THE ANALYTIC CLASS NUMBER FORMULA AND L-FUNCTIONS

AKSHAY VENKATESH
NOTES BY TONY FENG

CONTENTS

1. Overview

Let K be a number field. Then we have an associated *Dedekind ζ function*

$$\zeta_K(s) = \sum_{I \subset \mathcal{O}_K} (\operatorname{Nm} I)^{-s}$$

generalizing the Riemann zeta function (which is the special case $K = \mathbb{Q}$), and possessing the following basic properties:

- The series defining $\zeta_K(s)$ is convergent for $\operatorname{Re} s > 1$. (This is easy; it follows from the observation that there are very few ideals of a given norm.)
- $\zeta_K(s)$ has an *Euler product* factorization:

$$\zeta_K(s) = \prod_{\mathfrak{p} \text{ prime}} \left(\frac{1}{1 - (\operatorname{Nm} \mathfrak{p})^{-s}} \right).$$

- $\zeta_K(s)$ has a meromorphic continuation to $s \in \mathbb{C}$, with the only pole being a simple pole at $s = 1$.
- We have a *class number formula*

$$\operatorname{Res}_{s=1} \zeta_K(s) = \frac{hR}{w\sqrt{\operatorname{disc} K}} 2^{r_1} (2\pi)^{r_2} \tag{1}$$

  Here R is the regulator, which can be thought of as the "volume" of $\mathcal{O}_K^*$, h is the class number, w is the number of roots of unity, and r_1 and r_2 are the number of real and complex places of K.
- It satisfies a *functional equation.* An elegant way to phrase this is that

  $(\operatorname{disc} K)^{s/2} \Gamma_{\mathbb{R}}(s)^{r_1} \Gamma_{\mathbb{C}}(s)^{r_2} \zeta_K(s)$ is symmetric under $s \longleftrightarrow 1 - s$.

  Here $\Gamma_{\mathbb{R}}(s) = \pi^{-s/2}\Gamma(s/2)$ and $\Gamma_{\mathbb{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$.

We will begin by discussing the proofs of these properties, following the classical method of Hecke.

If K is Galois over $\mathbb{Q}$, then

$$\zeta_K(s) = \prod_{\substack{\rho \text{ irred. rep'n.} \\ \text{of } \operatorname{Gal}(K/\mathbb{Q})}} L(s, \rho)$$

where $L(s, \rho)$ is the Artin L-function attach to ρ. Each $L(s, \rho)$ has (at least conjecturally) analogous properties to those mentioned above.

Example 1.0.1. If $K = \mathbb{Q}(i)$, then

$$\zeta_K(s) = \sum_{n+mi} \frac{1}{(m^2 + n^2)^s}$$

and the corresponding factorization into Artin L-functions is

$$\zeta_K(s) = \underbrace{\zeta(s)}_{\rho=\text{triv}} \underbrace{(1^{-s} - 3^{-s} + 5^{-s} - \ldots)}_{\rho=\text{sign}}.$$

The main point of the course is to discuss the following question: since ζ_K factors, there should be a corresponding factorization of the class number formula (1): so the right hand side should be expressible as

$$\frac{hR}{w} = \text{“}\prod_{\rho} \frac{h(\rho)R(\rho)}{w(\rho)}\text{”}.$$

Even to formulate precisely what this factorization should be is a little tricky.

Example 1.0.2. Even for the case $K = \mathbb{Q}(i)$ (when the Galois group has size 2, and there are two representations), what goes on is subtle because of issues at 2. The problem is analogous to the following general issue: if A is a finite abelian group with an action of $\mathbb{Z}/2$, then one can define "eigenspaces" A^+, A^-. The map $A^+ \times A^- \to A$ fails to be an isomorphism in the presence of 2-torsion, which appears in both "eigenspaces."

This idea motivates many things in number theory, like Stark's conjecture and the Main Conjecture of Iwasawa theory.

2. L-FUNCTIONS: ANALYTIC PROPERTIES

We'll now prove some of the basic properties of ζ_K, following Hecke.

2.1. **Analytic continuation.** We first digress briefly about what it *means* to have analytic continuation (we use this blanket term even when referring to what might technically be called meromorphic continuation). When one says that the standard ζ function

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

has "analytic continuation," it usually comes with connotations of complex analysis. For us, complex analysis is not important; what's important is that there is a way to "make sense" of the expression for any s.

Example 2.1.1. How might we make sense of $\sum \frac{1}{\sqrt{n}}$? Observe that at least for $\operatorname{Re} s > 1$, $\sum \frac{1}{n^s}$ can be approximated by an integral. Therefore, we consider

$$\sum_{n=1}^{\infty} \frac{1}{n^s} = \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n^s} - \int_0^N \frac{dx}{x^s}.$$

You can prove that this has a limit for $\operatorname{Re} s > 0$. The integral is $\frac{N^{1-s}}{1-s}$, so

$$\lim_{N\to\infty} \sum_{n=1}^{N} \left(\frac{1}{n^s} - \frac{N^{1-s}}{1-s} \right)$$

exists for $\operatorname{Re} s > 0$. It equals $\zeta(s)$ for $\operatorname{Re} s > 1$, and we can "make sense" of $\zeta(s)$ by simply setting it to be the above expression for all $\operatorname{Re} s > 0$. You can play similar games to extend it further "by hand."

Example 2.1.2. Let $K = \mathbb{Q}(i)$. Then

$$\sum_{m+ni} \frac{1}{(m^2+n^2)^s} = \frac{1}{4} \sum_{(m,n)\neq(0,0)} \frac{1}{(m^2+n^2)^s}.$$

We can interpret this as the sum of a particular binary quadratic form over the lattice $\mathbb{Z}^2 \subset \mathbb{C}$. We can generalize this by asking about the sum of any (say definite) quadratic form Q over a lattice $\Lambda \subset \mathbb{C}$.

$$\left\{ \begin{array}{c} \text{positive definite} \\ \text{binary quadratic forms} \\ ax^2+bxy+cy^2 \\ \text{disc}=1 \end{array} \right\} / \mathrm{SL}_2(\mathbb{Z}) \longleftrightarrow \mathbb{H}/\mathrm{SL}_2(\mathbb{Z})$$

$$\updownarrow$$

$$\left\{ \begin{array}{c} \text{lattices } \Lambda \subset \mathbb{C} \\ \text{of covolume 1} \\ \text{up to rotation} \end{array} \right\}$$

♠♠♠ TONY: [discriminant should be -1 to be positive-definite, eh?] The horizontal map sends Q to the root of $Q(z,1)$ lying in $\mathbb{H}$. The vertical map sends $z \mapsto \langle 1, z\rangle / \operatorname{Im}(z)^{1/2}$, with the associated quadratic form Λ (choosing a positively oriented basis).

We actually show that there is a kind of functional equation for a general sum of the form

$$\sum_{z \in \Lambda - \{0\}} \|z\|^{-2s}, \quad \|z\| = Q(z).$$

This is (for fixed s) a real-analytic *Eisenstein series* on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. More precisely, we'll show that if $\Lambda \subset \mathbb{R}^n$ is a lattice of covolume 1,

$$\Gamma_{\mathbb{R}}(s) \sum_{v \in \Lambda} \|v\|^{-s}$$

has an analytic continuation with simple poles at $s = 0$ and $s = n$, and is symmetric under $s \mapsto n - s$. Then we'll recover the results for $\zeta_{\mathbb{Q}(i)}$ by taking $\Lambda = \mathbb{Z}[i] \subset \mathbb{C}$. However, it's not so obvious what to do for *real quadratic fields* like $\zeta_{\mathbb{Q}(\sqrt{2})}$.

Example 2.1.3. Just to convince you that this is not a mysterious object, let's write it in a more concrete way.

$$\zeta_{\mathbb{Q}(\sqrt{2})}(s) = \sum_{a + b\sqrt{2} \in \mathbb{Z}[\sqrt{2}]/\text{units}} \frac{1}{(a^2 - 2b^2)^s}.$$

The unit group $\mathbb{Z}[\sqrt{2}]^*$ is generated by $-1, \sqrt{2}-1$. Let $\alpha = \sqrt{2}-1$; then α is positive under one embedding and negative under the other. Therefore, any $\beta = a + b\sqrt{2}$ can be made totally positive by multiplying by a unit. If $\beta = a + b\sqrt{2}$ is totally positive, then by multiplying by something in $(\alpha^2)^{\mathbb{Z}}$ we can assume that

$$1 < \frac{\beta}{\overline{\beta}} \leq \alpha^4.$$

You can check that in terms of a and b, this is equivalent to $b > 0, 3b \leq 2a$. This gives a fundamental domain in $\mathbb{R}^2$ for $\mathbb{Z}[\sqrt{2}]/\mathbb{Z}[\sqrt{2}]^*$. So the zeta function is the sum over (a,b) in some *cone* in $\mathbb{Z}^2 \subset \mathbb{R}^2$ of $\frac{1}{(a^2-2b^2)^s}$. This is again a sum of a quadratic form over lattice points in a cone in $\mathbb{R}^2$, but the quadratic form is indefinite. However, the cone is away from the indefinite locus, so it doesn't "see" the indefiniteness.

This idea is due to Shintani, and we'll discuss it again later. The cones are not simplicial, i.e. the integer points on the interior are not an integer linear combination of those on the boundary. Shintani works with simplicial subdivisions to resolve this, but it isn't important for us.

For any arbitrary imaginary quadratic field K, $\zeta_K(s)$ is a sum of sums of the form $\sum_{z\in\Lambda-\{0\}}|z|^{-2s}$ as Λ ranges over the ideal classes of K. Hecke realized that there is a way to extend this to real quadratic fields as well. For imaginary K, we get a *finite* subset of $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ associated to K (namely, the points corresponding to the lattices of the ideal classes). For real K, we get instead a finite collection of *closed geodesics* on $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. For real quadratic K, one gets that $\zeta_K(s)$ is the sum of the *integrals* of essentially the same expression $\sum_{z\in\Lambda}|z|^{-2s}$ over these associated geodesics.

There is a uniform way to say this. For *any* K, one gets a set $\mathscr{L}_K$ of lattices in $(K\otimes\mathbb{R})$ stable by $\mathscr{O}_K$ (which has the structure of a torsor for a compact abelian group). Then ζ_K is obtained by integrating over this set. The result is

$$\zeta_K(s)=(\ldots)\int_{\substack{\text{lattices in } K\otimes\mathbb{R}\\ \text{stable by } \mathscr{O}_K\\ \text{of volume 1}}}\left(\sum_{v\in\Lambda-\{0\}}\|v\|^{-s}\right).$$

This has been an overview of the strategy. Now we'll carry it out it.

2.2. **Epstein ζ functions.**

Definition 2.2.1. If $\Lambda\subset\mathbb{R}^n$ is a lattice of covolume 1, define

$$E_\Lambda(s)=\sum_{v\in V}\|v\|^{-s}.$$

Definition 2.2.2. Let $\mathscr{L}_K$ be the set of lattices in $K\otimes\mathbb{R}$ stable by $\mathscr{O}_K$, and $\mathscr{L}_K^{(1)}\subset\mathscr{L}^K$ the subset of lattices of covolume 1.

We want to show that $E_\Lambda(s)$ analytically continues to $s\in\mathbb{C}$, with a simple pole at $s=n$, and $\Gamma_{\mathbb{R}}(s)E_\Lambda(s)$ is symmetric under $s\longleftrightarrow n-s$. Then we'll deduce similar properties for ζ_K (e.g. K a real quadratic number field) by writing

$$\zeta_K(s)=c(s)\int_{\mathscr{L}_K^1}E_\Lambda(s)$$

for an appropriate measure on $\mathscr{L}_K^1$.

Example 2.2.3. This generalizes

$$\zeta_{\mathbb{Q}(i)}=\sum\frac{1}{(m^2+n^2)^s}=E_\Lambda(s)\text{ for }\Lambda=\mathbb{Z}^2\subset\mathbb{C}=\mathbb{R}^2.$$

If you "unwind" our process then you will arrive at Tate's thesis, but this formalism is that of Hecke.

The first thing has nothing to do with number fields; it is purely a question of analysis. The trick is that instead of summing $\|v\|^{-s}$ over Λ, we

sum a nicer function Φ and then make it homogeneous. The key observation is that if Φ is any function on $\mathbb{R}^n$ depending only on $||v||$, and $\Phi_t(x) = \Phi(tx)$, then $\int \Phi_t(x) t^s \frac{dt}{t}$ will necessarily be proportional to $||x||^{-s}$ because it is radial and homogeneous of the right degree. (The point is that $||v||^{-s}$ has bad behavior at either 0 or ∞, depending on s, so we want to "smooth it out.")

More specifically, we'll take Φ to be a Schwartz function. Let

$$E_\Phi(\Lambda) = \sum_{v\in\Lambda} \Phi(v).$$

We'll first show that $\int E_{\Phi_t} t^s \frac{dt}{t}$ has analytic continuation. The key is to use *Poisson summation.*

Proposition 2.2.4 (Poisson summation)**.** *Let V be an n-dimensional real vector space, V^* the dual space, and Φ a Schwarz function on V. Define the Fourier transform*

$$\widehat{\Phi}(k) = \int \Phi(x) e^{2\pi i \langle k,x\rangle}\, dx$$

where dx is normalized so $\mathrm{vol}(V/\Lambda) = 1$. *Then*

$$\sum_{v\in\Lambda} \Phi(v) = \sum_{w\in\Lambda^*} \widehat{\Phi}(w)$$

where $\Lambda^* = \{\eta \in V^* : \langle y, \Lambda\rangle \subset \mathbb{Z}\}$ is the dual lattice.

Proof sketch. The proof is to expand the function $x \mapsto \sum_{v\in\Lambda} \Phi(x+v)$ in Fourier series on V/Λ and then evaluate at 0. □

Applying this to Φ_t, we find that

$$\widehat{\Phi_t} = t^{-n} \widehat{\Phi}_{1/t}.$$

So by Poisson summation,

$$E_{\Phi_t}(\Lambda) = t^{-n} E_{\widehat{\Phi}_{1/t}}(\Lambda^*).$$

What is this anyway? For $\mathrm{Re}\, s \gg 0$, we may define

$$G_{\Phi,\Lambda}(s) := \sum_{v\in\Lambda-\{0\}} \underbrace{\int_0^\infty \Phi_t(v) t^s \frac{dt}{t}}_{\propto ||v||^{-s} \text{ if } \Phi \text{ radial}} = \int_0^\infty t^s \frac{dt}{t} \underbrace{\sum_{v\in\Lambda-\{0\}} \Phi_t(v)}_{E_{\Phi_t} - \Phi(0)}.$$

(We claimed that $G(s) = E_\Lambda(s)$, by the uniqueness of radially symmetric, homogeneous smooth functions.) Note that $E_{\Phi_t} - \Phi(0) = \sum_{v\in\Lambda-\{0\}} \Phi(tv)$ decays rapidly if t is large (as Φ is Schwartz, and making t larger squishes the mass to the origin), so there's no problem with $\sum_{v\in\Lambda-\{0\}} \Phi_t$ converging when t is large. The only problem with the integral occurs when t

is small. In that range, we use Poisson summation to write it in another way.

We split up

$$G_{\Phi,\Lambda}(s) = \int_1^\infty t^s \frac{dt}{t}(E_{\Phi_t} - \Phi(0)) + \int_0^1 t^s \frac{dt}{t}(E_{\Phi_t} - \Phi(0))$$

and then apply Poisson summation to the second term:

$$\int_0^1 t^s \frac{dt}{t} E_{\Phi_t} = \int_0^1 t^s \frac{dt}{t} t^{-n} E_{\widehat{\Phi}_{1/t}}.$$

Therefore, the second term may be rewritten as

$$\begin{aligned}
\int_0^1 t^s \frac{dt}{t}(E_{\Phi_t} - \Phi(0)) &= \int_0^1 t^s \frac{dt}{t} t^{-n} E_{\widehat{\Phi}_{1/t}} - \frac{\Phi(0)}{s} \\
&= \int_0^1 t^{s-n} \frac{dt}{t}(E_{\widehat{\Phi}_{1/t}} - \widehat{\Phi}(0)) + \frac{\widehat{\Phi}(0)}{s-n} - \frac{\Phi(0)}{s} \\
(t \mapsto 1/t) &= \int_1^\infty t^{n-s} \frac{dt}{t}(E_{\widehat{\Phi}_t} - \widehat{\Phi}(0)) + \frac{\widehat{\Phi}(0)}{s-n} - \frac{\Phi(0)}{s}.
\end{aligned}$$

Now *both* of the integrals converge for all s, because they decay rapidly and the limits are bounded away from 0. We have proved:

Theorem 2.2.5. *Let Φ be any Schwartz function on $\mathbb{R}$. Then $G_{\Phi,\Lambda}(s)$ has an analytic continuation to $s \in \mathbb{C}$, with simple poles at $s = 0$ (residue $-\Phi(0)$) and n (residue $\widehat{\Phi}(0)$). Moreover, we have the functional equation*

$$G_{\Phi,\Lambda}(s) = G_{\widehat{\Phi},\Lambda^*}(n-s).$$

Now take $\Phi(x) = e^{-\pi\langle x,x\rangle}$ (so now we are finally making use of our quadratic form), where the volume associated to $\langle x,x\rangle$ gives Λ value 1. Identify $V \cong V^*$ by this quadratic form. With these normalization, $\widehat{\Phi} = \Phi$. Then

$$\begin{aligned}
G_{\Phi,\Lambda}(s) &= \sum_{v \in \Lambda - \{0\}} \int (e^{-\pi t^2 \langle v,v\rangle}) t^s \frac{dt}{t} \\
&= \frac{1}{2}\pi^{-s/2}\Gamma(s/2) \sum_{v \in \Lambda - \{0\}} \|v\|^{-s}.
\end{aligned}$$

This has analytic continuation, and is symmetric under $s \mapsto n-s$ and $\Lambda \mapsto \Lambda^*$. Don't ignore the Λ^*; it is important!

Remark 2.2.6. There is a way to prove this by "pure thought." The Eisenstein series $E_\Lambda(s) = \sum_{v \in \Lambda - \{0\}} \|v\|^{-s}$ is characterized some property, in terms of the spectral theory of the Laplace operator. One can check that the

other side of the functional equation also has this property, hence the identity.

2.3. **The Dedekind zeta function.** Let K be an imaginary quadratic field. Let I be an ideal of $\mathcal{O}_K$, and

$$\zeta_{[I]}(s) = \sum_{\substack{\mathfrak{a}\subset\mathcal{O}_K \\ \mathfrak{a}\sim I}} (\operatorname{Nm}\mathfrak{a})^{-s}.$$

(Here the equivalence relation is in the ideal class group.) Then

$$\zeta_K(s) = \sum_{[I]\in \mathrm{Cl}(K)} \zeta_I(s).$$

Every such $\mathfrak{a}$ is of the form $\lambda\cdot I$ where $\lambda\in I^{-1}$. Therefore,

$$\zeta_{[I]}(s) = \frac{(\operatorname{Nm} I)^{-s}}{w} \sum_{\lambda\in I^{-1}-\{0\}} (\operatorname{Nm}\lambda)^{-s}$$

where w is the number of units. We already know that this has analytic continuation, functional equation, etc. as it is a sum of the form described in Theorem 2.2.5 with I interpreted as a lattice in $\mathbb{C}$ with the obvious norm. Therefore, ζ_K does too.

More precisely, apply Theorem 2.2.5 with

- $V = K\otimes\mathbb{R}(\cong\mathbb{C})$.
- V and V^* identified by the trace form $(z_1,z_2)\mapsto \operatorname{Tr}(z_1z_2)$.
- The measure on V is given by $|dz\wedge d\overline{z}| = 2\cdot$ Lebesgue.
- $\Phi(z) = e^{-2\pi|z|^2}$.

With these normalizations,

$$\sum_{v\in\Lambda-\{0\}} \left(\int \Phi_t(v) t^{2s}\frac{dt}{t}\right) = \frac{1}{4}\Gamma_{\mathbb{C}}(s)\sum_{v\in\Lambda}||v||^{-s}.$$

(Recall that $\Gamma_{\mathbb{C}}(s) = 2(2\pi)^{-s}\Gamma(s)$.) Note that here $||v||^{-s}$ is the *complex modulus*, which explains the discrepancy of a factor of 2. For $\Lambda = \Lambda_I$, i.e. the ideal I considered as a lattice in V, what is the dual? By definition, that is $\Lambda_{\mathscr{D}^{-1}I^{-1}}$ where $\mathscr{D}$ is the different of $\mathcal{O}_K$. (By definition, the inverse different is the dual to the ring of integers under the trace pairing.) If $K = \mathbb{Q}(\sqrt{-d})$ then $\mathscr{D} = (\sqrt{-d})$.

The covolume of $\mathcal{O}_K$ is $\sqrt{D}$ where D is the discriminant. For instance, the discriminant of $\mathbb{Q}(i)$ is 4, and that's why we take *twice* the Lebesgue measure ♠♠♠ TONY: [???]. Then the volume of I^{-1} is $\sqrt{D}/\operatorname{Nm} I$, and $\operatorname{vol}(I)\operatorname{vol}(\mathscr{D}^{-1}I^{-1}) = 1$. Therefore, the theorem implies that $D^{s/2}\Gamma_{\mathbb{C}}(s)\zeta_{[I]}(s)$ is symmetric under $s\mapsto 1-s$, $I\mapsto \mathscr{D}^{-1}I^{-1}$. Summing over I, we obtain that $D^{s/2}\Gamma_{\mathbb{C}}(s)\zeta_K(s)$ is symmetric under $s\mapsto 1-s$.

Remark 2.3.1. Hecke proved that $\mathscr{D}$ is a square in the ideal class group. We'll prove this by exploiting the presence of the $\mathscr{D}^{-1}$ here. The corresponding statement in the function field is that the square root of the canonial bundle over a finite field exists and is *rational* over that field.

Recall that we defined $\mathscr{L}_K$ to be the set of lattices in $V = K \otimes \mathbb{R}$ stable by $\mathscr{O}_K$, and $\mathscr{L}_K^{(1)}$ to be the subset with covolume 1.

Proposition 2.3.2. *If $[K:\mathbb{Q}] = n$, and $V = K \otimes \mathbb{R}$, then*

$$\zeta_K(s) = \underbrace{c(s)}_{explicit} \int_{\mathscr{L}_K^{(1)}} E_\Lambda(s)\, d\Lambda.$$

We first have to explain what this formula even means! For the right hand side to make sense, we must define a measure on $\mathscr{L}_K^{(1)}$. We claim that every element of $\mathscr{L}_K$ is of the form $\Lambda_I y$ (the lattice in V attached to an ideal I), where $y \in (K \otimes \mathbb{R})^*$.

Why? First think about the quadratic imaginary case: it says that any lattice stable by $\mathscr{O}_K$ is an ideal times some complex number, which is a familiar fact. The general argument is easy. There is at least one element lying in $(K \otimes \mathbb{R})^*$, because that is the complement of a hypersurface. Then you can multiply by its inverse to move it to 1. Then the claim reduces to the assertion that any lattice containing 1 and stable by $\mathscr{O}_K$ is a fractional ideal. But now this is clear, as it already contains all of $\mathscr{O}_K$, which has full rank in $K \otimes \mathbb{R}$.

Therefore, we may write

$$\mathscr{L}_K = \coprod_{[I] \in \mathrm{Cl}(K)} \Lambda_I \cdot (K \otimes \mathbb{R})^* / \mathscr{O}_K^*.$$

Once we've fixed a Haar measure on $K \otimes \mathbb{R}$, $\mathscr{L}_K^{(1)}$ will be the volume one elements of $\mathscr{L}_K$, i.e.

$$\mathscr{L}_K^{(1)} = \coprod_{[I] \in \mathrm{Cl}(K)} (\Lambda_I \text{ scaled to volume 1}) \cdot (K \otimes \mathbb{R})^{(1)} / \mathscr{O}_K^*$$

where $(K \otimes \mathbb{R})^{(1)} = \{x \in K \otimes \mathbb{R} : ||x|| = 1\}$.

By the unit theorem, there are enough units to make $(K \otimes \mathbb{R})^{(1)} / \mathscr{O}_K^*$ compact, so $\mathscr{L}_K^{(1)}$ is also. As each $(K \otimes \mathbb{R})^{(1)} / \mathscr{O}_K^*$ is a torus, it looks like a finite union of tori. We then get a measure on $\mathscr{L}_K$ by fixing a Haar measure on $(K \otimes \mathbb{R})^*$, and similarly we get a measure on $\mathscr{L}_K^{(1)}$ by fixing a Haar measure on $(K \otimes \mathbb{R})^{(1)}$, because $\mathscr{L}_K$ is a disjoint union of quotients of $(K \otimes \mathbb{R})^*$ by discrete subgroups, etc.

Remark 2.3.3. Also, when you take residues, you get factors of the volume of $\mathscr{L}_K$, which is essentially the regulator. That explains why the regulator enters into the analytic class number formula.

Proof Sketch of Proposition 2.3.2. We first try to sketch *why* this should be true. Consider averaging $||v||^{-s}$ over $(K\otimes\mathbb{R})^{(1)}$, i.e.

$$\int_{y\in(K\otimes\mathbb{R})^{(1)}} ||yv||^{-s}\,dy.$$

We can figure out what this is by pure thought. Since this is averaged over norm 1 elements, it can only depend on the norm. Also, since it has the right homogeneity properties, it *must* be $a(s)|\operatorname{Nm} v|^{-s/n}$. Unfortunately, it's a bit messy to work out $a(s)$. For example, for a real quadratic field we have $K\otimes\mathbb{R}\cong\mathbb{R}^2$, and the orbits of $(K\otimes\mathbb{R})^*$ under $\mathcal{O}_K^*$ are hyperbolas, so you have to integrate $\frac{1}{(x^2+y^2)^s}$ over hyperbolas.

□

Remark 2.3.4. In this sketch derivation, we used the fact that $(K\otimes\mathbb{R})^{(1)}$ acts on $K\otimes\mathbb{R}$ with a single invariant, namely the norm. This fits into the following more general framework.

A *pre-homogeneous vector space* is the data of a reductive group $G/\mathbb{Q}$, together with a representation of G on V such that the ring of invariants of G on V is $\mathbb{Q}[f]$ for a single element f. Morally, this means that there is only a one-parameter family of orbits. Shintani proved meromorphic continuation for $\sum_{v\in V_{\mathbb{Z}}/G_{\mathbb{Z}}}|f(v)|^{-s}$ in such a situation. The proof in this special case replaces V by $K\otimes\mathbb{R}$ and G by $(K\otimes\mathbb{R})^{(1)}$. The general object is similar analytically, but it lacks some interesting arithmetic properties like an Euler product.

There are many sporadic examples of interesting pre-homogeneous vector spaces. SL_n acts on $\operatorname{Sym}^2\mathbb{Q}^n$, with a single invariant - the determinant (discriminant). Also, SL_n acts on $\bigwedge^2\mathbb{Q}$ with one invariant, the Pfaffian. SL_n acts on $\bigwedge^3\mathbb{Q}^n$, and this is prehomogeneous when $n\le 8$ (see that this is reasonable by dimension count), $\mathrm{SL}_4\times\mathrm{SL}_5$ acts on $\mathbb{Q}^4\otimes\bigwedge^2\mathbb{Q}^5$, etc.

Theorem 2.3.5. *Let K be a number field. Then ζ_K admits analytic continuation and a functional equation.*

Proof. One could give a direct proof, but it's easier to go back to Schwartz functions. Recall that for $\Phi\in\mathscr{S}(\mathbb{R}^n)$ a radial Schwartz function,

$$E_\Lambda(s)\propto\int_0^\infty E_{\Phi_t}(\Lambda)t^s\frac{dt}{t}$$

where $\Phi_t(x)=\Phi(tx)$, $E_\Phi=\sum_{v\in\Lambda-\{0\}}\Phi(v)$. Thus $E_{\Phi_t}(\Lambda)=E_\Phi(t\Lambda)$, so we

have

$$\begin{aligned}\int_{\mathscr{L}_K^{(1)}} E_\Lambda(s)\,d\Lambda &\propto \int_{\mathscr{L}_K^{(1)}\times\mathbb{R}_{>0}} (E_\Phi(t\Lambda)t^s-\Phi(0))\frac{dt}{t}\,d^*y \\ &= \sum_{[I]\in \mathrm{Cl}(K)} \int_{\underbrace{(K\otimes\mathbb{R})^{(1)}/\mathscr{O}_K^*}_{y}\times\underbrace{\mathbb{R}_{>0}}_{t}} (E_\Phi(ty\Lambda_I)-\Phi(0))\,d^*y\,dt \\ &= \sum_{[I]\in \mathrm{Cl}(K)} \int_{(K\otimes\mathbb{R})^*/\mathscr{O}_K^*} (E_\Phi(y\Lambda_I)-\Phi(0))|\operatorname{Nm} y|^s\,d^*y\end{aligned}$$

as our starting point. (Remark: except for making things adelic, this is almost the starting point of Tate's thesis).

Now we're finally ready to start the computation. Letting d^*y be a Haar measure on $(K\otimes\mathbb{R})^*/\mathscr{O}_K^*$, we have

$$\begin{aligned}\int_{(K\otimes\mathbb{R})^*/\mathscr{O}_K^*} E_\Phi(y\Lambda_I)|\operatorname{Nm} y|^s\,d^*y &= \int_{(K\otimes\mathbb{R})^*/\mathscr{O}_K^*} \sum_{v\in I-\{0\}} \Phi(yv)|\operatorname{Nm}(y)|^s\,d^*y \\ (\text{ for } \operatorname{Re} s\gg 0) &= \sum_{v\in I-\{0\}/\mathscr{O}_K^*} \int_{(K\otimes\mathbb{R})^*} \Phi(yv)|\operatorname{Nm} y|^s\,d^*y \\ &= \sum_{v\in I-\{0\}/\mathscr{O}_K^*} |\operatorname{Nm} v|^{-s}\int_{K\otimes\mathbb{R}^*} \Phi(y)|\operatorname{Nm}(y)|^s\,d^*y.\end{aligned}$$

We can re-arrange this as

$$\sum_{v\in I-\{0\}/\mathscr{O}_K^*} |\operatorname{Nm}(v)|^{-s} = \frac{\int_{(K\otimes\mathbb{R})^*/\mathscr{O}_K^*}(E_\Phi(y\Lambda_I)-\Phi(0))|\operatorname{Nm} y|^s\,d^*y}{\int_{(K\otimes\mathbb{R})^*}\Phi(y)|\operatorname{Nm} y|^s\,d^*y}.$$

Therefore,

$$\sum_{\mathfrak{a}\sim I}|\operatorname{Nm}\mathfrak{a}|^{-s} = |\operatorname{Nm} I|^{-s}\frac{\int_{(K\otimes\mathbb{R})^*/\mathscr{O}_K^*}(E_\Phi(y\Lambda_I)-\Phi(0))|\operatorname{Nm} y|^s\,d^*y}{\int_{(K\otimes\mathbb{R})^*}\Phi(y)|\operatorname{Nm} y|^s\,d^*y}.$$

♠♠♠ TONY: [is this right?] As before, Poisson summation relates $E_\Phi(\Lambda)$ and $E_{\widehat{\Phi}}(\Lambda^*)$. You can use this to show that the numerator has meromorphic continuation and a functional equation, and similarly for the denominator.

Specifically, we apply Theorem 2.2.5 with:

(1) $V=(K\otimes\mathbb{R})$, identified with V^* via the trace pairing $(x,y)\mapsto \operatorname{tr}(xy)$.
(2) $\Lambda_I^* = \Lambda_{\mathscr{D}^{-1}I^{-1}}$.
(3) $K\otimes\mathbb{R}\cong\mathbb{R}^{r_1}\times\mathbb{C}^{r_2}$, and we take the Lebesgue measure on $\mathbb{R}$ and twice the Lebesgue measure on $\mathbb{C}$.

(4) $\Phi = e^{-\pi(x_1^2+\ldots+x_{r_1}^2+2|z_1|^2+\ldots+2|z_{r_2}|^2)}$. Actually, the choice of Φ doesn't really matter (any choice should give you the same L-function), but this choice makes $\widehat{\Phi} = \Phi$, hence is more conducive to showing the meromorphic continuation and functional equation.

For this normalized Gaussian choice of Φ,

$$\int \Phi(y)|\mathrm{Nm}(y)|^s d^*y = \Gamma_{\mathbb{R}}(s)^{r_1}\Gamma_{\mathbb{C}}(s)^{r_2}$$

and the functional equation says that

$$D^{s/2}\Gamma_{\mathbb{R}}(s)^{r_1}\Gamma_{\mathbb{C}}(s)^{r_2}\sum_{\mathfrak{a}\sim I} N(\mathfrak{a})^{-s}$$

is symmetric under $(s, I) \mapsto (1-s, \mathscr{D}^{-1}I^{-1})$. The add over ideal classes to obtain the result for ζ_K. □

Remark 2.3.6. Unraveling this gives an equation of the form

$$\int \Phi(y)\mathrm{Nm}(y)^s d^*y = a(s)\int \widehat{\Phi}(y)\mathrm{Nm}(y)^{1-s}\, d^*y.$$

For $K = \mathbb{Q}$, this says that

$$\int \Phi(y)y^s d^*y = a(s)\int \widehat{\Phi}(y)y^{1-s}\, d^*y.$$

One can see this by "pure thought." Indeed, view $\Phi(s) \mapsto \int \Phi(s)y^s \frac{dy}{y}$ as the tempered distribution associated to y^s. Then the claim is that this is proportional to the distribution $\mathscr{F}(y^{1-s})$. Since $\mathscr{F}(y^{1-s})$ is homogeneous of degree s, the claim follows from:

Lemma 2.3.7. *There is a unique distribution on* $\mathbb{R}$ *up to scalar which is homogeneous of degree* s.

Proof. It is easy to show that this is the case for functions supported away from 0. One has to compute to see what happens to functions supported near 0. □

2.4. **Generalizations.** We just showed that if

$$\zeta_{[I]} = \sum_{\mathfrak{a}\sim I}(\mathrm{Nm}\,\mathfrak{a})^{-s}$$

then $D^{s/2}\Gamma_{\mathbb{R}}(s)^{r_1}\Gamma_{\mathbb{C}}(s)^{r_2}\zeta_{[I]}(s)$ is symmetric under $s \mapsto 1-s$, $I \mapsto \mathscr{D}^{-1}I^{-1}$. We'll use this to give a proof of:

Proposition 2.4.1. $\mathscr{D}$ *is a square in* $\mathrm{Cl}(K)$.

Definition 2.4.2. Let $\chi\colon \mathrm{Cl}(K) \to \mathbb{C}^*$ be a character. Then we define the L-function

$$L(s,\chi) = \sum_{a \subset \mathcal{O}_K} \chi(a)(Na)^{-s}.$$

This has very similar properties to ζ_K:

(1) It has an analytic continuation to $s \in \mathbb{C}$. Moreover, it has *no poles* if χ is non-trivial because we can write it as

$$L(s,\chi) = \sum_{[I] \in \mathrm{Cl}(K)} \chi(I)\zeta_{[I]}$$

and the key point is that all the $\zeta_{[I]}$ have the *same* residue at $s=1$, so if χ is non-trivial then the residues cancel out.

(2) It has a functional equation:

$$D^{s/2}\Gamma_{\mathbb{R}}(s)^{r_1}\Gamma_{\mathbb{C}}(s)^{r_2}L(s,\chi) = D^{(1-s)/2}\Gamma_{\mathbb{R}}(s)^{r_1}\Gamma_{\mathbb{C}}(s)^{r_2}L(1-s,\chi^{-1})\chi(\mathscr{D})$$

Note the $\chi(\mathscr{D})$ here; it will be important later!

(3) It has an Euler product:

$$L(s,\chi) = \prod_{\mathfrak{p}} \left(1 - \chi(\mathfrak{p})(\mathrm{Nm}\,\mathfrak{p})^{-s}\right)^{-1}.$$

Proof of Proposition 2.4.1. Suppose $\chi\colon \mathrm{Cl}(K) \to \mathbb{C}^*$ is a quadratic character. By class field theory, there is an *unramified* quadratic extension L/K such that $\mathfrak{p}$ is split in L/K if and only if $\chi(\mathfrak{p}) = 1$.

We claim that $\zeta_L(s) = L(s,\chi)\zeta_K(s)$. The right hand side can be factored as

$$\prod_{\mathfrak{p} \subset \mathcal{O}_K} \begin{cases} (1-(\mathrm{Nm}\,\mathfrak{p})^{-s})^{-2} & \chi(\mathfrak{p}) = 1, \\ (1-(\mathrm{Nm}\,\mathfrak{p})^{-2s})^{-1} & \chi(\mathfrak{p}) = -1. \end{cases}$$

Now we compare the functional equations for $\zeta_L, \zeta_K, L(s,\chi)$. The discriminants cancel out, and the Γ functions cancel out. But the extra $\chi(\mathscr{D})$ factor appears *only* on the right hand side, so it must be 1. Since this holds for all quadratic characters χ, $\mathscr{D}$ is a square in $\mathrm{Cl}(K)$.

□

The same theorem with the same proof goes through if we replace K by a finite extension of $\mathbb{F}_q(t)$, i.e. K is the function field of a curve $C/\mathbb{F}_q$. Then the different is $K_C \cong \mathscr{L}^{\otimes 2}$ where $\mathscr{L}$ is *defined over* $\mathbb{F}_q$. (That it exists over some larger extension is clear from the divisibility of the Jacobian; the non-trivial assertion is that it is rational.)

Tate's thesis is a generalization of Hecke's proof. If you go back to our expression for ζ_K as a sum over $[I] \in \mathrm{Cl}(K)$ and an integral over $(K \otimes \mathbb{R})^*/\mathcal{O}_K^*$, you can view $L(s,\chi)$ as obtained by introducing a character of

$\mathrm{Cl}(K)$, but we could *also* have introduced a character of $(K\otimes\mathbb{R})^*/\mathcal{O}_K^*$. A combination of these is an *idele class character*.

Definition 2.4.3. Let $\chi:\mathbb{A}_K^*/K^*\to\mathbb{C}^*$ be a character. Define

$$L(s,\chi)=\prod_{\mathfrak{p}:\,\chi|_{K_\mathfrak{p}}\text{ unramified}}\left(1-\chi(\pi_\mathfrak{p})(\mathrm{Nm}\,\mathfrak{p})^{-s}\right)^{-1}$$

where $\pi_\mathfrak{p}$ is a uniformizer of $K_\mathfrak{p}^*\hookrightarrow\mathbb{A}_K^*$.

Similarly you get analytic continuation and functional equation:

$$L(s,\chi)=(\epsilon-\text{factor})L(1-s,\chi^{-1}).$$

Example 2.4.4. For any field, we have a surjection

$$\mathbb{A}_K^*/K^*\to\mathbb{A}_K^*/K^*(\mathbb{A}_K^f)^*\prod_\mathfrak{p}\mathcal{O}_\mathfrak{p}^*=\mathrm{Cl}(K).$$

So any character of the class group induces a character of $\mathbb{A}_K^*/K^*$ by pull-back, and the corresponding L-function is the one we just discussed.

For $k=\mathbb{Q}$, the inclusion of $\prod_p\mathbb{Z}_p^*$ into $\mathbb{A}_\mathbb{Q}^*$ gives

$$\mathbb{A}_\mathbb{Q}^*/\mathbb{Q}^*\mathbb{R}_{>0}\cong\prod_p\mathbb{Z}_p^*=\varprojlim_N(\mathbb{Z}/N)^*.$$

So you can view any Dirichlet character as a character on the idele class group, and the corresponding L-function is a Dirichlet L-function.

Example 2.4.5. For $K=\mathbb{Q}(i)$ (which has class number 1),

$$\mathbb{A}_K^*=K^*\mathbb{C}^*\prod_\mathfrak{p}\mathcal{O}_\mathfrak{p}^*.$$

The inclusion $\mathbb{C}^*\hookrightarrow\mathbb{A}_K^*$ induces an isomorphism

$$\mathbb{C}^*/\langle i\rangle\cong\mathbb{A}_K^*/K^*\prod_\mathfrak{p}\mathcal{O}_\mathfrak{p}^*.$$

So a character of $\mathbb{C}^*$ trivial on i gives a character of $\mathbb{A}_K^*/K^*$. One such character is

$$\chi:z=re^{i\theta}\mapsto e^{4i\theta},$$

and the corresponding L-function is

$$L(s,\chi)=\prod_\mathfrak{p}\left(1-\frac{e^{4i\theta_\mathfrak{p}}}{(\mathrm{Nm}\,\mathfrak{p})^s}\right)^{-1}$$

where if you write $\mathfrak{p}=(a+bi)$, $a+bi=re^{i\theta_\mathfrak{p}}$. This is equal to

$$L(s,\chi)=\sum_{a+bi\in\mathbb{Z}[i]/\text{units}-\{0\}}\frac{e^{4i\theta_{a+bi}}}{(a^2+b^2)^s}.$$

Remark 2.4.6. If the character was instead $re^{i\theta} \mapsto r^t e^{4i\theta}$, then r^t, then you would find instead

$$\prod_{\mathfrak{p}} \left(1 - \frac{(\mathrm{Nm}\,\mathfrak{p})^t e^{4i\theta_{\mathfrak{p}}}}{(\mathrm{Nm}\,\mathfrak{p})^s}\right)^{-1}$$

which correspond to just a translation.

For context, there is an *adelic norm* $\mathbb{A}_K^*/K^* \xrightarrow{|\cdot|_{\mathbb{A}}} \mathbb{R}_{>0}$ sending $(x_v) \mapsto \prod |x_v|_v$. In general, $L(s,\chi|\cdot|_{\mathbb{A}}^t) = L(s+t,\chi)$.

3. L-FUNCTIONS: ARITHMETIC PROPERTIES

3.1. **Rationality.** The Riemann zeta function $\zeta_{\mathbb{Q}}$ has the interesting property that $\zeta_{\mathbb{Q}}(2k) \in \pi^{2k}\mathbb{Q}$ for $k > 0$. By the functional equation, this implies tha $\zeta(1-2k) \in \mathbb{Q}$ for $k > 0$. Also $\zeta(-2k) \in \mathbb{Q}$ for trivial reasons, as they're all 0: $\zeta(0) = -\frac{1}{2}$, while $0 = \zeta(-2) = \ldots = 0$. (The fact that $\zeta(0) \neq 0$ is somewhat of an anomaly, coming from the pole at 1.)

Similarly, if $\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}$, then $L(-m,\chi) \in \mathbb{Q}$ for $m > 0$ (and $m = 0$ if χ is nontrivial). Again, $L(-m,\chi)$ will vanish often: if $\chi(-1) = 1$, i.e. χ is even, then it vanishes for $m \geq 0$ even; if $\chi(-1) = -1$, i.e. χ is odd, then $\chi(-m)$ vanishes for $m \geq 1$.

Set $L^{(p)}$ to be L omitting the Euler factor at p, i.e.

$$\begin{aligned} L^{(p)}(s,\chi) &= L(s,\chi)\left(1 - \frac{\chi(p)}{p^s}\right) \\ &= \sum_{\substack{n=1 \\ (n,p)=1}}^{\infty} \frac{\chi(n)}{n^s}. \end{aligned}$$

It is a general fact that for any idele class character of finite order, the ζ values at negative integers will be rational. This interest phenomenon prompts us to investigate the p-adic behavior of the ζ function.

3.2. **p-adic continuation.** Let χ be a non-trivial Dirichlet character with conductor N, and p be a prime such that $\gcd(N,p) = 1$. Then as a general principle,

> the function $k \mapsto L^{(p)}(k,\chi)$ inherits the p-adic properties of $k \mapsto n^k$ for $(n,k) = 1$.

What does this mean? If $n^k \equiv n^{k'} \pmod{p^r}$ for all $(n,p) = 1$, then $L^{(p)}(k,\chi) \equiv L^{(p)}(k',\chi) \pmod{p^r}$. For example, if $k \equiv k' \pmod{p-1}$ then $L^{(p)}(k,\chi) \equiv L^{(p)}(k',\chi) \pmod{p}$. The intuition is that the L-function behaves like a *finite* sum.

Example 3.2.1. Suppose $\chi : (\mathbb{Z}/4\mathbb{Z})^* \to \mathbb{C}^*$ is the character with $\chi(1) = 1$, $\chi(3) = -1$. If $p = 5$, then we our claim is that $L^{(5)}(\chi,-2) \equiv L^{(5)}(\chi,-6) \pmod{5}$. This is difficult to see because these values lie outside the range where the zeta function can be evaluated by the series. While we know that the function extends by analytic continuation, this is not a robust way of thinking about it (at least for computation).

The two sums "are" $1^2 - 3^2 - 7^2 + 9^2 - \ldots$ and $1^6 - 3^6 - 7^6 + 9^6 - \ldots$. The content of the hypothesis is that they are *termwise* congruent mod 5, which motivates the assertion that their "values" are congruent mod

5, but we need a concrete way to evaluate these to know that this isn't screwed somehow up by analytic continuation.

Recall the philosophy we mentioned earlier that the precise analytic properties involved in extending ζ are not essential; any way of "making sense" of the sum must be "right."

For this example, there are two methods (which are in some sense the same). The first is a really clever trick is due to Euler: introduce a variable q, and write

$$1^2q - 3^2q^3 + 5^2q^5 - 7^2q^7 + 9^2q^9 - \ldots = \frac{q(q^2+2q-1)(q^2-2q-1)}{(1+q^2)^3} \in \mathbb{Q}(q).$$

Evaluating at $q = 1$, the right hand side is $-1/2$, confirming that $\zeta(-2) = -1/2$. Now this isn't quite what we want, as we've included the terms divisible by 5, but it illustrates the point.

Here's another, even more hands-on way to arrive at the same result. Write

$$\begin{array}{lccccccc} S = & 1 & -9 & +25 & -49 & +81 & -\ldots \\ S = & & 1 & -9 & +25 & -49 & +\ldots \end{array}$$

Adding this with a copy shifted to the right by 1, we get

$$2S = [-8+16-24+\ldots]+1 = -8(\underbrace{1-2+3-4\ldots}_{})+1.$$

Similarly, $2T = 1-1+1-1+\ldots$. This last thing is $1/2$ by the same reasoning, so $T = 1/4$. Then $-8T+1 = -1$, so $S = -1/2$.

3.3. **Abstract sequence spaces.** *Why* is it the case that these much more "algebraic" methods give the same answer as analytic continuation?

We can answer this question with some general abstractions on sequences and series.

Definition 3.3.1. Let K be a field of characteristic 0, and let V be the set of functions $f\colon \mathbb{N} \to K$ ($\mathbb{N} = \{1,2,\ldots\}$)such that for all n sufficiently large,

$$f(n) = \sum_i a_i \alpha_i^n n^{k_i}, \quad a_i \in \overline{K}, \alpha_i \in \overline{K}, k_i \in \{0,1,\ldots\}.$$

Let V_c be the subset of functions with compact support, i.e. $f(n) = 0$ for all large enough n.

Definition 3.3.2. Let $S\colon V \to V$ be the right shift operator, $Sf(n) = f(n-1)$, and $Sf(0) = 0$.

Note that another way of viewing V is as functions that eventually satisfy a linear recurrence. (It suffices to check this for $f(n) = \alpha^n n^k$, and this is clear by considering successive differences.) Said differently, for any $f \in V$, the span of $f, Sf, S^2f, \ldots$ in V/V_c is finite-dimensional. That gives a more *intrinsic* characterization of V.

Definition 3.3.3. For $f \in V$, we call the generalized eigenvalues of S on $V_f := \mathrm{Span}(f, Sf, S^2 f, \ldots) \subset V/V_C$ are the *exponents* of f.

Example 3.3.4. If f is eventually $\alpha^n n^k$, then $Sf(n)$ is eventually $\alpha^{n-1}(n-1)^k$, so $(1-\alpha S)f(n)$ is eventually $\alpha^n(n-(n-1)^k)$. This decreases the degree of the factor which is a polynomial in n. Thus we see that the exponent of f is $1/\alpha$.

In general, if

$$f(n) = \sum_i a_i \alpha_i^n n^{k_i} \quad n \gg 0$$

then the exponents are $\{1/\alpha_i\}$.

Definition 3.3.5. Let $D_m : V \to V$ be "dilation by m," i.e.

$$D_m f(n) = \begin{cases} f(n/m) & m \mid n \\ 0 & \text{otherwise} \end{cases}.$$

Exercise 3.3.6. Check that this preserves V.

Example 3.3.7. $D_2(1,2,3,\ldots) = (0,1,0,2,0,3,0\ldots)$.

Definition 3.3.8. Let $V^{\neq 1} = \{f \in V : 1 \text{ is not an exponent of } f\}$. (This means that f doesn't contain a term that is a pure polynomial.)

Proposition 3.3.9. *Let $\Sigma : V_c \to K$ be the summation map*

$$\Sigma(f) = \sum_n f(n).$$

Then

(1) Σ extends uniquely to an S-invariant functional $V^{\neq 1} \to K$.
(2) The extended Σ on $V^{\neq 1}$ is D_m-invariant (for all m), and it extends uniquely to a D_m-invariant $V \to K$.

Example 3.3.10. Note that Σ *cannot* extend to an S-invariant functional on all of V - consider $f = (1,1,1,\ldots)$. Then $Sf = (0,1,1,\ldots)$. Any extension would have to be equal on these f and Sf, but at the same time satisfy $\Sigma(f - Sf) = 1$.

Proof. (1) An S-invariant functional on W is the same as a functional $W/(S-1)W \to K$. But we claim the inclusion $V_c \hookrightarrow V^{\neq 1}$ induces an isomorphism

$$V_c/(S-1) \cong V^{\neq 1}/(S-1).$$

Why? We have an exact sequence

$$0 \to V_c \to V^{\neq 1} \to V^{\neq 1}/V_c \to 0.$$

But $(S-1)$ induces an isomorphism $V^{\neq 1}/V_c \to V^{\neq 1}/V_c$, because $V^{\neq 1}/V_c$ is a sum of generalized eigenspaces and the eigenvalues are $\neq 1$. Now applying the snake lemma to

$$\begin{array}{ccccccccc} 0 & \longrightarrow & V_c & \longrightarrow & V^{\neq 1} & \longrightarrow & V^{\neq 1}/V_c & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle S-1} & & \downarrow{\scriptstyle S-1} & & \downarrow{\scriptstyle S-1} & & \\ 0 & \longrightarrow & V_c & \longrightarrow & V^{\neq 1} & \longrightarrow & V^{\neq 1}/V_c & \longrightarrow & 0 \end{array}$$

yields the result. This immediately implies (1).

In all examples, there's a more concrete way to see this. Every $f \in V^{\neq 1}$ is of the form $f = (S-1)f' + h$ where $h \in V_c$. Then define $\Sigma(f) = \Sigma(h)$.

For example, if $f = (1,-2,3,-4,\ldots) \in V$, the exponent of f is -1, so there should be a unique shift-invariant way to define its sum. We have $Sf = (0,1,-2,3,-4,\ldots)$, and adding this back to f and using S-invariance shows that $2\Sigma(f) = \Sigma(1,-1,1,-1,\ldots)$. By the same reasoning, this value is $\frac{1}{2}$, so $\Sigma(f) = 1/4$.

(2) We're only interesting in $m = 2$, so we'll work it out in this case, but the general case is similar. On V_c, $\Sigma(D_2 f) = \Sigma f$. We have to check that $\Sigma(D_2 f) = \Sigma(f)$ for $f \in V^{\neq 1}$. First note that the exponents of $D_2(f)$ are $\sqrt{\lambda}$ for λ an exponent of f (with either sign), so $D_2(f) \in V^{\neq 1}$. Now, it suffices to check that $\Sigma(D_2 f)$ is shift-invariant by the uniqueness of the characterization in (1). But $D_2(Sf) = S^2(D_2 f)$, so $\Sigma(D_2 Sf) = \Sigma(D_2 f)$, so $\Sigma \circ D_2 = \Sigma$. This shows that Σ on $V^{\neq 1}$ is indeed D_2-invariant.

To show that there is a unique extension of Σ from $V^{\neq 1}$ to V as D_2-invariant functionals, we need

$$D_2 - 1 \colon V/V^{\neq 1} \cong V/V^{\neq 1}.$$

If we could establish this, then the result follows from the same argument as we gave for S. But $V/V^{\neq 1}$ is simply the space of polynomials, so it suffices to examine

$$D_2(n^k) = \underbrace{\frac{1+(-1)^n}{2}}_{=\chi_{2\mathbb{Z}}} \left(\frac{n}{2}\right)^k \equiv \frac{n^k}{2^{k+1}} \pmod{V^{\neq 1}}$$

because $(-1)^n$ is a function of exponent -1.

So $V/V^{\neq 1}$ is a sum of D_2-eigenspaces, with eigenvalues 2^{-k-1} for $k \geq 0$. This implies that $D_2 - 1$ is invertible. □

3.4. **Application to rationality.** Now let's apply this to get the rationality of ζ values. Suppose K is a p-adic field, with $\mathcal{O}_K$ its ring of integers. Let

$$\Lambda = \{f \colon \mathbb{N} \to \mathcal{O}_K \colon \text{exponents } \lambda \text{ satisfy } |\lambda - 1| = 1\}.$$

(i.e. the exponents are units in $\mathcal{O}_K$, and no exponent is congruent to 1). This is certainly a subset of $V^{\neq 1}$.

Proposition 3.4.1. *For $f \in \Lambda$, we have $\Sigma(f) \in \mathcal{O}_K$.*

Proof. For $f \in \Lambda$, let V_f be the span of $S^i f$ in V/V_c. On V_f, all the eigenvalues of $S-1$ are units by the assumptions. So as endomorphisms on V_f, we have $(S-1)^{-1} = P(S)$ where $P \in \mathcal{O}_K(T)$ (by Cayley-Hamilton, its entries are integral).

Set $f' = P(S)f \in V$. Then $(S-1)f' = f + h$ where $h \in V_c$, since $(S-1)P(S)$ is the identity on V/V_c. So $\Sigma(f) = -\Sigma(h)$, but by the equation h takes values in $\mathcal{O}_K$, so the right hand side is in $\mathcal{O}_K$. □

An immediate consequence of this is an analogue of the "p-adic continuity" result mentioned earlier.

Corollary 3.4.2. *For $f, g \in \Lambda$, if $f \equiv g \pmod{\pi^r}$ then $\Sigma(f) \equiv \Sigma(g) \pmod{\pi^r}$.*

That means we can extend the functional Σ to the *closure* of Λ for the uniform p-adic topology. The closure is much larger, in a way that we will use crucially later.

Now we want to compare our abstract results with those obtained by analytic continuation, so let $K = \mathbb{C}$.

Proposition 3.4.3. *Let $f \in V$ be such that all inverse exponents α_i satisfy $|\alpha_i| \leq 1$ (which is satisfied when f is a Dirichlet character, for instance). Then $\sum \frac{f(n)}{n^s}$ has a meromorphic extension to $s \in \mathbb{C}$, and its value is $\Sigma(f)$.*

Remark 3.4.4. From this you can deduce results for $s < 0$ by replacing $f(n)$ by $f(n)n^k$, which doesn't change the exponents.

Proof. Write f as a polynomial plus f', where $f' \in V^{\neq 1}$. Any n^k monomial in f contributes $\sum \frac{n^k}{n^s} = \zeta(s-k)$. Therefore, it suffices to study the case of $f' \in V^{\neq 1}$.

First we show the existence of an analytic continuation. If $f' \in V^{\neq 1}$, we may write $f' = (S-1)f'' + h$ where $h \in V_c$. So

$$\sum \frac{f'(n)}{n^s} = \sum f''(n)\left(\frac{1}{(n+1)^s} - \frac{1}{n^s}\right) - \sum_n \frac{h(n)}{n^s}.$$

The term $\sum \frac{h(n)}{n^s}$ is obviously analytic, so we have to show that the first term on the right hand side has a meromorphic extension. The point here is that it is"more convergent" because $\frac{1}{(n+1)^s} - \frac{1}{n^s}$ decays faster than $\frac{1}{n^s}$: indeed, $\frac{1}{(n+1)^s} - \frac{1}{n^s} \approx \frac{1}{n^{s+1}}$.

By iterating this argument, the expression on the right hand side converges for larger and larger half-plane. For instance, the second difference

$$\frac{1}{n^s} - \frac{2}{(n+1)^s} + \frac{1}{(n+2)^s}$$

is equal to n^{-s} times the second difference of x^{-s} at 1 with step $1/n$, which can be controlled by the Taylor series.

An important technical point is that f' is bounded by a polynomial since it's in V with $|\alpha_i| \leq 1$, and f'' is bounded by the *same* polynomial (otherwise this argument couldn't work), because $f' \in V^{\neq 1}$.

Now that we see that analytic continuation gives one method of "evaluating" the series at $s = 0$, we can sensibly claim that it agrees with $\Sigma(f)$. To prove this, it suffices to check that evaluation at 0 is dilation-invariant, and shift-invariant function on $V^{\neq 1}$. The dilation invariance is easy: at least for $\operatorname{Rep} s \gg 0$, D_m takes the series to

$$\sum \frac{f(n)}{(nm)^s} = m^{-s} \sum \frac{f(n)}{n^s}.$$

As both sides are analytic, this holds true for all s, and the factor of m^{-s} clearly doesn't affect the evaluation at $s = 0$.

Shift invariance is a little more involved. We only have to check it for $f \in V^{\neq 1}$. (The problem with 1 is that if some $\alpha_i = 0$, then you'll get a pole.) We are interested in

$$I(s) := \sum \frac{f(n)}{n^s} - \sum \frac{f(n+1)}{n^s} = \sum \frac{f(n)}{n^s} \left(\frac{1}{(1+1/n)^s} - 1 \right).$$

You can expand this multiplier term as

$$\left(\frac{1}{(1+1/n)^s} - 1 \right) = -\frac{s}{n} + \frac{(-s)(-s-1)}{2!} \frac{1}{n^2} + \ldots + \frac{\operatorname{poly}(s)}{n^k} + R_k(s,n).$$

The point is that all remainder terms are divisible by s, so you get 0 at $s = 0$. (That is what breaks down if there are poles.) To make this rigorous, we can write

$$\frac{(1+1/n)^{-s} - 1}{s} = -\frac{1}{n} + \ldots + \frac{P_{k-1}(s)}{n^{k-1}} + R_k(s,n).$$

As f is bounded by a polynomial, we may choose k large enough so that $|f(n)| \leq C n^{k-2}$. Then $\sum \frac{f(n)}{n^s}$ is absolutely convergent for $\operatorname{Rep} s \geq k$. Let's go back to the expression

$$I(s) = s \sum_n \frac{f(n)}{n^s} \left(-\frac{1}{n} + \ldots + \frac{P_{k-1}(s)}{n^{k-1}} + R_k(s,n) \right).$$

For $|s| \leq k+1$, we have a Taylor estimate $|R_k(s,n)| \leq \frac{C_k}{n^k}$. So the series is absolutely convergent near $s = k$. Hence, at least in an open neighborhood of $s = k$, we have

$$I(s) = s\left(-\sum \frac{f(n)}{n^{s+1}} + \ldots + P_{k-1}(s)\sum \frac{f(n)}{n^{s+k-1}} + \sum f(n)R_k(s,n)\right)$$

and the term $\sum f(n)R_k(s,n)$ is analytic in $|s| \leq k+1, \operatorname{Re} s > -1/2$. The other terms all admit analytic continuations, by our preceding discussion. So finally we can *rigorously* say that the right hand side is "divisible by s," hence $I(0) = 0$. □

Corollary 3.4.5. *For any $\chi : (\mathbb{Z}/n\mathbb{Z})^* \to \mathbb{C}^*$, we have $L(-k,\chi) \in \mathbb{Q}(\chi)$ (the field generated by values of χ) for $k \geq 0$.*

4. p-ADIC L-FUNCTIONS

4.1. **Analyticity for real quadratic fields.** Let χ be a Dirichlet character mod Q, with $(Q,p)=1$ and $\chi \neq 1$. (So this does *not* apply to the zeta function, and we'll have to revisit and correct that later.) The exponents of $n \mapsto \chi(n)$ are all non-trivial Qth roots of unity. (This is easier to see from the description of "eventually satisfying a linear recurrence." You can see that 1 is not an exponent, since the sum of χ over a period is 0, which would not be the case of 1 was an exponent.) Removing the Euler factor at p, we get

$$L^{(p)}(s,\chi)=\left(1-\frac{\chi(p)}{p^s}\right)L(s,\chi)=\sum_{(n,p)=1}\frac{\chi(n)}{n^s}.$$

So

$$L(-k,\chi)\text{“}=\text{”}\sum_n \chi(n)n^k g(n), \quad \text{where} \quad g(n)=\begin{cases}1 & p\nmid n,\\ 0 & p\mid n,\end{cases}$$

The exponents of $g(n)$ are pth roots of 1, so the exponents of $n\mapsto\chi(n)n^k g(n)$ are all of the form $\zeta_Q\zeta_p$ where ζ_Q is a non-trivial Qth root of unity and ζ_p is a pth root of unity.

In order to apply Proposition 3.4.1, we need to show that this is not congruent to 1 mod p. To that end, note that

$$|\zeta_Q\zeta_p-1|=|\zeta_Q-\zeta_p^{-1}|$$

so it's enough to show that $|\zeta_Q-1|=1$ because $|\zeta_p^{-1}-1|<1$. But notice that

$$\prod_{\zeta_Q\neq 1}(\zeta_Q-1)=\left.\frac{x^Q-1}{x-1}\right|_{x=1}=Q\in\mathbb{Z}_p^\times.$$

Therefore, $\chi(n)n^k g(n)\in\Lambda$. Therefore, Proposition 3.4.1 says that if $n^k\equiv n^{k'}\pmod{p^r}$ for all n (e.g. $k\equiv k'\pmod{(p-1)p^{r-1}}$), then $L^{(p)}(-k,\chi)\equiv L^{(p)}(-k',\chi)(p^r)$.

So for $m\in\mathbb{N}$, $m\mapsto L^{(p)}(-k_0-(p-1)m,\chi)$ extends from $\mathbb{N}$ to a p-adically continuous function $\mathbb{Z}_p\to K$ (since $m\mapsto n^{k_0+(p-1)m}$) is p-adically continuous for all $(n,p)=1$. Unfortunately, p-adic continuity is a nearly useless condition because it is so weak, but something much stronger is true: it is even given by a power series in M, convergent for $|m|\leq 1+\epsilon$.

Namely, if $(n,p)=1$ then we can write $n^{p-1}=1+pn'$ where $n'\in\mathbb{Z}$, so then

$$(n^{p-1})^m=(1+pn')^m=1+(pn')m+\frac{m(m-1)}{2}(pn')^2+\ldots.$$

This converges as a power series in m because of the terms have increasing p-adic valuation. This may not be so clear from the expression above

(because it is not written as a power series in m), but it can be expressed alternately as

$$(1+pn')^m = e^{m\log(1+pn')} = \sum \frac{(\log(1+pn'))^k}{k!} m^k.$$

Since $v_p(k!) \approx \frac{k}{p-1}$ and $v_p(\log(1+pn')) \geq 1$, this is convergent when $v_p(m) > -1+\frac{1}{p-1}$.

Proposition 4.1.1. *Suppose that for each $k \in \mathbb{N}$, we have*

$$f_k(n) = a_0(n) + a_1(n)k + a_2(n)k^2 + \ldots \in \Lambda$$

such that for fixed n, $k \mapsto f_k(n)$ converges uniformly in n when $|k| \leq R$ for some $R > 1$, i.e. $|a_i(n)|R^{-i} \ll C$. Then

$$k \mapsto \sum_n f_k(n) = b_0(n) + b_1(n)k + \ldots$$

also satisfies $|b_i(n)|R^{-i} \ll C$, i.e. converges uniformly in n when $|k| \leq R$.

This is a p-adic analogue analogue of the fact that a uniformly convergent sum of complex-analytic functions is complex-analytic.

Since we just saw that $m \mapsto n^{(p-1)m}$ converges when $v_p(m) > -1+\frac{1}{p-1}$, the Proposition implies that $m \mapsto L^{(p)}(-k_0-(p-1)m, \chi)$ extends to a p-adic analytic function if $v_p(m) > -1+\frac{1}{p-1}$.

Concretely, this means that for m such that the series expression does not converge,

$$L^{(p)}(-k_0-(p-1)m, \chi) \to L^{(p)}(-s, \chi)$$

as $-k_0-(p-1)m$ ranges over integers approximating s to high degree. As this evaluation is indirect, and limits are taken in a p-adic sense, it is a remarkable phenomenon (which we shall see) that one often gets the same answer as in the complex-analytic case!

Proof. Let Λ^* be the closure of Λ inside the set of functions $\mathbb{N} \to K$ for the uniform topology, i.e. the topology defined by the norm

$$||f-g|| = \sup_n |f(n)-g(n)|.$$

Then $\Sigma\colon \Lambda \to \mathcal{O}$ extends to $\Sigma\colon \Lambda^* \to \mathcal{O}$ by general properties on extensions of continuous functionals, since $f, g \in \Lambda$ and $f \equiv g \pmod{\pi^r}$ imply that $\Sigma(f) \equiv \Sigma(g) \pmod{\pi^r}$.

Example 4.1.2. Suppose $p \neq 2$. Then the following interesting function is in Λ^*:

$$n \mapsto \begin{cases} (-1)^n/n & p \nmid n \\ 0 & p \mid n \end{cases}$$

Indeed, $n^{-1} = \lim_{m\to\infty} n^{(p-1)p^m-1}$ for $(n,p) = 1$ (here is where we need $p \neq 2$), because $n^{p-1} \equiv 1 \pmod{p}$.

Note that up to some uniform constant multiple, $a_0(n), a_1(n), \ldots \in \Lambda^*$. For instance, $f_0(n) = a_0(n) \in \Lambda$, and

$$\begin{aligned}\frac{f_{p^m}(n) - f_0(n)}{p^m} &= p^{-m}(p^m a_1(n) + p^{2m} a_2(n) + \ldots) \\ &= a_1(n) + p^m a_2(n) + \ldots\end{aligned}$$

so $\lim_{m\to\infty} \frac{f_{p^m}(n)-f_0(n)}{p^m} = a_1(n)$ expresses $a_1(n)$ as a uniform limit of functions in Λ.

In the higher-order terms, there are some factorials that appear from the derivatives, but we've assumed $|a_i(n)| \leq R^i C$. Since $\Sigma\colon \Lambda^* \to \mathcal{O}_K$ is continuous,

$$\sum_n f_k(n) = \sum_n a_0(n) + k\sum_n a_1(n) + k^2 \sum_n a_2(n) + \ldots$$

but $a_i(n) \in \Lambda^*$ and $|a_i(n)| \leq CR^i$, hence $\sum_n a_i(n) \leq CR^{-i}$ (and exists by the extension property), which gives a convergent power series for $\sum_n f_k(n)$. □

This gives a p-adic analytic continuation for $L^{(p)}(s,\chi)$ when χ has conductor $Q > 1$ and $(Q,p) = 1$ (which latter restriction was to get the exponents not to be congruent to 1 (mod p)).

What about other cases, e.g. ζ itself? Then

$$\zeta(-k) = \sum_{n=1}^{\infty} n^k.$$

In this case the key is to use the dilation operator. Let $\ell \neq p$ be an auxiliary prime (e.g. $\ell = 2$). Then

$$\begin{aligned}(1-\ell^{1+k})\zeta(-k) &= \sum_{n=1}^{\infty} n^k - \ell \sum_{\substack{n \\ \ell | n}} n^k \\ &= \sum_{n=1}^{\infty} n^k \begin{cases} 1 & \ell \nmid n \\ -(\ell-1) & \ell \mid n \end{cases}\end{aligned}$$

For instance, if $\ell = 2$ then you get $1^k - 2^k + 3^k \ldots$ and the only exponent is -1. Now you can proceed as before. You can see that the exponents in general are ζ_ℓ for ζ_ℓ a non-trivial ℓth root of 1, because the "average" over a period is 0 (if 1 were an exponent, you would pick up a non-trivial contribution over every period). (This trick is analogous as the way we used D_m to extend Σ from $V^{\neq 1}$ to V.)

Now proceed as before and you get the same p-adic properties for the function $(1-\ell^{1+k})\zeta(-k)$. This is satisfactory *except* when $k-1$. What happens then?

Take $\ell=2$. Then

$$(1-2^{1+k})\zeta(-k)=1^k-2^k+3^k-\dots$$

and

$$(1-2^{1+k})\zeta^{(p)}(-k)=\sum_{\substack{n\\(n,p)=1}}\underbrace{n^k(-1)^{n+1}}_{\in\Lambda^*}.$$

Since $(1-2^{1+k})$ is 0 at $k=-1$, on might expect the factor $\zeta^{(p)}(-k)$ to have a pole.

Proposition 4.1.3. *The p-adic analytic continuation $\zeta^{(p)}(-k)$ has a simple pole at $k=-1$ with the same residue as the complex analytic continuation $\zeta^{(p)}(-s)$ at $s=-1$, i.e. if $s_M=-1+(p-1)p^M$ then*

$$\lim_{M\to\infty}\zeta^{(p)}(-s_M)(s_M+1)=\lim_{s\to-1,s\in\mathbb{C}}\zeta^{(p)}(-s)(s+1).$$

Remark 4.1.4. More precisely, this value is $-(1-\frac{1}{p})$, since the residue of the usual ζ function at 1 is 1, but we've negated and removed the Euler factor at p.

It's very interesting that the result is true (approximating -1 by integers and approximating it by complex numbers give the same result?!).

Let's first give a heuristic calculation. Formally, $(1-2^{1+k})$ has Taylor series expansion $-(k+1)\log 2+\dots$ about $k=-1$. The right hand side is

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\dots$$

but omitting the terms divisible by p. Formally, this is $\log 2-\frac{1}{p}\log 2$. Dividing, you get $-(1-1/p)$. To make this rigorous, we first need to define the p-adic logarithm.

Definition 4.1.5. There is a unique function

$$\log_p:\mathbb{C}_p\to\mathbb{C}_p$$

($\mathbb{C}_p$ is the completion of $\overline{\mathbb{Q}_p}$) which we call the *p-adic logarithm*, satisfying

(1) $\log_p(1+x)=x-\frac{x^2}{2}+\dots$ if $|x|<1$
(2) $\log_p(xy)=\log_p(x)+\log_p(y)$ (which extends it to all units in $\mathcal{O}_{\mathbb{C}_p}$)
(3) $\log_p(p)=0$ (this part is arbitrary)

Proof. It suffices to show that

$$\sum_{n=1,(n,p)=1}^{\infty} \frac{(-1)^{n+1}}{n} = \left(1-\frac{1}{p}\right)\log_p 2.$$

Again, we need to use some sort of trick to evaluate this. We'll think of the function $n \mapsto \begin{cases} \frac{(-1)^{n+1}}{n} & (n,p)=1 \\ 0 & \text{otherwise} \end{cases}$ as lying in Λ^*, and apply the regularized summation method.

Note that for $|u|<1$,

$$\begin{aligned}\sum_{\substack{n=1\\(n,p)=1}}^{\infty} \frac{(-1)^{n+1}u^n}{n} &= \log_p(1+u) - \frac{1}{p}\log_p(1+u^p) \\ &= \frac{1}{p}\log_p\left(\frac{(1+u)^p}{1+u^p}\right) \\ &= \frac{1}{p}\log_p\left(1+\frac{(1+u)^p-1-u^p}{1+u^p}\right) \\ &= \frac{1}{p}\log_p\left(1+\frac{pP(u)}{1+u^p}\right)\end{aligned}$$

for some $P(u)\in\mathbb{Z}[u]$. The original expression converged for $|u|<1$, but we've turned it into something better. Now, this last expression can be written as a convergent power series in u and $\frac{1}{1+u^p}$, i.e.

$$\sum a_{m,n}u^m\left(\frac{1}{1+u^p}\right)^n \quad |a_{m,n}|\to 0.$$

The function $n \mapsto \begin{cases} \frac{(-1)^{n+1}}{n}u^n & (n,p)=1 \\ 0 & \text{otherwise} \end{cases}$ is in Λ^* as long as $|u|\le 1$ and $|u+1|=1$, since

(1) $(-1)^{n+1}u^n$ has exponent $-u$, and $|-u-1|=1$ by assumption,
(2) $1_{(n,p)}=1$ is periodic, with exponents ζ_p for ζ_p a pth root of unity, so the exponents of $(-1)^{n+1}u^n 1_{(n,p)=1}$ has exponents $-u\zeta_p$ (since $|\zeta_p-1|<1$, $|1-u\zeta_p-1|=1 \iff |u+1|=1$)
(3) So $\sum\frac{(-1)^{n+1}u^n}{n}$ makes sense for $|u|\le 1$, $|1+u|=1$, coincides with the previous when $|u|<1$.

This shows that $\Sigma^{\Lambda^*}_{(n,p)=1}\frac{(-1)^{n+1}u^n}{n}$ makes sense for $|u|\le 1$ and $|1+u|=1$, and coincides with the previous definition when $|u|<1$.

In fact, we claim that $\Sigma^{\Lambda^*}_{(n,p)=1}\frac{(-1)^n u^n}{n}$ is also given by a convergent power series in u, $\frac{1}{1+u^p}$. Once this is established, the two notions must coincide: a convergent power series in u, $\frac{1}{1+u^p}$ is a rigid analytic function on (say)

$|u| \leq 1, |1+u^p| = 1$ i.e. $|1+u| = 1$. Such a function has only finitely many zeroes (see Fresnel and van der Put Chapter 1, Chapter 3). So once we show that the claim is true, then the difference of the two definitions agrees at infinitely many points, hence is 0.

So to evaluate the sum on Λ^*, we compute

$$\sum_{(n,p)=1} \frac{(-1)^{n+1}u^n}{n} = \lim_{M\to\infty} \underbrace{\sum_{(n,p)=1} n^{p^M(p-1)-1}(-1)^{n+1}u^n}_{A_M}$$

We claim that $A_M(1+u^p)^{N_M} = P_M(u) \in \mathbb{Z}_p[u]$. Indeed, A_M is evidently a rational function. Multiplying by $(1+u^p)$ has the effect of difference the series against the pth translate by it, which you can see will have the effect of taking a pth successive difference in the polynomial term, and hence eventually kill it (up to a finite number of terms).

Now, we write

$$\lim_{M\to\infty} A_M = A_1 + (A_2 - A_1) + (A_3 - A_2) + \dots$$

We've just seen that $A_{M+1} - A_M$ is a polynomial in u and $(1+u^p)^{-1}$, so this is a power series in $u, (1+u^p)^{-1}$. Moreover, A_{M+1} and A_M become congruent modulo higher and higher powers of p, since $n^{p^M(p-1)-1}$ and $n^{p^{M+1}(p-1)-1}$ do, and Proposition 3.4.1, i.e.

$$A_{M+1} - A_M = p^{k_M} \frac{Q_M(u)}{(1+u^p)^{N_M}} \text{ with } k_M \to \infty.$$

That shows that they fit into a convergent power series.

The point here is that $(1+u^p)$ is congruent to 1 modulo p, so multiplying by it doesn't change the p-adic valuation. Then you want to check that for large M, A_{M+1} and A_{M+1} are are very p-adically close. □

4.2. **Totally real fields.**

Example 4.2.1. Let $K = \mathbb{Q}(\sqrt{2})$. This has class number 1, and the unit group is $\mathcal{O}_K^\times = \{\pm 1, 1+\sqrt{2}\}$. So every ideal class $I = (\alpha)$ where α is totally positive, because the units take all positive signs, which is unique up to totally positive units, i.e. multiplication by powers of $(1+\sqrt{2})^2 = 3+2\sqrt{2}$.

In fact, I has a unique generator $\alpha = a + b\sqrt{2}$ in the cone $C = \{(a,b) \mid b \geq 0, 3b \leq 2a\}$. Therefore,

$$\zeta_{\mathbb{Q}(\sqrt{2})}(s) = \sum_{(a,b)\in\mathbb{Z}^2\cap C} \frac{1}{(a^2-2b^2)^s}.$$

Let $u = 3+2\sqrt{2}$. The reason is that by multiplying α by a power of u, we can arrange that $1 \leq \frac{\alpha}{\bar{\alpha}} \leq u^2$. The condition that $\frac{\alpha}{\bar{\alpha}} > 1$ is equivalent to

$a+b\sqrt{2} > a-b\sqrt{2}$, i.e. $b>0$. The condition that $\frac{\alpha}{\bar{\alpha}} \leq u^2$ basically shifts this calculation by u^2, hence corresponds to $3b \leq 2a$.

There's a unique prime ideal of K above 2, namely $(\sqrt{2})$. So

$$(1-2^{1-s})\zeta_K(s) = \sum_{a,b\in C} (a^2-2b^2)^{-s}(-1)^a$$

This is the analogue of the expression

$$(1-2^{1-s})\zeta(s) = 1^{-s} - 2^{-s} + 3^{-s} - \dots.$$

Exercise 4.2.2. How might you do this for a real cubic field?

Now let's say you want to evaluate $\zeta_K(-1)$. According to this,

$$-3\zeta_K(-1)\text{“} = \text{”} \sum_{(a,b)\in C\cap\mathbb{Z}^2} (-1)^a(a^2-2b^2).$$

Here we mean holomorphic continuation, but one can use the same formal tricks as before: this equals the value of the rational function

$$\sum_{(a,b)\in C} (-1)^a(a^2-2b^2)x^a y^b$$

at $(x,y)=(1,1)$. This rational function is always of the form

$$\frac{\text{poly}(x,y)}{(1+x)^A(1+x^3y^2)^B}.$$

(Basically because multiplying by the denominator differences the sequence against the walls of the cone many times.) For one thing, this is evidently rational.

Remark 4.2.3. If we hadn't put in the $(1-2^{1-s})$ term, then we would have encountered a pole in certain cases (e.g. ζ_K).

Cones in $\mathbb{R}^n$. The preceding example motivates the following discussion.

Definition 4.2.4. A *polyhedral cone* in $\mathbb{R}^n$ is the convex hull of rays $\mathbb{R}_{\geq 0}v_i$ for a finite collection of vectors $v_i \in \mathbb{R}^n$. This is equivalent to the locus determined by fintely many linear inequalities, $\{x \in \mathbb{R}^n \mid \ell_\alpha(x) \geq 0\}$.

A *rational polyhedral cone* is the convex hull of rays $\mathbb{R}_{\geq 0}v_i$ for a finite collection of vectors $v_i \in \mathbb{Q}^n$. This is equivalent to the locus determined by finitely many linear inequalities, $\{x \in \mathbb{R}^n \mid \ell_\alpha(x) \geq 0\}$ where ℓ_α has rational coefficients.

A *smooth cone* is the convex hull of $\mathbb{R}_{\geq 0}v_i$, $i=1,\dots k$ where $v_i \in \mathbb{Z}^n$ form part of a $\mathbb{Z}$-basis for $\mathbb{Z}^n \subset \mathbb{R}^n$.

Remark 4.2.5. The terminology "smooth" comes from the theory of toric varieties.

Example 4.2.6. Our cone from the previous example is not smooth, as it is spanned by $(1,0)$ and $(3,2)$. However, it can be cut into smooth cones (in fact, any cone can), by drawing lines through $(1,0)$ and $(2,1)$, and $(2,1)$ and $(3,2)$.

The notion of smooth cones makes sense with $\mathbb{Z}^n, \mathbb{Q}^n, \mathbb{R}^n$ replaced by $\mathcal{O}_K, K, K \otimes \mathbb{R}$. Let $(\mathcal{O}_K)_+$ denote the subset of $\mathcal{O}_K$ which is totally, etc.

Theorem 4.2.7 (Shintani)**.** *There is a finite collection of smooth cones C_i such that $\coprod (C_i^{interior} \cap \mathcal{O}_+)$ is a fundamental domain for $(\mathcal{O}_K)_+/(\mathcal{O}_K^\times)_+$.*

The goal is to generalize the preceding results for the Dedekind zeta function of a real quadratic field. More precisely, let K be a totally real field and $\chi : \mathbb{A}_K^\times / K^\times \to \mathbb{C}^\times$ a *finite order* character (the analogue of a Dirichlet character). Then we can form $L(s,\chi)$ as discussed before.

Theorem 4.2.8. *We have $L(-m,\chi) \in \mathbb{Q}(\chi)$. Moreover, if $\ell \subset K$ is a prime ideal with $\mathrm{Nm}(\ell)$ prime and relatively prime to the conductor of χ, then*

$$(1 - \chi(\ell)(\mathrm{Nm}\,\ell)^{m+1})L(-m,\chi)$$

is integral away from $\mathrm{Nm}\,\ell$.

In addition, if p is a prime of $\mathbb{Q}(\chi)$ relatively prime to $\mathrm{Nm}(\mathrm{cond}(\chi))$ and $\mathrm{Nm}(\ell)$, then

$$m \mapsto (1 - \chi(\ell)(\mathrm{Nm}\,\ell)^{m+1})L(-m,\chi)$$

extends from $k_0 + (\mathrm{Nm}\,p - 1)\mathbb{Z}$ to a p-adic analytic function $\mathbb{Z}_p \to \mathbb{Q}(\chi)_p$.

Example 4.2.9. For $\mathbb{Q}$, this says that $(1-\ell^{1+k})\zeta(-k) \in \mathbb{Z}[1/\ell]$ for all primes ℓ. What happens if you try to apply this to multiples primes at once? If we apply it to distinct primes $\ell_1 \neq \ell_2$, then we find that $(1-\ell_1^{1+k})(1-\ell_2^{1+k})\zeta(-k) \in \mathbb{Z}$. This doesn't necessarily imply integrality: if $k = -1$, we have $\zeta(-1) = -\frac{1}{12}$. Indeed, you can check that $\frac{1}{12}(\ell^2 - 1) \in \mathbb{Z}[\frac{1}{\ell}]$ for all ℓ.

This gives a bound on the denominators appearing in $L(-m,\chi)$, by

$$\gcd_{\ell \gg 0}(\mathrm{Nm}\,\ell - 1) = \#\mu_K = H^0(G_K, \mathbb{Q}/\mathbb{Z}(1))$$

and for higher invariants you get $H^0(G_K, \mathbb{Q}/\mathbb{Z}(k))$. This gives an interpretation of the denominators of Bernoulli numbers in terms of the torsion of the algebraic K-theory. From the point of view of number theory, this is less interesting; the numerators, which correspond to some H^1 group, are more interesting.

Remark 4.2.10. Siegel was the first to prove that $\zeta_K(-m) \in \mathbb{Q}$. In fact, there is an interpretation

$$\chi(\mathrm{SL}_2\,\mathbb{Z}) = \zeta_{\mathbb{Q}}(-1) = -\frac{1}{12}.$$

Here, the Euler characteristic means to take group cohomology Euler characteristic of a torsion-free finite index subgroup of $\mathrm{SL}_2\mathbb{Z}$ in $\mathbb{C}$ (with the trivial action), and then divide by that index. (It's an "orbifold Euler characteristic" of $K(\mathrm{SL}_2\mathbb{Z},1)$.)

More generally, we have the interesting equations:

$$\chi(\mathrm{Sp}_{2n}\mathbb{Z}) = \zeta(-1)\zeta(-3)\dots\zeta(1-2n)$$
$$\chi(\mathrm{Sp}_{2n}\mathcal{O}_K) = \zeta(K,-1)\dots\zeta(K,1-2n).$$

Ideas of Proof. The proof is essentially the same as for $\mathbb{Z}$, but we'll go through the detaills. For simplicity let's just consider $\zeta_K = \sum(\mathrm{Nm}\,\mathfrak{a})^{-s}$. As before, split up into ideal classes.

$$\zeta_K = \sum_{\mathfrak{a}\sim I}(\mathrm{Nm}\,\mathfrak{a})^{-s} = \sum_{\lambda\in I^{-1}/\mathcal{O}_K^\times}\mathrm{Nm}(\lambda I)^{-s}.$$

For s a negative integer, $\mathrm{Nm}(\lambda I)^{-s}$ is a *polynomial* in the coordinates of λ. Choose (as discussed) smooth cones $C_1,\dots,C_r$ giving a fundamental domain for $\mathcal{O}_K^\times$ in I^{-1}. Thus, we have to analyze (regularized) sums of polynomials over lattice points in smooth cones.

A smooth cone C in $\mathbb{Z}^n$ is just the convex hull of some collection of vectors $(v_1,\dots,v_r)$ where the v_i are part of a $\mathbb{Z}$-basis. By changing basis, we may as well assume that $\{x_1\geq 0,\dots x_r\geq 0, x_{r+1}=\dots=x_n=0\}\subset\mathbb{R}^n$. Then $C^{\text{interior}}\cap\mathbb{Z}^n\cong\mathbb{N}^r$, i.e. $(a_1,\dots,a_r,0,\dots 0)$ where $a_i\in\{1,2,3,\dots\}$. The point of smoothness is to give this kind of parametrization.

We want a similar story for $\mathbb{N}^r$ as we had for $\mathbb{N}$. Recall that we defined a subspace

$$V^{\neq 1}\subset\{\text{functions }\mathbb{N}\to\mathbb{C}\}.$$

You can view

$$\underbrace{V^{\neq 1}\otimes\dots\otimes V^{\neq 1}}_{r\text{ copies}}\subset\{\text{ functions }\mathbb{N}^r\to\mathbb{C}\}.$$

Remark 4.2.11. Be warned that functions in the tensor product can have different asymptotics along the different axes.

This gives a summation on functions on $\mathbb{N}^r$ of the form

$$\sum_\alpha f_1^\alpha(x_1)\dots f_r^\alpha(x_\sigma)\text{ where }f_i\in V^{\neq 1}.$$

The regularized summation of this is

$$\sum_\alpha \Sigma(f_1^\alpha)\Sigma(f_2^\alpha)\dots\Sigma(f_r^\alpha).$$

In particular, this contains $P(x_1,\dots,x_r)\alpha_1^{x_1}\dots\alpha_r^{x_r}$ as long as all $\alpha_i\neq 1$.

This last point is important, though it may seem like an artifact. Go back to $\zeta_K = \sum(a^2-2b^2)^m$. We extended Σ to a dilation-invariant functional by using the dilation operators. Concretely, that means

$$\begin{aligned}(1-(\mathrm{Nm}\,\ell)^{1-s})\zeta_K(s) &= \sum_{\mathfrak{a}}(\mathrm{Nm}\,\mathfrak{a})^{-s} - \mathrm{Nm}(\ell)\sum_{\mathfrak{a}}\mathrm{Nm}(\ell\mathfrak{a})^{-s} \\ &= \sum_{\mathfrak{a}}(\mathrm{Nm}\,\mathfrak{a})^{-s}\begin{cases}1 & (\mathfrak{a},\ell)=1\\ 1-\mathrm{Nm}\,\ell & \ell\mid\mathfrak{a}\end{cases}.\end{aligned}$$

What are the "exponents" of this along a cone? It has average 0, but not along every wall. Given $v_1,\dots,v_r\in\mathscr{O}_K$, the function on the cone spanned by $\begin{cases}1 & \ell\nmid x\\ 1-\mathrm{Nm}\,\ell & \ell\mid x\end{cases}$ has exponents all different from 1 if

(1) $\mathrm{Nm}\,\ell$ is prime, and
(2) ℓ doesn't divide any v_i.

Example 4.2.12. What goes wrong without these assumptions? The last one is obviously necessary for the exponents along the v_i-wall to not include 1. For an example of what happens when $\mathrm{Nm}\,\ell$ isn't prime, take $[K:\mathbb{Q}]=2$ and $\ell=(2)$, hence $\mathrm{Nm}\,\ell=4$. The function looks like

1	-3	1	-3	1	-3
1	1	1	1	1	1
1	-3	1	-3	1	-3
1	1	1	1	1	1
1	-3	1	-3	1	-3
1	1	1	1	1	1

You can see that the exponents are problematic along the walls.

5. Hurwitz zeta functions

5.1. **Interlude on analysis.** We begin with some purely analytic results that we'll need later.

Lemma 5.1.1. *Suppose $\varphi(x)$ is a function on $\mathbb{R}_{>0}$ which is rapidly decreasing as $x \to \infty$ (i.e. for all N, there exists c_N such that $|\varphi| < c_N|x|^{-N}$) such that φ has an asymptotic near 0:*

$$\varphi(x) \sim \sum_i a_i x^{\alpha_i}, \quad \alpha_i \to \infty$$

(i.e. for all M, there exists d_M such that $\varphi - \sum_{i=1}^k a_i x^{\alpha_i} < c_M x^M$ for $x \in (0,1)$). Then

$$\int_0^\infty \varphi(x) x^s \frac{dx}{x}$$

extends from $\operatorname{Re} s \gg 0$ *to a meromorphic function of $s \in \mathbb{C}$.*

Proof. There is no problem with the integral for x large, since $\varphi(x)$ decays rapidly. To address the divergence near 0, we write

$$\varphi = \left[\varphi - (\sum_{i=1}^k a_i x^{\alpha_i}) 1_{[0,1]}\right] + (\sum_{i=1}^k a_i x^{\alpha_i}) 1_{[0,1]}.$$

The first term can be made to behave well at 0 by choosing k large enough, by the asymptotic approximation near 0. The second term integrated against $x^s \frac{dx}{x}$ gives

$$\sum_{i=1}^k \int_0^1 a_i x^{s+\alpha_i} \frac{dx}{x} = \sum_i \frac{a_i}{s+\alpha_i} \text{ for } \operatorname{Re} s \gg 0$$

and can therefore by meromorphically continued by the expression on the right hand side. □

Remark 5.1.2. From the proof, we see that the meromorphic continuation has simple poles at $s = -\alpha_i$, with residue a_i.

Consider the space of functions with asymptotics $\varphi \sim \sum a_i x^{\alpha_i} (\log x)^{b_i}$ near 0 and ∞, where $\alpha_i \neq 0$. This is analogous to $V^{\neq 1}$. The lemma shows that on this space, there exists a unique functional $\varphi \mapsto$ "$\int \varphi \frac{dx}{x}$" invariant by $\mathbb{R}^\times$, and which coincides with $\int \varphi$ when the latter converges (analogous to the extension of Σ to $V^{\neq 1}$).

5.2. **Hurwitz zeta functions.**

Definition 5.2.1. For $\alpha > 0$ (this discussion applies more generally for $\operatorname{Re}\alpha > 0$, but we won't use that), we define the *Hurwitz zeta function*

$$\zeta(s,\alpha) = \sum_{n=0}^{\infty} \frac{1}{(n+\alpha)^s}.$$

Example 5.2.2. Notice that $\zeta(s,1) = \zeta(s)$, the usual Riemann zeta function.

Remark 5.2.3. $\zeta(s,\alpha)$ also has a functional equation. Basically, the function

$$\sum_{n=0}^{\infty} \frac{e^{2\pi i\beta}}{(n+\alpha)^s}.$$

has a functional equation that roughly speaking interchanges the role of α and β, because it comes from a Fourier transform.

Proposition 5.2.4. *For fixed α, $\zeta(s,\alpha)$ is meromorphic in s, with a simple pole at $s = 1$ having residue* 1.

Proof. Note that if $\varphi(x) = e^{-\alpha x}$, then we have

$$\int_0^{\infty} \varphi(x) x^s \frac{dx}{x} = \Gamma(s)\alpha^{-s}.$$

Therefore, if

$$\varphi = \sum_{n\geq 0} e^{-(n+\alpha)x}$$

then

$$\int_0^{\infty} \varphi(x) x^s \frac{dx}{x} = \Gamma(s)\zeta(s;\alpha) \text{ for } \operatorname{Re} s \gg 0.$$

(The hypothesis that $\operatorname{Re} s \gg 0$ is needed to ensure that you can indeed interchange the order of summation and integration.)

Now, $\varphi(x) = \frac{e^{-\alpha x}}{1-e^{-x}}$ decays rapidly as $x \to \infty$. Moreover, it has a nice asymptotic near $x = 0$:

$$\frac{1 - \alpha x + \frac{\alpha^2 x^2}{2} - \ldots}{x - \frac{x^2}{2} + \frac{x^3}{6} + \ldots} = \frac{1}{x} + (\frac{1}{2} - \alpha) - \ldots$$

By Remark 5.1.2 $\Gamma(s)\zeta(s,\alpha)$ has simple poles at $s = 1, 0, -1, -2, \ldots$ (at least for "general" α), but $\Gamma(s)$ has poles at $0, -1, -2, \ldots$, so dividing by it to obtain $\zeta(s,\alpha)$ leaves only the pole at $s = 1$.

□

The residue of $\Gamma(s)\zeta(s;\alpha)$ at $s=-k$ is the coefficient of x^k in the above expansion. You can see that it will be a polynomial in α of degree $k+1$, e.g.

$$\begin{aligned}\zeta(-0,\alpha) &= \frac{1}{2}-\alpha \\ \zeta(-1;\alpha) &= -\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12} \\ \vdots \qquad &= \qquad \vdots\end{aligned}$$

Up to normalization, these are the Bernoulli polynomials.

How can we calculate them? Set $P_k(\alpha)=\zeta(-k,\alpha)$, which is basically the residues of $\Gamma(s)\zeta(s;\alpha)$ at $s=-k$ (up to the residue of $\Gamma(-k)$, which is $(-1)^k/k!$). Then $P_k(\alpha+1)-P_k(\alpha)=-\alpha^k$, because the series defining these sums are "shifted" by one term, namely $-\alpha^k$. Explicitly, from the series definition (valid for $\operatorname{Re} s \gg 0$) it is evident that

$$\zeta(s,\alpha+1)-\zeta(s,\alpha)=\sum_{n=0}^{\infty}(n+1+\alpha)^{-s}-(n+\alpha)^{-s}=\alpha^{-s}$$

for $\operatorname{Re} s \gg 0$, and the same holds for all s by analytic continuation. Therefore, we see that

$$1^k+2^k+\ldots+n^k=P_k(0)-P_k(n+1).$$

That determines P_k up to an additive constant, which is pinned down by:

Exercise 5.2.5. Check that for all k,

$$\int_0^1 P_k(\alpha)\,d\alpha=0.$$

There is a formal way to define $\sum(n+\alpha)^{-s}$. Crucially, the sequences considered here are *no longer* shift/dilation invariant. The special values discussed here are defined by analytic continuation, and are *different* from what one would define by the abstract sequence spaces (because of a failure of shift invariance).

5.3. **Explicit evaluations.** Let χ be a Dirichlet character with modulus q. Then

$$\begin{aligned} L(-k,\chi) &= \sum_{\substack{1\le a\le q-1\\ n\ge 0}} \chi(a)(nq+a)^k \\ &= \sum_{1\le a\le q-1} \chi(a)q^k \sum_{n\ge 0}(n+a/q)^k \\ &= \sum_{1\le a\le q-1} \chi(a)q^k P_k(a/q). \end{aligned}$$

From this, it is clear that the value lies in $\mathbb{Q}(\chi)$. Some other things are less clear, like the integrality/p-adic continuity (because the polynomials that appear have different degrees...)

Example 5.3.1. For $k=0$, we get

$$\begin{aligned} L(0,\chi) &= \sum_{1\le a\le q-1} \chi(a)\left(\frac{1}{2}-\frac{a}{q}\right) \\ (\text{if } \chi\neq 1) &= \sum_{1\le a\le q-1} -\chi(a)\frac{a}{q} \end{aligned}$$

Suppose χ is the quadratic character modulo 7 (corresponding to the field extension $\mathbb{Q}(\sqrt{-7})$). Then

$$\sum_{1\le a\le q-1} -\chi(a)\frac{a}{q} = \frac{1}{7}(3+5+6-1-4-2) = 1.$$

This exhibits the interesting general fact that

$$\frac{1}{q}\left(\sum \text{quadratic non-residues} - \sum \text{quadratic residues}\right) > 0.$$

The reason for this is that the functional equation relates $L(0,\chi)$ to $L(1,\chi)$, and the series converges for $\operatorname{Re} s > 1$, where it is evidently positive (e.g. by the Euler product).

By the class number formula for $\mathbb{Q}(\sqrt{-q})$, if χ is the corresponding character then $L(0,\chi) = h_{(\mathbb{Q}\sqrt{-q})}$ unless $\operatorname{disc}\mathbb{Q}(\sqrt{-q}) = -4,-3$.

Real fields. The real quadratic case is more interesting, but if $q<0$, then one just gets $L(0,\chi)=0$. There are two explanations. One is that in relating $L(0,\chi)$ to $L(1,\chi)$, you use $\zeta_{\mathbb{Q}(\sqrt{-q})} = \zeta(s)L(s,\chi)$ and the functional equations for each factor, but this automatically brings in a factor of 0.

Alternatively, you can see from the series definition that the "opposite" a values will cancel out because $\chi(-1)=1$. Indeed, we computed above

that

$$L(0,\chi) = \sum_{1 \le a \le q-1} -\chi(a)\frac{a}{q}$$

and if $\chi(-1) = 1$ then we can pair off a and $q-a$.

In this case, the more interesting value is $L'(0,\chi)$. This is a *very* important computation, and its answer much more interesting. To summarize, we have

$$L(s,\chi) = q^{-s} \sum_{1 \le a \le q-1} \zeta(s; a/q)\chi(a)$$

and we want to know $\frac{d}{ds}\zeta(s;\alpha)|_{s=0}$.

We noticed above that for imaginary quadratic fields, the class number formula was "nicer" when phrased in terms of values at 0. There is a similar story here: the answer is more suggestive in terms of $L'(0,\chi)$, although one could prove it by going over to $L(1,\chi)$.

Proposition 5.3.2. *We have*

$$\frac{d}{ds}|_{s=0}\zeta^*(s,\alpha) = -\frac{1}{2}\log[(1+u)(1-u^{-1})], \quad u = e^{2\pi i\alpha}$$

where

$$\zeta^*(s;\alpha) = \zeta(s;\alpha) + \zeta(s;1-\alpha) = \sum_{n\in\mathbb{Z}} \frac{1}{|n+\alpha|^s}.$$

Proof sketch. We'll prove this up to a constant independent of α, which doesn't affect our application.

The right hand side is easily differentiated (in α). For the left hand side,

$$\frac{d}{d\alpha}\left(\frac{d}{ds}|_{s=0}\zeta(s,\alpha)\right)$$

is meromorphic for $\operatorname{Re}\alpha > 0$ and $s \in \mathbb{C}$. You can see this just by unwinding the proof we gave at the beginning. If you switch the order of differentiation, then you get

$$\frac{d}{ds}\frac{d}{d\alpha}\zeta(s;\alpha) = \frac{d}{ds}(-s\zeta(s+1,\alpha))$$

by differentiating $(n+\alpha)^{-s}$ term-by-term.

Therefore,

$$\begin{aligned}\frac{d}{d\alpha}\frac{d}{ds}|_{s=0}\zeta^*(s,\alpha) &= \frac{d}{ds}|_{s=0}(-s)(\zeta(s+1,\alpha) - \zeta(s+1,1-\alpha)) \\ &= -(\zeta(1,\alpha) - \zeta(1,1-\alpha)) \\ &= -\sum_{n\in\mathbb{Z}} \frac{1}{n+\alpha}.\end{aligned}$$

(In this computation we took the difference of two zeta functions both having simple poles at $s=0$, which canceled out to give a finite answer.)

By considering the poles, we see that this must be proportional to $\frac{\pi}{\tan(\pi\alpha)}$. □

We apply this result to calculating $\frac{d}{ds}|_{s=0}\zeta(s,\alpha)$.

Theorem 5.3.3. *We have*

$$L'(0,\chi) = -\frac{1}{2}\log\prod_{a=1}^{q-1}(1-e^{2\pi i a/q})^{\chi(a)}.$$

Example 5.3.4. For $K=\mathbb{Q}(\sqrt{5})$ and $\chi:(\mathbb{Z}/5)^\times\to\{\pm1\}$ the corresponding quadratic character, and $\xi=e^{2\pi i/5}$ the formula says that

$$L'(0,\chi) = \frac{1}{2}\log\left(\frac{(1-\xi^2)(1-\xi^3)}{(1-\xi)(1-\xi^4)}\right).$$

In fact, $\left(\frac{(1-\xi^2)(1-\xi^3)}{(1-\xi)(1-\xi^4)}\right) = \frac{3+\sqrt{5}}{2} = u^2$, where $u=\frac{1+\sqrt{5}}{2}$ is a fundamental unit for $\mathbb{Q}(\sqrt{5})$. So the formula can be written succintly as $L'(0,\chi)=\log u$.

Visibly, u^2 lies in the quadratic fixed field of $\mathbb{Q}(\xi)$, which is K. As a sanity check, can we see why u^2 is a unit? We can clearly write $\frac{1-\xi^2}{1-\xi}$ as an algebraic integer, and apply the same to $\frac{1-\xi^3}{1-\xi^4}=\frac{(1-(\xi^4)^2}{1-\xi^4}$. Applying the same logic to the inverse, we see that $u\in\mathcal{O}_K^\times$.

More generally, for any prime p, $\frac{1-\zeta_p^a}{1-\zeta_p^b}$ is a unit in $\mathbb{Z}[\zeta_p]$, and is called a *cyclotomic unit*. If n is not a prime power, then we get something even better: $1-\zeta_n$ is already a unit.

In general, if χ is a quadratic character associated to $\mathbb{Q}(\sqrt{q})$, and u is a fundamental unit, then

$$L'(0,\chi) = \log(u^h)$$

(taking a real embedding so u^h is positive) where h is the class number of $\mathbb{Q}(\sqrt{q})$. Therefore,

$$\frac{\prod_{a\in\mathrm{QNR}}(1-\xi^a)}{\prod_{b\in\mathrm{QR}}(1-\xi^b)} = u^{2h}.$$

Again, for sanity let's check that $\frac{\prod_{a\in\mathrm{QNR}}(1-\xi^a)}{\prod_{b\in\mathrm{QR}}(1-\xi^b)}\in\mathbb{Q}(\sqrt{q})$. If we apply the automorphism $\xi\mapsto\xi^\alpha$, where $\alpha\in(\mathbb{Z}/q)^\times$, then it is fixed as long as α is a quadratic residue, so it lies in the (unique!) quadratic subfield $\mathbb{Q}(\sqrt{q})$.

6. Artin L-functions

6.1. **Motivation.** One of the initial clues for Artin L-functions was Hecke's observation concerning divisibility relations between zeta functions. In particular, Hecke proved that if $L/\mathbb{Q}$ is a cubic extension, then $\zeta(s)$ divides $\zeta_L(s)$, i.e. $\zeta(s)=0 \implies \zeta_L(s)=0$.

Example 6.1.1. Let L be the cubic field generated by $\alpha^3=\alpha+1$. This L has discriminant -23, and is in sense the *simplest* cubic field. Then

$$\zeta_L(s)=\prod_p \zeta_{L,p}(s)$$

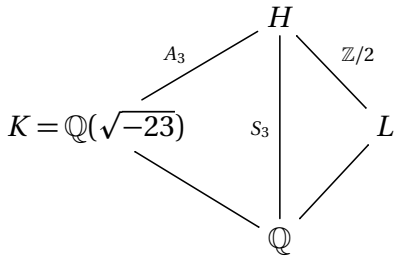
where

$$\zeta_{L,p}(s)^{-1}=\begin{cases}(1-p^{-s})^3 & (p)=\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\\ (1-p^{-s})(1-p^{-2s}) & (p)=\mathfrak{p}_1\mathfrak{p}_2\\ (1-p^{-3s}) & (p)=\mathfrak{p}\\ (1-p^{-s})^2 & (p)=(23)=\mathfrak{p}_1^2\mathfrak{p}_2\end{cases}$$

It's remarkable how well things work out for the ramified places. Now recall that $\zeta(s)=\prod_p(1-p^{-s})^{-1}$, so if we divide by $\zeta(s)$, then that amounts to stripping out a factor of $(1-p)^{-s}$ everywhere. Therefore,

$$\left(\frac{\zeta_{L,p}(s)}{\zeta_p(s)}\right)^{-1}=\zeta_{L,p}(s)^{-1}=\begin{cases}(1-p^{-s})^2 & (p)=\mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3\\ (1-p^{-2s}) & (p)=\mathfrak{p}_1\mathfrak{p}_2\\ (1-\zeta_3 p^{-s})(1-\zeta_3^{-1}p^{-s}) & (p)=\mathfrak{p}\\ (1-p^{-s}) & (p)=(23)=\mathfrak{p}_1^2\mathfrak{p}_2\end{cases} \tag{2}$$

Let H be the Galois closure of L. This is an S_3 extension of $\mathbb{Q}$, with the following subfield lattice.



Now the interesting fact is that H/K is *unramified*. This is a rather general phenomenon in this situation (i.e. a cubic non-Galois extension with square-free discriminant), as we'll see later. You can check this "by hand" by examining the ramification at the discriminant and 2 and 3. So Cl_K is divisible by 3, and in fact we have $\mathrm{Cl}_K\cong\mathbb{Z}/3\mathbb{Z}$.

Rephrasing

$$\left(\frac{\zeta_{L,p}(s)}{\zeta_p(s)}\right)^{-1} = \zeta_{L,p}(s)^{-1} = \begin{cases} (1-p^{-s})^2 & (p) = \mathfrak{p}_1\mathfrak{p}_2\mathfrak{p}_3 \\ (1-p^{-2s}) & (p) = \mathfrak{p}_1\mathfrak{p}_2 \\ (1-\zeta_3 p^{-s})(1-\zeta_3^{-1}p^{-s}) & (p) = \mathfrak{p} \\ (1-p^{-s}) & (p) = (23) = \mathfrak{p}_1^2\mathfrak{p}_2 \end{cases} \tag{3}$$

in terms of the field K:

(1) The first case occurs if and only if p splits in K and $p = \mathfrak{p}_1\mathfrak{p}_2$ with each $\mathfrak{p}_i$ principal. That's because H is the Hilbert class field of K, so a prime ideal $\mathfrak{q}$ of K splits in H if and only if $\mathfrak{q}$ is principal.
(2) The second case happens if and only if p is inert in K.
(3) The third happens if and only if p splits in K but the $\mathfrak{p}_i$ are both not principal (either both are principal or not, since they are inverses in the ideal class group).
(4) The fourth happens if and only if $p = 23$.

Let $\theta\colon \mathrm{Cl}_K \to \mathbb{C}^\times$ be a non-trivial class group character. Then

$$L(s,\theta) = \sum_{I \subset \mathcal{O}_K} \frac{\theta(I)}{(\mathrm{Nm}\, I)^s} = \prod_{\mathfrak{p} \subset \mathcal{O}_K} \left(1 - \frac{\theta(\mathfrak{p})}{(\mathrm{Nm}\,\mathfrak{p})^s}\right)^{-1}.$$

Now, by inspection

$$\prod_{\mathfrak{p}|p} \left(1 - \frac{\theta(\mathfrak{p})}{(\mathrm{Nm}\,\mathfrak{p})^s}\right)^{-1} = \begin{cases} (1-p^{-s})^2 & p \text{ split into principals} \\ (1-p^{-2s}) & p \text{ inert} \\ (1-\zeta_3 p^{-s})(1-\zeta_3^{-1}p^{-s}) & p \text{ split into non-principal} \\ (1-p^{-s}) & p = 23 \end{cases}$$

So we observe that $\zeta_L(s) = \zeta(s)L(K,\theta)$.

Artin knew this, and realized that something more general was going on. He realized that this identity of L-functions came from some identity of *representations* for S_3.

6.2. **Artin's conjecture.**

Definition 6.2.1. Let E/K be Galois, $\rho\colon \mathrm{Gal}(E/K) \to \mathrm{GL}_n(\mathbb{C})$ (or more invariantly $\mathrm{GL}(V)$, $V \cong \mathbb{C}^n$) be a representation. The *Artin L-function* associated to ρ is

$$L(\rho, s) = \prod_{\mathfrak{p} \subset \mathcal{O}_K} \det\left(1 - (\mathrm{Frob}_\mathfrak{p}|_{V^{I_\mathfrak{p}}})(\mathrm{Nm}\,\mathfrak{p})^{-s}\right)^{-1}.$$

For $\mathfrak{p} \subset \mathcal{O}_K$ downstairs, fixing a prime of E above $\mathfrak{p}$ gives $I_\mathfrak{p} \subset D_\mathfrak{p} \subset \mathrm{Gal}(E/K)$ such that $D_\mathfrak{p}/I_\mathfrak{p} \cong \langle \mathrm{Frob}_\mathfrak{p} \rangle$. Technically this depends on a choice of prime above $\mathfrak{p}$, but changing $\mathfrak{p}$ conjugates the element Frob_p in the

Galois group. So for almost all $\mathfrak{p}$ (namely unramified ones) $\mathrm{Frob}_\mathfrak{p}$ is a well-defined conjugacy class in $\mathrm{Gal}(E/K)$, hence the local factor

$$\det\left(1-\mathrm{Frob}_\mathfrak{p}(\mathrm{Nm}\,\mathfrak{p})^{-s}\right)^{-1}$$

is well-defined. The product is absolutely convergent for $\mathrm{Rep}\, s \gg 1$.

Conjecture 6.2.2 (Artin's Conjecture)**.** *$L(s,\rho)$ extends to a meromorphic function of $S \in \mathbb{C}$, which is holomorphic if ρ doesn't contain the trivial representation, and satisfies a functional equation.*

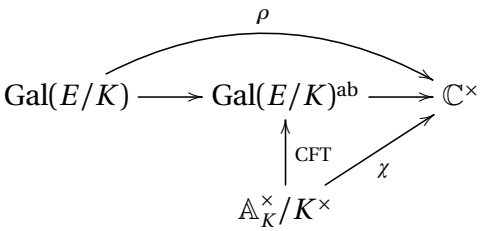
Remark 6.2.3. The shape of the equation involves Γ-functions as before, but also a new phase factor which is very interesting in its own right.

Theorem 6.2.4 (Artin-Brauer)**.** *Artin's conjecture is true, except possibly the clause about the holomorphicity.*

The proof expresses L as a product of ratios of L-functions attached to characters. So conceivably there could many many poles.

Formal properties of $L(s,\rho)$. Let $\rho\colon \mathrm{Gal}(E/K) \to \mathrm{GL}(V)$ be a Galois representation.

(1) If V is 1-dimensional, $\rho\colon \mathrm{Gal}(E/K) \to \mathbb{C}^\times$ factors through $\mathrm{Gal}(E/K)^{\mathrm{ab}}$, and by class field theory you can view it as coming from an idele class character:



Then $L(s,\rho) = L(s,\chi)$ (the Hecke L-function). In particular, it is holomorphic if $\chi \neq 1$ (and in any case, we understand its poles).

(2) $L(s,\rho_1 \oplus \rho_2) = L(s,\rho_1)L(s,\rho_2)$.

(3) Suppose you have a tower of field extensions $K \subset L \subset E$. If σ is a representation of $\mathrm{Gal}(E/L)$, then

$$L(s,\sigma) = L(s,\mathrm{Ind}_L^K \sigma).$$

Here $\mathrm{Ind}_L^K \sigma = \mathrm{Ind}_{\mathrm{Gal}(E/L)}^{\mathrm{Gal}(E/K)} \sigma$.

Example 6.2.5. (Regular representation) If E/K is Galois, then applying (4) to the tower $K \subset E \subset E$, we have

$$\begin{aligned}\zeta_E(s) &= L(s, \text{trivial representation of } \mathrm{Gal}(E/E)) \\ &= L(s, \text{regular representation of } \mathrm{Gal}(E/K)) \\ &= \prod_{\rho \text{ irred.}} L(s,\rho)^{\dim \rho}\end{aligned}$$

If E/K is abelian, then this is a product of $L(s,\chi)$ over χ the 1-dimensional characters of $\mathrm{Gal}(E/K)$.

Recall that we posed as motivation a hypothetical splitting of the analytic class number formula corresponding to this factorization of the L-function.

Example 6.2.6. Referring back to Example 6.1.1, we have

$$\zeta_L(s) = L(s, \mathbf{1} \text{ as a representation of } \mathrm{Gal}(H/L))$$

By (3), this is in turn equal to $L(s, \mathrm{Ind}_{\mathrm{Gal}(H/L)}^{\mathrm{Gal}(H/\mathbb{Q})} \mathbf{1})$, and $\mathrm{Ind}_{\mathrm{Gal}(H/L)}^{\mathrm{Gal}(H/\mathbb{Q})} \mathbf{1}$ is the standard representation of S_3 on $\mathbb{C}^3$, which splits as $\mathbf{1} \oplus \mathbf{2}$ (the irreducible 2-dimensional representation). So by (2), this is a product of $L(s,\mathbf{1})$ and $L(s,\mathbf{2})$ (with $\mathbf{1}$ and $\mathbf{2}$ regarded as representation of $\mathrm{Gal}(H/\mathbb{Q})$. The first factor is of course equal to $\zeta(s)$. Next, we use the fact that the 2-dimensional irreducible of S_3 is isomorphic to the induction of a non-trivial character of A_3, so $L(s,\mathbf{2}) = L(s,\theta)$, recovering the decomposition

$$L(s, \mathbf{1}_{\mathrm{Gal}(H/L)}) = L(s, \mathbf{1}_{\mathrm{Gal}(H/\mathbb{Q})}) L(s, \theta_{\mathrm{Gal}(K/\mathbb{Q})})$$

6.3. **The conductor-discriminant formula.** The functional equation of $L(s,\rho)$ looks like

$$N_\rho^{s/2}(\Gamma - \text{factors}(s))L(s,\rho) = \epsilon_\rho N_\rho^{\frac{1-s}{2}}(\Gamma - \text{factors}(1-s))L(1-s,\widetilde{\rho})$$

where ϵ_ρ has absolute value 1. This factor is very subtle. Thanks to great effort of Langlands and Dwork, we have a local definition. Deligne gave a much more concise global definition.

Here N_ρ is the *Artin conductor* of ρ,

$$\prod_{\text{ramified } \mathfrak{p}} \mathfrak{p}^{n_\mathfrak{p}}$$

where $n_\mathfrak{p}$ is a measure of the ramification. For example, if ρ is *tamely ramified* at $\mathfrak{p}$, then $n_\mathfrak{p} = \dim V - \dim V^{I_\mathfrak{p}}$. If there is higher inertia, then the sum is more complicated, and not obviously an integer. This agrees with the usual conductor if ρ is 1-dimensional.

Theorem 6.3.1 (Conductor-discriminant formula)**.** *Let $E/\mathbb{Q}$ be a number field. Then*

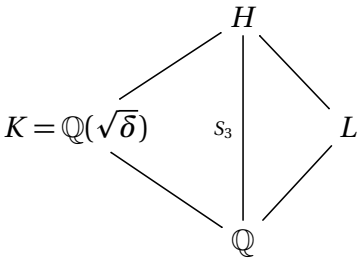
$$\operatorname{disc} E = \prod_{\rho} N_{\rho}^{\dim \rho}.$$

Example 6.3.2. Let's use the conductor-discriminant formula to compute $\operatorname{disc}(E = \mathbb{Q}(\zeta_{p^2}))$. According to the formula, the discriminant should be

$$\prod_{\rho\colon \operatorname{Gal}(E/K) \to \mathbb{C}^\times} N_\rho.$$

The ρ correspond to the Dirichlet characters of $(\mathbb{Z}/p^2)^\times$. There are $(p-1)$ having conductor $n_p = 1$, and the remaining $(p-1)^2$ have conductor $n_p = 2$. Therefore, the discriminant ideal is $(p^{(p-1)+2(p-1)^2})$ Over $\mathbb{Q}$, that means we have pinned down the discriminant up to a sign.

Example 6.3.3. Let L be a non-cyclic cubic extension of $\mathbb{Q}$ of discriminant δ. Then we have the subfield lattice:



We saw, using the factorization formula and induction, that

$$\begin{aligned} \operatorname{disc} H &= (N_{\mathbf{2}})^2 N_{\mathbf{1}} N_{\mathrm{sgn}}^1 \\ \operatorname{disc} L &= N_{\mathbf{1}} N_{\mathbf{2}} \\ \operatorname{disc} K &= N_{\mathbf{1}} N_{\mathrm{sgn}} \end{aligned}$$

Of course, we always have $N_{\mathbf{1}} = 1$. So the conductor-discriminant formula implies that $\operatorname{disc} H = (\operatorname{disc} L)^2 \operatorname{disc} K$.

If $\operatorname{disc} L = \operatorname{disc} K$ (which is automatic if the δ is squarefree), then this says that

$$\operatorname{disc} H = (\operatorname{disc} K)^3.$$

On the other hand, transitivity of discriminants says that

$$\operatorname{disc} H = (\operatorname{disc} K)^3 \operatorname{Nm}_{K/\mathbb{Q}}(\operatorname{disc} H/K).$$

This implies that $\operatorname{disc} H/K = (1)$, i.e. H/K is everywhere unramified.

6.4. **Analytic properties.** We'll prove that $L(s,\rho)$ is meromorphic and has a functional equation, as predicted by Artin. The holomorphicity, as mentioned earlier, is still wide open .

Theorem 6.4.1 (Brauer)**.** *If G is a finite group, then any (complex) representation ρ of G can be written as*

$$\rho = \sum m_i \operatorname{Ind}_{H_i}^G \psi_i$$

where $m_i \in \mathbb{Z}$, $H_i \subset G$, and $\psi_i : H \to \mathbb{C}^\times$ are characters. Concretely, this means that the characters of both sides are equal.

Why does this imply what we want? If $G = \mathrm{Gal}(E/K)$, then

$$\begin{aligned} L(s,\rho) &= \prod L(s, \operatorname{Ind}_{H_i}^G \psi_i \text{ as } \mathrm{Gal}(E/K)\text{-representation})^{m_i} \dots \\ &= \prod L(s, \psi_i \text{ as } \mathrm{Gal}(E/K_i)\text{-representation})^{m_i} \end{aligned}$$

where the ψ_i are considered as characters of $\mathrm{Gal}(E/K_i)$, where $K_i = \mathrm{Fix}(H_i)$. This expresses $L(s,\rho)$ as a product of ratios of Dirichlet L-functions, for which we know the functional equation and meromorphicity (and even the location and residues of poles). However, it is important to note that since some of the m_i may be negative, we have no control of poles of this product. If Artin's conjecture is true, then there must be non-trivial cancellation between zeros and poles.

The rest of the section is devoted to the proof of Theorem 6.4.1. Let $\rho : G \to \mathrm{GL}(V)$ be a representation. Then G acts on $\mathbb{P}(V)$. For each subgroup $H \leq G$, we can consider the fixed point locus $\mathrm{Fix}(H) \subset \mathbb{P}(V)$. A point $x \in \mathrm{Fix}(H) \subset \mathbb{P}(V)$ corresponds to a line $\ell_x \subset V$ stable under H. The action of H on this line determines a character $\psi_x : H \to \mathbb{C}^\times$.

For $H \leq G$ and $\psi : H \to \mathbb{C}^\times$ a character, let

$$X_{H,\psi} = \{x \in \mathbb{P}(V) \mid \mathrm{Stab}_G(x) = H, \psi_x = \psi\}.$$

Clearly these partition $\mathbb{P}(V)$:

$$\mathbb{P}(V) = \coprod_{H,\psi} X_{H,\psi}.$$

Let $N_{H,\psi}$ be the normalizer of the pair (H,ψ). We claim that

$$\boxed{\rho = \sum_{(H,\psi)/\text{conjugacy}} \frac{\chi(X_{H,\psi})}{[N_{H,\psi} : H]} \cdot \operatorname{Ind}_H^G \psi.} \tag{4}$$

This is remarkable because it even says that we can make this construction *functorially*, which was not clear in Brauer's original proof. Look up "canonical Brauer induction" for more about this.

Write χ_ρ for the character associated to ρ. Let's first check that (4) holds at least when evaluated on $g = 1$:

$$\chi_\rho(1) \stackrel{?}{=} \sum_{(H,\psi)/\text{conjugacy}} \frac{\chi(X_{H,\psi}) \cdot [G:H]}{[N_{H,\psi}:H]}.$$

The left hand side is $\dim V$. The right hand side is

$$\begin{aligned}\sum_{(H,\psi)/\text{conjugacy}} \chi(X_{H,\psi})[G:N_{H,\psi}] &= \sum_{(H,\psi)} \chi(X_{H,\psi}) \\ &= \chi(\mathbb{P}V) \\ &= \dim V.\end{aligned}$$

This shows that (4) holds at least when evaluated at $1 \in G$. But we've brushed something under the rug: χ is not additive in general. For example, the Euler characteristic of a line segment is 1. If you write the line segment as a disjoint union of a point and two half-segments, then each piece has Euler characteristic 1, so their sum is 3.

Compactly supported cohomology. However, the *compactly supported* Euler characteristic *is* additive: if X is locally compact, then we define the compactly supported Euler characteristic

$$\chi_c(X) = \sum_{i \geq 0} (-1)^i \dim H^i_c(X, \mathbb{C})$$

where $H^i_c(X, \mathbb{C})$ is the cohomology of *compactly supported cochains* (this is originally due to Borel-Moore). In "nice" situations, this compactly supported Euler characteristic is addivie, e.g. if $Z \subset X$ is closed, then

$$\chi_c(X) = \chi_c(Z) + \chi_c(X - Z).$$

If M is a manifold, then Poincaré duality implies $H^i(M, \mathbb{C}) \cong H^{\dim M - i}_c(M, \mathbb{C})^*$. In particular, if M is even-dimensional then $\chi(M) = \chi_c(M)$. This legitimizes the calculations we made above with $M = \mathbb{P}(V)$.

Remark 6.4.2. $X_{H,\psi}$ is smooth because $\text{Fix}(H)$ smooth, and $X_{H,\psi}$ is a connected component of it. However, this is not important for our purposes.

Continuing on with the proof, we must show that (4) holds for all $g \in G$, and then prove that the coefficients $\frac{\chi(X_{H,\psi})}{[N_{H,\psi}:H]}$ are integers. Let $g \in G$, and suppose that the eigenvalues of g are $\alpha_1, \ldots, \alpha_k$ with multiplicities $m_1, \ldots, m_k$ (since finite-dimensional complex representations of finite groups are unitary, these reallly are eigenvalues and not generalized eigenvalues). Let $U_1, \ldots, U_k$ be the corresponding eigenspaces, so $\dim U_i = m_i$. Then $\chi_\rho(g) = \sum m_i \alpha_i$. We are basically going to repeat the above argument evaluating instead at g and examining things at the level of eigenspaces.

Now we decompose

$$\mathbb{P}U_i = \coprod_{\substack{H \ni g \\ \psi: H \to \mathbb{C}^\times \\ \psi(g)=\alpha_i}} X_{H,\psi}.$$

Taking Euler characteristics, we get

$$m_i = \sum_{\substack{H \ni g \\ \psi: H \to \mathbb{C}^\times \\ \psi(g)=\alpha_i}} \chi(X_{H,\psi}).$$

Evaluating the right hand side of (4) at g, we get (by similar reasoning as before)

$$\chi_\rho(g) = \sum_{H \ni g, \psi} \chi(X_{H,\psi})\psi(g).$$

This shows that

$$\chi_\rho = \sum_{(H,\psi)} \chi(X_{H,\psi}) e_{H,\psi}$$

where $e_{H,\psi}(g) = \begin{cases} \psi(g) & g \in H \\ 0 & \text{otherwise} \end{cases}$. Now it only remains remains to compare this with the characters of induced representations. As is "well-known,"

$$\chi_{\mathrm{Ind}_H^G \psi} = \sum_{g \in G/H} e_{gHg^{-1}, g\psi g^{-1}}.$$

To express the previous formula in terms of this, simply group together conjugates

$$\begin{aligned}
\sum_{(H,\psi)} \chi(X_{H,\psi}) e_{H,\psi} &= \sum_{(H,\psi)/\text{conjugacy}} \sum_{g \in G/N_{H,\psi}} \chi(X_{H,\psi}) e_{gHg^{-1}, g\psi g^{-1}} \\
&= \sum_{(H,\psi)/\text{conjugacy}} \frac{\chi(X_{H,\psi})}{[N_{H,\psi} : H]} \sum_{g \in G/H} e_{gHg^{-1}, g\psi g^{-1}} \\
&= \sum_{(H,\psi)/\text{conjugacy}} \frac{\chi(X_{H,\psi})}{[N_{H,\psi} : H]} \cdot \chi_{\mathrm{Ind}_H^G \psi}.
\end{aligned}$$

This finally establishes the claim. Now we just have to argue why the coefficients $\frac{\chi(X_{H,\psi})}{[N_{H,\psi} : H]}$ are integers. But $N_{H,\psi}/H$ acts *freely* on $X_{H,\psi}$ (because by definition it is the full stabilizer of any point) so its order divides $\chi(X_{H,\psi})$.

6.5. **Positive characteristic speculation.** There are hints of a theory of Artin L-functions in positive characteristic, but this theory is very much underdeveloped. For instance, nobody knows the answer to:

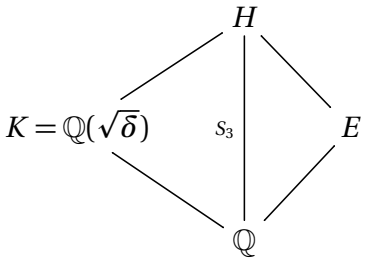
Question: are there "mod p Artin L-functions?"

In particular, given $\rho\colon \mathrm{Gal}(E/K) \to \mathrm{GL}_n(\mathbb{F}_p)$, can you (sometimes) make sense of something like "$L(0,\rho) \in \mathbb{F}_p$"?

It is a fact that any representation of a finite group over $\overline{\mathbb{F}_p}$ lifts *virtually* to $\overline{\mathbb{Q}_p}$. That gives a way of getting started, but for instance it is not clear that it is independent of the lift. (In fact, it probably is not, stated in this crude way.)

One hint is that over a function field, it is clear that there "should be" such a theory. For instance, there is a way of describing the L-function in terms of cohomology via the Grothendieck-Lefschetz trace formula, and if one just takes $\mathbb{F}_p$-coefficients, then one gets the "right" thing.

Example 6.5.1. Here is one "shadow" of a mod-p L-function in the number field case. Consider again



Suppose H/E is unramified ♠♠♠ TONY: [can i get rid of this?] and $\operatorname{disc} E = \operatorname{disc} K$, so that H/K is unramified.

A theorem of Gerth says that

$$\boxed{3-\text{rank of } \mathrm{Cl}_K = 3-\text{rank of } \mathrm{Cl}_E + 1}$$

i.e. $\dim_{\mathbb{F}_3}(\mathrm{Cl}_K/3\,\mathrm{Cl}_K) = \dim_{\mathbb{F}_3}(\mathrm{Cl}_E/3\,\mathrm{Cl}_E)$. The left hand side is at least 1, because we know that H/K is an unramified degree 3 extension. The first case where $\operatorname{rank}_3 \mathrm{Cl}_E > 0$ is $\operatorname{disc} = -3299$.

This is quite striking. Morally, there is "no relation" between Cl_K and Cl_E, but there *is* a relation between the 3-parts. Morally, that comes from the factorization

$$\zeta_E(s) = \zeta(s)L(s,\theta)$$

where θ is a non-trivial character of Cl_K. Why? Modulo 3, θ is trivial. Then one might expect that the equality of "mod 3" L-functions reads

$$\zeta_E = \zeta \cdot \zeta_K.$$

Now, evaluating this at 1 (or 0) should reflect this relation between the class groups. That is, Gerth's theorem should reflect an equality of mod 3 representations of $\mathrm{Gal}(H/\mathbb{Q})$.

7. Stark's conjectures

We will discuss a circle of ideas concerning the interplay between the factorization of the Dedekind zeta function into Artin L-functions, and a corresponding hypothetical interpretation of the special values of Artin L-functions.

7.1. **The class number formula.** If E/K is Galois, then the decomposition of the regular representation of $G = \mathrm{Gal}(E/K)$ into irreducible representations induces the factorization

$$\zeta_E(s) = \prod_{\rho \text{ irred.}} L(s,\rho)^{\dim \rho} \tag{5}$$

and the class number formula reads

$$\mathrm{Res}_{s=1}\, \zeta_E(s) = \frac{2^{r_1}(2\pi)^{r_2}}{w_E\sqrt{\mathrm{disc}\, E}} h_E R_E$$

where $E \otimes \mathbb{R} \cong \mathbb{R}^{r_1} \times \mathbb{C}^{r_2}$. By the functional equation, this corresponds to

$$\boxed{\zeta_E(s) \sim -s^{r_1+r_2-1} \frac{h_E R_E}{\omega_E} \text{ near } s = 0.}$$

Interpreting $r_1 + r_2 - 1 = \mathrm{rank}\, \mathcal{O}_E^*$, we can view this as relating the behavior of ζ_E at 0 with the class number and the "size" of the unit group.

Since the left hand sides factorizes according to (5) into a product of $L(s,\rho)$, we want to have a corresponding "factorization" on the right hand side.

Example 7.1.1. Suppose E/K were Galois with $\mathrm{Gal}(E/K) \cong \mathbb{Z}/3\mathbb{Z}$. Also assume that K is totally real. There are two non-trivial Galois characters $\chi : \mathrm{Gal}(E/K) \to \mathbb{C}^\times$.

Pretend that $R_E = R_K = 1$, and ignore w_E and w_K for the moment. (This never actually happens, but you would be in this case if you were examining the value at $s = -1$, and replacing the class numbers by K_2.) Anyway, the point is that you get

$$h_E \sim \zeta_E(0) = \zeta_K(0) L(0,\chi) L(0,\overline{\chi}) \sim h_K L(0,\chi) L(0,\overline{\chi}).$$

Now, $L(0,\chi) = a + b\zeta_3 \in \mathbb{Q}(\zeta_3)$ (though we actually know its value is 0), and also $L(0,\overline{\chi}) = a + b\overline{\zeta_3}$. So we are imagining some factorization

$$\frac{h_E}{h_K} = (a + b\zeta_3)(a + b\overline{\zeta}_3).$$

One imagines $\frac{h_E}{h_K}$ as the order of a kind of relative class group. There are natural maps between class groups: the norm induces $C_E \to C_K$ and extension of ideals induces $C_K \to C_E$. It turns out that this is relative in the

first sense: "away from 3" we have

$$\#\frac{h_E}{h_K} = \#\ker(C_E \xrightarrow{\mathrm{Nm}} C_K).$$

So (ignoring issues at 3), we want to express $\#(C_E \to C_K)$ as $(a+b\zeta_3)(a+b\overline{\zeta_3})$. How might we obtain such a factorization?

The key observation is that we have a $\mathbb{Z}[\mathbb{Z}/3] = \mathbb{Z}[\sigma]$ action C_E. Now, the elements of C_E that are killed by the norm are killed by $1+\sigma+\sigma^2$, so $\mathbb{Z}[\sigma]/(1+\sigma+\sigma^2) = \mathbb{Z}[\zeta_3]$ acts on $\ker(C_E \xrightarrow{\mathrm{Nm}} C_K)$. By the classification of finitely generated modules over a PID, we have an isomorphism

$$\ker(C_E \xrightarrow{\mathrm{Nm}} C_K) \cong \bigoplus \mathbb{Z}[\zeta_3]/(\alpha_i)$$

and we can say that the "order in $\mathbb{Z}[\zeta_3]$" is $\prod \alpha_i$. Now, there is definitely an ambiguity up to units here, which may not be able to be pinned down exactly, but it is Akshay's opinion that there is clearly "something more" that we aren't seeing. In order to do the factorization for $L(s,\rho)$, we'll have to define versions of h_E, R_E, w_E in $\mathbb{Z}[G], \mathbb{Z}[G]\otimes\mathbb{R}$, etc.

There are many problems with formulating this in general. For instance, in general the base ring will *not* be principal. Indeed, $\mathbb{Z}[\mathbb{Z}/3]$ was not principal; we only got a nice result because we considered the submodule killed by the norm. But we do expect that if the $\ker(C_E \to C_K)$ are not described principally, then this will be matched by a correspoding failure in the regulators (which we ignored anyway here).

You can work around this principality issue by localizing $\mathbb{Z}[G]$ at a prime, which is basically what's done in Iwasawa theory.

7.2. **Aside: how to compute $L(s,\rho)$.** We'll say a few words about computation. One way to compute is to go through Brauer's theorem, but you almost never want to do that.

You can write

$$L(s,\rho) = \prod_{\mathfrak{p}} \prod (\ldots) = \sum_n \frac{a_n}{n^s}.$$

Proposition 7.2.1. *Let ϕ be a C^∞ function on $\mathbb{R}$ of rapid decay (i.e. $\varphi(x) \ll (1+|x|)^{-N}$ for any N) with $\phi(0)=1$. If $L(s,\rho)$ is holomorphic (expected from the Artin conjecture) then its value at s_0 is*

$$L(s_0,\rho) = \lim_{x\to\infty} \sum_n \frac{a_n}{n^{s_0}} \varphi(n/x).$$

This is a basic piece of intuition from analysis. You would like to say that you can get the value by taking partial sums. In general that doesn't work because the "cutoff" is too sharp. If you do it in a smoother way, then that does work. In practice, the sum will tend to converge once x is bigger than the conductor of ρ.

Example 7.2.2. Let $K = \mathbb{Q}(\alpha)$ where $\alpha^3 = \alpha + 1$ and H be the Galois closure of $K/\mathbb{Q}$. Then $\zeta_K(s) = \zeta(s)L(s,\rho)$ where $\rho : S_3 \to \mathrm{GL}_2(\mathbb{C})$ is the irreducible two-dimensional representation. Then $L'(0,\rho) = \pm 0.2811996 = \log|\alpha|$ because α is a fundamental unit. Now, if $L(s,\rho) = \sum \frac{a_n}{n^s}$, then taking $x = 1000$ and $\phi(t) = e^{-t^2}$ we obtain the estimate

$$L'(0,\rho) \approx \sum_{n=1}^{5000} a_n \log n \, e^{-(n/1000)^2} = -0.2811986.$$

Proof. Let

$$F(s) = \int_{x>0} \varphi(x) x^s \frac{dx}{x}.$$

(This is the Fourier transform on $(\mathbb{R}_{>0}, \times)$, i.e. the Mellin transform). This is convergent for $\operatorname{Re} s > 0$, because φ decays rapidly. Fourier inversion gives

$$\varphi(x) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} F(s) x^{-s} \, ds \quad \text{for } \sigma > 0.$$

So

$$\begin{aligned} \sum \frac{a_n}{n^{s_0}} \varphi(n/x) &= \sum_n \frac{a_n}{n^{s_0}} \frac{1}{2\pi} \int_{\sigma - i\infty}^{\sigma + i\infty} F(s) \left(\frac{n}{x}\right)^{-s} ds \\ &= \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} F(s) x^s \underbrace{\sum_n \frac{a_n}{n^{s+s_0}}}_{L(s+s_0,\rho)} ds. \end{aligned}$$

Now we shift the contour from $[\sigma - i\infty, \sigma + i\infty]$ to $[\sigma' - i\infty, \sigma' + i\infty]$ where $\sigma' < 0$. To do this, you have to be careful about the growth at the "edges" of the rectangle," which we'll leave as an exercise.

By Cauchy's theorem, you pick up terms from the residues. The factors x^s and $L(s + s_0, \rho)$ are holomorphic, but $F(s) = \int \varphi(x) x^s \frac{dx}{x}$ might have poles. In fact, we showed in Lemma 5.1.1 that it *does* have poles at $s = 0, -1, -2, \ldots$ with the pole at 0 having residue $\varphi(0)$.

So the above is

$$= \frac{1}{2\pi} i \int_{-\frac{1}{2} - i\infty}^{-\frac{1}{2} + i\infty} (\ldots) + \underbrace{2\pi i \frac{1}{2\pi i} \varphi(0) L(s_0, \rho)}_{L(s_0,\rho)}.$$

It remains only to estimate the error term. The integrand has magnitude $x^{-1/2(\ldots)}$ so

$$\sum_n \frac{a_n}{n^{s_0}} \varphi(n/x) = L(s_0, \rho) + O(x^{-1/2}).$$

We can try to make the convergence faster by shifting further, and we pick up more terms

$$L(s_0,\rho)+\frac{L(s_0+1,\rho)\varphi'(0)}{x}+\dots \tag{6}$$

so if all the derivatives of φ at 0 are 0, then

$$\sum -\frac{a_n}{n^{s_0}}\varphi(n/x)=L(s_0,\rho)+O(x^{-N}).$$

(Of course, in practice this will be offset by growth of the implicit constants.) □

Remark 7.2.3. From the proof, we see that this works even if $L(s,\rho)$ has a pole, but then you have to subtract off a leading term. It also works if φ is "smooth enough."

Example 7.2.4. Recall that we continued

$$\zeta(s)=\lim_{N\to\infty}\left(\sum_{n=1}^{N}\frac{1}{n^s}-\frac{N^{1-s}}{1-s}\right)$$

for $\operatorname{Re} s>0$. How do you continue further? One way is to subtract off more stuff, but a better way is to smooth the sum. For $\zeta(0)$, we can just use the function $\varphi(x)=\begin{cases}1-x & x\le 1,\\ 0 & x>1.\end{cases}$ This gives

$$\sum_{n=1}^{x}(1-\frac{n}{x})=x-\frac{x(x+1)}{2x}\sim\frac{x}{2}$$

which reflects that $\zeta(s)$ has a pole with residue 1 at 1, and value $-1/2$ at 0 (taking note of Remark 7.2.3).

7.3. **Stark's conjectures.** Let E/K be Galois with Galois group G and $\rho: G\to \mathrm{GL}(V)$ an irreducible representation. We know that

$$\zeta_E=\prod_{\rho}L(s,\rho)^{\dim\rho}.$$

We have the class number formula

$$\zeta_E(s)\sim-\frac{h_E R_E}{w_E}s^{r_1+r_2-1}\text{ near }s=0$$

and our goal was to split up this formula in a manner corresponding to the factorization of ζ_E into Artin L-functions. For that, we need refined (equivariant) versions of $h_E, R_E,\dots$ in (something like) $\mathbb{Z}[G]$. The problem is that $\mathbb{Z}[G]$ can be a very nasty ring, and in Stark's conjecture the solution is to work instead with $\mathbb{Z}[G]\otimes\mathbb{R}$.

Let's recall the definition of R_E. We have a map

$$U_E = \mathcal{O}_E^\times/\text{torsion} \xrightarrow{\log} \mathbb{R}^{r_1+r_2}$$

sending $\epsilon \mapsto (\log|\epsilon|_v)_{v \text{ archimedean}}$. The image $\log(U_\epsilon)$ is a lattice inside the hyperplane $\{(x_i) \mid \sum x_i = 0\}$ and R_E is the covolume ofthis lattice.

Remark 7.3.1. The issue of which normalization to pick for the volume form on this hyperplane is a little tricky. There are several reasonable possibilites:

(1) the volume form $dx_1 \dots \widehat{dx_i} \dots dx_n$ (i.e. project from one coordinate)
(2) The Riemannian form induced from $\mathbb{R}^n$
(3) Parametrization by the coordinates $x_2 - x_1, \dots, x_n - x_{n-1}$.

These differ by $\sqrt{n}$ in arithmetic proression. The first one turns out to be correct, so $R_E = \det(\log|\epsilon_i|_{v_j})$ where ϵ_i is a $\mathbb{Z}$-basis for U_E and v_j are all places but one. Note that if v is a complex place, then $|\epsilon|_v = |\epsilon|^2$ (we aways normalize the measure to grow with how scaling changes volume).

We want to factorize the regular as $R_E = \prod R_\rho^{\dim \rho}$ (up to an element of $\mathbb{Q}^\times$). We could try to decompose according to the G-action on the unit lattice. If G is something very nice, like $\mathbb{Z}/2$, then it would split up the group into $+1$ and -1 parts. If $G = \mathbb{Z}/3$, then it doesn't split over $\mathbb{Q}$, which is confusing. And there is a second source of confusion: the group algebra over $\mathbb{Q}$ might contain a division algebra rather than a (commutative) field. We'll elaborate on this shortly.

Let's ignore the G-action for now. Let

$$X = \left\{ \sum a_i v_i \mid {}^{v_i \text{ archimedean for } E}_{a_i \in \mathbb{Q}, \sum a_i = 0} \right\},$$

which is a $\mathbb{Q}$-vector space of dimension $r_1 + r_2 - 1$. Then we can rephrase the regulator map as saying that

$$U \otimes_{\mathbb{Q}} \mathbb{R} \xrightarrow{\log} X \otimes_{\mathbb{Q}} \mathbb{R} \text{ is an isomorphism.}$$

Let us now abstract a bit. Given two $\mathbb{Q}$-vector spaces V_1, V_2 and an isomorphism $\alpha\colon V_1 \otimes \mathbb{R} \to V_2 \otimes \mathbb{R}$, it doesn't quite make sense to talk about $\det\alpha$. However, we can define "$\det\alpha$" in $\mathbb{R}^\times/\mathbb{Q}^\times$ by choosing any volume forms in V_1, V_2, which are defined up to $\mathbb{Q}^\times$. Explicitly, "$\det\alpha$" is $\det(\alpha \circ \varphi^{-1})$ where $\varphi\colon V_1 \xrightarrow{\sim} V_2$ is an arbitrary isomorphism over $\mathbb{Q}$ (ambiguous up to $\mathrm{GL}(V_2)$, so the determinant is ambiguous up to $\mathbb{Q}^\times$).

To get an equivariant version of this, we're going to replacing $\mathbb{Q}$ by $\mathbb{Q}[G]$, we're going to make sense of the determinant as a value in $Z(\mathbb{R}[G])^\times/Z(\mathbb{Q}G)^\times$.

Remark 7.3.2. Although ambiguity up to $\mathbb{Q}^\times$ may not sound very satisfying, it is usually the case that the rational number turns out to be something simple. ♠♠♠ TONY: [can this be quantified in any way?]

The map we produced earlier

$$U \otimes \mathbb{R} \xrightarrow{\log} X \otimes \mathbb{R}.$$

is even an isomorphism of $\mathbb{R}[G]$-modules. We claim that this is even an isomorphism *rationally* as $\mathbb{Q}[G]$-modules. This is a special case of the more general (and elementary!) result:

Lemma 7.3.3. *If V, V' are finite-dimensional G-representations over k and $V \otimes K \cong V' \otimes K$, then $V \cong V'$.*

Proof. A G-equivariant map is a linear map commuting with a bunch of operators, which are just *linear* constraints on its coefficients. If this linear system admits a solution over K, then it admits a solution over k. This shows that $\operatorname{Hom}_{K[G]}(V, V') = \operatorname{Hom}_{k[G]}(V, V') \otimes_k K$.

Now, we have to check that if there is an *invertible* map over K, then there is one over k. It is easy to see that the determinant when restricted to $\operatorname{Hom}_{k[G}[V, V']$ cannot vanish identically if k is infinite. This is less clear over a finite field (which of course we don't need), although it's still true, so that part is left as an exercise. ♠♠♠ TONY: [todo] □

By the lemma, we may choose an isomorphism $\varphi : U \xrightarrow{\sim} X$ as $\mathbb{Q}[G]$-modules. Then $\varphi^{-1} \circ \log : U \otimes \mathbb{R} \xrightarrow{\sim} U \otimes \mathbb{R}$ as $\mathbb{R}[G]$-modules. We want to make sense of the "determinant" of this in $\mathbb{R}[G]$.

Abstracting again, given a semisimple algebra A (e.g. $\mathbb{R}[G], \mathbb{Q}[G]$) over k and a homomorphism of A-modules $\alpha : V \to W$, we want to define $\det_A(\alpha) \in Z(A)$ with the property that $N_{A/k}(\det_A \alpha) = \det_k \alpha$ as a k-vector space map. Here $N_{A/k}$ is the *reduced norm* from A to k.

Definition 7.3.4. We define the *reduced norm* of a semisimple algebra over k as follows. First, if A is a central simple algebra of dimension n^2, then over $\overline{k}$ we know that A splits as a matrix algebra, whose left regular representation decomposes as a direct sum of n copies of the standard representation V. Then $N_{A/k}(a)$ is the usual norm of a acting by multiplication on V. Although we have defined this over $\overline{k}$, it is a fact that it descends to k.

Now if A is a general semisimple algebra, then by the classification of such algebras we have $A = \bigoplus M_{n_i}(D_i)$ where D_i a division algebra over k with center E_i. For $a = (a_i)$, we then define

$$N(a_i) = \prod_i \operatorname{Nm}_{E_i/k}(N_{M_{n_i}(D_i)/E_i} a_i).$$

Applying this construction to $\varphi^{-1} \circ \log\colon U \otimes \mathbb{R} \to U \otimes \mathbb{R}$ will give an element $\mathscr{R} \in Z(\mathbb{R}[G])^\times / Z(\mathbb{Q}[G])^\times$.

In terms of this $\mathscr{R} \in Z(\mathbb{R}[G])$, *Stark's conjecture* predicts that

$$L(s, \rho^*) \sim s^{r_\rho} \cdot \rho(\mathscr{R}) \text{ as } s \to 0.$$

Note that the left hand sides lies in $\mathbb{Q}(\rho)$, the field of traces of ρ, and $\rho(\mathscr{R})$ is in $\mathbb{C}^\times / \mathbb{Q}(\rho)^\times$ where ρ^* the dual representation to ρ.

Determinants over A. If A is a semisimple algebra over k, then

$$A \cong \bigoplus M_n(D_i)$$

where D_i is a division algebra over the base field k (with possibly bigger center than k). The determinant of the direct sum will be defined as the product of the determinants for each factor, so it suffices to handle the case where $A = M_n(D)$.

Let's first consider the simplest case $A = M_n(k)$. Let S be a simple, nonzero A-module, so $S \cong k^n$. Then $\alpha\colon V \xrightarrow{\sim} W$ induces $\alpha_S\colon \operatorname{Hom}(S, V) \to \operatorname{Hom}(S, V)$. This is now simply a map of k-vector spaces, and we put

$$\boxed{\det{}_A \alpha = \det{}_k \alpha_S} \in k = Z(A).$$

The fact that $(\det_A \alpha)^n = \det_k \alpha$ follows from the observation that $\alpha = \alpha_S \otimes_A \mathrm{Id}_S$.

In general, if A is simple then we can pick a Galois extension E/k such that $A \otimes E$ splits as a matrix algebra over E:

$$A \otimes_k E \cong M_n(E).$$

Given $\alpha\colon V \xrightarrow{\sim} V$, we can define $\det_{A \otimes E}(\alpha \otimes E) \in Z(A) \otimes E$. In fact, this is invariant by $\mathrm{Gal}(E/k)$ so we get that it actually lies in $Z(A)$.

Exercise 7.3.5. Check this.

Example 7.3.6. Suppose $A = D$ and $V = A^{\oplus s}$. Then $\alpha \in \operatorname{End}_A(V) \cong M_s(D)$ acting by right multiplication, and $\det_A \alpha$ is the "familiar" *reduced norm* on $M_s(D)$, which has the property that if $\dim_k D = n^2$, then $(\det_A \alpha)^n = \det_k \alpha$.

Example 7.3.7. Suppose G is *abelian* and E/K is a Galois field extension with $\mathrm{Gal}(E/K) = G$. For simplicity, just assume that $K = \mathbb{Q}$ and E is totally real. Then $\mathscr{R}_E \in \mathbb{R}[G]^\times / \mathbb{Q}[G]^\times$. Since $U_E \cong X$ as $\mathbb{Q}[G]$-modules, recalling $X = \{\sum a_i v_i \mid \sum a_i = 0\}$, there's a unit $\epsilon \in U_E$ such that $(g\epsilon)_{g \in G}$ generate $U_E/\mathbb{Q}$ (because it's true for X, taking e.g. an elementary vector). Without loss of generality, replacing ϵ by ϵ^2, we may assume that $\prod_g (g\epsilon) = 1$ (it might have be 1 or -1 originally).

An explicit isomorphism $\varphi: U_E \otimes \mathbb{Q} \to X \otimes \mathbb{Q}$ sends $\epsilon \mapsto v - \frac{\sum_w w}{n}$, e.g. if $n = 3$ we send $\epsilon \mapsto (2/3, -1/3, -1/3)$. We want to compare these two maps

$$\log: U_E \otimes \mathbb{R} \to X \otimes \mathbb{R}$$

and

$$\varphi \otimes \mathbb{R}: U_E \otimes \mathbb{R} \to X \otimes \mathbb{R}.$$

They differ (i.e. $\varphi^{-1} \circ \log$) by the following element of $\mathbb{R}[G]$ (which we will see later)

$$\mathscr{R} := \theta := \sum_{g \in G} \log |g\epsilon|_v \, g.$$

Fixing an isomorphism $\alpha: U \xrightarrow{\sim} X$, we can finally state a precise version of Stark's conjecture.

Conjecture 7.3.8 (Stark)**.** *Let $\rho: G = \mathrm{Gal}(E/K) \to \mathrm{GL}_n(V)$. Then*

$$L(s, \rho^*) \sim \alpha_\rho s^{r_\rho} \rho(\mathscr{R}) \text{ near } 0$$

where

- $\alpha \in \mathbb{Q}(\rho)^\times$ *(the field generated by traces),*
- $\mathscr{R} = \det_{\mathbb{R}[G]}(\alpha_{\mathbb{R}}^{-1} \log) \in Z(\mathbb{R}[G])^\times / Z(\mathbb{Q}[G])^\times$*, and*
- $r_\rho = \sum_{v \text{ arch. of } K} \dim(V^{G_v}) - \dim(V^G)$.

Remark 7.3.9. The philosophy here is that we forget about the class numbers and care only about the formula up to rational numbers.

Example 7.3.10. For ρ irreducible and non-trivial, $r_\rho = 0 \iff K$ is totally real, and every complex conjugation acts by $-I$.

Example 7.3.11. For χ non-trivial, $L(0, \chi) \neq 0 \iff \chi(-1) = -1$.

How could you go about checking this? We know that $L(s, \rho)$ is uniquely determined by compatibility with direct sums and induction (and induced from characters), so if we can verify that the right hand side is true for *characters* and compatible with direct sum and induction, then we're done.

7.4. **Compatibility with class number formula.** We check that this "factorization" is compatible with the class number formula:

$$\boxed{\zeta_E(s) \sim -s^{r_1 + r_2 - 1} \frac{h_E R_E}{w_E}.}$$

Now,

$$\prod L(s, \rho^*)^{\dim \rho^*} \sim \alpha s^{\sum \dim \rho \cdot r_\rho} \prod_\rho (\rho(\mathscr{R}))^{\dim \rho}.$$

Now $\alpha = \prod \alpha_{\rho^*}^{\dim \rho^*} \in \langle \mathbb{Q}(\rho^*) \rangle$ and $\prod_\rho (\rho(\mathscr{R}))^{\dim \rho} = N(\mathscr{R}) = \det_{\mathbb{R}}(\alpha_{\mathbb{R}}^{-1} \circ \log)$, which is the usual regulator.

It's annoying that we have this ambiguous constant factor α. There is a refinement of Stark's conjecture that can pin it down more, which says that

$$L(s, \rho^*) \sim \alpha_\rho s^{r_\rho} \rho(\mathscr{R})$$

and α_ρ is Galois-equivariant: for $\sigma \in \mathrm{Aut}(\mathbb{C})$,

$$\alpha_{\rho^\sigma} = (\alpha_\rho)^\sigma.$$

The refined version implies compatibility with the class number formula up to $\mathbb{Q}^\times$, not just $\langle \mathbb{Q}(\rho) \rangle^*$.

Remark 7.4.1. Although there is an ambiguity of $\mathscr{R}$ up to $Z(\mathbb{Q}[G])^\times$, that relative ambiguity is settled by choosing the same rational isomorphism $U \cong X$ for all ρ.

Remark 7.4.2. Although such a prediction seems natural, you have to be a little cautious. You can view this as taking an L-value and dividing by something trascendental to get an algebraic number, which is hypothesized to be rational. If you try the analogous thing for an L-function of an elliptic curve, dividing by a period gives $L(1/2, E \times \chi)/\Omega_E$ which is algebraic but *not* Galois-equivariant.

Example 7.4.3. Let $E = \mathbb{Q}(\zeta_m)$ and $K = \mathbb{Q}$. Then an irreducible representation $\rho \longleftrightarrow \chi$ corresponds to a character of conductor m, i.e. $(\mathbb{Z}/m\mathbb{Z})^\times \to \mathbb{C}^\times$. Let $\epsilon \in U = \mathscr{O}_E^\times \otimes \mathbb{Q}$ be such that $\{g\epsilon\}_{g \in G}$ generate U and $\prod (g \cdot \epsilon) = 1$. So we have two maps

$$U \underset{\alpha}{\overset{\log}{\rightleftarrows}} X$$

such that $\log(\epsilon) = \sum_v \log|\epsilon|_v \cdot v$ and $\alpha(\epsilon) = v_0 - \frac{1}{n}\sum_v v$ where $n = [E : \mathbb{Q}] = \varphi(m)$. Then $\log = \theta\alpha$ for some $\theta \in \mathbb{R}[G]$,

$$\theta = \sum_{g \in G} \log|\epsilon|_{g \cdot v_0} g.$$

Exercise 7.4.4. Check this: v_0 gets spread out into $\log \epsilon$, and kills the second term.

Then $\theta = \det_{\mathbb{R}[G]}(\alpha^{-1} \circ \log)$. Then as characters of $\mathbb{R}[G]$,

$$\begin{aligned}\chi(\theta) &= \sum_{g \in G} \log|\epsilon|_{g \cdot v_0} \chi(g) \\ &= \sum_{g \in G} \log|g^{-1}\epsilon|_{v_0} \chi(g) \\ &= \sum_{g \in G} \log|g\epsilon|_{v_0} \chi^{-1}(g).\end{aligned}$$

Stark's conjecture predicts that

$$L(s, \chi^{-1}) \sim \alpha_\chi s \sum_{g \in G} \log|g\epsilon|_{v_0} \chi^{-1}(g) \text{ near } s = 0$$

i.e.

$$\boxed{L'(0, \chi) = \alpha_{\chi^{-1}} \sum_g \log|g\epsilon|_{v_0} \chi(g),}$$

where (using the refined version) $\alpha_{\chi^\sigma} = (\alpha_\chi)^\sigma$.

How does this compare with our earlier computation of the L-value in Theorem 5.3.3? There we found

$$L'(0, \chi) = -\frac{1}{2} \sum_{i \in (\mathbb{Z}/m\mathbb{Z})^\times} \chi(i) \log|1 - \zeta^i|$$

Put $\epsilon = 1 - \zeta \in E$. Then we can rewrite the above as

$$L'(0, \chi) = -\frac{1}{2} \sum_{g \in (\mathbb{Z}/m\mathbb{Z})^\times} \chi(g) \log|g \cdot \epsilon|$$

which verifies Stark's conjecture with the $\alpha_\chi \in \mathbb{Q}$.

There is a slight problem here. This $\epsilon = 1 - \zeta \in E$ is not always a unit. Indeed, $\mathrm{Nm}(1 - \zeta_p) = p$. However, it is if m is not a prime power. For example, consider $1 - \zeta_p \zeta_q$. It is easy to see that this is a unit away from p, q. If $\widetilde{p}$ lies above p, note that

$$|1 - \zeta_p \zeta_q|_{\widetilde{p}} = |\zeta_q^{-1} - \zeta_p|_{\widetilde{p}} = |\zeta_q^{-1} - 1|_{\widetilde{p}}$$

because $\zeta_p \equiv 1 \pmod{p}$, but the norm of this last guy is q.

What if m *is* a prime power? Then we take instead the algebraic unit

$$\epsilon = \frac{1 - \zeta^j}{1 - \zeta} \quad (j, m) = 1.$$

Then

$$\sum_{g \in (\mathbb{Z}/m\mathbb{Z})^\times} \chi(g) \log|g\epsilon| = (1 - \chi(j)) \sum \chi(g) \log|g \cdot (1 - \zeta)|$$

so this is compatible with Stark's conjecture, with $\alpha_\chi = (-1/2?)\frac{1}{(1-\chi(j))}$. Note that this depends on j and is no longer rational, but it's visibly Galois-equivariant.

7.5. **Imaginary quadratic fields.** Now let's look at the case where $K = \mathbb{Q}(\sqrt{-d})$ and E/K is abelian. We're going to prove Stark's conjecture in this case by studying a character $\chi : \mathrm{Gal}(E/K) \to \mathbb{C}^\times$. [A reference is Stark's paper "Derivatives of L-functions... IV," but our argument will be a little different.]

Recall that for an abelian extension over $\mathbb{Q}$ of the form $\mathbb{Q}(\zeta_m)/\mathbb{Q}$, our verification of Stark's conjecture came down to studying

$$\frac{d}{ds}|_{s=0} \sum_{n\in\mathbb{Z}} |n+\alpha|^{-s} = -\frac{1}{2}\log((1-u)(1-u^{-1})), \quad u = e^{2\pi i\alpha}.$$

Over $\mathbb{Q}(\sqrt{-d})$, in order to evaluate $L'(0,\chi)$ we have to similarly evaluate

$$\frac{d}{ds} \sum_{z\in\Lambda\subset\mathbb{C}} |z+\alpha|^{-s} = ?$$

for $\alpha \in K$ and Λ an ideal in $\mathbb{Q}(\sqrt{-d})$. This turns out to be $\log|\varphi|$ for some $\varphi \in \overline{\mathbb{Q}}^\times$ (analogous to the result over $\mathbb{Q}$). In fact, it's a special value of some modular function. This reflects a broader analogy between $\mathbb{Q}$ and a quadratic imaginary field, in which results over $\mathbb{Q}$ extend to quadratic imaginary fields after replacing $\mathbb{G}_m$ by an appropriate elliptic curve.

More precisely, we'll show:

Theorem 7.5.1. *For $z \in \mathbb{H}$, $\alpha = pz+q \neq 0$ where $p,q \in \mathbb{Q}$, there exists M such that*

$$H(z) := \frac{d}{ds}|_{s=0} \sum_{(c,d)\in\mathbb{Z}^2} \frac{1}{|cz+d+\alpha|^s} = \frac{1}{M}\log|\varphi|$$

where φ is a modular function on $\Gamma(N)\backslash\mathbb{H}$ and $N \in \mathbb{Z}$ is such that $Np, Nq \in \mathbb{Z}$.

Remark 7.5.2. This formula, together with the explicit description of φ, is called *Kronecker's second limit formula.* Kronecker's first limit formula deals with the case $\alpha = 0$.

In fact, the q-expansion of φ at every cusp has coefficients inside $\mathbb{Q}(\zeta_N)$. Now, $X(N)$ has a model over $\mathbb{Q}(\zeta_N)$, so this implies that $\varphi \in \mathbb{Q}(\zeta_N)(X(N))$. Therefore, for any CM-point z we have $\varphi(z) \in \overline{\mathbb{Q}}$, because z corresponds to a point of $X(N)$ defined over $\overline{\mathbb{Q}}$.

We won't discuss this fact about the Fourier coefficients; it follows from a finite computation (of the first few coefficients). The point is that being holomorphic is very close to being algebraic.

Remark 7.5.3. This proof is quite remarkable. To evaluate a sum, it recognizes that a family of variations is actually a *modular function*, and then specializes. It's "obvious" from modern perspectives.

Stark proves Theorem 7.5.1 by an explicit computation. We'll sketch that, and then we'll try to explain *why* this should be true by "pure thought."

Sketch of explicit proof. First sum over d, using the evaluation of

$$\frac{d}{ds}|_{s=0} \sum \frac{1}{|d+\beta|^s}$$

that we performed earlier. (Recall that we only addressed this for real α earlier, so one first has to extend this formula.) Then one gets a sum over logarithms, which (after addressing convergence issues) can be recast as the logarithm of an infinite product, which turns out to be the same as that which shows up in the Jacobi triple product formula. This is a highly non-trivial analysis, but ends up working out. ♠♠♠ TONY: [never understood JTP] □

Now we give the conceptual proof that $H = \log|\varphi|$. The outline is as follows.

(1) First show that H is $\Gamma(N)$-invariant, i.e. descends to a function on $\Gamma(N)\backslash\mathbb{H}$. We write

$$H(z) = \frac{d}{ds}|_{s=0} \sum_{(c',d')\in\mathbb{Z}^2+(p,q)} \frac{1}{|c'z+d'|^s}.$$

(2) Check that $\partial\overline{\partial}H = 0$, i.e. H is harmonic.

(3) Argue that for a suitable integer M, $M \cdot H$ is *locally* of the form $\log|\varphi|$ near every point of $X(N)$. (Any harmonic function is the real part of a holomorphic function, which you exponentiate. The integer is needed at the cusps, because you can exponentiate certain things.)

(4) There is an obstruction to globalizing φ, which lies in $H^1(X(N), S^1)$. This is *torsion* by an argument with Hecke operators (the "Manin-Drinfeld" trick, which is usually used in a different context).

Proof Sketches. (1) Note that for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{R})$ and $z \in \mathbb{H}$,

$$\mathrm{Im}(\gamma z) = \frac{\mathrm{Im}\, z}{|cz+d|^2}. \tag{7}$$

Set

$$Z_s(z) = \sum_{(c',d')\in\mathbb{Z}^2+(p,q)} \frac{\mathrm{Im}(z)^s}{|c'z+d'|^{2s}}.$$

If there were no shift, and we omitted the origin, then this would be just the standard real-analytic Eisenstein series. By (7) we can rewrite this as

$$Z_s(z) = \sum_{(c',d')\in\mathbb{Z}^2+(p,q)} \mathrm{Im}(\gamma_{c',d'}(z))^s$$

where $\gamma_{c'd'}$ is any matrix in $\mathrm{SL}_2(\mathbb{R})$ of the form $\begin{pmatrix} * & * \\ c' & d' \end{pmatrix}$. For $\gamma \in \Gamma(N)$,

$$Z_s(\gamma z) = \sum_{(c',d')\in\mathbb{Z}^2+(p,q)} \mathrm{Im}(\gamma_{c'd'}\gamma z)^s$$

and $\gamma_{c'd'}\gamma$ has the same property as the $\gamma_{c'd'}$ (this was the key property of choosing N to be a common denominator). So we've established that $Z_s(z)$ is Γ_N-invariant. Now this isn't quite the same as $Z_s(z)$, but at $s=0$ they basically coincide. More precisely, we claim that

$$\boxed{H(z) = \frac{1}{2}\frac{d}{ds}|_{s=0} Z_s(z).}$$

It is easy to see that

$$H(z) = \frac{1}{2}\frac{d}{ds}|_{s=0} Z_s(z) - \log(\mathrm{Im}\, z) Z_0(z)$$

so the result follows from the computation that $Z_0 \equiv 0$.

Exercise 7.5.4. Check this. (It's similar to how L-functions for non-trivial characters vanish at 0). This *fails* if $\alpha = 0$.

(2) $Z_s(z)$ is convergent for $\mathrm{Re}\, s \gg 0$, but it extends by meromorphic continuation in the s-variable. In fact, $(s-1)Z_s(z)$ is holomorphic in s and smooth in z. Let Δ be the Laplacian $-y^2(\partial_{xx} + \partial_{yy})$ on $\mathbb{H}$. (This is the Laplacian associated to the hyperbolic metric, or alternativley the one that is SL_2 invariant.) Then you can check that

$$\boxed{\Delta(\mathrm{Im}(z)^s) = s(1-s)\mathrm{Im}(z)^s.}$$

Since Δ is $\mathrm{SL}_2(\mathbb{R})$-invariant, $\Delta Z_s = s(1-s)Z_s$ by a term-by-term comparison. Technically, this requires some things about absolute convergence, and then analytic continuation - that's ultimately a matter of the level of smoothness.

Therefore,

$$\Delta(\frac{d}{ds}Z_s) = s(1-s)\frac{d}{ds}Z_s + (1-2s)Z_s.$$

Evaluating at $s=0$, we get $\Delta(\frac{d}{ds}|_{s=0} Z_s) = Z_0$. But as mentioned above, for $s=0$ we have $Z_s \equiv 0$. (When $\alpha = 0$, i.e. of the first Kronecker Limit Formula, $\frac{d}{ds}Z_s$ is *not* quite harmonic.)

Remark 7.5.5. There is a general story underlying this - see the paper of Quillen, "Determinants of Cauchy-Riemann operators on Riemann surfaces." There, $|\phi|$ is interpreted as the determinant of the $\overline{\partial}$-operator acting on a line bundle on the elliptic curve determined by z.

We're going to mostly skip (3) because it's just a local computation at the cusps. It has no content at the interior, since that's just a general property of harmonic functions: if H is harmonic, then $H = \operatorname{Re}\psi = \log|e^{\psi}|$. The real content is why you can still do this at the cusps, and the question boils down to understanding the asymptotic behavior at the cusps of H. This is one of the first things treated. Basically, you study the asymptotic version of the differential equation determined by harmonicity, and at the cusps you find that you should get a linear combination of some simple fundamental solutions.

(4) Now we argue that $Z_s(z)$ is globally $\frac{1}{M}\log|\varphi|$. The obstruction to globalizing φ is represented by a class $\alpha \in H^1(X(N), \mathbb{R}/\mathbb{Z})$. We claim that α is torsion.

Manin and Drinfeld proved that for any cusps α, β for $X(N)$, $[\alpha]-[\beta]$ is torsion in $\mathrm{Jac}(X(N))$, i.e. there exists M such that $M([\alpha]-[\beta]) = \mathrm{Div}(f)$. Their original motivation was to construct rational points on an elliptic curve by projecting points from $X(N)$ via modular correspondences. (We now know that for an elliptic curve E of conductor N, then there is a uniformization $X_0(N) \to E$, hence also $J_0(N) \to E$.) The first thing you try is to use the cusps, and this is telling you that you can only get torsion points.

Example 7.5.6. Consider $X_0(11)$. Then $10([\infty]-[0]) = \frac{\Delta(11z)}{\Delta(z)}$. Akshay says that $[\infty]-[0]$ is even actually 5-torsion. In fact, $X_0(11)$ is an elliptic curve of conductor 11. There are three such: one is $X_0(11)/\langle[0]-[\infty]\rangle$. The other is $X_1(11)$, and this admits a degree 5 isogeny *to* $X_0(11)$.

$$X_1(11) \to Z_0(11) \to X_0(11)/\langle[0]-[\infty]\rangle.$$

Going down the isogenies, the equations get bigger and bigger, because the Faltings height gets bigger.

Claim 1. There exists a prime p such that T_p has eigenvalue $p+1$ acting on $H(z)$. (In fact any prime $p \equiv 1 \pmod{N}$ will work.)

Claim 2. All eigenvalues λ of T_p on $H^1(X(N), \mathbb{R})$ satisfy $|\lambda| < p+1$.

These two results put together imply that the class of $H(z)$, α must be torsion (after also ironing out some issue with S^1-coefficients). Indeed, if

α were not torsion, then its image in $H^1(X(N),\mathbb{R}/\mathbb{Q})$ would be non-zero. If you trace through the derivation of the obstruction class, you get that $T_p\alpha = (p+1)\alpha$, hence $T_p\overline{\alpha} = (p+1)\overline{\alpha}$. By composing with a linear map $\mathbb{R}/\mathbb{Q} \to \mathbb{Q}$, we get $\overline{\overline{\alpha}} \in H^1(X(N),\mathbb{Q})$ with $T_p\overline{\overline{\alpha}} = (p+1)\overline{\overline{\alpha}}$. Therefore, $\overline{\overline{\alpha}} = 0$. That implies that $\overline{\alpha}$ is 0, i.e. α is torsion.

Remark 7.5.7. The Ramanujan conjecture, proved in this case by Shimura, implies that all the eigenvalues λ of T_p satisfy the stronger bound $|\lambda_p| \leq 2\sqrt{p}$, using an interpretation of the eigenvalues of Hecke in terms of point counts on the Jacobian $\mathrm{Jac}(X(N))$.

Proof of Claim 2. We observe that if f is a function on $\Gamma(N)\backslash\mathbb{H}$, and $f \to 0$ at the cusps, and $T_p f = (p+1)f \implies f = 0$. Why? Choose $x \in \Gamma(N)\backslash\mathbb{H}$ maximizing $|f|$. Then

$$(p+1)f(x) = T_p f(x) = \underbrace{\sum_{y \in T_p x}}_{\text{set size } p+1} f(y).$$

Here we think of T_p as a "multivalued function in x." By the maximality assumption of $|f(x)|$, we must have $f(y) = f(x)$ for all $y \in T_p x$. Similarly, $f(y) = f(x)$ for all $y \in T_p(T_p(x))$, etc. But $\bigcup T_p^n(x)$ is *dense* in $\mathbb{H}$, because it is the orbit of x under $\mathrm{PGL}_2(\mathbb{Z}[1/p])$ ♠♠♠ TONY: [why?], which is dense in $\mathrm{PGL}_2(\mathbb{R})$. Hence f is constant, and this constant must be 0.

♠♠♠ TONY: [Akshay says that this should still true for mod ℓ modular forms, although nobody knows how to do it]

For $H^1(X(N),\mathbb{R})$ each class is represented uniquely by $f(z)\,dz + \overline{g(z)\,dz}$ for f, g weight 2 cusp forms. We want to view this as a function, so we can apply the preceding result. Now $f(z)\,dz + \overline{g(z)\,dz}$ is a 1-form on $\Gamma(N)\backslash\mathbb{H}$, i.e. a function on the tangent bundle of $\Gamma(N)\backslash\mathbb{H}$. $\mathrm{PSL}_2(\mathbb{R})$ acts on $\mathbb{H}$, hence on its unit tangent bundle (using the natural Riemannian metric on $\mathbb{H}$), and this action is *simply transitive* (its stabilizer in $\mathbb{H}$ was $\mathrm{SO}_2(\mathbb{R})$). So we get a function on $\Gamma(N)\backslash\mathrm{PSL}_2(\mathbb{R})$, which is the aforementioned unit tangent bundle, and we can apply the same argument (noting that because f, g are cusp forms, this function goes to 0 at ∞). □

Proof of Claim 2. We have to compute $T_p Z_s$. Z_s is a function on $\Gamma(N)\backslash\mathbb{H}$, but for this argument we will prefer to think of Z_s as a function on the set $\{\text{lattices } \Lambda \subset \mathbb{C}, \alpha \in N^{-1}\Lambda/\Lambda)\}$ and Z_s takes this to

$$\mathrm{vol}(\Lambda)^s \sum_{z \in (\Lambda+\alpha)} |z|^{-2s}.$$

Then

$$T_p Z_s(\Lambda,\alpha) = \sum_{\substack{\Lambda' \\ [\Lambda':\Lambda]=p}} Z_s(\Lambda',\alpha)$$

$$= p^{-s}\operatorname{vol}(\Lambda)^s \sum_{\substack{\Lambda' \\ [\Lambda':\Lambda]=p}} \left(\sum_{z\in\Lambda'} |z+\alpha|^{-2s} \right)$$

The z that appear in this sum are one of the following form:

- If $z \in \Lambda$, then $z \in \Lambda'$ for all $(p+1)\Lambda's$.
- If $z \in p^{-1}\Lambda$ but $z \notin \Lambda$, then $z \in \Lambda'$ for a unique Λ',

Therefore, the above is

$$= p^{-s}\operatorname{vol}(\Lambda)^s \left(p \sum_{z\in\Lambda} \frac{1}{|z+\alpha|^{2s}} + \sum_{z\in p^{-1}\Lambda} \frac{1}{|z+\alpha|^{2s}s} \right).$$

The first part is $p^{1-s}Z_s(\Lambda,\alpha)$ and the second is $p^s Z_s(\Lambda, p\alpha)$. If $p \equiv 1 \pmod N$, then we get $(p^s + p^{1-s})Z_s(\Lambda,\alpha)$. At $s=0$, we get $p+1$, as desired. □

Reference: Siegel, "... advanced analytic number theory."

□

8. Class numbers of cyclotomic fields

8.1. **Reformulating Stark's conjecture.** Let E/K be a field extension with Galois group G and $\rho: G \to \mathrm{GL}(V)$ a representation. Stark's conjecture predicts that

$$L(s,\rho^*) \sim \alpha_\rho s^{r_\rho} \rho(\mathscr{R}) \quad \text{near } s=0$$

where $\rho(\mathscr{R}) \in Z(\mathbb{R}[G])^*/Z(\mathbb{Q}[G])^*$ and $\alpha_\rho \in \mathbb{Q}(\rho)^*$. If we admit the refined conjecture (which we shall always do in the future), then we get that moreover $\rho \mapsto \alpha_\rho$ is Galois equivariant.

We can reformulate this by using a $\mathbb{C}[G]$-valued L-function. This is just a matter of packaging - we replaced the regulator with something in the group algebra, and now we want to do the same for everything else. Let $e_\rho \in \mathbb{C}[G]$ be the idempotent associated to ρ, i.e. if α is an irreducible representation (viewed on the group algebra), then we have an equality of *endomorphisms*

$$\alpha(e_\rho) = \begin{cases} 0 & \alpha \ncong \rho, \\ \mathrm{Id}_\rho & \alpha \cong \rho. \end{cases}$$

Exercise 8.1.1. Check that $e_\rho = \dim\rho \cdot \chi_{\rho^*}$.

Solution. We recall the orthogonality of matrix coefficients:

$$\frac{1}{|G|}\sum_{g\in G}\langle \rho(g)x,y\rangle\overline{\langle \alpha(g)w,z\rangle} = \begin{cases} 0 & \alpha \ncong \rho, \\ \frac{1}{\dim\rho}\langle x,w\rangle\overline{\langle y,z\rangle} & \alpha \cong \rho. \end{cases}$$

This immediately shows that if $\alpha \ncong \rho$ then

$$\alpha(e_\rho^*) = \sum_{g\in G}\overline{e_\rho(g)}\alpha(g) = 0$$

since $e_\rho \propto \chi_\rho$ is a matrix coefficient of ρ.

On the other hand, if $\alpha \cong \rho$ and x_i in an orthonormal basis for the space of ρ, then the jk-matrix coefficient of $\alpha(e_\rho)$ is

$$\begin{aligned}\left\langle \sum_{g\in G}\overline{\langle \rho(g)x_i,x_i\rangle}\alpha(g)x_j,x_k \right\rangle &= \sum_{g\in G}\overline{\langle \rho(g)x_i,x_i\rangle}\langle \alpha(g)x_j,x_k\rangle \\ &= \frac{\delta_{ijk}|G|}{\dim\rho}\end{aligned}$$

Summing over i, we find that

$$\langle \alpha(\chi_{\rho^*})x_j,x_k\rangle = \frac{\delta_{jk}}{\dim\rho}.$$

This shows that $e_\rho = \dim\rho\cdot\chi_{\rho^*}$. □

Set

$$\mathscr{L}(s) = \sum_{\text{irred. } \rho} L(s, \rho^*) e_\rho \in Z(\mathbb{C}[G]).$$

(This is in the center because evaluating at idempotents identifies $\mathbb{C}[G]$ with a direct sum of matrix algebras, and the e_ρ are identity matrices in their respective components.)

The projection of $\mathscr{L}(s)$ to different ρ have different vanishing orders, so it's best to talk about vanishing orders of $\mathscr{L}(s)$ after projecting to various factors. Let $\{\rho_1, \dots, \rho_r\}$ be an orbit of $\mathrm{Aut}(\mathbb{C})$ on the irreducible representations of G. This corresponds to a $\mathbb{Q}$-simple factor A of $\mathbb{Q}[G]$. That is, since $\mathbb{Q}[G]$ is a semisimple $\mathbb{Q}$-algebra, we have

$$\mathbb{Q}[G] \cong \bigoplus M_{n_i}(D_i).$$

If we pick a simple summand $A = M_N(D)$ where $[Z(D) : \mathbb{Q}] = r$, then after tensoring up to $\mathbb{C}$ it will break into r different matrix algebras), corresponding to the representations $\rho_1, \dots, \rho_r$.

By the explicit formula for r_ρ, we get that r_ρ is constant on $\{\rho_1, \dots, \rho_r\}$.

Exercise 8.1.2. Check this.

Write $\mathscr{L}_A(s)$ for the projection of $\mathscr{L}(s)$ to $A \otimes \mathbb{C}$ and $\mathscr{R}_A$ for the projection of $\mathscr{R}$ to $A \otimes \mathbb{C}$.

Conjecture 8.1.3 (Reformulation of Stark's conjecture)**.** *For each A,*

$$\mathscr{L}_A(s) \sim \alpha_A s^{r_A} \mathscr{R}_A \text{ near } s = 0$$

for some $\alpha \in Z(\mathbb{Q}[G])^*$.

We recover the old conjecture by applying $\rho \in \{\rho_1, \dots, \rho_r\}$:

$$L(s, \rho^*) \sim \rho(\alpha_A) s^{r_\rho} \rho(\mathscr{R}_A)$$

where $\rho(\alpha)$ is the scalar by which α acts in ρ. Here $r_\rho = r_A$ for all ρ in the equivalence class determined by A. Note that now the Galois equivariance is now *packaged into* the statement. We can absorb $\rho(\alpha_A)$ into $\rho(\mathscr{R}_A)$ to write this more concisely as

$$L(s, \rho^*) \sim s^{r_\rho} \rho(\mathscr{R}_A).$$

Note that this is a slight abuse of notation since $\rho(\mathscr{R}_A)$ is a linear transformation, but since $\mathscr{R}_A$ is in the center of the group algebra it is a scalar multiple of the identity, and we identify it with that scalar.

8.2. **Stickelberger's Theorem.** Let E/K be an abelian extension of number fields with Galois group G.

Fantasy. We would like to have an "equivariant class number formula"

$$\mathscr{L}_A \sim s^r \frac{h_A \mathscr{R}_A}{w_A}, \quad \text{near } s=0$$

where $h_A \in \mathbb{Z}[G]$. What would h_A be? Since we are in a fantasy, we can dream that there is a decomposition

$$\mathrm{Cl}_E \cong \mathbb{Z}[G]/\alpha_1 \oplus \mathbb{Z}[G]/\alpha_2 \oplus \ldots \oplus \mathbb{Z}[G]/\alpha_r.$$

Then we set $h_A = (\alpha_1 \cdot \ldots \cdot \alpha_r)_A$. This is still only defined up to $\mathbb{Z}[G]^\times$. So at the very least, $h_A \in \mathbb{Z}[G]$ should kill $(\mathrm{Cl}_E)_A$ as an endomorphism. Now that is a statement that does make sense, without all this fantasized structure.

Conjecture 8.2.1 (Brumer-Stark-Stickelberger)**.** *If E/K is abelian, then*

$$w_E \cdot \mathscr{L}_B(0) \textit{ annihilates } \mathrm{Cl}_E.$$

Here w_E is the number of roots of unity in E and $\mathscr{L}_B$ is a slightly modified version of the L-function to make this non-trivial, which we will shortly explain. As a first approximation, you can just think of it as $\mathscr{L}$.

We have

$$\mathscr{L}(s) = \sum_\chi L(s, \chi^{-1}) e_\chi$$

where $e_\chi = \frac{1}{|G|} \sum_{g \in G} \chi(g)^{-1} g$ (Check that this is compatible with Exercise 8.1.2). Now A corresponds a Galois orbit $\{\chi_1, \ldots, \chi_r\}$. All the χ_i have a common conductor $\mathfrak{f}$ (an ideal of K). That was the point of restricting to things in a single Galois orbit. For $\chi \in A$,

$$L(s, \chi^{-1}) = \sum_{(I,\mathfrak{f})=1} \frac{\chi^{-1}(I)}{(\mathrm{Nm}\, I)^s}$$

where $\chi(I = \prod \mathfrak{p}_i^{a_i}) \to G$ takes $\mathfrak{p}_i \mapsto \mathrm{Frob}_{\mathfrak{p}_i}$ and extends by multiplicativity. So

$$\begin{aligned}
\mathscr{L}(s)_A &= \left(\sum_\chi L(s, \chi^{-1}) e_\chi \right)_A \\
&= \left(\sum_\chi \frac{1}{|G|} \sum_{g \in G} L(s, \chi^{-1}) \chi(g)^{-1} \cdot g \right)_A \\
&= \left(\frac{1}{|G|} \sum_\chi \sum_{g \in G} \sum_{(I,\mathfrak{f})=1} \frac{\chi^{-1}(I)\chi^{-1}(g)}{(\mathrm{Nm}\, I)^s} \cdot g \right)_A.
\end{aligned}$$

Hence by the usual cancellation of summing over χ, we have

$$\boxed{\mathscr{L}(s)_A = \left(\sum_{(I,\mathfrak{f})=1} \frac{1}{(\operatorname{Nm} I)^s} [I^{-1}] \right)_A}$$

where here $[I^{-1}]$ is the class of I in G.

Example 8.2.2. Let $E = \mathbb{Q}(\zeta_m), K = \mathbb{Q}$, and A any orbit of a character of $(\mathbb{Z}/m\mathbb{Z})^\times$. Then

$$\mathscr{L}(s)_A = \left(\sum_a [a^{-1} \mod m] \left(\sum_{n \equiv a \pmod m} \frac{1}{n^s} \right) \right)_A \in \mathbb{C}[(\mathbb{Z}/m\mathbb{Z})^\times].$$

Then $\mathscr{L}_B$ is basically defined by this formula without the projection to A. Let $f = \operatorname{disc}(E/K)$, an ideal of K. Then

$$\mathscr{L}_B = \sum_{(I,\mathfrak{f})=1} \frac{1}{(\operatorname{Nm} I)^s} [I^{-1}].$$

So this is like $\mathscr{L}$ but removing the same Euler factors from everything, whereas the characters have different levels of ramification (conductor).

Example 8.2.3. Let $E = \mathbb{Q}(\zeta_m)$, etc. Then $\mathscr{L}_B = \mathscr{L}$ from now on is

$$\boxed{\mathscr{L}_B(s) = \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^\times} [a^{-1}] \sum_{n \equiv a} \frac{1}{n^s}.}$$

Then $\mathscr{L}_B(0)$ (from Hurwitz zeta function) is

$$\sum_a [a^{-1}] \left(\frac{1}{2} - \left\{ \frac{a}{m} \right\} \right).$$

The number of roots of unity in $\mathbb{Q}(\zeta_m)$ is m if m is even and $2m$ if m is odd. So if m is odd, then Brumer's conjecture in this predicts that

$$m \sum_a [a^{-1}] - 2 \sum_{a \in (\mathbb{Z}/m\mathbb{Z})^\times} [a^{-1}] a \text{ kills } \operatorname{Cl}_{\mathbb{Q}(\zeta_m)}.$$

Remark 8.2.4. $\mathscr{L}_B$ is present in order to make a clean integral conjecture possible, since the splitting occurs over $\mathbb{Q}$ and not $\mathbb{Z}$.

Theorem 8.2.5 (Stickelberger)**.** *The element*

$$\sum_{1 \le a < m, (a,m)=1} a[a^{-1}] \in \mathbb{Z}[(\mathbb{Z}/m\mathbb{Z})^\times].$$

kills $\operatorname{Cl}(\mathbb{Q}(\zeta_m))$.

This proves Brumer because $\sum_a [a^{-1}]$ realizes the ideal norm (to $\mathbb{Q}$, and then extended back) and so kills $\mathrm{Cl}(\mathbb{Q}(\zeta_m))$. In fact, we even win by an extra factor of 2.

Remark 8.2.6. $\mathrm{Cl}(\mathbb{Q}(\zeta_m))$ is trivial until $m = 23$ and then starts to increase rapidly in size.

Note that this says nothing about $\mathrm{Cl}(\mathbb{Q}(\zeta_m)^+ = \mathbb{Q}(\zeta_m + \zeta_m^{-1})$. That's because on this subfield you can pair up the action of a and $-a$, and so the thing is formally 0 again.

Exercise 8.2.7. Check this.

Next we'll prove Stickelberger's theorem, and prove Herbrand's theorem (which is Stickelberger plus ϵ). We'll finish up by doing the converse to Herbrand's theorem. This involves Euler systems plus more (both real and p-adic L-functions).

Proof. The idea is to factorize Gauss sums. Let $K = \mathbb{Q}(\zeta_m)$ and $\ell \equiv 1 \pmod{m}$ be a prime. In $\mathcal{O}_K$, (ℓ) splits into $\lambda_1 \cdot \ldots \cdot \lambda_d$. Let $d = [K : \mathbb{Q}]$, where λ_i is a prime of degree 1. The class of λ_i, as ℓ varies, generates Cl_K (basically Dirichlet/Cebotarev for this field, because the density of primes of degree > 1 is 0.)

Let $\theta = \sum_{1 \le a < m, (a,m)=1} a[a^{-1}]$. So it's enough to show that θ kills all such λ_i. But for that it's enough to show that θ kills *one* of them.

In other words, we want some element of K whose associated ideal factors as $\theta(\lambda_1)$. If we index λ_i by $(\mathbb{Z}/m\mathbb{Z})^\times$, then

$$\theta(\lambda_i) = \prod_{a \in (\mathbb{Z}/m\mathbb{Z})^\times} \lambda_{a^{-1}}^a.$$

For example, for $m = 5$ this is

$$\lambda_1^1 \lambda_2^3 \lambda_3^2 \lambda_4^4.$$

Let $\chi : (\mathbb{Z}/\ell\mathbb{Z})^\times \to \mu_m$ be a non-trivial character. Fix ζ_ℓ a primitive ℓth root of unity. Let g_χ be the Gauss sum

$$\boxed{g_\chi = \sum_{x \in (\mathbb{Z}/\ell\mathbb{Z})^\times} \zeta_\ell^x \chi(x).}$$

A priori this lies in $K^* := K(\zeta_\ell)$, but we claim that in fact $g_\chi^m \in K$, and has the desired factorization.

Exercise 8.2.8. As a plausibility check, compute the norms down to $\mathbb{Q}$ and check that they agree.

Properties of Gauss Sum.

(1) Let $\tau \in \mathrm{Aut}(K(\zeta_\ell)/K)$. Then we claim that

$$\boxed{g_\chi^\tau = g_\chi \chi([\tau]^{-1})}$$

where $[\tau]$ is the corresponding element of $(\mathbb{Z}/\ell\mathbb{Z})^\times$. In other words,

$$\frac{g_\chi^\tau}{g_\chi} = \chi([\tau]^{-1}) \in \mu_m.$$

If you think about this carefully, you'll see that this is the *only* property of the Gauss sum that we need.

Proof. For $\tau \in \mathrm{Aut}(K^*/K) \cong (\mathbb{Z}/\ell\mathbb{Z})^*$, we have

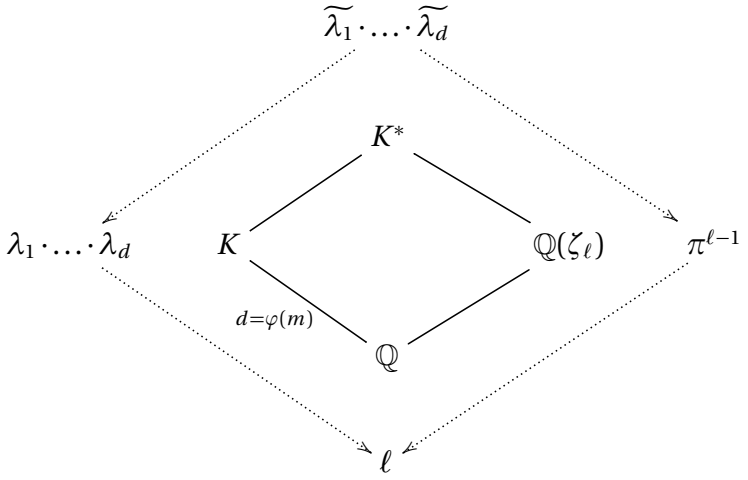
$$\tau(g_\chi) = \sum_{i \in (\mathbb{Z}/\ell)^\times} \chi(i)\zeta_\ell^{i[\tau]} = \chi([\tau])^{-1} g_\chi.$$

This shows that g_χ^m is τ-invariant, since χ has image in μ_m by definition. □

(2) The norm of g_χ to $\mathbb{Q}$ is a power of ℓ, so the only primes dividing g_χ are above ℓ. In fact, $|g_\chi|^2 = \ell$, and the same is true for any conjugate.

Exercise 8.2.9. Check this.

So ℓ splits completely K as $\lambda_1, \dots, \lambda_d$, and each of these is totally ramified in K^*. We'll factorize g_χ in K^* and then pass to K.



where $\pi = 1 - \zeta_\ell$.

So let's compute

$$g_\chi \pmod{\widetilde{\lambda_i}^{\ell-1}} \equiv \sum_{x \in (\mathbb{Z}/\ell)^\times} \chi(x) \underbrace{\zeta_\ell^x}_{(1+\pi)^x} \pmod{\widetilde{\lambda_i}^{\ell-1}}$$

Now, viewing

$$\chi : (\mathbb{Z}/\ell)^\times \to \mathcal{O}_K^\times \to (\mathcal{O}_K/\lambda_i)^\times \cong (\mathbb{Z}/\ell)^\times$$

as a map $(\mathbb{Z}/\ell)^\times \to (\mathbb{Z}/\ell)^\times$, it must be of the form $x \mapsto x^q$ for some $0 \le q \le \ell - 1$. So

$$\begin{aligned} g_\chi \pmod{\widetilde{\lambda_i}^{\ell-1}} &\equiv \sum_{x=1}^{\ell-1} x^q (1+\pi)^x \pmod{\widetilde{\lambda_i}^{\ell-1}} \\ &= \sum_{x=1}^{\ell-1} x^q \left(1 + \pi x + \frac{\pi^2 x(x-1)}{2} + \dots\right) \pmod{\widetilde{\lambda_i}^{\ell-1}} \end{aligned}$$

expanding by the binomial theorem in the last equality. We're doing this by brute force for now; we'll see a better way later. Now, $\sum_{x=1}^{\ell-1} x^i \equiv 0 \bmod \ell$ unless $x^i = 1$ for all x, i.e. $\ell - 1 \mid i$. So all the terms vanish mod $\widetilde{\lambda}_i^{\ell-1}$ until you hit $i = \ell - 1$. For the purposes of computing the valuation, we can focus on the lowest-order contributing term:

$$\pi^{\ell-1-q} \sum_{x=1}^{\ell-1} x^q \binom{x}{\ell-1-q} sim \pi^{\ell-1-q} \sum_{x'=1}^{\ell-1} x^{\ell-1} \sim -\pi^{\ell-1-q}.$$

(Here $\sim$ means up to units.) Now $v_{\widetilde{\lambda_i}}(\pi) = 1$, so the upshot is that

$$\boxed{v_{\widetilde{\lambda_i}}(g_\chi) = \ell - 1 - q \text{ if } \chi = (x \mapsto x^q) : (\mathbb{Z}/\ell)^\times \to (\mathbb{Z}/\ell)^\times.}$$

Ok, so we've found that

$$v_{\widetilde{\lambda_i}}(g_\chi^m) = m(\ell - 1 - q).$$

Therefore,

$$v_{\lambda_i}(g_\chi^m) = \frac{m(\ell-1-q)}{\ell-1} = m - \frac{q}{(\ell-1)/m}.$$

How does q vary with λ_i? Take $\sigma \in \mathrm{Gal}(K/\mathbb{Q}) \cong (\mathbb{Z}/m\mathbb{Z})^\times$. Suppose

$$\chi(x) \equiv x^q \pmod{\lambda}$$

for $x \in (\mathbb{Z}/\ell)^\times$. Then

$$\chi(x)^{[\sigma]} \equiv x^q \pmod{\sigma\lambda}.$$

Therefore,

$$\chi(x) \equiv q^{q[\sigma]^{-1}} \pmod{\sigma\lambda}.$$

If we choose λ_1 so that $q_{\lambda_1} = \frac{\ell-1}{m}$, then $q_{\sigma\lambda_1} = \frac{\ell-1}{m}[\sigma^{-1}]$ (choosing $1 < [\sigma^{-1}] < m$).

So

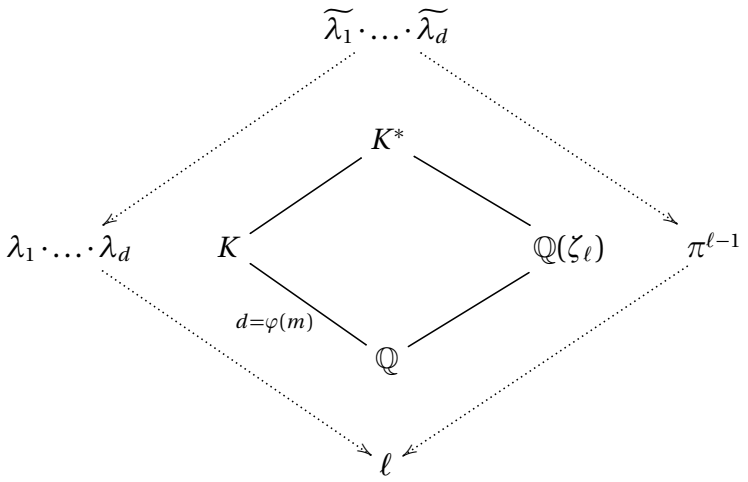
$$(g_\chi^m) = \prod_{\sigma \in \mathrm{Gal}(K/\mathbb{Q})} (\sigma\lambda_1)^{m-[\sigma^{-1}]}.$$

Therefore,

$$\prod (\sigma\lambda_1)^{[\sigma^{-1}]}$$

is principal, because we know that $\prod_\sigma (\sigma\lambda_i)^m$ is principal (and in fact equal to ℓ^m). □

Remark 8.2.10. As alluded to earlier, if you think about this carefully then you'll realize that we didn't *really* need to know about the Gauss sum. We could have worked out $v_{\lambda_i}(g_\chi^m) \pmod{\ell-1}$ by pure thought. How?



Now, $K_{\widetilde{\lambda_i}}^\times / K_{\lambda_i}$ is a totally ramified extension of degree $\ell-1$. So $g_\chi \in K_{\widetilde{\lambda_i}}^\times$. We know what $\tau(g_\chi)/g_\chi$ looks like, for τ in the Galois group $(\mathbb{Z}/\ell)^*$ (namely $\chi([\tau]^{-1})$).

More abstractly, say L/E is a totally, tamely ramified extension of local fields with Galois group G. Then there's a homomorphism $G \to k_L^\times = k_E^\times$ given by $\tau \mapsto \frac{\tau(\pi_L)}{\pi_L} \in k_L^\times$, where π_L is a uniformizer. This is independent of the choice of π_L because the extension is totally ramified (G acts trivially on the residue field). In our case $L = K_{\widetilde{\lambda_i}}^\times, E = K_{\lambda_i}$ this gives the usual identification of $\mathrm{Gal}(L/E)$ with $(\mathbb{Z}/\ell)^\times$.

Given $x \in L$, we can determine $v(x) \mod [L:E]$ from the knowledge of $\tau(x)/x$ for $\tau \in G$. The reason is that

$$\frac{\tau(x)}{x} \pmod{\pi_L} \equiv \left(\frac{\tau\pi_L}{\pi_L}\right)^{v(x)} \pmod{\pi_L}.$$

This determines $v(x)$ modulo $|G|$, i.e.

$$\frac{\tau(x)}{x} \mod \pi_L = [\tau]^{v(x)}$$

where $[\tau]$ is the image of $\tau \in G$ in $k_E^\times$.

In our case, we know that $\frac{\tau(g_\chi)}{g_\chi} = \chi([\tau])^{-1}$, so

$$\chi([\tau])^{-1} \mod \lambda_i = [\tau]^{\mathrm{val}_{\tilde{\lambda}_i}(g_\chi)}.$$

This matches what our earlier computation: $\chi([\tau]) \mod \lambda_i = [\tau]^q$ of $\chi = (x \mapsto x^q)$.

♠♠♠ TONY: [can we recover what we want by varying ℓ? Ah yes]

So we could have reconstructed the element g_χ as following. Since $\chi([\tau])^{-1}$ is a unit, it's of the form $\frac{\tau(g_\chi)}{g_\chi}$ for some g_χ by Hilbert's Theorem 90.

Stickelberger's Theorem is usually phrased in the following stronger way. Let

$$\theta = \sum_{(\mathbb{Z}/m)^\times} [i^{-1}] \left\{ \frac{i}{m} \right\} \in \mathbb{Q}[(\mathbb{Z}/m\mathbb{Z})^\times]$$

We just showed that $m\theta$ kills the class group. The stronger version is that if $\tau \in \mathbb{Z}[G]$ is such that $\tau\theta \in \mathbb{Z}[G]$ then $\tau\theta$ kills $\mathrm{Cl}(\mathbb{Q}(\zeta_m))$. This refinement basically has to do with formulating things in a more refined way in terms of the group algebra.

Example 8.2.11. For any $c \in (\mathbb{Z}/m)^\times$, the element $\tau = c[1] - [c] \in \mathbb{Z}[(\mathbb{Z}/m)^\times]$ satisfies $\tau\theta \in \mathbb{Z}[G]$, because

$$\begin{aligned} \tau\theta &= c\sum_i [i^{-1}] \left\{ \frac{i}{m} \right\} - \sum_i [ci^{-1}] \left\{ \frac{i}{m} \right\} \\ &= \sum_i [i^{-1}] \left(c \left\{ \frac{i}{m} \right\} - \left\{ \frac{ic}{m} \right\} \right). \end{aligned}$$

Let $\sigma_c \in \mathrm{Gal}(K/\mathbb{Q}) \cong (\mathbb{Z}/m)^\times$ have image c. But also $g_\chi^c / g_\chi^{\sigma_c} \in K$, since for $\gamma \in \mathrm{Aut}(K^*/K)$,

$$\gamma(g_\chi^c) = \chi([\gamma]^{-1})^c g_\chi^c$$

and

$$\gamma(g_\chi^{\sigma_c}) = g_\chi^{\gamma\sigma_c} = \chi([\gamma]^{-1})^{\sigma_c} g_\chi^{\sigma_c} = \chi([\tau]^{-1})^c g_\chi^{\sigma_c}.$$

Factorizing $g_\chi^c / g_\chi^{\sigma_c}$ shows that $\tau\theta$ kills $\mathrm{Cl}_{\mathbb{Q}(\zeta_m)}$. In other words, the role of m in our argument before was used in two places: to multiply θ and to exponentiate g_χ, and that is replaced by something else (namely τ) in the group algebra.

Exercise 8.2.12. Work this out.

You can check that any element taking θ into $\mathbb{Z}[G]$ is a linear combination of these.

We remind you that this is suggested by looking at the equivariant L-function $\mathscr{L}(s) \in \mathbb{C}([\mathbb{Z}/m\mathbb{Z})^\times]$. If something the equivariant analytic class number formula is true, then $\mathscr{L}(0)$ to kill $\mathrm{Cl}(\mathbb{Q}(\zeta_m))$.

8.3. **Herbrand's Theorem.** Now we specialize the previous setting to the case where $m = p$ is a prime. We want to look at the p-part of the class group of $\mathbb{Q}(\zeta_p)$, call it C_p. Why would we do that? This has a special significance which we'll talk about in a moment. Now, C_p has an action of $(\mathbb{Z}/p)^\times$ via Galois, which induces a decomposition

$$C_p = \bigoplus C_{p,i}$$

where $\mathrm{Gal}(K/\mathbb{Q}) \cong (\mathbb{Z}/p)^\times$ acts on $C_{p,i}$ by $x \mapsto x^i$.

Since we will use this repeatedly, we highlight this construction. More generally, if A is an $\mathbb{Z}_p$ module with an action of $(\mathbb{Z}/p)^\times$, then

$$A \cong \bigoplus_i A_i, \quad A_i = \{a \in A \mid \sigma(a) = \omega^i(\sigma)a\}$$

where we view $\sigma \in (\mathbb{Z}/p)^\times$ and $\omega : (\mathbb{Z}/p)^\times \to \mathbb{Z}_p^\times$ is the *Teichmüller character*, with image μ_{p-1}. (It furnishes the unique *multiplicative* inverse to the reduction map $\mathbb{Z}_p^\times \to (\mathbb{Z}/p)^\times$.) It is important to note that this splitting occurs *integrally*.

Theorem 8.3.1 (Herbrand, Ribet)**.** *For i even and $2 \le i \le p-3$*

$$C_{1-i} \neq 0 \text{ if and only if } p \mid \zeta(1-i).$$

The direction $\implies$ is due to Herbrand, and is much easier. The converse was due to Ribet, but also proved by Kolyvagin using Euler systems, and that is the proof we'll follow.

Example 8.3.2. For $p = 37$, $p \mid \zeta(-31)$. This is the first p for which an interesting example occurs.

Digression on the motivation. The C_i arise naturally when one computes Galois cohomology of $\mathbb{Z}_p(r)$. Through this, they arise in algebraic K-theory. (Recall that the "Tate twist" $\mathbb{F}_p(r)$ is defined by $\mathbb{F}_p(r) := \mu_p^{\otimes r}$ as a $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-module.)

Let Γ be the Galois group of the largest extension of $\mathbb{Q}$ unramified outside p. What comes up often is the group cohomology $H^1(\Gamma, \mathbb{F}_p(r))$, in its equivalent guise $H^1_{\text{ét}}(\mathbb{Z}[1/p], \mathbb{F}_p(r))$. The Chern character of algebraic K-theory is valued in this group.

How do you compute the Galois cohomology? You restrict to a group where the action becomes trivial. The action on $\mathbb{F}_p(r)$ becomes trivial when restricted to the subgroup $\Gamma' = \ker(\Gamma \to \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}))$, since this

fixes μ_p. Let Γ' be the Galois group of the maximal extension of $\mathbb{Q}(\zeta_p)$ unramified outside p. Then we have

$$H^1(\Gamma, \mathbb{F}_p(r)) \cong H^1(\Gamma', \mathbb{F}_p(r))^{\Gamma/\Gamma' = (\mathbb{Z}/p)^\times}.$$

(Here we assume that $p-1 \nmid r$. If $p-1 \mid r$, then the module is trivial. ♠♠♠ TONY: [still true though?]) This is $(\mathrm{Hom}(\Gamma', \mathbb{F}_p) \otimes \mathbb{F}_p(r))^{\Gamma/\Gamma'}$ because Γ' acts trivially on $\mathbb{F}_p(r)$. This in turn is $\mathrm{Hom}(C'_r, \mathbb{F}_p)$ where $C' = (\Gamma')^{\mathrm{ab}}$, the (Ray) class group of $\mathbb{Q}(\zeta_p)$ and C'_r is that part on which $(\mathbb{Z}/p)^\times$ acts by $x \mapsto x^r$.

By class field theory, C' maps to C with kernel described by local/global units by class field theory.

Proof of Herbrand's Theorem. Let $2 \le i \le p-3$ be even and k be such that $i+k = p$. Then $k = (p-1)+(1-i) \equiv (1-i) \pmod{p-1}$. We suppose that $C_{p,k} \neq 0$ (the summand of the p-part of the class group where σ acts by $\sigma(a) = \omega^k(\sigma)a$). We're going to plug this into Stickelberger's theorem and see what comes out.

The statement we'll use is that for any $c \in (\mathbb{Z}/p)^\times$, if $\theta = \sum \{\frac{i}{m}\}[i^{-1}]$ then $(c[1]-[c])\theta \in \mathbb{Z}[G]$ and annihilates $\mathrm{Cl}(K)$. Since the group action is described by the Teichmüller character ω, this translates into the assertion that $\omega^k((c-[c])\theta)$ lies in $\mathbb{Z}_p$ and kills $C_{p,k}$, hence must be divisible by p. (We are extending $\omega\colon (\mathbb{Z}/p)^\times \to (\mathbb{Z}/p)^\times$ to $\mathbb{Z}[(\mathbb{Z}/p)^\times] \to \mathbb{Z}_p^\times$.)

It remains to relate this to an L-function. Let $\chi = \omega^k\colon (\mathbb{Z}/p)^\times \to \mathbb{Z}_p^\times$. Then $(c-\chi(c))\sum \left\{\frac{i}{p}\right\} \chi(i)^{-1} \in \mathbb{Z}_p$ (it is a priori in $\mathbb{Q}_p$, but we showed it's actually in $\mathbb{Z}_p$) kills $C_{p,k}$.

Now recall that we showed earlier

$$L(0, \chi^{-1}) = \sum_{i=1}^{p-1} \left(\frac{1}{2} - \left\{\frac{i}{p}\right\}\right) \chi(i)^{-1}.$$

(You might complain that L is valued in $\mathbb{Z}_p$, but recall we developed an algebraic theory for this and the values agreed.) Since χ is nontrivial, the sum over the constant term vanishes and we get

$$L(0, \chi^{-1}) = -\sum_{i=1}^{p-1} \left\{\frac{i}{p}\right\} \chi(i)^{-1}.$$

So $(c-\chi(c))L(0,\chi^{-1})$ kills $C_{p,k}$. Modulo p we have $\chi(c) \equiv c^k \pmod{p}$, so $c-\chi(c)$ is a p-unit because c is a generator. Then $L(0,\chi^{-1}) = L(0,\omega^{-k}) \in \mathbb{Z}_p$ kills C_k, so $p \mid L(0,\omega^{-k})$.

Finally, it remains to relate $L(0,\omega^{-k})$ with $\zeta(1-i)$. We have

$$L(0,\omega^{-k}) = \sum_{(n,p)=1} \omega^{-k}(n).$$

Now, $\omega^{-k}(n) \pmod p \equiv n^{-k} \pmod p \equiv n^{p-1-k} \equiv n^{i-1} \pmod p$. We discussed how termwise congruences give rise to congruences of the regularized sum. So

$$L(0,\omega^{-k}) \equiv \sum_{(n,p)=1} n^{i-1} = \zeta(1-i)(1-p^{i-1}).$$

Remark: the condition on the exponents for deducing this kind of stuff is not satisfied here, so you have to do the dilation tricks. □

9. Converse to Herbrand's Theorem

We have proved that

$$p \mid C_{1-i} \implies p \mid \zeta(1-i)$$

or more precisely $p \mid L(0, \omega^{-k})$, which is a factor of the zeta function. You can think of this as saying that the splitting up of the class group is compatible with the splitting up of the zeta function.

9.1. **Ribet's proof.** We're going to sketch Ribet's proof. We start with the fact that $p \mid \zeta(1-i)$, and then produce element of the class group. Producing elements of the class group is difficult; it is easier to bound from above. In order to bound from below, in Kolyvagin's proof you use the class number formula.

Suppose $p \mid \zeta(1-i)$, where i is even. We can consider the Eisenstein series of weight i:

$$E_i(z) = (?) + \sum_{n=1}^{\infty} \sigma_{i-1}(n) q^n$$

What's the constant term? You remember the constant term by the fact that it should give you 0 when $q = 1$.

Exercise 9.1.1. Why is this the case?

Now, we have

$$\sum \frac{\sigma_{i-1}(n)}{n^s} = \zeta(s)\zeta(s-i+1)$$

so plugging in $s = 0$ shows that the constant term of $E_i(z)$ is $-\zeta(0)\zeta(1-i)$.

Now, $p \mid \zeta(1-i)$ means that the Eisenstein series mod p is a cusp form. So by Riemann Roch, you can lift to a cusp form of weight i such that $f \equiv E_i$, i.e. $a_n \equiv \sigma_{i-1}(n) \pmod{p}$.

Exercise 9.1.2. Figure out this lifting step.

Then f has a Galois representation $\rho_f \colon G_{\mathbb{Q}} := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{GL}_2(\overline{\mathbb{Q}_p})$. The congruence $f \equiv E_i$ implies that the ℓth Hecke eigenvalue is $\equiv \sigma_{i-1}(\ell) = 1 + \ell^{i-1} \pmod{p}$. Thus

$$\operatorname{tr} \rho_f(\mathrm{Frob}_\ell) \equiv \operatorname{tr} \begin{pmatrix} \omega^{i-1} & 0 \\ 0 & 1 \end{pmatrix} (\mathrm{Frob}_\ell) \pmod{p}.$$

So we have determined the semisimplification of the reduction:

$$\overline{\rho_f}^{ss} \pmod{p} = \begin{pmatrix} \omega^{i-1} & 0 \\ 0 & 1 \end{pmatrix}.$$

Exercise 9.1.3. Why can you reduce a Galois representation?

Ribet shows that you can choose the reduction to be of the form $\begin{pmatrix} \omega^{i-1} & * \\ 0 & 1 \end{pmatrix}$ where $*$ is non-zero; i.e. the corresponding representation is an extension of the form

$$0 \to \overline{\mathbb{F}_p} \to ? \to \overline{\mathbb{F}_p}(i-1) \to 0.$$

The restriction of $*$ to $\mathrm{Gal}(\mathbb{Q}(\zeta_p)) := \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}(\zeta_p))$, defines a Galois-equivariant homomorphism $\mathrm{Gal}(\mathbb{Q}(\zeta_p)) \to \overline{\mathbb{F}_p}(i-1)$. We claim that this is *unramified*, which by class field theory implies that it factors through a Galois-equivariant homomorphism $\mathrm{Hom}(C_p, \overline{\mathbb{F}_p}(i-1))$. By the description of the $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ action on C_p, this homomorphism even factors through $\mathrm{Hom}(C_{p,1-i}, \overline{\mathbb{F}_p})$, which shows that $C_{p,1-i}$ is non-zero. ♠♠♠ TONY: [It's supposed to be clear that this extension can't be split.]

It's easy to see this away from p; the key is that this is unramified at p, i.e. descends to $* \in \mathrm{Hom}(C_{p,1-i}, \overline{\mathbb{F}_p})$. Now, E_k being weight k relates to p-adic Hodge theory in some form, which wasn't available to Ribet, so he replaces E_k with a congruent weight 2 Eisenstein series (i.e. run the whole argument using something of weight 2 with the right constant term, instead of E_k). This is a useful trick! So it turns out that there exists a weight 2 cusp form $f' \in \mathscr{S}_2(\Gamma_1(p))$ with the property that

$$E'_k \equiv f' \pmod{p}.$$

This f' has an associated Galois representation $\rho_{f'}$, which occurs in the Tate module of $T_p(J_1(p))$, and even in the Tate module of $T_p(J_1(p)/J_0(p))$. This is the point of making congruences to weight 2 forms - they have associated Galois representations which we understand well. These Jacobians have bad reduction at p only, which already shows you unramifiedness outside p.

The miraculous fact is that $J_1(p)/J_0(p)$ acquires *good* reduction over $\mathbb{Q}(\zeta_p)$ (this is not true of $J_1(p)$, which has semistable reduction). This plus the theory of finite flat group schemes implies that $*$ is unramified at p. This key fact can be guessed from Langlands predictions. ♠♠♠ TONY: [how?] (for f a form on $\mathscr{S}_2(\Gamma_1(p))$ and not on $\mathscr{S}_2(\Gamma_0(p))$, the base-change of f to $\mathbb{Q}(\zeta_p)$ has level 1).

Remark 9.1.4. There are some subtleties to the construction, which are not apparent in this sketch. As is, it is not clear why this doesn't work for $i-1$ as well. This has to do with the choice of lattice used to make the reduction.

9.2. **Cyclotomic units.** Let p be an odd prime and $K = \mathbb{Q}(\zeta_p)^+ = \mathbb{Q}(\zeta_p + \zeta_p^{-1})$.

Definition 9.2.1. The group $\mathscr{C}$ of *cyclotomic units* of K is the subgroup of $\mathcal{O}_K^\times$ generated by

$$\operatorname{Nm}\left(\frac{1-\zeta_p^i}{1-\zeta_p}\right) = \frac{(1-\zeta_p^i)(1-\zeta_p^{-i})}{(1-\zeta_p)(1-\zeta_p^{-1})}.$$

These have really nice properties.

Definition 9.2.2. Define the *regulator* of $\mathscr{C}$ to be the covolume of $\log(\mathscr{C})$ inside $\mathcal{O}_K^\times$:

$$\operatorname{reg}(\mathscr{C}) := \operatorname{vol}(\log \mathcal{O}_K^\times / \log \mathscr{C}).$$

Proposition 9.2.3. *Up to powers of* 2 *and sign,* $\operatorname{reg}(\mathscr{C})$ *is the leading term of* $L(K,s)$ *at* $s=0$*, i.e.*

$$\begin{aligned}\operatorname{reg}(\mathscr{C}) &= (\pm 2^?) \cdot \textit{leading term of } L(K,s) \textit{ at } s=0 \\ &= (\pm 2^?) \prod_{\substack{\chi:(\mathbb{Z}/p)^\times \to \mathbb{C}^\times \\ \chi(-1)=1}} \textit{(leading term of } L(\chi,s) \textit{ at } s=0).\end{aligned}$$

Now, we know from the analytic class number formula that the leading term of $L(K,s)$ is also equal to $\frac{-h_K \operatorname{reg}(\mathcal{O}_K^\times)}{2}$. Therefore, putting these together gives

$$[\mathcal{O}_K^\times : \mathscr{C}] \stackrel{2^?}{=} h_K.$$

♠♠♠ TONY: [exericse: ratio of regulators is index] As an indication of how nontrivial this is, we remark that $\mathcal{O}_K^\times/\mathscr{C}$ need not be isomorphic to the class group, even when the orders are the same.

Now there's a p-adic analogue of Proposition 9.2.3. Basically, you consider instead the p-adic logarithm $\log_p : \mathscr{C} \to (K \otimes \mathbb{Q}_p)$ instead of the usual logarithm. Let $\mathcal{O}_{K,p}$ be the ring of integers of $K \otimes \mathbb{Q}_p$.

Proposition 9.2.4. *We have*

$$[\log_p(\mathcal{O}_{K,p}^\times) : \log_p \mathscr{C}] = \left| \prod_{\chi \neq 1} L_p(1,\chi) \right|_p^{-1}.$$

where L_p *is the* p*-adic* L*-function interpolating* $L(-k,\chi)$ *for* $-k \equiv 1 \pmod{p-1}$.

Euler systems show that the equality $[\mathcal{O}_K^\times : \mathscr{C}] = h_K$ holds equivariantly, i.e. $\#(\mathcal{O}_K^\times/\mathscr{C})_{p,k} = \#(\operatorname{Cl}(K))_{p,k}$ (broken up via characters of $(\mathbb{Z}/p)^\times$). That shows the converse to Herbrand.

Remark 9.2.5. The main conjecture of Iwasawa theory is basically this carried out for ζ_p^n, for n growing to ∞.

Proof. (1) We showed that the leading term of $L(\chi,0)$ is

$$\frac{1}{2}\sum_{i=1}^{p-1}\chi(i)\log(1-\zeta_p^i)(1-\zeta_p^{-i})=\sum_{i=1}^{p-1}\chi(i)\log|1-\zeta_p^i|.$$

So the right hand side is

$$\prod_{\chi\neq 1}\sum_{i=1}^{p-1}\chi(i)\log|1-\zeta^i|.$$

Let $\gamma\in(\mathbb{Z}/p)^\times$ be a generator. We're going to write down a basis of the cyclotomic units which is better adapted to the group action (hence easier for the computation):

$$\left\{\frac{1-\zeta^\gamma}{1-\zeta},\frac{1-\zeta^{\gamma^2}}{1-\zeta^\gamma},\ldots,\frac{1-\zeta^{\gamma^{(p-1)/2}}}{1-\zeta^{\gamma^{(p-3)/2}}}\right\}.$$

Let $u=\mathrm{Nm}_{\mathbb{Q}(\zeta_p)/K}\left(\frac{1-\zeta^\gamma}{1-\zeta}\right)$ so the norms of the basis are

$$\{u,u^\gamma,\ldots u^{\gamma^{(p-3)/2}}\}.$$

These generate because if you want to get a particular numerator, you multiply the right things and the denominators cancel appropriately. They are subject to the relation

$$u\cdot u^g\cdot\ldots\cdot u^{g^{(p-3)/2}}=\prod_{\sigma\in\mathrm{Gal}(K/\mathbb{Q})}u^\sigma=1.$$

(If you multiply upstairs, you get something like $-\zeta^{-1}$, which is killed by the norm.)

This shows that as a module over the group algebra, $\mathscr{C}$ is the integral group algebra generated by these powers of u, modulo the relation that their sum is 0.

So we have the logarithm

$$\log\colon\mathscr{C}\to(K\otimes\mathbb{R})\cong\mathbb{R}^d,\quad d=[K:\mathbb{Q}]$$

sending $u\mapsto(\log|u|_v)_{v\text{ arch.}}$. We choose a basis on the target so that the logarithm takes the form

$$x\mapsto(\log|x|_{gv_0})_{g\in\mathrm{Gal}(K/\mathbb{Q})}=(\log|g^{-1}x|_{v_0})_{g\in\mathrm{Gal}(K/\mathbb{Q})}$$

for some fixed archimedean place v_0. The image is contained in the subspace $\sum x_i=0$, where $x_1,\ldots,x_d$ are the coordinates on $\mathbb{R}^d$. Recall that we're using the volume form on this subspace given by projecting away from one coordinate: $dx_1\wedge\ldots\wedge\widehat{dx_i}\wedge\ldots\wedge dx_n$ (which is independent of i, up to sign). We want to figure out the covolume of the lattice

$$\log\mathscr{C}=(\log(g\cdot u)\colon g\in\mathrm{Gal}(K/\mathbb{Q}))$$

This lattice is spanned by $\{\log u, \log u^\gamma, \ldots \log u^{\gamma^{(p-3)/2}}\}$. These are not linearly independent, so we can't evaluate the volume using the naïve determinant formula. Instead, we invoke the following lemma.

Lemma 9.2.6. *Given $v_1, \ldots, v_d \in \mathbb{R}^d$ lying on the hyperplane $\sum x_i = 0$, which satisfy $\sum v_i = 0$, the volume of the lattice spanned by any $d-1$ of the v_i in $\sum x_i = 0$ is*

$$\frac{1}{d} \det\left(\left(\frac{(1,1,\ldots,1)}{d} + v_i\right)_{1 \le i \le d}\right).$$

Let $e = (\frac{1}{d}, \ldots, \frac{1}{d})$. Then by the Lemma,

$$\begin{aligned}\operatorname{reg}(\mathscr{C}) &= \frac{1}{d} \det(e + \log u, e + \log u^\gamma, \ldots, e + \log u^{\gamma^{(p-3)/2}}) \\ &= \frac{1}{d} \det(e + \log |u^{gh}|_{v_0})_{g,h \in G}.\end{aligned}$$

We can evaluate this by viewing it as a special case of a general calculation on the group algebra. In general, if G is an abelian group and $\sum x_g g \in \mathbb{C}[G]$, then

$$\det(x_{gh})_{g,h \in G} = \operatorname{Norm}_{\mathbb{C}[G]/\mathbb{C}}\left(\sum_{g \in G} x_g g\right).$$

Now, as G is abelian we have a decomposition $\mathbb{C}[G] \cong \prod_{\chi : G \to \mathbb{C}^\times} \mathbb{C}\chi$ as G-modules, hence the above is

$$\boxed{\det(x_{gh})_{g,h \in G} = \prod_{\chi : G \to \mathbb{C}^\times} \left(\sum_G x_g \chi(g)\right).}$$

We applying this with $x_g = \frac{1}{d} + \log |u^g|_{v_0}$. For $\chi = 1$, we get

$$\sum_g x_g \chi(g) = \sum_g \left(\frac{1}{d} + \log |u^g|_{v_0}\right) = 1$$

since $|\operatorname{Nm}_{K/\mathbb{Q}} u^g| = 1$. For $\chi \neq 1$, the sum over $1/d$ cancels out and we obtain

$$\begin{aligned}\sum_{g \in G} x_g \chi(g) &= \sum_{g \in G} \chi(g) \log \left|\operatorname{Nm}_{\mathbb{Q}(\zeta_p)/K}\left(\frac{1 - \zeta^{g \cdot \gamma}}{1 - \zeta^g}\right)\right|_{v_0} \\ &= \sum_{g \in (\mathbb{Z}/p)^\times} \chi(g) \log \left|\frac{1 - \zeta^{g \cdot \gamma}}{1 - \zeta^g}\right|_{\widetilde{v_0}}\end{aligned}$$

for some $\widetilde{v_0}$ a place of $\mathbb{Q}(\zeta_p)$ above v_0, where the last step follows from unwinding definition of the norm of $\mathbb{Q}(\zeta_p)/K$. Then by splitting the logarithm into a difference and re-indexing one of the sums, the above is

$$= (\chi(\gamma)^{-1} - 1) \sum_{g \in (\mathbb{Z}/p)^\times} \chi(g) \log |1 - \zeta^g|_{\widetilde{v_0}}.$$

We've evaluated the term $\sum_{g \in (\mathbb{Z}/p)^\times} \chi(g) \log |1-\zeta^g|_{\widetilde{v_0}}$ before: it is the leading term of $L(0,\chi)$ up to a factor of 2.

So putting this all together, we have

$$\operatorname{reg}(\mathscr{C}) = \frac{1}{d} \prod_{\chi \neq 1} (1 - \chi(\gamma^{-1})) \underbrace{L^*(0,\chi)}_{\text{leading term}}.$$

Now, the product over characters of $(1-\chi(\gamma^{-1}))$ is

$$\prod_{1 \neq x \in \mu_d} (1-x) = \frac{x^d - 1}{x-1}\Big|_{x=1} = d.$$

♠♠♠ TONY: [in lecture, d was $d/2$...?] Therefore, up to powers of 2 we have

$$\prod_{\chi} L^*(0,\chi) = \zeta_K^*(0)$$

□

Now we move on to the p-adic version. Let K_p be the completion of K above p and L_p the completion of $\mathbb{Q}(\zeta_p)$ above p, so L_p/K_p has degree 2.

Proposition 9.2.7. *We have*

$$[\log_p \mathscr{O}_{K,p}^\times : \log_p \mathscr{C}]_p = \operatorname{val}_p \left(\prod_{\chi} L_p(1,\chi) \right).$$

Strictly speaking, we have not interpolated the p-adic L-function $L_p(1,\chi)$. However, the same method as we used to evaluate the p-adic zeta function $L_p(1,\chi)$ applies here.

In fact, we'll prove something more precise, which gives an equality at level of each character χ. Let $\mathscr{O}_{L,p}^\times \widehat{\otimes} \mathbb{Z}_p := \varprojlim_N \mathscr{O}_{L,p}^\times / p^N$. This is a finitely generated $\mathbb{Z}_p$-module.

Example 9.2.8. $\mathbb{Z}_p^\times \cong \mathbb{Z}/(p-1) \times (\mathbb{Z}_p, +)$ via the p-adic logarithm. Tensoring kills the $\mathbb{Z}/(p-1)$ factor, so $\mathbb{Z}_p^\times \widehat{\otimes} \mathbb{Z}_p \cong \mathbb{Z}_p$ via the p-adic logarithm.

Proposition 9.2.9. *With the notation above,*

(1) *There is a decomposition of G-modules*

$$\mathscr{C} \otimes \mathbb{Z}_p \cong \bigoplus_{\chi \neq 1} \mathbb{Z}_p u_\chi$$

where $g \cdot u_\chi = \chi(g) u_\chi$.

(2) *There is a decomposition of G-modules*

$$\mathscr{O}_{L,p}^\times \widehat{\otimes} \mathbb{Z}_p / \text{torsion} \cong \bigoplus_{\chi \neq \omega} \mathbb{Z}_p v_\chi$$

where $g \cdot v_\chi = \chi(g) v_\chi$ and χ ranges over characters $\chi : (\mathbb{Z}/p)^\times \to \mathbb{Z}_p^\times$ distinct from the Teichmüller character ω.

(3) *We have*

$$\log_p u_\chi = L_p(1, \chi)\alpha \cdot \log_p v_\chi$$

for some p-adic unit α.

These statements compare the local units $\mathscr{O}_{L,p}^\times$ and the global units $\mathscr{C}$. The proposition says that Proposition 9.2.7 holds even at the level of characters, interpreted appropriately: after tensoring with $\mathbb{Z}_p$, everything splits and matches up piece-by-piece.

Proof. (1) We can write down explicitly

$$u_\chi = \sum_{g \in G} (g \cdot u) \otimes \chi^{-1}(g) \in \mathscr{C} \otimes \mathbb{Z}_p.$$

Exercise 9.2.10. Check that this is non-zero for all χ non-trivial (because we know the $\mathbb{Q}_p$-module) and spans integrally. Work out where this came from: since $p \nmid |G|$, it should be analogous to representation theory over characteristic 0 fields.

For any $x \in \mathscr{C} \otimes \mathbb{Z}_p$, we can define the projector

$$P_\chi(x) = \frac{1}{|G|} \sum_{g \in G} (g \cdot x) \otimes \chi^{-1}(g).$$

Then $g \cdot P_\chi(x) = \chi(g) P_\chi(x)$, and we claim that $x = \sum_\chi P_\chi(x)$. Indeed,

$$\begin{aligned} \sum_\chi P_\chi(x) &= \frac{1}{|G|} \sum_\chi \sum_g (g \cdot x) \chi^{-1}(g) \\ &= \frac{1}{|G|} \sum_g (g \cdot x) \sum_\chi \chi^{-1}(g) \\ &= x. \end{aligned}$$

(2) Let π be a uniformizer of $\mathscr{O}_{L,p}$, say $\pi = 1-\zeta_p$. We filter $\mathscr{O}_{L,p}^\times$ in the standard way:

$$\ldots \overset{\mathbb{Z}/p}{\subset} 1+\pi^3\mathscr{O}_L \overset{\mathbb{Z}/p}{\subset} 1+\pi^2\mathscr{O}_L \overset{\mathbb{Z}/p}{\subset} 1+\pi\mathscr{O}_L \overset{(\mathbb{Z}/p)^\times}{\subset} \mathscr{O}_{L,p}^\times.$$

Since $\zeta_p \equiv 1 \pmod{\pi}$, we have $\langle\zeta_p\rangle \xrightarrow{\sim} (1+\pi\mathscr{O}_L)/(1+\pi^2)\mathscr{O}_L$, which tells us that

$$(1+\pi\mathscr{O}_L) \cong \mu_p \times (1+\pi^2\mathscr{O}_L).$$

The power series defining the p-adic logarithm just fails to converge on $1+\pi$, but it does on $1+\pi^2\mathscr{O}_L$ and this allows us to define it on $1+\pi$. Note that $\sigma \in (\mathbb{Z}/p)^\times$ acts on π via

$$\sigma(\pi) = \sigma(1-\zeta) = 1-\zeta^{[\sigma]} = 1-(1-\pi)^{[\sigma]} = [\sigma]\pi + (\text{higher order})$$

viewing $[\sigma] \in \{1,\ldots,p\}$. This tells us that $(\mathbb{Z}/p)^\times$ acts on $(1+\pi^j\mathscr{O}_{L,p})/(1+\pi^{j+1}\mathscr{O}_{L,p}) \cong \mathbb{F}_p$ through the character $[\sigma] \mapsto [\sigma]^j$. Recall that the Teichmüller character $\omega\colon (\mathbb{Z}/p)^\times \to \mathbb{Z}_p^\times$ is the unique charactering lifting of the identity map $(\mathbb{Z}/p)^\times \to (\mathbb{Z}/p)^\times$. By successively lifting and averaging, we can find $v_{\omega^j} \in (1+\pi^j\mathscr{O}_L)$ which transforms via ω^j, and is non-zero in $(1+\pi^j\mathscr{O}_{L,p})/(1+\pi^{j+1}\mathscr{O}_{L,p})$.

Exercise 9.2.11. Show this.

We claim that

$$(1+\pi\mathscr{O}_L) \cong \mu_p \times \mathbb{Z}_p v_{\omega^2} \times \mathbb{Z}_p v_{\omega^3} \times \ldots \times \mathbb{Z}_p v_{\omega^p}.$$

The v_{ω^j} are certainly independent, since they transform differently under G. We haven't checked that they span. The point is that our claimed basis generates $(1+\pi^j\mathscr{O}_L)/(1+\pi^{j+1}\mathscr{O}_L)$ for $j \le p$. For $j = p+1$, that step of the filtration transforms again like the reduction of ω^2 so you come back to $v_{\omega^2}^p$.

Exercise 9.2.12. Prove it.

Remark 9.2.13. If you reduce mod p, then see that $\mathbb{Z}_p v_{\omega^p}$ is distinguished in the sense that its pth roots give rise to unramified extensions.

(3) We want to compute $\log_p(v_{\omega^k})$ for $2 \le k \le p-1$. This is $\pi^k z_k$ plus higher order terms, since $v_{\omega^k} = 1+\pi^k z_k$ for some $z_k \in \mathscr{O}_{L,p}$. So the valuation of $\log_p v_{w^k}$ is k if we normalize so that $v(\pi) = 1$, or $\frac{k}{p-1}$ for $v(p) = 1$. We'll use the latter normalization.

If $\chi = \omega^k$, then a generator for $(\mathscr{C} \otimes \mathbb{Z}_p)_\chi$ is

$$u_k := \sum_{g \in G} v^g \otimes \chi(g)^{-1} \in \mathscr{C} \otimes \mathbb{Z}_p.$$

So

$$\begin{aligned}\log_p(u_k) &= \sum \chi(g)^{-1}\log_p \chi(g)^{-1}\log_p(u^g)\\ &= \sum_{g\in(\mathbb{Z}/p)^\times} \chi(g)^{-1}\log_p\left(\frac{1-\zeta^{\sigma\cdot g}}{1-\zeta^g}\right)\\ &= (\chi(\sigma)-1)\sum_{g\in(\mathbb{Z}/p)^\times}\chi(g)^{-1}\log_p(1-\zeta^g).\end{aligned}$$

We want to relate this to $L_p(1,\chi)$, and more specifically we want an equality

$$\log_p u_k = L_p(1,\chi)\log_p v_k\cdot\underbrace{\alpha}_{\in\mathbb{Z}_p^\times}.$$

The valuation of $\log_p v_k$ is $\frac{k}{p-1}$, so we want to show that if $\chi=\omega^k$, then

$$\boxed{\operatorname{val}_p(L_p(1,\chi)) = \frac{-k}{p-1} + \operatorname{val}_p\left(\sum_{g\in(\mathbb{Z}/p)^\times}\chi(g)^{-1}\log_p(1-\zeta^g)\right).}$$

The main thing to understand is where $\frac{-k}{p-1}$ comes from, and it turns out that it comes from a Gauss sum. We want to write $L_p(1,\chi)$ as a sum of logarithms, i.e. in terms of the series $\sum\frac{u^n}{n}$ for various u. To do this, we want to expand χ into characters of $(\mathbb{Z}/p,+)$. By the Fourier transform, as a function on $\mathbb{Z}/p$,

$$\chi = \frac{1}{p}\sum_{\psi:(\mathbb{Z}/p,+)\to\mathbb{C}^\times}\left(\sum_{i\in\mathbb{Z}/p}\chi(i)\psi(i)\right)\psi^{-1}.$$

These ψ are all of the form $i\mapsto\zeta^{ij}$ for ζ a pth root of unity and $j\in\mathbb{Z}/p$, so the inner sums are of the form

$$\sum_{i\in\mathbb{Z}/p}\chi(i)\psi(i) = \sum\chi(i)\zeta^{ij} = \chi(j)^{-1}g_\chi,$$

where $g_\chi=\sum\chi(i)\zeta^i$. So

$$\chi(i) = \frac{g_\chi}{p}\sum_j\chi(j)^{-1}\zeta^{-ij}.$$

Substituting this above, we formally write

$$\begin{aligned} L_p(1,\chi) &= \sum_{\frac{\chi(n)}{n}} \\ &= \frac{g_\chi}{p} \sum_j \chi(j)^{-1} \left(\sum_n \frac{\zeta^{-jn}}{n} \right) \\ &= \frac{g_\chi}{p} \sum_j \chi(j)^{-1} \log_p(1-\zeta^{-j}). \end{aligned}$$

Now we computed $\mathrm{val}_p(g_\chi)$ before, and it was $\frac{p-1-k}{p-1}$ (because we are using a different normalization here). So $\mathrm{val}_p(g_\chi/p) = \frac{-k}{p-1}$, and this proves the equality that we wanted.

Remark 9.2.14. This actually shows that $[(\mathscr{O}_{K,p}^\times \otimes \mathbb{Z}_p)_\chi : (\mathscr{C} \otimes \mathbb{Z}_p)_\chi]$ is a p-unit times $L_p(1,\chi)$.

The gap is that we haven't justified the power series expansion for the p-adic logarithm. Go back to that computation for the residue of the p-adic zeta function, and repeat the same dilation that we used in analyzing

$$(1-2^{1-s})\zeta(s) = \frac{1}{1^s} - \frac{1}{2^s} + \frac{1}{3^s} - \frac{1}{4^s} + \dots$$

□

9.3. **Converse to Herbrand.** We know that (up to powers of 2)

$$\#\mathscr{O}_K^\times/\mathscr{C} \stackrel{2?}{=} \#\mathrm{Cl}_K.$$

The input from Euler systems will be the following refinement:

$$\#\left(\mathscr{O}_K^\times/\mathscr{C}\right)_{p,\chi} \geq \#(\mathrm{Cl}_K)_{p,\chi}.$$

We indicate how to use this to show the converse to Herbrand:

$$p \mid \zeta(1-k) \implies p \mid C_{\omega^{1-k}}$$

for $2 \leq k \leq p-3$ even.

Firstly, note that because we have an equality after taking a product over χ, we have that

$$\#(\mathscr{O}_K^\times/\mathscr{C})_{p,\chi} = \#(\mathrm{Cl}_K)_{p,\chi}.$$

Now, let i be such that $i+k=p$ (so i is odd and $3 \leq i \leq p-2$). By one of the termwise congruences we discussed,

$$p \mid \zeta(1-k) \iff p \mid L(0,\omega^{-i}).$$

Indeed, since $k-1 = p-1-i$ we have

$$\zeta(1-k)\text{“}=\text{”}\sum n^{k-1} \equiv \sum n^{-i} \pmod{p}.$$

and that is (formally) $L_p(0,\omega^{-i})$ and also $L_p(1,\omega^{1-i})$. So we find that $p \mid L_p(1,\omega^{1-i})$, hence by Proposition 9.2.7 we have $p \mid [\mathscr{O}_{K,p}^\times \otimes \mathbb{Z}_p : \mathscr{C} \otimes \mathbb{Z}_p]_{\omega^{1-i}}$. Now, we have nested inclusions

$$(\mathscr{C} \otimes \mathbb{Z}_p)_{\omega^{1-i}} \subset_A (\mathscr{O}_K^\times \otimes \mathbb{Z}_p)_{\omega^{1-i}} \subset_B (\mathscr{O}_{K,p}^\times \otimes \mathbb{Z}_p)_{\omega^{1-i}}$$

so either inclusion A or inclusion B is proper. We analyze these separately.

If inclusion A is proper, then $p \mid \#(\mathrm{Cl}_K)_{p,\omega^{1-i}}$. This is good, but not what we wanted (the field and character are different). By Kummer theory, it implies that $(C_{\mathbb{Q}(\zeta_p)})_{\omega^{1-k}}$ is non-trivial, by something analogous to the Weil pairing over number fields, by which one eigenspace must be dual to another.

If inclusion B is proper, then

$$(\mathscr{O}_K^\times/p)_{1-i} \to (\mathscr{O}_{K,p}^\times/p)_{1-i}$$

is not surjective. Therefore, it is not injective either, because the right hand side is rank 1 over $\mathbb{Z}_p$, and the left hand side is of rank at least 1 (possibly bigger). Then there exists $u \in \mathscr{O}_K^\times$ such that u is locally (at p) a pth power and the class of u in $(\mathscr{O}_K^\times/p)$ transforms under ω^{1-i}.

Then $\mathbb{Q}(\zeta_p)(u^{1/p})$ is everywhere unramified (split at p because it's locally a pth power). It defines a homomorphism $C_{\mathbb{Q}(\zeta_p)} \to \mathbb{Z}/p$ which exhibits $(C_{\mathbb{Q}(\zeta_p)})_{1-k} \neq 0$. (Translating the Galois action on the extension to that on the class group is an exercise in Kummer theory.)

The only loose end is the “pairing” of class groups.

Example 9.3.1. The Scholz reflection principle says that the 3-rank of $\mathrm{Cl}_{\mathbb{Q}(\sqrt{d})}$ and the 3-rank of $\mathrm{Cl}_{\mathbb{Q}(\sqrt{-3d})}$ differ by at most 1.

9.4. **Euler Systems.**