, P_h> = c_k h^{1-k} a_f(h)$$

for some (explicit) constant $c_k > 0$. It follows that the Poincaré series span $S_k(q, \chi)$. However, the relations they satisfy are quite mysterious.

EXAMPLE 3.4.2. The Fourier expansion at ∞ of the Eisenstein series of weight k (even) for $SL(2, \mathbf{Z})$ is given by

(3.13)
$$E_k(z) = 2\zeta(k) + \frac{2(2i\pi)^k}{(k-1)!} \sum_{n \geqslant 1} \sigma_{k-1}(n) e(nz)$$

with

$$\sigma_k(n) = \sum_{d \mid n} d^k.$$

For the non-holomorphic Eisenstein series for $SL(2, \mathbf{Z})$ we have

$$E(z, s) = y^s + \varphi(s) y^{1-s} + 4\sqrt{y} \sum_{n \geqslant 1} \eta_{s-1/2}(n) K_{s-1/2}(2\pi ny) \cos(2\pi nx),$$

where

$$\eta_u(n) = \sum_{ab=n} \left(\frac{a}{b}\right)^u,$$

and $\varphi(s)$ is the *scattering matrix* for $SL(2, \mathbf{Z})$, reduced to a single function:

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

This expansion shows that (thanks to the properties of the zeta function) $E(z, s)$ can be analytically continued to all of $\mathbf{C}$ and satisfies

$$E(z, s) = \varphi(s) E(z, 1 - s)$$

or symmetrically, by the functional equation of $\zeta(s)$

(3.14) $$\tilde{\varphi}(s) E(z, s) = \tilde{\varphi}(1 - s) E(z, 1 - s) \text{ with } \tilde{\varphi}(s) = \pi^{-s} \Gamma(s) \zeta(2s).$$

Analogues of this are true in much greater generality; this was proved by Selberg for finite covolume subgroups of $SL(2, \mathbf{R})$, and then extended to higher rank groups by Langlands. The analytic continuation of Eisenstein series is one of the keys to the spectral theory of automorphic forms and to their applications in analytic number theory (see e.g. [**I2**, Ch. 6]).

EXAMPLE 3.4.3. The n-th Fourier coefficient of Δ at the only cusp ∞ of $SL(2, \mathbf{Z})$ is denoted $\tau(n)$; the arithmetic function τ is usually called the Ramanujan function.

Because the modular forms considered are of moderate growth, it is easily shown that the Fourier coefficients $a_f(n)$ must be of (at most) polynomial size as functions of n: for we have

$$a_f(n) = \int_x^{x+1} f(x + t + iy) e(-nt) dt$$

for any $x \in \mathbf{R}$, $y > 0$ if f is holomorphic.

For Eisenstein series, the exact Fourier expansions reveals that (up to arithmetic fluctuations, see (3.13)), $a_f(n)$ is of size about n^{k-1}.

For the same reason, the coefficients of cusp forms should be smaller since f then vanishes at infinity. Such is indeed the case and the Parseval formula applied to the bounded function $y^{k/2} f(z)$ (resp. $f(z)$ in the non-holomorphic case) proves that

(3.15) $$\sum_{n \leqslant N} |a_f(n)|^2 \ll N^k \text{ for } N \geqslant 1$$

(taking $y = N^{-1}$), resp. $\ll N$ if f is non-holomorphic.

This turns out to be close to the truth (as an average result), as the Rankin-Selberg method shows, and it also indicates that $a_f(n) \ll n^{(k-1)/2}$, on average at least (resp. $a_f(n) \ll 1$ on average).

That this is actually individually true was first conjectured by Ramanujan for the τ function, in the strikingly precise form

(3.16) $$|\tau(n)| \leqslant d(n) n^{11/2}$$

$(d(n) = \sigma_0(n)$ is the number of divisors of n).

Petersson generalized this conjecture to other modular forms (in a less precise form): one should have

(3.17) $$a_f(n) \ll_\varepsilon n^{(k-1)/2+\varepsilon}, \text{ for } n \geqslant 1,$$

for any $f \in S_k(q, \chi)$, the implied constant depending on f and $\varepsilon > 0$.

Indeed, besides the evidence on average described above, one can justify this expectation quickly by applying the "square-rooting" philosophy to the sum of Kloosterman sums in (3.11): since the Poincaré series span $S_k(q, \chi)$ (for $k > 2$ at least), one derives very strong evidence for (3.17). However, no proof has been found along those lines and it is

through a remarkable application of the Riemann Hypothesis for varieties over finite fields (involving the link between modular forms and Galois representations or "motives") that P. Deligne [**De**] proved the Ramanujan-Petersson conjecture for cusp forms of weight $\geqslant 2$, in the precise form (3.17) for the τ function, or more generally

$$|a_f(n)| \leqslant d(n)n^{(k-1)/2}$$

for the coefficients of a primitive form (see Section 3.5 below). For weight 1, the analogue was proved by Deligne and Serre.

One shouldn't neglect the approach through (3.11), however: its analogues still apply for instance in the case of Maass forms or of half-integral weight forms for which the corresponding conjectures remain open (see [**I1**, 5.3] for the half-integral case which has remarkable applications to quadratic forms in 3 variables).

3.4.3. L-series attached to modular forms. Let f be a modular form of weight k or Maass form, so that is has the Fourier expansion (3.8) or (3.10) at ∞ (one can work at the other cusps also). To study the properties of the coefficients, consider the associated L-series (it seems better to reserve the terminology "L-function" to Dirichlet series with Euler products)

(3.18) $$L(f,s) = \sum_{n \geqslant 1} a_f(n)n^{-s} \text{ (for } f \text{ holomorphic)}$$

(3.19) $$L(f,s) = \sum_{n \neq 0} a_f(n)|n|^{-s} \text{ (for Maass forms).}$$

Hecke (resp. Maass[7]) showed that those L-series still carried most analytic properties of the L-functions considered in the previous lectures. First observe that by the trivial bound $a_f(n) \ll n^k$ (resp. $a_f(n) \ll |n|$) for Fourier coefficients, the L-series converges absolutely for $\mathrm{Re}(s)$ large enough.

Let

$$W = \begin{pmatrix} 0 & -1/\sqrt{q} \\ \sqrt{q} & 0 \end{pmatrix} \in SL(2, \mathbf{R})$$

acting by $z \mapsto -1/(qz)$. This W normalizes $\Gamma_0(q)$ (it is in $SL(2, \mathbf{Z})$ for $q = 1$), so a simple computation shows that if $f \in M_k(q, \chi)$ then $g = f \mid_k W \in M_k(q, \overline{\chi})$ (resp. if $f \in M^\lambda(q, \chi)$, then $g = f \mid W \in M^\lambda(q, \overline{\chi})$).

PROPOSITION 3.4.4. *Let*

$$\Lambda(f,s) = (2\pi)^{-s}\Gamma(s)L(f,s)$$

for f holomorphic,

(3.20) $$L(f,s) = \pi^{-s}\Gamma\Big(\frac{s+it}{2}\Big)\Gamma\Big(\frac{s-it}{2}\Big)L(f,s)$$

for Maass forms with eigenvalue $\lambda = 1/4 + t^2$.

---

[7] If the Maass form f is odd, i.e. $f(-\bar{z}) = -f(z)$, this L-series vanishes; a slight modification still works which we will not discuss.

Then $\Lambda(f, s)$ extends to a meromorphic function on $\mathbf{C}$. If f is a cusp form it is entire, otherwise it has poles at $s = 1$ and $s = 0$. Moreover we have

$$\Lambda(f, s) = i^k q^{k/2-s} \Lambda(g, k - s)$$

where $g = f|_k W$ for f holomorphic or

$$\Lambda(f, s) = q^{1/2-s} \Lambda(g, 1 - s)$$

where $g = f|W$ for f a Maass form.

PROOF. Consider the holomorphic case: the Maass form case is similar, except one has to appeal to the formula for the Mellin transform of the Bessel K-functions. Also we assume that f is a cusp form to avoid dealing with the constant term (which is not hard either).

We proceed much as in proving the analytic continuation of Hecke L-functions, f replacing the theta function: one has for $\mathrm{Re}(s) > 1$

$$(2\pi)^{-s} \Gamma(s) n^{-s} = \int_0^{+\infty} e^{-2\pi n y} y^s \frac{dy}{y},$$

hence we find for $\mathrm{Re}(s)$ large enough that

(3.21) $$\Lambda(f, s) = \int_0^{\infty} f(iy) y^s \frac{dy}{y}.$$

Again we split the integral at $\alpha > 0$ to get

$$
\begin{aligned}
\Lambda(f, s) &= \int_0^{\alpha} f(iy) y^s \frac{dy}{y} + \int_{\alpha}^{\infty} f(iy) y^s \frac{dy}{y} \\
&= \int_{1/(q\alpha)}^{+\infty} f\left(\frac{i}{qu}\right) (qu)^{-s} \frac{du}{u} + \int_{\alpha}^{\infty} f(iy) y^s \frac{dy}{y} \\
&= \int_{1/(q\alpha)}^{+\infty} q^{k/2} i^k u^k g(iu) (qu)^{-s} \frac{du}{u} + \int_{\alpha}^{\infty} f(iy) y^s \frac{dy}{y}
\end{aligned}
$$

since $g(iu) = q^{-k/2} i^{-k} u^{-k} f(i/(qu))$

$$= i^k q^{k/2-s} \int_{1/(q\alpha)}^{+\infty} g(iu) u^{k-s} \frac{du}{u} + \int_{\alpha}^{\infty} f(iy) y^s \frac{dy}{y}.$$

This yields already the analytic continuation, and the functional equation follows if $\alpha = q^{-1/2}$. $\qquad\square$

Compared to the various results seen in the first two lectures about analytic continuation and functional equation, one notices that the gamma factors are different (especially for Maass forms), and the functional equation relates s to $k - s$, putting the center of the critical strip at $s = k/2$. Moreover, the functional equation relates f to $g = f|_k W$. If $q = 1$, then $W \in SL(2, \mathbf{Z})$ so that $f = g$, but in general there is no reason for such a relation to hold.

In analytic number theory, it is often convenient to renormalize the coefficients, putting

$$a_f(n) = n^{(k-1)/2} \lambda_f(n)$$

and using

$$\sum_{n \geqslant 1} \lambda_f(n) n^{-s}$$

instead of $L(f, s)$, so that the functional equation for this other series relates s to $1 - s$.

3.5. Hecke operators and applications

Proposition 3.4.4 gives us more examples of Dirichlet series with functional equations. However there doesn't seem to be any reason to have an Euler product. Yet, Ramanujan also conjectured that the τ function is multiplicative (i.e. $\tau(mn) = \tau(m)\tau(n)$ if $(m, n) = 1$) and more precisely that

$$(3.22) \qquad L(\Delta, s) = \sum_{n \geqslant 1} \tau(n) n^{-s} = \prod_p (1 - \tau(p) p^{-s} + p^{11-2s})^{-1}.$$

Similarly, we know that the Fourier coefficient of $\theta(2z)^2$ is $4r(n)$ where $r(n)$ is the coefficient of the Dedekind zeta function of $\mathbf{Q}(i)$, so

$$\sum_{n \geqslant 1} r(n) n^{-s} = \zeta(s) L(\chi_4, s) = \prod_p (1 - p^{-s})^{-1} (1 - \chi_4(p) p^{-s})^{-1}$$

$$(3.23) \qquad\qquad = \prod_p (1 - (1 + \chi_4(p)) p^{-s} + \chi_4(p) p^{-2s})^{-1}.$$

Notice in both cases the denominator is of degree 2 in p^{-s} (except for $p = 2$ in the second case, when $\chi_4(2) = 0$).

Mordell first proved Ramanujan's multiplicativity conjecture, but it was Hecke who did most to create a coherent theory, revealing a really remarkable arithmetic structure for the Fourier coefficients of holomorphic forms. However, he couldn't obtain a completely satisfactory answer to some problems and it was not until Atkin and Lehner developped the theory of "newforms" that the situation got really clarified. The adèlic theory of Jacquet-Langlands also throws much light on these matters.

Hecke's idea is to obtain "good" arithmetic modular forms by finding an algebra acting on spaces of modular forms in such a way that it is diagonalizable: the eigenfunctions then inherit much of the structural property of the algebra. This had no "classical" counterpart for Dirichlet characters, although analogues can be constructed *a posteriori*.

The *Hecke operators* can be defined in a number of ways. We mention a few:

- One can define an abstract algebra generated by *double cosets* $\Gamma\gamma\Gamma$ for some γ in the *commensurator* of Γ. For congruence subgroups, the latter is larger than Γ. Then this algebra is shown to act on $M_k(q, \chi)$ in an appropriate way, essentially

$$f \,|_k \Gamma\gamma\Gamma = \sum_{\alpha_i} f \,|_k \alpha_i$$

where $\Gamma\gamma\Gamma = \cup_i \Gamma\alpha_i$ is a decomposition of the double coset. By local-global principles (the Chinese Remainder Theorem), the algebra is shown to admit generators

$$T_p = \Gamma \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} \Gamma \text{ for prime } p$$

$$R_p = \Gamma \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix} \Gamma \text{ for prime } p$$

and those satisfy simple multiplicativity relations (see below). See [**Sh**] for a detailed study based on this approach; to study the interactions between Hecke operators defined on various subgroups, it is often the most precise.

- If one sees modular forms as functions on the space of lattices, one can define the Hecke operator T_p as follows:

$$(f \mid_k T_p)(\Lambda) = \sum_{[\Lambda:\Lambda']=p} f(\Lambda'),$$

where the sum is over all sublattices of index p in Λ (see [**Se1**], VII-5]).

- To get a quick definition, although not a very practical one for proving the properties of the operators, one can just give the action on the Fourier coefficients, using the fact that the expansion at ∞ suffices to recover modular forms. We'll take this approach for simplicity.

Fix q, the weight k and the character χ. For p prime we let

$$T(p) = \frac{1}{p} \sum_{b \, (\mathrm{mod} \, p)} \begin{pmatrix} 1 & b \\ 0 & p \end{pmatrix} + \chi(p)p^{k-1} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$$

(in the rational group ring of $GL(2, \mathbf{R})^+$). For p dividing q, notice that the second term vanishes (even for the trivial character: χ is seen as defined modulo q). When dealing with Maass forms, a different definition must be taken, where the factor p^{k-1} is removed; the same change applies to the various formulae below.

For a given weight k, we let these $T(p)$ act on 1-periodic functions on $\mathbf{H}$ in the obvious way (with the $\mid_k$ action or the $\mid$ action), and we denote this $f \mid_k T(p)$ or $f \mid T(p)$. (Notice $T(p)$ still depends on χ and q).

LEMMA 3.5.1. *The above operator $T(p)$ on 1-periodic function induces an endomorphism on $M_k(q, \chi)$, resp. on $M^\lambda(q, \chi)$, and preserves cusp forms. If $a_f(n)$ are the Fourier coefficients of f, then the n-th Fourier coefficient of $f \mid_k T(p)$ is equal to*

$$a_f(n/p) + \chi(p)p^{k-1}a_f(np)$$

with the convention $a_f(x) = 0$ for $x \in \mathbf{Q}$ not an integer.[8]

The first part of this lemma is of course the crucial assertion and can not be proved without, in effect, relating the ad-hoc definition of $T(p)$ we have given to one of the more intrinsic definitions. The second part, on the other hand, is quite simple (orthogonality of the characters of $\mathbf{Z}/p\mathbf{Z}$).

---

[8] As previously this formula is for holomorphic forms and must be changed (remove p^{k-1}) for Maass forms.

Now we extend the definition to define $T(n)$, $n \geqslant 1$ by multiplicativity and induction

$$T(nm) = T(n)T(m) \text{ for } (n,m) = 1$$
$$T(p^{i+1}) = T(p)T(p^i) + \chi(p)p^{k-1}T(p^{i-1}), \text{ for } i \geqslant 0$$
$$T(1) = 1.$$

A simple computation shows that the m-th Fourier coefficient of $f|_k T(n)$ is

(3.24)
$$\sum_{d|(n,m)} \chi(d)d^{k-1}a_f\left(\frac{nm}{d^2}\right),$$

in particular the first Fourier coefficient of $f|_k T(n)$ is equal to $a_f(n)$, the n-th coefficient of f and the 0-th Fourier coefficient is $\sigma_{k-1}(n)a_f(0)$.

LEMMA 3.5.2. *The subalgebra of* $\mathrm{End}\, M_k(q,\chi)$ *generated by the Hecke operators* $T(p)$, *or* $T(n)$, *is commutative and one has*

$$\sum_{n \geqslant 1} T(n)n^{-s} = \prod_p (1 - T(p)p^{-s} + \chi(p)p^{k-1-2s})^{-1}.$$

Compare the shape of this Euler product with that conjectured by Ramanujan for the τ function (3.22) and also with that in (3.23). In fact:

COROLLARY 3.5.3. *The Euler product (3.22) holds for* $\mathrm{Re}(s)$ *large enough.*

PROOF. Consider the Hecke operators for weight 12, level 1 and $\chi = 1$. One has $\dim S_{12}(1) = 1$, so $S_{12}(1) = \Delta\mathbf{C}$. It follows that $\Delta|_{12} T(n) = \lambda(n)\Delta$ for any n. By the lemma, we have tautologically

$$\sum_{n \geqslant 1} \lambda(n)n^{-s} = \prod_p (1 - \lambda(p)p^{-s} + p^{11-2s})^{-1}.$$

Comparing the first Fourier coefficient using (3.24), one finds

$$\tau(n) = \lambda(n)\tau(1) = \lambda(n)$$

hence the result. $\qquad\qquad\square$

For the same reason, there will be an Euler product for any f which is an eigenfunction of all Hecke operators $T(n)$. But do these exist? When the weight or the level is large there are many linearly independent holomorphic modular forms so the simple argument of the corollary can not extend much. However Hecke proved:

LEMMA 3.5.4. *Let* $n \geqslant 1$ *be coprime with* q. *Then the operator* $T(n)$ *acting on* $S_k(q,\chi)$, *resp.* $S^\lambda(q,\chi)$, *is* normal *with respect to the Petersson inner product (3.12), resp. with respect to the inner product in* $L^2(\Gamma_0(q)\backslash\mathbf{H})$, *in fact the adjoint of* $T(n)$ *is*

$$T(n)^* = \overline{\chi}(n)T(n).$$

The proof of this lemma is actually quite subtle. Most importantly, the condition $(n,q) = 1$ is necessary: the operators $T(p)$ for $p \mid n$ are usually *not* normal. In any case Hecke could deduce:

COROLLARY 3.5.5. *There is an orthonormal basis of $S_k(q, \chi)$, resp. $S^\lambda(q, \chi)$, consisting of forms f which are eigenfunctions of all Hecke operators $T(n)$ with $(n, q) = 1$. Such a modular form is called a* Hecke form, *and we denote by $\lambda(n)$ its Hecke-eigenvalues.*

For a Hecke form f, we derive from the equation

$$f\,|_k\,T(n) = \lambda_f(n)f \text{ for } (n, q) = 1$$

that $a_f(n) = \lambda_f(n)a_f(1)$ for $(n, q) = 1$. If $a_f(1) \neq 0$, we deduce from the Euler product for the Hecke operators that the form $g = a_f(1)^{-1}f$ is such that

$$L_q(g, s) = \sum_{(n,q)=1} a_g(n)n^{-s} = \sum_{(n,q)=1} \lambda_f(n)n^{-s} = \prod_{p \nmid q} (1 - \lambda_f(p)p^{-s} + \chi(p)p^{k-1-2s})^{-1}.$$

Two difficulties remain: one might have $a_f(1) = 0$, in which case the reasoning breaks down, and we would prefer an Euler product involving all primes for $L(f, s)$ itself.

Hecke was able to show that in some cases, a basis consisting of eigenfunctions of all the $T(n)$ existed, in which case the L-functions of forms in this basis would have an Euler product (after normalizing the first coefficient to be 1). In particular if the character χ is primitive modulo q, this is so, but this excludes the important case $\chi = 1$ for $q > 1$. Hecke's idea was to prove that there is "multiplicity one" for the Hecke algebra: the space of modular forms with given eigenvalues $\lambda(n)$, $(n, q) = 1$, is at most 1-dimensional. If this is true, then since Hecke operators $T(p)$ with $p \mid q$ act on this common eigenspace (they commute with all the other Hecke operators), it would follow that $f\,|_k\,T(p)$ is a multiple of f for any p, and the theory would proceed as before. However, multiplicity one fails in general.

Atkin and Lehner [**AL**] showed how to correct Hecke's theory by introducing an analogue of the notion of primitive character. This is reasonable since we know that only primitive Dirichlet characters satisfy a nice functional equation.

The starting point of the theory is the following way of inducing modular forms to a higher level: let $q \geqslant 1$, χ a character modulo q and $a > 1$. Let $f \in S_k(q, \chi)$. Then for any integer $d \mid a$, the function

$$g_d(z) = f(dz)$$

is an element of $S_k(aq, \chi')$, χ' being induced modulo aq from χ, and if f is a Hecke form then g_d is a Hecke form *with the same eigenvalues for $(n, aq) = 1$*. This result is very simple to check. In particular, f and g_d provide an example showing the failure of multiplicity one. (Note also that

$$g_d(z) = \sum_{n \equiv 0 \,(\mathrm{mod}\,d)} a_f(n/d)e(nz),$$

so the first Fourier coefficient of g_d is $= 0$ if $d > 1$.) This example motivates the following:

DEFINITION. Let $q \geqslant 1$, χ a character modulo q and k an integer. Let $q' \mid q$ be the conductor of χ, χ' the primitive character inducing χ. The *old space* of $S_k(q, \chi)$, denoted $S_k(q, \chi)^\flat$, is the subspace spanned by all functions of the form $f(az)$ where
(i) $f \in S_k(q'', \chi')$ for some q'' such that $q' \mid q'' \mid q$, $q'' \neq q$;
(ii) $a \mid q/q''$.
The *new space* $S_k(q, \chi)^*$, is the orthogonal of the old-space in $S_k(q, \chi)$ for the Petersson inner product.

EXAMPLE 3.5.6. (i) If $q = 1$ or if χ is primitive, then $S_k(q, \chi) = S_k(q, \chi)^*$.

(ii) If $k < 12$ and $q = p$ is prime, then since $S_k(1) = 0$, it follows that $S_k(p)^* = S_k(p)$. If $k \geqslant 12$, then

$$S_k(p) = S_k(1) \oplus pS_k(1) \oplus S_k(p)^*,$$

denoting $pS_k(1)$ the space of cusp forms of type $f(pz)$ with $f \in S_k(1)$. In general, Möbius inversion can be used to compute the dimension of the new-space.

One shows easily that all the Hecke operators $T(n)$ with $(n, q) = 1$ act on the old-space and the new-space.

THEOREM 3.5.7 (Atkin-Lehner). *The multiplicity one principle holds for $S_k(q, \chi)^*$, i.e. if $(\lambda(n))$ is any sequence of complex numbers defined for $(n, q) = 1$, the space of $f \in S_k(q, \chi)^*$ such that*

$$f \,|_k\, T(n) = \lambda(n)f \ \ for \ (n, q) = 1$$

is at most 1-dimensional.

There is a unique orthogonal basis of $S_k(q, \chi)^*$ made of primitive *forms (also called* newforms*), i.e. forms f such that $a_f(1) = 1$ and $f \,|_k\, T(n) = a_f(n)f$ for any $n \geqslant 1$.*

The first part is the crucial statement and the second part follows readily following the sketch above.

COROLLARY 3.5.8. *Let f be a primitive form. The L-function $L(f, s)$ has an Euler product*

(3.25) $$L(f, s) = \prod_p (1 - a_f(p)p^{-s} + \chi(p)p^{k-1-2s})^{-1}.$$

Moreover there exists $\eta(f) \in \mathbf{C}$ with modulus 1 such that $f \,|_k\, W = \eta(f)\bar{f} \in S_k(q, \bar{\chi})^*$, where

$$\bar{f}(z) = \sum_{n \geqslant 1} \overline{a_f(n)}e(nz),$$

hence $L(f, s)$ satisfies the functional equation

$$\Lambda(f, s) = \varepsilon(f)q^{k/2-s}\Lambda(\bar{f}, k - s)$$

with $\varepsilon(f) = i^k\eta(f)$.

The functional equation is now perfectly analogous to that (1.9) for primitive Dirichlet characters. The argument $\varepsilon(f)$ is a very subtle invariant for which there is no simple formula in general. If q is squarefree, then one can relate $\eta(f)$ to the q-th Fourier coefficient of f. If $q = 1$, for example, $f \,|_k\, W = f$ and the sign of the functional equation is i^k.

It also follows from this that in fact for any Hecke form f of level q, there is a unique primitive form g, of some lower level in general, with the same eigenvalues for the operators $T(n)$, $(n, q) = 1$. Hence it is indeed possible to use the primitive forms to gain information about the whole space of cusp forms, using primitive forms of lower levels if necessary.

Also, one may find that the original definition of a primitive form, involving as it does that $f \in S_k(q, \chi)^*$, which is defined "negatively", is not very convenient. However, W. Li has shown that the converse of the corollary holds: if f is such that $f \,|_k\, W = \eta f$ and $L(f, s)$ has an Euler product (it automatically also satisfies the required functional equation), then

f is a primitive form. This agrees perfectly with the larger "automorphic" philosophy that "good" L-functions are associated to the correct "primitive" objects.

3.6. Other "openings"

We close this lecture as we opened it, with three brief remarks about further topics in this classical theory which have had a great influence on the evolution of the subject of L-functions and automorphic forms, and that will be presented, more or less transmogrified, in the other lectures.

3.6.1. Converse theorem. The simple steps leading to the proof of Proposition 3.4.4 can be obviously reversed and lead to a criterion for two 1-periodic holomorphic functions $f(z)$ and $g(z)$ to be related by the Fricke involution, $g = f \mid_k W$ (similarly for Maass forms with f and g having Fourier expansions of type (3.10)).

For level 1, where $W \in SL(2, \mathbf{Z})$, it is also known that the matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

generate $SL(2, \mathbf{Z})$, hence one deduces a *characterization* of modular forms of level 1 from the functional equation of their L-functions.

One would like to have a similar characterization for higher levels, but the group theory is not so simple. It was Weil who found the correct generalization, using the functional equations not only of the L-function of f itself, but of the *twists* of f by Dirichlet characters. Roughly speaking, one wants information not only about the L-series of f, but also about the Dirichlet series generated by the coefficients $a_f(n)$ of f in a given congruence class (although those are seen "dually" using Dirichlet characters).

We discuss the holomorphic case: the non-holomorphic one is similar, with functional equations of the type (3.20) to "specify" the eigenvalue.

Let $f \in S_k(q, \chi)$ for instance, and let ψ be a Dirichlet character modulo m. One defines

$$(f \otimes \psi)(z) = \sum_{n \geqslant 1} a_f(n)\psi(n)e(nz).$$

Decomposing ψ into additive characters (which introduces the Gauss sum of ψ), one shows that this is again a modular form, precisely $f \otimes \psi \in S_k(qm^2, \chi\psi^2)$. If $(q, m) = 1$, computing $(f \otimes \psi) \mid_k W$ yields the following functional equation for $L(f \otimes \psi, s)$, where $g = f \mid_k W$:

$$\Lambda(f \otimes \psi, s) = \varepsilon(f \otimes \psi)(qm^2)^{k/2-s}\Lambda(g \otimes \overline{\psi}, k - s),$$

with

(3.26) $$\varepsilon(f \otimes \psi) = i^k \chi(m)\psi(q)\tau(\psi)^2 r^{-1}.$$

Weil's theorem says that having such functional equations for "enough" twists ensures that a Dirichlet series comes from a modular form. For simplicity we state the version for cusp forms. For a proof, see e.g. [**I1**, 7.4].

THEOREM 3.6.1. *Let*

$$L_1(s) = \sum_{n \geqslant 1} a(n)n^{-s} \text{ and } L_2(s) = \sum_{n \geqslant 1} b(n)n^{-s}$$

be two Dirichlet series absolutely convergent for $\mathrm{Re}(s) > C$ for some $C > 0$.

Assume that there exists integers $k \geqslant 1$, $q \geqslant 1$ and $M > 0$ such that for any ψ primitive modulo m, with $(m, Mq) = 1$, the Dirichlet series

$$L(f \otimes \psi, s) = \sum_{n \geqslant 1} a(n)\psi(n)n^{-s} \ \text{and} \ L(g \otimes \psi, s) = \sum_{n \geqslant 1} b(n)\psi(n)n^{-s}$$

admit analytic continuation to entire functions bounded in vertical strips such that the functions

$$\Lambda(f \otimes \psi, s) = (2\pi)^{-s}\Gamma(s)L(f \otimes \psi, s) \ \text{and} \ \Lambda(g \otimes \psi, s) = (2\pi)^{-s}\Gamma(s)L(g \otimes \psi, s)$$

are entire and satisfy the functional equation

$$\Lambda(f \otimes \psi, s) = \varepsilon(f \otimes \psi)(qm^2)^{k/2-s}\Lambda(g \otimes \overline{\psi}, k - s),$$

with $\varepsilon(f \otimes \psi)$ given by (3.26).

Then there exists a cusp form $f \in S_k(q, \chi)$ for some χ modulo q such that $L(f, s) = L_1(s)$ and $L(f|_k W, s) = L_2(s)$.

Generalizations of this converse theorem have a long history and have been of great importance in many developpments of automorphic forms and for the functoriality conjectures of Langlands in particular: one may mention the construction of the symmetric square of $GL(2)$-forms by Gelbart and Jacquet. The more recent results of Cogdell and Piatetski-Shapiro are also used in Lafforgue's proof of the Global Langlands Correspondance for $GL(n)$ over function fields.

3.6.2. Rankin-Selberg L-function. Let f be a weight k cusp form of level q. We have mentioned in (3.15) that the upper bound

$$\sum_{n \leqslant N} |a_f(n)|^2 \ll N^k \ \text{for} \ N \geqslant 1,$$

is quite easy to obtain. Rankin and Selberg independently proved that it is indeed sharp; as for the proof of the Euler product (3.22), their method had considerable influence on the later developpement of the theory of L-functions and automorphic forms.

Given two modular forms f and g (of arbitrary weights and levels, one or both being possibly non-holomorphic), define first the (naive) *Rankin-Selberg L-function*

$$L(f \times \bar{g}, s) = \sum_{n \geqslant 1} a_f(n)\overline{a_g(n)}n^{-s},$$

so that $L(f \times \bar{f}, s)$ has $|a_f(n)|^2$ as coefficients.

By the polynomial bound for Fourier coefficients, the series converges for $\mathrm{Re}(s)$ large enough. The upshot of the work of Rankin and Selberg is that this new type of L-function satisfies much of the good analytic properties of Hecke L-functions (abelian) and automorphic L-functions: namely they possess analytic continuation, functional equations, and in privileged cases, an Euler product expansion. We state here only the simplest case:

PROPOSITION 3.6.2. *Let $f \in S_k(1)$ and $g \in S_\ell(1)$ with $k \geqslant 2$, $\ell \geqslant 2$ be two holomorphic modular forms. In addition to $L(f \times \bar{g}, s)$ define*

$$L(f \otimes \bar{g}, s) = \zeta(2s - (k + \ell) + 2)L(f \times \bar{g}, s)$$

and

$$\Lambda(f \otimes \bar{g}, s) = (2\pi)^{-2s}\Gamma(s)\Gamma(s - (k + \ell)/2 + 1)L(f \otimes \bar{g}, s).$$

Then $\Lambda(f \otimes \bar{g}, s)$ admits analytic continuation to $\mathbf{C}$, entire if $f \neq g$ and with simple poles at $s = 0$ and $s = k + \ell$ otherwise, satisfying the functional equation

$$\Lambda(f \otimes \bar{g}, s) = \Lambda(g \otimes \bar{f}, k + \ell - s).$$

Hence $L(f \otimes \bar{g}, s)$ is entire if $f \neq g$ and has a simple pole at $s = k = \ell$ if $f = g$. The residue is given by

$$\operatorname{Res}_{s=k} L(f \times \bar{f}, s) = \frac{3}{\pi}\frac{(4\pi)^k}{\Gamma(k)}<f, f>.$$

SKETCH OF PROOF. We consider the simplest case $k = \ell$. The proof depends crucially on the non-holomorphic Eisenstein series $E(z; s)$ defined in (3.7): indeed the key is the following *Rankin-Selberg integral representation*

$$(4\pi)^{-s}\Gamma(s)L(f \times g, s) = \int_{\Gamma_0(1)\backslash\mathbf{H}} f(z)\overline{g(z)}y^k E(z; s - k + 1)d\mu(z)$$

involving the non-holomorphic Eisenstein series (for $k \neq \ell$, and in more general situations, one has to use a different Eisenstein series; see [**I1**, Ch. 13] for instance). To prove this, notice that

$$(4\pi)^{-s}\Gamma(s)n^{-s} = \int_0^{+\infty} e^{-4\pi ny}y^s\frac{dy}{y}$$

and

$$\int_0^1 f(z)\overline{g(z)}dx = \sum_{n \geqslant 1} a_f(n)\overline{a_g(n)}e^{-4\pi ny},$$

hence since $]0, 1[\times]0, +\infty[\subset \mathbf{R}^2$ is a fundamental domain for the stabilizer Γ_∞ of $+\infty$ in $SL(2, \mathbf{Z})$, we obtain

$$(4\pi)^{-s}\Gamma(s)L(f \times \bar{g}, s) = \int_{\Gamma_\infty\backslash\mathbf{H}} f(z)\overline{g(z)}y^{s-1}dxdy$$

$$= \sum_{\gamma \in \Gamma_\infty\backslash SL(2,\mathbf{Z})} \int_{SL(2,\mathbf{Z})\backslash\mathbf{H}} f(z)\overline{g(z)}y^k(\operatorname{Im}\gamma z)^{s-k+1}d\mu(z),$$

using the fact that $z \mapsto f(z)\overline{g(z)}y^k$ is an $SL(2, \mathbf{Z})$-invariant function. Exchanging the order of summation, the integral formula follows from the definition of $E(z; s)$.

Next multiply both sides by $\pi^{-s}\Gamma(s - k + 1)\zeta(2s - 2k + 2)$, finding $\Lambda(f \otimes \bar{g}, s)$ on the left-hand side. On the right-hand side, the functional equation (3.14) of $E(z; s)$ means that we obtain an expression invariant under $s \mapsto k - s$, hence the functional equation for $\Lambda(f \otimes \bar{g}, s)$.

Since $E(z, s)$ has a simple pole at $s = 1$ with residue equal to $1/\operatorname{Vol}(SL(2, \mathbf{Z})\backslash\mathbf{H}) = 3/\pi$, the other statements are also consequences of this integral representations.

Note in particular that it is indeed $L(f \otimes \bar{g}, s)$ which has "good" analytic properties: the simpler $L(f \times \bar{g}, s)$ is indeed meromorphic on $\mathbf{C}$ but it has infinitely many poles at the points $s = \rho/2 + k + 1$ where ρ is a zero of the Riemann zeta function. □

COROLLARY 3.6.3. *Let $f \in S_k(1)$ be a modular form. We have*

$$\sum_{n \leqslant X} |a_f(n)|^2 \sim c(f)X^k$$

as $X \to +\infty$, where

$$c(f) = \mathrm{Res}_{s=k} L(f \times \bar{f}, s).$$

PROOF. Apply the standard methods of contour integration to $L(f \times \bar{f}, s)$: the simple pole at $s = k$ gives the leading term. $\square$

REMARK 3.6.4. In addition to the above, Selberg observed that if f and g have multiplicative Fourier coefficients, then the coefficients of $L(f \times g, s)$ are also multiplicative, hence this L-function also has an Euler product expansion. If $f \in S_k(1)$ and $g \in S_\ell(1)$ are primitive, write (using (3.25))

$$a_f(p^k) = \frac{\alpha_1^{k+1} - \alpha_2^{k+1}}{\alpha_1 - \alpha_2}, \text{ and } a_g(p^k) = \frac{\beta_1^{k+1} - \beta_2^{k+1}}{\beta_1 - \beta_2}$$

where (α_1, α_2) (resp. (β_1, β_2)) are the roots of the quadratic equation

$$1 - a_f(p)X + p^{k-1}X^2 = 0 \text{ (resp. } 1 - a_g(p)X + p^{\ell-1}X^2 = 0).$$

Then one finds that

$$\sum_{k \geqslant 0} a_f(p^k)a_g(p^k)X^k == \frac{1 - p^{\ell+k-2}X^2}{(1 - \alpha_1\beta_1 X)(1 - \alpha_1\beta_2 X)(1 - \alpha_2\beta_1 X)(1 - \alpha_2\beta_2 X)}$$

hence $L(f \otimes g, s)$ has the following Euler product expansion valid for $\mathrm{Re}(s)$ large enough

$$L(f \otimes g, s) = \prod_p \left((1 - \alpha_{1,p}\beta_{1,p}p^{-s})(1 - \alpha_{1,p}\beta_{2,p}p^{-s})(1 - \alpha_{2,p}\beta_{1,p}p^{-s})(1 - \alpha_{2,p}\beta_{2,p}p^{-s}) \right)^{-1}.$$

In the general case, if f and g are primitive forms, the same formal computation shows that $L(f \otimes g, s)$ has an Euler product expansion, with all but finitely many Euler factors of the same form. However, the functional equation becomes quite a delicate matter if one proceeds classically: since it should be inherited from the Eisenstein series, and for these, whenever the group involved has more than one cusp, the functional equation takes a matrix form, involving the Eisenstein series at all cusps. In the adèlic language, the results are again much cleaner and the proofs conceptually probably simpler.

REMARK 3.6.5. In keeping with the analogy with abelian L-functions, one should think of $L(f \otimes g, s)$ as the analogue of the L-function $L(\chi_1\chi_2, s)$ for χ_i some Dirichlet characters: of course in the latter case, this remains an abelian L-function since $1 \times 1 = 1$. In particular one should think of $L(f \otimes g, s)$ has giving a way of measuring the "distance" between two modular forms. That the order of the pole at $s = 1$ gives a way of deciding whether $f = \bar{g}$ or not is very useful, as is the fact that the residue at $s = k$ is related to the Petersson norm of f (this is the way the Petersson norm of primitive forms can often be best studied, for instance).

A difference is that the notation $f \otimes g$ is rather formal. It is only recently that Ramakrishnan has proved that the Rankin-Selberg L-functions are also attached to some automorphic object, namely an automorphic form on $GL(2) \times GL(2)$, as predicted by the Langlands functoriality conjectures.

3.6.3. Theta functions and quadratic fields, revisited. It has already been mentioned that the kind of theta functions used to prove the analytic properties of abelian L-functions are themselves modular forms. Hecke and later Maass considered a particularly interesting special case which can be seen as a first instance of relating automorphic forms on $GL(1)$ over a quadratic field K to automorphic forms on $GL(2)$ over $\mathbf{Q}$. The result can be stated as follows, in the special case of class group characters (i.e. trivial weight and modulus $\mathfrak{m} = \mathcal{O}$):

PROPOSITION 3.6.6. *Let $K/\mathbf{Q}$ be a quadratic extension with discriminant D. Let χ be a character of the ideal class group of K. There exists a primitive modular form f_χ of level $|D|$ with nebentypus ε_D equal to the quadratic character associated to K, which is holomorphic of weight 1 if $D < 0$ and a Maass form with eigenvalue $\lambda = 1/4$ if $D > 0$, such that*

$$L(f_\chi, s) = L(\chi, s),$$

and this holds even locally for every p-factor. Moreover, f_χ is a cusp form if and only if the character χ is not real.

One can construct f as a theta function with character, given by

$$f(z) = \sum_{\mathfrak{a} \neq 0} \chi(\mathfrak{a}) e(zN\mathfrak{a})$$

(if χ is not real and K is imaginary).

Or, one can easily see that the functional equation of the L-function of χ is indeed compatible with the existence of f. Expanding

$$L(\chi, s) = \sum_{n \geqslant 1} a_\chi(n) n^{-s}$$

one can then apply Weil's converse theorem: the functional equations of abelian L-functions over K will show that the hypothesis is satisfied. The primitivity of f_χ is easy to derive since by construction $L(f_\chi, s)$ has an Euler product.

In the case of real fields, this gives essentially the only known explicit examples of Maass forms. Notice they have algebraic Fourier coefficients. It is conjectured that all primitive Maass forms with eigenvalue $\lambda = 1/4$ are similarly algebraic (arising from even Artin representations of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ into $GL(2, \mathbf{C})$; those which dihedral image in $PGL(2, \mathbf{C})$ should correspond to real quadratic fields as above.)

A similar argument can be applied with characters of other weights, yielding in particular the analytic continuation and functional equation of the Hasse-Weil zeta functions of elliptic curves *with complex multiplication* over number fields (Deuring, Weil; see e.g. [**I1**, 8] for a specific case in complete details).

Bibliography

[AL]   Atkin, A. and Lehner, J.: *Hecke operators on Gamma$_0$(m)*, Math. Ann. 185 (1970), 134–160.

[Bu]   Bump, D.: *Automorphic forms and representations*, Cambridge Univ. Press, 1996.

[CF]   Cassels, J.W.S. and Fröhlich, A.: *Algebraic number theory*, Academic Press, 1967.

[CP]   Cogdell, J. and Piatetski-Shapiro, I.: *The arithmetic and spectral analysis of Poincaré series*, Perspectives in Mathematics, 13, Acad. Press, (1990).

[Co]   Cohen, H.: *Advanced Topics in Computational Number Theory*, GTM 193, Springer-Verlag 2000.

[CS]   Conrey, B. and Soundararajan, K.: *Real zeros of quadratic Dirichlet L-functions*, AIM Preprint 2001-26.

[De]   Deligne, P.: *Les conjectures de Weil, I*, Inst. Hautes Études Sci. Publ. Math. No. 43 (1974), 273–30.

[Di]   Dirichlet, P.G.: *Beweis des Satzes, dass jede unbegrenzte arithmetische Progression, deren erstes Glied und Differenz ganze Zahlen ohne gemeinschaftlichen Factor sind, unendlich viele Primzahlen enthält*, in Collected Works, vol. I, 313–342, Chelsea 1969.

[Du]   Duke, W.: *The dimension of the space of cusp forms of weight one*, Internat. Math. Res. Notices 2 (1995), 99–109.

[Ge]   Gelbart, S.: *Automorphic forms on adèle groups*, Annals of Math. Studies 83, Princeton Univ. Press (1975).

[Go]   Goldfeld, D.: *The class number of quadratic fields and the conjectures of Birch and Swinnerton-Dyer*, Ann. Scuola Norm. Sup. Pisa 3, 4 (1976), 623–663.

[I1]   Iwaniec, H.: *Topics in classical automorphic forms*, Grad. Studies in Math. 17, A.M.S 1997.

[I2]   Iwaniec, H.: *Introduction to the spectral theory of automorphic forms*, Biblioteca de la Rev. Mat. Iberoamericana, 1995.

[IR]   Ireland, K. and Rosen, M.: *A classical introduction to modern number theory*, Second edition, GTM 84, Springer Verlag (1990).

[KS]   Katz, N. and Sarnak, P.: *Random matrices, Frobenius eigenvalues, and monodromy*, Colloquium Publications, A.M.S (1998).

[L]   Landau, E.: *Bemerkungen zum Heilbronnschen Satz*, Acta Arithmetica 1 (1936), 1–18.

[La]   Lang, S.: *Algebraic number theory*, GTM 110, Springer-Verlag 1986.

[LRS]   Luo, W., Rudnick, Z. and Sarnak, P.: *On Selberg's eigenvalue conjecture*, Geometric and Functional Analysis 5 (1994), 387–401.

[Mi]   Miyake, T.: *Modular forms*, Springer-Verlag 1989.

[PS]   Phillips, R. and Sarnak, P.: *On cusp forms for cofinite subgroups of $PSL(2, \mathbf{R})$*, Invent. math. 80 (1985), 339–364.

[Rie]   Riemann, B.: *Über die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monatsber. Berlin. Akad. 1859, 671–680, or in *Gesammelte mathematische Werke, wissenschaftlicher Nachlass und Nachtrage*, edited by R. Narasimhan, Springer-Verlag, Berlin, 1990, 177–185.

[RS]   Rudnick, Z. and Sarnak, P.: *Zeros of principal L-functions and random matrix theory*, Duke Math. J. 81 (1996), 269–322.

[Sa]   Sarnak, P.: *Quantum Chaos, Symmetry and Zeta Functions*, Current Developments in Math., International Press 1997.

[Se1]   Serre, J-P.: *Cours d'arithmétique*, 3ème édition, P.U.F (1988).

[Se2]   Serre, J-P.: *Modular forms of weight one and Galois representations*, in Algebraic Number Fields (A. Frölich, ed.), Acad. Press (1977), 193–268.

[Sh]   Shimura, G.: *Introduction to the arithmetic theory of automorphic functions*, Iwanami Shoten and Princeton Univ. Press (1971).

[Si]   Siegel, C.L.: *Über die Classenzahl quadratischer Zahlkörper*, Acta Arithmetica 1 (1936), 83–86.

[Ta]   Tate, J.: *Fourier analysis in number fields and Hecke's zeta function*, in [**CF**].

[Tu]   Tunnell, J.: Rutgers University graduate course (1995–96).

[Wa]   Washington, L.: *Cyclotomic fields*, 2nd edition, GTM 83, Springer (1997).

[We]   Weil, A.: *Prehistory of the zeta function*, in "Number Theory, Trace Formulas and Discrete Groups", K.E. Aubert, E. Bombieri and D. Goldfeld eds., Acad. Press (1989).

[Ti]   Titchmarsh, A.C.: *The theory of the Riemann Zeta-function*, Second edition (revised by D. R. Heath-Brown), Oxford University Press, 1986.