Arithmetic incarnations of zeta in Iwasawa theory

Peter Xu

September 14, 2016

Abstract

We give a brief exposition of the Iwasawa theory of cyclotomic extensions, so as to discuss its relationship with p-adic zeta functions. We give an overview of these connections and the arithmetic significance of the theory, leading up to a statement of the main conjecture.

Contents

1 Overview

We are interested in the structures of ideal class groups of number fields. This is an enormous question to tackle in general, but an easier but still very interesting case is that of cyclotomic extensions of a given base field. This study was initiated by Kenkichi Iwasawa in the 1950s, and has led to connections with a vast world of arithmetic and analytic objects, which we will explore the beginnings of here.

The treatment we give of the core of Iwasawa theory will be somewhat technical but as brief as possible, attempting to give proofs of all the core statements of the classical theory in full generality. It owes much to the accounts given in [5], [19], and [17], which each give partial treatments. We will eschew any cohomological tools in the entirety of this essay.[1]

1.1 First definitions

Let F be a number field, and fix a prime p. Our basic object of study will be a tower of field extensions

$$F = F_0 \subset F_1 \subset \ldots \subset F_n \subset \ldots$$

such that $\mathrm{Gal}(F_n/F) \cong \mathbb{Z}/p^n\mathbb{Z}$. Denote $F_\infty = \bigcup F_i$; we will call F_∞/F a **$\mathbb{Z}_p$-extension**. Note that by higher ramification theory it is immediate that F_∞/F is unramified except at the primes above p.

In particular, the basic object of study we will be interested in is the **cyclotomic $\mathbb{Z}_p$-extension**, which is the unique $\mathbb{Z}_p$-extension inside the extension obtained by adjoining μ_{p^∞} - for example, when $F = \mathbb{Q}$ and p is odd, this is the extension given by setting F_n as the unique field of index $p-1$ inside $F(\mu_{p^n})$; here, p ramifies totally. Most of the following will hold for arbitrary $\mathbb{Z}_p$-extensions; we will note when we need the cyclotomic hypothesis.

We are interested in understanding the ideal class groups of the fields in this tower, but even this is too much to ask for. Instead, we will set $A_n := \mathrm{Cl}(F_n)_p$ the p-Sylow subgroup. In other terms, if M_n is the maximal unramified abelian p-extension of F_n, then global class field theory tells us that $A_n \cong \mathrm{Gal}(M_n/F_n)$. We set $M_\infty := \bigcup M_i$, which is the maximal unramified abelian p-extension of F_∞. We then can define A_∞ as $\mathrm{Gal}(M_\infty/F_\infty) \cong \mathrm{Cl}(F_\infty)$; we have $A_\infty \cong \varprojlim A_n$ under the norm maps.

Understanding the A_n was our original motivation, but there is a related set of modules which will also be important, especially in the formulation of the relation to zeta functions later. Let N_n be the maximal abelian p-extension of F_n unramified away from p, and take $N_\infty = \bigcup N_n$, the maximal abelian p-extension

[1] We phrase a little Kummer theory in a cohomological manner since it is convenient, but of course these ideas predate cohomology.

of F_∞ unramified outside the primes above p. Define $X_\infty = \mathrm{Gal}(N_\infty/F_\infty)$. Then let $X_n \cong \mathrm{Gal}(N_n/F_\infty)$; again, $\varprojlim X_n \cong X_\infty$.

Iwasawa's insight, developed in his series of papers [7] [8] [9] [10] between 1956 and 1959, was that it is possible to understand the structure of "infinity-objects" like A_∞ and X_∞, from which information about the otherwise intractable finite layers can be retrieved as quotients.

We will, as is standard, use the letter Γ for $\mathrm{Gal}(F_\infty/F) \cong \mathbb{Z}_p$. Then X_∞, under its interpretation as $\mathrm{Gal}(M_\infty/F_\infty)$, has a continuous Γ-action via inner automorphisms: consider it as a subgroup of $\mathrm{Gal}(M_\infty/F)$; the action passes to the quotient $\mathrm{Gal}(M_\infty/F)/X_\infty \cong \Gamma$ because X_∞ is abelian. Further, it is a pro-p-group, and hence is naturally $\mathbb{Z}_p$-linear.

Define the **Iwasawa algebra**, then, to be

$$\Lambda = \mathbb{Z}_p[[\Gamma]] := \varprojlim \mathbb{Z}_p[\Gamma/H]$$

where H runs over the finite-index subgroups of Γ; the previous discussion shows that X_∞ has a continuous Λ-action.

1.2 Structure theory

The entire theory rests essentially on the observation that, amazingly, we can identify Λ as a very familiar object:

Theorem 1.1. *There exists a topological isomorphism of rings $\varphi : \Lambda \xrightarrow{\sim} \mathbb{Z}_p[[T]]$ sending a fixed topological generator γ of Γ to $1 + T$.*[2]

Proof. We have a topological generator γ of Λ as a $\mathbb{Z}_p$-algebra, so there is at most one map sending $\gamma \mapsto 1 + T$. We will show such a map does exist. First, one can extend to γ^t for any $t \in \mathbb{Z}_p$, because $(1+T)^t$ will be a well-defined power series with coefficients in $\mathbb{Z}_p$ by p-adic continuity of binomial coefficients.[3]

The Iwasawa algebra is, additively, the formal (potentially infinite) $\mathbb{Z}_p$-combinations of $\mathbb{Z}_p$-powers of γ, where only finitely many terms in the sum are outside any p-adic neighborhood of zero. We can extend the map to such sums as well, since two $\mathbb{Z}$-linear combinations of powers of γ which are congruent modulo a high power of p will map to power series whose coefficients are congruent modulo the same high power of p. Hence we have a well-defined map φ.

[2] Why is this so amazing? It is the reason why we had to restrict to the Sylow-p: if the coefficient ring were $\mathbb{Z}$ or its completion at some $l \neq p$, the analogous attempted definition for the Iwasawa algebra would degenerate into an infinite product which is not even noetherian.

[3] Explicitly, two integers t_1, t_2 are congruent modulo p^k, the respective binomial coefficients $\binom{t_1}{n}$ and $\binom{t_2}{n}$ will be as well. This is more-or-less equivalent also to the existence of the p-adic exponential and logarithm maps, whose convergence properties as formal series are a formal exercise to verify.

To construct the inverse, note that $1+T$ is also a topological generator, as truncated polynomials (taking T-adic approximations) are polynomials having Taylor expansions at -1 whose coefficients are p-integral, since they are evaluations of derivatives. We can send construct an inverse by sending $1+T$ to γ, and again we need to show that everything can be extended. Indeed, we can again extend to $\mathbb{Z}_p$-powers, since we can do so for γ, and we can take arbitrary (possibly infinite) linear combinations of such powers so long as only a finite number of the coefficients are outside any given p-adic neighborhood of zero. □

Our new description of the ring Λ gives us a very nice structure to work with. Call a monic polynomial $f(T) \in \mathbb{Z}_p[[T]]$ **distinguished** if all its non-leading coefficients are divisible by p. We then have the following description of the prime ideals of the power series ring:

Theorem 1.2. *$\mathbb{Z}_p[[T]]$ has as prime ideals (p), $(f(T))$ for $f(T)$ irreducible distinguished, and the unique maximal ideal (p,T).*

Proof. This is a consequence of a pseudo-division algorithm on $\mathbb{Z}_p[[T]]$ which can be stated as follows:

Lemma 1.3 (Division algorithm)**.** *For any power series $f, g \in \mathbb{Z}_p[[T]]$ such that the coefficients of f up to degree $n-1$ are divisible by p but not the degree-n coefficient, there is a unique expression*

$$g(T) = q(T)f(T) + r(T)$$

with r a polynomial of degree $\leq n-1$.

Proof. The following is an argument due to Manin. Let the head H and the tail T be linear operators defined by

$$H\left(\sum_{i=0}^{\infty} a_i T^i\right) = \sum_{i=0}^{n-1} a_i T^i$$

and

$$T\left(\sum_{i=0}^{\infty} a_i T^i\right) = \sum_{i=n}^{\infty} a_i T^{i-n}.$$

We wish to find q so that $T(g) = T(qf)$, which a few manipulations turn into

$$T(g) = \left(1 + T \circ \frac{H(f)}{T(f)}\right)(qT(f))$$

since $T(f)$ is evidently invertible. The operator on the RHS is unipotent mod p since $p|H(f)$, hence invertible mod p, hence invertible; thus we obtain a q.

Finally, such an expression is unique because if $qf + r = 0$, we may assume that q and r are not both divisible by p. r must be divisible by p, since modulo p, it has the only low-degree terms. But then $p|fq$ so $p|q$; contradiction. □

Given this lemma, by tne analogue of the classical Euclidean algorithm argument we obtain that $\mathbb{Z}_p[[T]]$ is a unique factorization domain with irreducibles p and $f(T)$ for irreducible distinguished f. Then for general reasons, (p) and such $(f(T))$ are precisely the height-1 primes, and (p, T) can be the only height-2 prime. □

Iwasawa discovered an important structure theorem for modules over Λ, analogous to the structure theorem for finitely generated modules over a PID. We say that two Λ-modules M, N are **pseudo-isomorphic**, and write $M \sim N$, if there exists a Λ-module morphism $M \to N$ with finite kernel and cokernel.

Theorem 1.4 (Iwasawa structure theorem)**.** *Every finitely generated Λ-module is pseudo-isomorphic to a direct sum of the form*

$$\Lambda^r \oplus \bigoplus \Lambda/(p^{r_i}) \oplus \bigoplus \Lambda/(f_j(T)^{s_j})$$

where the $f_j(T)$ are irreducible distinguished polynomials, and these canonical forms are unique up to order of the factors.

Proof. First, note $\mathbb{Z}_p[[T]]$ is a 2-dimensional regular local ring. All of its localizations at height-1 prime ideals are then discrete valuation rings.

Let M be a finitely generated Λ-module, and let T be its torsion submodule. We claim that there exists a map $M \to T$ which has finite cokernel and such that the composite $T \hookrightarrow M \to T$ has finite kernel. Indeed, consider $\hom(M, T)$; in particular, its localization at an arbitrary height-1 prime ideal $\mathfrak{p}$ is $\hom(M_\mathfrak{p}, T_\mathfrak{p})$; furthermore, $T_\mathfrak{p} = 0$ for almost all $\mathfrak{p}$ almost by definition. If T is nonzero and not finite (otherwise the assertion is trivial), we can find some $\mathfrak{p}$ so that $\hom(M, T)_\mathfrak{p}$ is nontrivial, as by structure theory over the DVR $\Lambda_\mathfrak{p}$, M is a finite direct sum of free terms and $T_\mathfrak{p}$, which is in turn is a finite direct sum of terms of the form $\Lambda_\mathfrak{p}/\mathfrak{p}^n$. Hence there is always an element with cokernel zero, and kernel zero after composition: take a projection onto $T_\mathfrak{p}$ or any multiple of it by an element which goes to a unit in $\Lambda/\mathfrak{p}$.

We claim there is an element in $\hom(M, T)$ whose image in each height-1 localization has cokernel zero, and kernel zero after composition. Indeed, we only have to worry about finitely many height-1 primes, and further, the property of mapping to a (co)kernel zero map in $\hom(M, T)_\mathfrak{p}$ can be subsumed under a congruence condition modulo some sufficiently high power of $\mathfrak{p}$, since adding an element of $\mathfrak{p}^k \hom(M, T)$ for $k \gg 0$ will not change the projection map.

That at least one such congruence class will have this property for each $\mathfrak{p}$ is guaranteed because if we represent the localization $\hom(M, T)_\mathfrak{p}$ in the standard way as pairs (a, s) for $a \in \hom(M, T)$ and s in the corresponding multiplicatively closed subset, we can note that the (co)kernels of (a, s_1) and (a, s_2) visibly will never differ. Since the image of $\hom(M, T)$ in the localization consists of elements of the form $(a, 1)$ (as we can take the multiplicatively closed subset to be saturated), this guarantees us the existence of at least one element in $\hom(M, T)$ with (co)kernel zero in the localization. Then by the Chinese remainder theorem over our finitely many congruence conditions corresponding to height-1 $\mathfrak{p}$ with nontrivial localizations, we obtain an element of $\hom(M, T)$ which has trivial (co)kernel in every height-1 localization. This element then must have finite corresponding (co)kernel, since both are supported on codimension 2, which is also the dimension of the ring.

We then can construct the map $M \to T \times M/T$, which is a pseudo-isomorphism. Hence we can separate into the torsion-free and the torsion cases. That in the torsion case we always have a pseudo-isomorphism $T \to \bigoplus \Lambda/\mathfrak{p}_i{}^{r_i}$ follows from the exact same localization argument we gave above - in fact, for two T, T' with the same corresponding "canonical" form of this type, it also gives pseudo-isomorphisms in both directions between T and T', establishing that for torsion modules, pseudo-isomorphism is an equivalence relation. Combined with the last portion of this proof, this shows that the representation as above is unique.

For M torsion-free, we take the natural map to the double dual $M \to \hom(\hom(M, \Lambda), \Lambda)$; this map must be an isomorphism localized at each height-1 prime, hence by the same considerations as above is a pseudo-isomorphism. Replacing M by its image, we obtain a pseudo-isomorphic module which is reflexive - that is, isomorphic to its double dual. It is a general result of commutative algebra that such modules on 2-dimensional regular local rings are free; see lemma 6 of [19] for details. □

Returning to the arithmetic content of the theory now that our analytic framework is in place, the key point is that we can extract knowledge of the finite class groups X_n from X_∞.

1.3 Quantitative results

We can now reap the fruits of all the structural knowledge we have obtained.

Theorem 1.5. *Let $\omega_n \in \Lambda$ correspond to the power series $(1+T)^{p^n} - 1$. Then $X_n \cong X_\infty/\omega_n X_\infty$.*

Proof. The group $X_n \cong \mathrm{Gal}(N_n/F_\infty)$ is the quotient of $\mathrm{Gal}(M_\infty/F_\infty)$ by $\mathrm{Gal}(M_\infty/N_n)$, which is the closure of its commutator subgroup, by general considerations. $\mathrm{Gal}(F_\infty/F_n)$ is topologically generated by γ^{p^n}, so if we choose a lift h in $\mathrm{Gal}(M_\infty/F_\infty)$, the latter group is topologically generated by h and X_∞. For any $x \in X_\infty$, we have that $[h, x] = \gamma^{p^n} x - x = \omega_n x$, so indeed $X_n \cong X_\infty/\omega_n X_\infty$. □

We can deal with A_n along the same lines:

Lemma 1.6. *We also have $A_\infty/\omega_n A_\infty \cong A_n$, and the Sylow-p of the class group of F_∞ is a finitely generated torsion Λ-module.*

Proof. Let H_∞ be the p-Hilbert class field of F_∞, and likewise H_n for F_n.

We have a map

$$A_\infty \cong \mathrm{Gal}(H_\infty/F_\infty) \to \mathrm{Gal}(H_n F_\infty/F_\infty) \cong \mathrm{Gal}(H_n/F_n) \cong A_n.$$

The kernel is $\mathrm{Gal}(H_\infty/H_n F_\infty)$, which is precisely the commutator of $\mathrm{Gal}(H_\infty/F_n)$. But this group is topologically generated by $\mathrm{Gal}(H_\infty/F_\infty)$ and $\gamma^{p^n} \in \mathrm{Gal}(F_\infty/F)$, so just as before we have $A_n \cong A_\infty/\omega_n A_\infty$. The LHS is a finite group for all n, so the result follows by the structure theory. □

We will come back to this to calculate the order of A_n at the end of the section. For now, however, we can use this to further understand the structure of X_n. In the next theorem, we let r_1, r_2 as usual be the number of real and complex places of F, respectively. Notice that the number of real and complex places of F_n is always $p^n r_1, p^n r_2$, since real places extend solely to real places - true by parity for p odd, and true by inspection for $p = 2$.

Theorem 1.7. *The $\mathbb{Z}_p$-free part of X_n has rank $r_2 p^n + \delta_n$ for a bounded series of constants δ_n.*

Proof. Write $U_{\mathfrak{p},n}$ for the local units of F_n at $\mathfrak{p}$, and $U_{\mathfrak{p},1,n}$ for those which are 1 mod $\mathfrak{p}$. Let E_n be the global units, and $E_{1,n}$ those which are 1 mod $\mathfrak{p}$ for each $\mathfrak{p}$ above p. Let $\varphi : E_n \to \prod_{\mathfrak{p}|p} U_{\mathfrak{p},n}$ be the natural diagonal map. Then Artin reciprocity tells us that $\mathrm{Gal}(N_n/M_n) \cong \prod_{\mathfrak{p}|p} U_{\mathfrak{p},1,n}/\overline{\varphi(E_{1,n})}$. What is the $\mathbb{Z}_p$ rank of this thing? Via the p-adic logarithm map, each term on the top is of rank equal to the degree of the extension of local fields induced by F_n/F, so the total $\mathbb{Z}_p$-rank of the top is $[F_n : \mathbb{Q}] = p^n[F : \mathbb{Q}] = p^n(r_1 + 2r_2)$.

If we take $\delta_n = \mathrm{rk}_{\mathbb{Z}}(E_n) - \mathrm{rk}_{\mathbb{Z}_p}(\overline{\varphi(E_{1,n})})$, then, using Dirichlet's unit theorem, we see that the rank of $\mathrm{Gal}(N_n/M_n)$ is $p^n r_2 + 1 + \delta_n$. There is an exact sequence

$$0 \to X_n \to \mathrm{Gal}(N_n/F_n) \to \mathrm{Gal}(F_\infty/F_n) \to 0$$

where the third term is $\mathbb{Z}_p$, so X_n has rank $p^n r_2 + \delta_n$.

It remains to show that δ_n is bounded. Indeed, certainly $E_{n,1}$ is finite-index in E_n and hence has the same rank; thus, we may pick δ_n free generators in $E_{n,1}$ which map under φ to p^m-powers in $\prod_{\mathfrak{p}|p} U_{\mathfrak{p},n}$ for arbitrarily high m. Adjoining the mth roots of these generators to $F(\mu_{p^\infty})$ then gives an unramified extension by Kummer theory, whose Galois group $B_{m,n}$ is a product of δ_n cyclic p-groups whose order is exponential in our choice of m, i.e. $|B_{m,n}| \sim p^{m\delta_n}$.

To conclude, we only need observe that in fact $B_{m,n}$ is a quotient of $\mathrm{Gal}(H(F(\mu_{p^{\max\{m,n\}}}))/F(\mu_{p^{\max\{m,n\}}}))$, where H denotes the p-Hilbert class field construction. This is because we only need the base field to contain units of F_n and the p^mth roots of unity for the Kummer theory to work. The result of theorem 1.6 applies just as well to the "completed" cyclotomic $\mathbb{Z}_p$-extension over the base $F(\mu_q) \subset \ldots \subset F(\mu_{p^\infty})$, for $q = p$ when p odd and $q = 4$ when $p = 2$. Hence $\mathrm{Gal}(H(F(\mu_{p^{\max\{m,n\}}}))/F(\mu_{p^{\max\{m,n\}}}))$ is one of the ω_k-quotients of the torsion module $\mathrm{Cl}(F(\mu_{p^\infty}))_p$, where k differs from $\max\{m,n\}$ by a constant (since F may already contain some of the roots of unity, causing the "completed" extension to start at an offset).

Suppose the canonical form of $\mathrm{Cl}(F(\mu_{p^\infty}))_p$ as a Λ-module has r components corresponding to quotients by height-1 primes. Now fix some m and let n grow. Then for large n, by some calculations by an application of the division algorithm, summands corresponding to $\Lambda/(p^r, \omega_n)$ have cardinality p^{rp^n} by an application of the division algorithm, and summands corresponding to $\Lambda/(f(T), \omega_n)$, $\deg f = s$, have cardinality p^{sn}. Thus the total cardinality is asymptotically either $\sim p^{rp^n}$ or $\sim p^{sn}$, for some fixed constants r, s. Then in particular, it is impossible for it to be $\sim p^{m\delta_n}$ if δ_n is unbounded. Unboundedness excludes the latter case with linear exponent, but δ_n cannot possibly grow exponentially since the $\mathbb{Z}$-rank of E_n does not even grow exponentially, so it is impossible for it to be the former case either. Hence the result. □

Addendum 1.8. Leopoldt conjectured in [15] that $\delta_n = 0$ always. The more well-known form of the conjecture is in terms of the p-adic regulator R_p, defined in analogy to the usual regulator: choose a $\mathbb{Z}$-basis $\{e_i\}$ for the rank-$(r_1 + r_2 - 1)$ module E_1. There are $r_1 + r_2$ ways of embedding E_1 into $\mathbb{C}_p$; then R_p is the absolute value of the determinant of the matrix whose (i, j)th entry is the image of $\mathrm{Log}_p(e_i)$ under the jth embedding with a column removed, where we have written down the Iwasawa logarithm on $\mathbb{C}_p$. That this is independent of the choices we have made is proven in much the same way as in the archimedean case. Equivalence of the two formulations of the conjecture is formal. Also as in the global case, the p-adic regulator is closely tied to p-adic zeta functions, and a third, analytic, formulation can be stated (see later).

The conjecture is known when the F_n are abelian over $\mathbb{Q}$, due to Brumer in [1], and we will actually prove a specific case later in this section.

Theorem 1.9. *X_∞ is a finitely generated Λ-module, and its rank is r_2.*

Proof. Finite generation is a local property, and there is only one closed point to check; namely, (p, T). But the reduction of X_∞ at (p, T) is X_0/pX_0 from above, which is finite.

The free rank follows immediately from the structure theorem and the ranks of the X_n from above. □

Let us discuss the case of the cyclotomic $\mathbb{Z}_p$-extension over $F = \mathbb{Q}$. Notice we have the exact sequence

$$0 \to \mathrm{Gal}(N_n/M_nF_\infty) \to X_n \to A_n \to 0,$$

Moreover, we know that $\mathrm{Gal}(N_n/M_nF_\infty) \oplus \mathbb{Z}_p \cong \mathrm{Gal}(N_n/M_n) \cong \prod_{\mathfrak{p}|p} U_{\mathfrak{p},1,n}/\overline{\varphi(E_{1,n})}$, coming from the split exact sequence

$$0 \to \mathrm{Gal}(N_n/M_nF_\infty) \to \mathrm{Gal}(N_n/M_n) \to \mathrm{Gal}(M_nF_\infty/M_n) \to 0.$$

With this setup, we obtain:

Corollary 1.10. *If $F = \mathbb{Q}$ and we take the cyclotomic $\mathbb{Z}_p$-extension, $A_n \cong X_n$. In particular, Leopoldt's conjecture holds for $\mathbb{Q}$.*

Proof. From the above, we need to show that M_nF_∞ is already the maximal abelian p-extension of F_n unramified outside of p. Stated differently, we need for F_∞ to be the maximally wildly ramified extension of F_n, as it is for $\mathbb{Q}$. Indeed, suppose otherwise; then there must be an abelian extension K/F_∞ of degree p such that $K/\mathbb{Q}$ is nonabelian. But there are no nonabelian group extensions $0 \to \mathbb{Z}/p \to (?) \to \mathbb{Z}_p \to 0$, so this is a contradiction.

From the previous discussion, we find also that $\prod_{\mathfrak{p}|p} U_{\mathfrak{p},1,n}/\overline{\varphi(E_{1,n})} \cong \mathbb{Z}_p$, which also establishes that $\delta_n = 0$ as promised, and hence that $R_p \neq 0$ for the fields in the cyclotomic $\mathbb{Z}_p$-extension of $\mathbb{Q}$. □

Our original goal, to understand the A_n, is also within reach. To conclude this section, we thus exhibit Iwasawa's theorem on the orders of the groups A_n.

Corollary 1.11. *There exist constants μ, λ, ν such that the power of p dividing the class group of F_n is $\mu p^n + \lambda n + \nu$ for $n \gg 0$.*

Proof. From theorem 1.6 A_∞ is a finitely-generated torsion Λ-module, and $A_\infty / \omega_n A_\infty \cong A_n$. This computation is essentially the one we already did in theorem 1.7; μ corresponds to the sum of the exponents of the summands $\Lambda/(p^{r_i})$ in the canonical form of X, λ to the sum of the degrees of the polynomials in the summands $\Lambda/(f_j(T))$. □

Iwasawa conjectured that $\mu = 0$ and reduced this problem to a congruence condition between certain Bernoulli numbers, though discovered it was false for some non-cyclotomic $\mathbb{Z}_p$-extensions. Ferrero-Washington proved the conjecture for cyclotomic $\mathbb{Z}_p$-extensions of abelian number fields in [4], giving us the remarkable result that the growth of $v_p(|\mathrm{Cl}(F_n)|)$ is linear. This is much stronger than any bound which can be obtained by analytic methods.

2 Characteristic power series

In this section, we will look at a similar situation as before, but will mostly deal with the "completed" cyclotomic $\mathbb{Z}_p$-extension over $F(\mu_q)$ (recall $q = p$ or 4) with maximal extension $F(\mu_\infty)$. The aim is to obtain an understanding of the Galois-module structure of the objects involved. We follow [5] closely.

2.1 Setup

We write $G_\infty = \mathrm{Gal}(F(\mu_\infty)/F)$, $\Delta = G(F(\mu_q)/F)$. $\Gamma = \mathrm{Gal}(F(\mu_\infty)/F(\mu_q))$, as consistent with before; G_∞ is a product of the latter two groups. Let χ be the p-adic cyclotomic character on G_∞, and let θ and κ be its restrictions to Δ, Γ.

Define N_∞ to be the maximal abelian p-extension of $F(\mu_\infty)$ as before, and $X_\infty = \mathrm{Gal}(N_\infty/F_\infty)$, which is a G_∞-representation by inner automorphisms. Define A_n, A_∞ analogously as in the previous section.

Denote by Y the torsion part of X_∞; considering it as a representation of Δ, we may decompose it into the subspaces on which Δ acts by θ^i, since Δ is an abelian group of order prime to p, so a semisimplicity argument can be constructed for the $\mathbb{Z}_p$-group algebra. The powers of the cyclotomic character θ^i are the characters of Δ; denote by e_i the orthogonal projector associated to that character in the group algebra. Then by the structure theory of the previous section, we can write

$$e_i Y \sim \bigoplus_{j=1}^{r_i} \Lambda/(f_{ji}).$$

Let $f_i = \prod_j f_{ji}$; we call this the **characteristic power series** of Y_i, well-defined up to a unit. This will be the analytic object coming from the arithmetic side of the theory, which will have the relation to zeta.

We can use Kummer theory to relate this to class groups. Gluing together the long exact sequences associated to the short exact sequences

$$0 \to \mu_{p^n} \to N_\infty^\times \to N_\infty^\times \to 0$$

yields the isomorphism of Galois modules

$$\hom(X_\infty, \mu_{p^\infty}) \cong V_\infty := \varinjlim (F(\mu_{p^\infty})^\times \cap (M_\infty^\times)^{p^n})/(F(\mu_{p^\infty})^\times)^{p^n}.$$

We can also write V_∞ as the kernel of the map $(\mathbb{Q}_p/\mathbb{Z}_p) \otimes F(\mu_{p^\infty})^\times \to (\mathbb{Q}_p/\mathbb{Z}_p) \otimes I^{(p)}(F(\mu_{p^\infty}))$, where the last group is the ideal group of $F(\mu_{p^\infty})$ away from primes dividing p.

Further, the long exact sequences associated to the Kummer sequences

$$0 \to \mu_{p^n} \to \mathcal{O}_{F_\infty}^\times \to \mathcal{O}_{F_\infty}^\times \to 0$$

afford us a short exact sequence

$$0 \to (\mathbb{Q}_p/\mathbb{Z}_p) \otimes \mathcal{O}_{F(\mu_{p^\infty})}^\times \to V_\infty \to A_\infty \to 0.$$

2.2 Arithmetic to analytic

We now restrict to the case of F **totally real**, and $p \neq 2$. In this case, $e_i \mathcal{O}_{F(\mu_{p^\infty})}^\times = 0$ for i odd: the action of Δ on the units of $F(\mu_{p^n} + \overline{\mu_{p^n}})$ factors through a quotient by the subgroup of index 2 corresponding to "complexifying", whereas θ^i does not factor through any such quotient for parity reasons. We use the following lemma of Hasse:

Lemma 2.1. *The units of a totally imaginary extension of a totally real field either are generated by the units of the totally real field and the roots of unity, or contain an index-2 subgroup generated by them.*

Proof. Analyze the action of $\sigma - 1$ on the units of the extension, where σ is complex conjugation. For details, this is Satz 14 in [6]. □

Corollary 2.2. *$e_i V_\infty \cong e_i A_\infty$ as G_∞-modules for i odd.*

Proof. The units of $F(\mu_{p^n})$ and $F(\mu_{p^n} + \overline{\mu_{p^n}})$ thus coincide after tensoring with $\mathbb{Q}_p/\mathbb{Z}_p$ since $p \neq 2$, so indeed $e_i((\mathbb{Q}_p/\mathbb{Z}_p) \otimes \mathcal{O}_{F(\mu_{p^\infty})}^\times) = 0$. This then follows from the earlier exact sequence. □

Putting it all together, we obtain:

Corollary 2.3. *$e_i X_\infty(-1) \cong \hom(e_{1-i} A_\infty, \mathbb{Q}_p/\mathbb{Z}_p)$ for all even integers i.*

Proof. The only part which is not immediate is that $e_i X_\infty = e_i Y$; i.e. that the free part of X_∞ contributes nothing. But the modules $\hom(e_{1-i}A_\infty, \mathbb{Q}_p/\mathbb{Z}_p)$ are Λ-torsion, so this is true. □

A slight variation of the power series $f_i \in \Lambda$ fits more naturally into the analytic world we will enter. First, let κ be the cyclotomic character on $\mathrm{Gal}(F(\mu_\infty)/F(\mu_p))$, and set $u = \kappa(\gamma)$, so that u is the unit which gives the action of γ on μ_{p^∞}. Then a short formal manipulation gives us

$$\hom(e_i A_\infty, \mathbb{Q}_p/\mathbb{Z}_p) \sim \bigoplus_{j=1}^{r_{1-i}} \Lambda/(f_{(1-i)i}(u(1+T)-1)).$$

Thus, we set $g_i(T) = f_{1-i}(u(1+T)-1)$ for i odd. The power series which will be compared to zeta functions are given by $G_i(T) = g_i((1+T)^{-1}-1) = f_{1-i}(u(1+T)^{-1}-1)$, which are those corresponding to the parts of the actual class group $e_i A_\infty$: to see this, note that taking the pro-p dual changes the action of γ to γ^{-1}, which corresponds to the power series substitution $1+T \mapsto (1+T)^{-1}$, i.e. $T \mapsto (1+T)^{-1}-1$.

For now, we note the very suggestive fact that we have the following analogue of the analytic class number formula:

Theorem 2.4. *Let F be totally real. $G_1(u^s-1)/(u^s-u)$ has a pole at $s=1$ if and only if $R_p \neq 0$. In this case, it is a simple pole whose residue differs by a p-adic unit from*

$$\frac{2^{d-1}h_F R_p}{\sqrt{\Delta_{F/\mathbb{Q}}}} \prod_{\mathfrak{p}|p} (1-(\mathbb{N}\mathfrak{p})^{-1})$$

where $d = [F:\mathbb{Q}]$, $h_F = |Cl(F)|$, $\Delta_{F/\mathbb{Q}}$ the discriminant of the extension $F/\mathbb{Q}$.[4]

Proof. The key step is to compute the index of the embedding of the 1-global units in the product of the 1-local units, and use the local-global discriminant relation. The proof itself is long and technical, involving chaining together many algebraic intermediary objects; it is given in its entirety in the appendix to [5]. □

This gives yet another way to state Leopoldt's conjecture, for totally real fields.

2.3 Analogy and motivation

Before moving on to the next section, we take a moment to mention the appearance here of an important analogy, which in some ways clarifies everything we have done. This is the number field/function field analogy which underlies vast swathes of number theory; its influence on arithmetic thought in the mid-20th century is captured in André Weil's famous 1940 letter from prison to his sister Simone, which can be found in [22].[5]

[4] Apologies for the overloaded notation, but the association of Δ to the discriminant is extremely strong.

[5] On a related note, André Weil perhaps has the distinction of being the greatest mathematician of all time who was less interesting than their sibling.

Indeed, our topic is the number field incarnation of work by Weil himself on the function field side. If K is the function field of a algebraic curve over $\mathbb{F}_q$,[6] the analogous construction to our cyclotomic towers is to consider is the base changes to $\mathbb{F}_{q^p}, \mathbb{F}_{q^{p^2}}, \ldots$; i.e., extensions of the field of coefficients. In this setting, every tower of extensions of the field of coefficients actually is very well-understood, thanks to Weil's proof of the conjectures which bear his name,[7] specifying the behavior of the local zeta function $\zeta_K(s)$ associated to the curve.

In particular, this zeta function can be proven to satisfy

$$\zeta_k(s) = \frac{P(q^{-s})}{(1 - q^{-s}(1 - q^{1-s}))}$$

where P is a polynomial of degree $2g$ for g the genus of the curve, and is the characteristic polynomial of the geometric Frobenius acting on the Jacobian of the curve.

In our context, the Jacobian (which, recall, classifies degree zero line bundles, so is analogous to the class group[8]) is replaced by $(X^\infty)^-$, the part of X where complex conjugation acts by -1.[9] The action of the Frobenius is replaced by the action of the generator γ of the Galois group.

The analogue of Iwasawa's conjecture that $\mu = 0$ is true in the geometric case, following from computations with the polynomial P; in fact, this was the motivation for the the conjecture in the first place. Further work has been done in this direction, building on the fact that the coefficient λ is analogous to the genus of the curve; see [11].

Most significantly, however, the characteristic power series also come from zeta/L-functions on the number field side; this is the topic to which almost the entirety of the rest of this essay is devoted. Apocryphally, Weil predicted the existence of such a result at the time he was formulating his conjectures, over a decade before Iwasawa's work.

3 p-adic L-functions

In this section, we will define the p-adic L-functions, going through several different constructions so as to provide a better understanding of these objects. The essential idea is to p-adically interpolate the rational special values of an L-function (more precisely, of different L-functions "twisted" by a power of a cyclotomic character corresponding to the argument) at the nonpositive integers. The result will turn out to be not

[6] For simplicity and fidelity to the classical treatment, we take a smooth projective model of the function field.

[7] These Weil "conjectures" are actually theorems for general varieties, the last having been proven by Deligne in 1974.

[8] In fact, it is the class group of the ring of integral elements of the function field, as is not too difficult to prove: classes of line bundles over this affine geometric object are in correspondence with degree zero line bundles over the projective closure since the effect of the extra points is to allow one to move around the degree. With reference to the analogy with the Jacobian, in the case where $\mu = 0$ the class group A_∞ even is quasi-isomorphic to the Tate module-like $(\mathbb{Q}_p/\mathbb{Z}_p)^\lambda$.

[9] This is another way of restricting ourselves to the odd characters, as we did above; the entire theory can be easily reformulated in terms of the submodules on which complex conjugation is -1, in place of what we did.

just continuous, but a p-adic meromorphic function (and analytic except in certain cases); the existence of this object helps understand the behavior of these special values, which are themselves related to class numbers, along with being an interesting arithmetic object in its own right which provides the "zeta" part of the statement of the main conjecture.

3.1 Motivation

We begin by showing that what we are doing is feasible and interesting. First, we define the values we wish to interpolate. Pick a totally real extension $F/\mathbb{Q}$ and a continuous multiplicative character χ of its absolute Galois group; we can identify it with a character of some ray class group. We can hence set

$$L_S(\chi, s) = L(\chi, s) \prod_{\mathfrak{p}|p} (1 - \chi(\mathfrak{p})(\mathbb{N}\mathfrak{p})^{-s}).$$

where S here is the set of primes $\mathfrak{p}$ dividing p; i.e. we are simply removing the Euler factors above p. The values $L_S(\chi, n)$ at integers $n \leq 0$ are those we wish to interpolate.

Theorem 3.1. *Each $L_S(\chi, n)$ for $n \leq 0$ is an algebraic number contained in $\mathbb{Q}(\chi)$.*

Proof. Certainly the removed Euler factors satisfy this, so we need to show that $L(\chi, n)$ does as well. Notice that for the Riemann zeta function, this follows from Euler's classical calculation for positive even integers paired with the functional equation, which gives us the explicit formula $\zeta(\chi, n) = B_{1-n}/(1-n)$ where B_k is the kth Bernoulli number, for negative odd integers n. The even values vanish from the functional equation; this is the analytic/zeta incarnation of the necessity of restricting to $(X^\infty)^-$.

Euler's method points to the technique in general, though complex analysis is required for rigor: the identity

$$\sum_{\mathfrak{a}} \chi(\mathfrak{a}) e^{-(\mathbb{N}\mathfrak{a})x} = \sum_{n=0}^{\infty} \frac{L(\chi, -n)(-x)^n}{n!},$$

where $\mathfrak{a}$ runs over ideals of F, can be deduced by calculating the inverse Mellin transform of $L(\chi, s)\Gamma(s)$ in two ways: one by moving the vertical line of integration to $-\infty$ and adding up the residues, the other the Dirichlet series/Fourier series dictionary.

By comparing power series coefficients it is evident that the L-values are contained in $\mathbb{Q}(\chi)$. In the case of base field $\mathbb{Q}$, we can make a few power series manipulations to obtain the formula

$$L(\chi, n) = -\frac{\mathfrak{f}(\chi)^{-n}}{1-n} \sum_{i=1}^{\mathfrak{f}(\chi)} \chi(i) B_{1-n}\left(\frac{i}{\mathfrak{f}(\chi)}\right)$$

where B_k is the kth Bernoulli polynomial; often, this is written using the notation

$$B_{n,\chi} = \mathfrak{f}(\chi)^{n-1} \sum_{i=1}^{\mathfrak{f}(\chi)} \chi(i) B_n\left(\frac{i}{\mathfrak{f}(\chi)}\right)$$

and are referred to as "generalized Bernoulli numbers." Then we can express this last formula as $L(\chi, n) = -B_{1-n,\chi}/(1-n)$, in parallel with the result of Euler's classical calculation.[10] □

Why do we localize away from p? The point is to make the values we desire to interpolate p-adically continuous, so that our task is actually feasible.

To illustrate, in the simple case where we are just taking the Riemann zeta function, Kummer discovered the congruences

$$(1-p^{m-1})\frac{B_m}{m} \equiv (1-p^{n-1})\frac{B_n}{n} \pmod{p^k}$$

whenever $m \equiv n \pmod{\varphi(p^k)}$. Kummer's work which led him to this result also sheds light on the interest of these values. The analytic class number formula allows one to deduce the following fact, which shows the arithmetic importance of our zeta functions:

Theorem 3.2 (Kummer's regularity criterion). *p divides the class number of $\mathbb{Q}(\zeta_p)$ if and only if p does not divide the numerators of B_{2k} for $1 \le k \le (p-3)/2$.*[11]

Proof. Write h_p for said class number; we can factor it as $h_p^+ h_p^-$, the parts of the class group which are acted on by complex conjugation by 1 and -1 respectively; alternately, h_p^+ is the class number of the maximal totally real subextension.

Kummer reduced this to the case of the negative class number; i.e., showed that $p|h_p^+$ only if $p|h_p^-$.[12] We omit the argument here, as it involves a detailed examination of cyclotomic units; it can be found in Kummer's original paper [14].

Having reduced to this case, we may apply the analytic class number formula to the full extension $\mathbb{Q}(\zeta_p)$ and its maximal real subextension $\mathbb{Q}(\zeta_p)^+ = \mathbb{Q}(\zeta_p + \overline{\zeta_p})$; the quotient of the two equations yields

$$\prod_{\chi(-1)=-1} L(\chi, 1) = \pi^{(p-1)/2} \cdot \frac{R_{\mathbb{Q}(\zeta_p)} \cdot \sqrt{|\Delta_{\mathbb{Q}(\zeta_p)^+}|} \cdot 2}{R_{\mathbb{Q}(\zeta_p)^+} \cdot \sqrt{|\Delta_{\mathbb{Q}(\zeta_p)}|} \cdot 2p} \cdot h_p^-.$$

The functional equation tells us that

$$\prod_{\chi(-1)=-1} L(\chi, 1) = \prod_{\chi(-1)=-1} \frac{i\pi\tau(\chi)}{\mathfrak{f}(\chi)} B_{1,\chi}.$$

[10] Notice that for odd characters, it is the even values which are the zeroes, and vice versa; naturally, twisting by an odd character flips the parity that contains the interesting information. This is why we insist on totally real base extensions, since otherwise the presence of both odd and even twisted L-functions destroys all zeta information at the special values we are considering. The resulting values, and hence the p-adic L function, would all be zero.

[11] A **regular** prime p is precisely one for which p does not divide the class number of $\mathbb{Q}(\zeta_p)$. Kummer's motivation here was that one can give a very direct argument for Fermat's last theorem for regular primes; unfortunately, there are infinitely many irregular primes.

[12] Vandiver conjectured that, in fact, p never divides $p|h_p^+$; this is still open.

where $\tau(\chi)$ is the Gauss sum associated to χ. Notice that $\tau(\chi)$ and $\tau(\overline{\chi})$ pair to give $\sqrt{\mathfrak{f}(\chi)}$, so this is actually

$$\pi^{(p-1)/2} \prod_{\chi(-1)=-1} \sqrt{\mathfrak{f}(\chi)^{-1}} = \pi^{(p-1)/2}(-1)^{(p-1)/4}(-1)^{(p-1)/2} \frac{\sqrt{|\Delta_{\mathbb{Q}(\zeta_p)}|}}{\sqrt{|\Delta_{\mathbb{Q}(\zeta_p)^+}|}}$$

the second equality by the conductor-discriminant formula and checking the signs.

An examination of Hasse's argument on unit groups mentioned earlier shows that the ratio of regulators $R_{\mathbb{Q}(\zeta_p)}/R_{\mathbb{Q}(\zeta_p)^+}$ is equal to $2^{\mathrm{rk}(U)} = 2^{(p-1)/3}$. Putting this all together, we obtain the formula

$$h_p^- = 2p \prod_{\chi(-1)=-1} \left(-\frac{1}{2}B_{1,\chi}\right).$$

By a manipulation of the Mellin transform computation we did earlier, one can show that the generalized Bernoulli numbers we are concerned with satisfy the power series relation

$$\sum_{k=0}^{\mathfrak{f}(\chi)} \frac{\chi(k)e^{kt}}{e^{\mathfrak{f}(\chi)t}-1} = \sum_{n=0}^{\infty} B_{n,\chi}\frac{t^n}{n!}$$

When χ is nonprincipal, by reading off coefficients one obtains that

$$B_{1,\chi} = \frac{1}{\mathfrak{f}(\chi)} \sum_{r=1}^{\mathfrak{f}(\chi)} \chi(r)r.$$

Pick a generator ω of the character group of $\mathrm{Gal}(\mathbb{Q}(\mu_p)/\mathbb{Q})$; if we consider the roots of unity as sitting inside $\mathbb{C}_p$ rather than $\mathbb{C}$, we can identify it with the Teichmüller character. We find hence that

$$B_{1,\omega^{p-2}} = \frac{1}{p}\sum_{r=1}^{p-1} r\omega^{-1}(r) \equiv \frac{p-1}{p} \pmod{\mathbb{Z}_p}.$$

We can thus write $2p(-B_{1,\omega^{p-2}}/2) \equiv 1 \pmod{p}$ and strike off those terms. For the other terms corresponding to $\omega, \omega^3, \ldots, \omega^{p-4}$, a generalized version of the Kummer congruences shows that

$$B_{1,\omega^k} \cong \frac{B_{k+1}}{n+1} \pmod{p}$$

from which the result follows. This, as mentioned, is possible to prove directly with congruence manipulations, using only the original Kummer congruences, as Kummer himself did. However, we do not cover this, because it is also a manifestation of a fact that we will see later in the analytic properties of p-adic L-functions: namely, that their power series expansions at $s = 1$ have all coefficients of nonconstant terms divisible by p (except in special cases).[13] □

[13] We see here hints of a much stronger arithmetic/analytic connection that we do not yet have the tools to deal with: notice the similarity in our expression of h_p^- to the breakup of zeta into L-functions and the decomposition of the class group into Galois eigenspaces from section 2. Indeed, a refinement of our argument shows that in fact the decomposition of the class group into these irreducible representations does correspond to the decomposition into L-functions, or equivalently to the values of zeta at negative integers if we local at the behavior modulo p. This is one direction of Herbrand's theorem, which follows from the main conjecture.

Analogues of the Kummer congruences also hold for the "generalized Bernoulli numbers," and this analogously allows the construction of p-adic L-functions in general over $\mathbb{Q}$.[14] It is hence theoretically possible to first prove our desired congruences (as Kummer did for the zeta case), and hence construct the desired L-functions; this was historically the first construction of the p-adic zeta function for $\mathbb{Q}$ by Kubota and Leopoldt. We will have to end up doing work equivalent to this by the conservation of difficulty in any case, though often in different guises.

How do these congruences relate to the construction of the L-functions, more explicitly? The point is that a function defined at a dense subset of $\mathbb{Z}_p$ can be p-adically interpolated if and only if it is p-adically continuous, by taking limits. The condition of being p-adically continuous is precisely that of the function $L_p(\chi, s)$ satisfying $L_p(\chi, m) \cong L_p(\chi, n)$ modulo a high power of p whenever m and n are congruent modulo a high power of p, i.e. exactly Kummer-like congruences. The $(p-1)$ factor in $\varphi(p^n)$ is necessary as well, because rather than interpolating the values $L_S(\chi, n)$, the values to be interpolated are the twisted L-values $L_S(\chi\omega^{n-1}, n)$, and ω is of order $p-1$; in this case, the dense subset of nonpositive integers which are 1 $(\text{mod } p-1)$ give us the values which the Kummer congruences allow us to interpolate. The essence of the construction, then, is that this leads to the in-between values to be twisted L-values (which can also be expressed using congruences), and that the result is p-adic analytic (except possibly at 1).

Maybe this is unsatisfying. Why should we *expect* that p-localization gives rise to Kummer-type congruences, allowing us to p-adically interpolate? As mentioned, Kummer stumbled across them when proving his regularity criterion, but the calculations do not provide much conceptual satisfaction. One way to think about it is that in removing the terms of a Dirichlet series corresponding to a prime, we are forcing the conductor of the associated character to be divisible by p. This has to be the case when we look at the arithmetic side of things, since the characters of the totally ramified extension $F(\mu_p)/F$ all satisfy this. Indeed, this can be viewed as a sort of inverted version of the l-adic philosophy; the total ramification of p in the tower of extensions we are considering means that we in some sense are throwing away all information about behavior at that prime, and thus we can use characteristic p objects (and inverse limits thereof) to fruitfully study the resulting arithmetic.

3.2 Analytic constructions

We first outline a very explicit and direct construction using the generalized Bernoulli formula derived above, though it has the disadvantage of only working over $\mathbb{Q}$. Let us write the formula down in the slightly modified form

$$L(\chi, n) = \frac{1}{\mathfrak{f}(\chi)(n-1)} \sum_{i=0}^{\mathfrak{f}(\chi)} \frac{\chi(i)}{i^{n-1}} \sum_{j=0}^{\infty} \binom{1-n}{j} \left(\frac{\mathfrak{f}(\chi)}{i}\right)^j B_j.$$

[14] In general, formulas are not known for arbitrary real base extensions. Siegel gave formulas for the real quadratic case in [20].

We have used the classic interpolative trick of replacing the finite sum with an infinite one which collapses to the finite case when $1-n$ is a positive integer. This almost looks like it can be directly interpolated by replacing n with an arbitrary p-adic variable s, but it cannot since the values are not even p-adically continuous; indeed, by inspection, the infinite sum cannot converge.

We need to multiply by the localization factor $1-\chi(p)p^{-n}$, but it turns out one other manipulation is necessary: instead of using the expression

$$B_{n,\chi} = \mathfrak{f}(\chi)^{n-1}\sum_{i=1}^{\mathfrak{f}(\chi)} \chi(i) B_n\left(\frac{i}{\mathfrak{f}(\chi)}\right)$$

for $B_{n,\chi}$, it turns out from a short power series computation that it is equivalent to replace the role of $\mathfrak{f}(\chi)$ in this sum with any integral multiple thereof; we thus write instead

$$B_{n,\chi} = F\sum_{i=1}^{F} \chi(i) B_n\left(\frac{i}{F}\right)$$

where F is the least common multiple of $\mathfrak{f}(\chi)$ and q, which is either p or 4 (for typical persnickety 2-reasons).[15] Then our total expression becomes

$$(1-\chi(p)p^{-n})L(\chi,n) = (1-\chi(p)p^{-n})\frac{1}{F(n-1)}\sum_{i=0}^{F}\frac{\chi(i)}{i^{n-1}}\sum_{j=0}^{\infty}\binom{1-n}{j}\left(\frac{F}{i}\right)^{j} B_j.$$

This gives us the presence of increasing powers of F, divisible by q, which is promising for convergence purposes. Likewise, the localization factor allows us to remove p from denominators. Knowing we have to twist, we can turn this by a change of characters into

$$(1-\chi(p)\omega^{n-1}(p)p^{-n})L(\chi\omega^{n-1},n) = \frac{1}{F(n-1)}\sum_{(i,p)=1}^{F}\chi(i)\langle i\rangle^{1-s}\sum_{j=0}^{\infty}\binom{1-s}{j}\left(\frac{F}{i}\right)^{j} B_j.$$

where ω is the Teichmüller character considered as a Dirichlet character, and $\langle - \rangle$ is defined by $\langle a\rangle\omega(a) = a$.

Theorem 3.3.

$$L_p(\chi,s) := \frac{1}{F(n-1)}\sum_{(i,p)=1}^{F}\chi(i)\langle i\rangle^{1-s}\sum_{j=0}^{\infty}\binom{1-s}{j}\left(\frac{F}{i}\right)^{j} B_j.$$

converges on the open disk of radius $qp^{1/(1-p)}$ in $\mathbb{C}_p$, and on that disk defines a p-adic analytic function when χ is nonprincipal, and a p-adic meromorphic function with a simple pole at $s=1$ with residue $(p-1)/p$ otherwise.

Proof. We have given all the main ideas for the necessary manipulations; the rest is a few formal results in p-adic analysis which we do not cover here. For details, consult chapter 6 of [17]. □

As promised, analytic properties of these functions allow recovery of congruences.

Corollary 3.4. *If χ is nonprincipal and pq does not divide $\mathfrak{f}(\chi)$, then the p-adic power series expansion of $L_p(\chi,s)$ at 1 has all nonconstant coefficients divisible by p, and integral constant coefficient.*

[15] Actually, F can be any multiple of $(\mathfrak{f}(\chi),p)$; it makes no difference to the final definition of the L-function.

Proof. Again, see chapter 6 of [17]; this is painstaking analysis of the explicit formula given above. □

The Kummer congruences follow formally from this, as do the congruences

$$B_{1,\omega^n} \equiv \frac{B_{n+1}}{n+1} \pmod{p}$$

which we needed earlier, for odd n such that $p-1$ does not divide $n+1$.

The second analytic construction, using p-adic integration, is due to Mazur [16], and is interesting for its conceptual clarity and the presence of a p-adic analogue of the Mellin transform. We will consider for now only the case over $\mathbb{Q}$ because it is easier to state with the tools that we have.

The main idea is that one can define a suitable p-adic measure on $G_\infty \cong \mathbb{Z}_p^\times$, i.e. a $\mathbb{Q}_p$-functional on the space of locally constant functions whose value on indicator functions of compact-open subspaces is p-adically bounded, so that Riemann sum integrals can be defined. The values $-(1-p^{k-1})B_k/k$ can be expressed as an integral against this measure, which gives the p-adic zeta function as a "Mellin transform" when n is replaced with a general p-adic variable.

Essentially, we are able to simply define the measure to do what we want in terms of the Bernoulli numbers thanks to Euler's formula. The work of proving the Kummer congruences (since of course somehow this work needs to be done) is hidden inside the analytic manipulations needed to prove that we have indeed defined a measure.

Indeed, we first define a family of p-adic distributions (that is, functionals on the space of locally constant functions) on the whole of $\mathbb{Z}_p$ by specifying that $\mu_k(a + p^n\mathbb{Z}) = p^{n(k-1)}B_k(a/p^n)$. This certainly is not p-adically bounded, but can be "regularized" by setting $\mu_{k,\alpha}(U) = u_K(U) - \alpha^{-k}\mu_k(\alpha U)$ for U compact-open, and α some arbitrary p-adic unit.

Theorem 3.5. *$\mu_{k,\alpha}$ defines a measure on $\mathbb{Z}_p$. Furthermore,*

$$\int_U d\mu_{k,\alpha} = k\int_U x^{k-1} d\mu_{1,\alpha}$$

for any compact-open U.

Proof. Manipulations of Bernoulli generating functions. See chapter 7 of [17]. The latter statement is essentially the Kummer congruences in disguise, and can be written in the form

$$d_k\mu_{k,\alpha}(a + p^N\mathbb{Z}_p) \cong d_k k a^{k-1}\mu_{1,\alpha}(a + p^N\mathbb{Z}_p) \pmod{p^N}$$

where d_k is a term which clears all the denominators of the coefficients of the kth Bernoulli polynomial. □

$\mu_{k,\alpha}$ does what we want it to: $\mu_{k,\alpha}(\mathbb{Z}_p) = (1-\alpha^{-k})B_k$, so we find that $\mu_{k,\alpha}(\mathbb{Z}_p^\times) = (1-\alpha^{-k})(1-p^{k-1})B_k$. Hence we have that

$$-(1-p^{k-1})B_k/k = \frac{1}{\alpha^{-k}-1}\int_{\mathbb{Z}_p^\times} d\mu_{k,\alpha} = \frac{1}{\alpha^{-k}-1}\int_{\mathbb{Z}_p^\times} x^{k-1} d\mu_{1,\alpha}.$$

Here is our p-adic Mellin transform. However, we can not simply replace $1-k$ with s; we have no guarantee the result is p-adically continuous, since we do not have such continuity results on varying the integrand as parameterized by a p-adic variable; in fact, it is not continuous. After all, recall that we do not ever construct a p-adic zeta function which interpolates all the untwisted special values.

The idea becomes clearer if we view it as imitating Tate's idea from his famous thesis of the argument of the zeta function coming from the character one is integrating over (up to some simple factor); in this case, instead of the complex characters of the ideles, the relevant character group is $\hom_c(\mathbb{Z}_p^\times, \mathbb{Q}_p^\times)$,[16] which is canonically isomorphic to $\mathbb{Z}/(p-1) \times \mathbb{Z}_p$, by specifying where the root of unity $\omega(1)$ goes, and where the topological generator $\exp(1)$ goes. The p-adic integral is a continuous operator on this group, so in this way, it gives a natural domain for the argument. Indeed, it is customary to define

$$\zeta_{p,s_0}(s) = \frac{1}{\alpha^{-s_0+(p-1)s} - 1} \int_{\mathbb{Z}_p^\times} x^{s_0+(p-1)s-1} d\mu_{1,\alpha}$$

with $(s_0, s) \in \mathbb{Z}/(p-1) \times \mathbb{Z}_p$, where s_0 is identified with its representative in $\{0, \dots, p-2\}$. For each fixed s_0, this function is referred to as a "branch" of the p-adic zeta function. This provides a conceptual reason why we must interpolate zeta/L-values which are "twisted" based on residue modulo $p-1$, instead of being able to interpolate all the values of a single complex L-function. This perspective is developed in [12].

This choice of representatives is kind of artificial in the context of the weight space, but is the convention since it aligns more naturally with the interpolation. To make the relationship to the weight space more comprehensible, the integrand should look like $\omega(x)^{s_0} \cdot \langle x \rangle^{s_1}$, in which case s_0 can genuinely be thought of as a class modulo $(p-1)$. Indeed, if we fix $s_0 = 0$ here and take $s = -s_1$ as the argument, this latter formulation recovers for $L_p(\chi_0, s)$ the formula

$$\frac{1}{\langle \alpha \rangle^{s-1} - 1} \int_{\mathbb{Z}_p^\times} \langle x \rangle^{-s} d\mu_{1,\alpha}.$$

The procedure for more general L-functions, over more general fields, is not very different, but requires one to explicitly prove strong Kummer-type congruences, which will be established in the next section. We will briefly mention this generalization at that point.

3.3 Algebraic theory

From this point forward, we will insist p is an odd prime, because the algebraic theory becomes rather finicky for $p = 2$. The construction of p-adic L-functions does in fact go through for $p = 2$ with only minor changes, as we saw via the analytic constructions, but the main conjecture needs significant reworking which we will not undertake, so there is no reason for us to accomodate for it.

[16] Iwasawa calls this the weight space. Notice that just as characters of the ideles are the same as characters of the absolute abelian Galois group, elements of the weight space are the characters of the Galois group of the total composite of our tower of extensions.

As before, F is a totally real field, and we will often refer to the $\mathbb{Z}_p$-extension given by the cyclotomic extension with base field $F(\mu_p)$. G_∞, Δ, and Γ are defined with respect to this cyclotomic extension as in section 2.

What are, actually, the p-adic L-functions? Our two analytic constructions defined them as meromorphic (analytic when χ is nonprincipal) functions from a disk in $\mathbb{C}_p$ to $\mathbb{C}_p$, but eventually they will be compared to the "characteristic power series" $G_i(T)$, which live in Λ.

Let us think of $\Lambda \cong \mathbb{Z}_p[[T]]$ as an algebra of functions on $\mathbb{Z}_p$, where the evaluation-at-s homomorphism is given by substituting s for T; note that these homomorphisms are only well-defined for $|s|_p < 1$ for convergence reasons.

The p-adic L-functions do not live in Λ, first because the zeta function has a pole, and second because as mentioned previously, the L-functions can take on non-$\mathbb{Z}_p$ values in $\mathbb{Q}_p(\chi)$ for some character χ. Hence, we will often want to consider the extended Iwasawa algebras $\Lambda_\chi \cong \Lambda \otimes_{\mathbb{Z}} \mathbb{Q}(\chi)$, as well as their fields of fractions $\mathrm{Q}(\Lambda_\chi)$.

Let ω be the cyclotomic character for Δ.[17] Axiomatically, $L_p(\chi, s)$ can be specified then as the unique element of $\mathrm{Q}(\Lambda_\chi)$ such that:

(a) Meromorphicity: if the degree $|\Delta|$ divides i, $(\gamma - u)L_p(\chi, s) \in \Lambda_\chi$. Else, $L_p(\chi, s) \in \Lambda_\chi$.

(b) Interpolation: under the evaluation homomorphism induced by κ^n, $\kappa^n_* : \Lambda_\chi \to \mathbb{C}_p$, $(\gamma - u)L_p(\chi, s)$ maps to $\kappa^n_*(\gamma - u)L_S(\chi\omega^{n-1}, n)$ (or, as the case may be, $L_p(\chi, s)$ maps to $L_S(\chi\omega^{n-1}, n)$ for all integers $n \leq 0$).[18]

As the names suggest, these correspond precisely to the two specifying conditions of meromorphicity and interpolating the special values of the localized complex L-function; the previous section's constructions show that these are enough to specify $L_p(\chi, s)$ uniquely.

Note that s is not the formal variable T. Our construction will produce functions $G(T, \omega^i) \in Q(\Lambda_\omega) \cong Q(\mathcal{O}_{\mathbb{Q}_p(\omega)}[[T]])$, and we will end up setting $L_p(\chi, s) = G(u^s - 1, \omega\chi^{-1})$. It is then visible that the factor $\gamma - u$ corresponds to removing a simple pole at $s = 1$, and that κ^n_* corresponds to evaluation at $s = n$, since sending $\gamma \mapsto u^n$ amounts to sending $T \mapsto u^n - 1$.[19]

Let $\mathfrak{m}$ be a modulus of F. It will be convenient to formulate things in terms of the partial zeta functions

[17] We earlier referred to this in the algebraic theory as θ, but now use ω to emphasize the connection with the Teichmüller character.

[18] Recall that κ is the cyclotomic character for Γ, and u the image of γ under the induced homomorphism. This condition can be equivalently stated as $L_p(\chi, s)$ maps to $L_S(\chi, n)$ under κ^n_* for $n \equiv 1 \pmod{[F(\mu_p) : F]}$, i.e., the behavior at the values where the twisting is trivial is sufficient to specify the function, as we mentioned briefly earlier. This follows from a continuity argument and the Chinese remainder theorem.

[19] It was apparent some sort of substitution of this sort would be necessary, because the evaluation homomorphisms only exist for $|T|_p < 1$, whereas the p-adic L-functions are defined on the disk $|s|_p < p^{(p-2)/(p-1)}$. Indeed, u^s is defined and is $\equiv 1 \pmod{p}$ precisely when s is in that disk by the properties of the p-adic exponential map.

associated to ray ideal classes C modulo $\mathfrak{m}$, which are defined by the series

$$\zeta_{\mathfrak{m}}(C,s) = \sum_{\alpha \in C} (\mathbb{N}\alpha)^{-s}$$

for $\mathrm{Re}(s) > 1$, the sum over integral ideals. Notice that these encode the same information as L-functions: Artin L-functions associated to characters of conductor $\mathfrak{m}$ are linear combinations of partial zeta functions with roots-of-unity coefficients, so by Fourier inversion we can write the partial zeta functions as linear combinations of the L-functions with similar such coefficients; this immediately gives us analytic continuation. Explicitly,

$$\zeta_m(C,s) = \frac{1}{|C_{\mathfrak{m}}|} \sum_{\chi \in \widehat{C_{\mathfrak{m}}}} \overline{\chi}(C) L(\chi, s).$$

Further, if $\sigma \in \mathrm{Gal}(M/F)$ for some abelian extension M, we will write

$$\zeta_F(\sigma,s) = \sum_C \zeta_{\mathfrak{m}}(C,s)$$

where $\mathfrak{m}$ is the conductor of the extension, and the sum is over ray ideal classes modulo $\mathfrak{m}$ mapping to σ under the Artin map. Notice in particular that the S-localized L-function can be equally written as

$$L_S(\chi, s) = \sum_C \chi(C) \zeta_{(p,\mathfrak{f}(\chi))}(C, -s).$$

Theorem 3.6. *Each $\zeta_M(\sigma, n)$ for $n \leq 0$ is a rational number.*

Proof. Recall the formula

$$\sum_{\mathfrak{a}} \chi(\mathfrak{a}) e^{-(\mathbb{N}\mathfrak{a})x} = \sum_{k=0}^{\infty} \frac{L(\chi, -k)(-x)^k}{k!}.$$

It suffices to prove the theorem for partial zeta functions associated to ray classes; hence let $[\mathfrak{b}]$ be a ray class in $C_{\mathfrak{m}}$ the ray group modulo $\mathfrak{m}$. Substituting in the Fourier-theoretic relation

$$\zeta_m(\sigma_a, -k) = \frac{1}{|C_{\mathfrak{m}}|} \sum_{\chi \in \widehat{C_{\mathfrak{m}}}} \overline{\chi}([\mathfrak{b}]) L(\chi, -k)$$

from above, we obtain

$$\sum_{k=0}^{\infty} \frac{L(\chi, -k)(-x)^k}{k!} = \frac{1}{|C_{\mathfrak{m}}|} \sum_{\chi \in \widehat{C_{\mathfrak{m}}}} \overline{\chi}([\mathfrak{b}]) \sum_{\mathfrak{a}} \chi(\mathfrak{a}) e^{-(\mathbb{N}\mathfrak{a})x} = \sum_{\mathfrak{a} \in [\mathfrak{b}]} e^{-(\mathbb{N}\mathfrak{a})x}$$

where $\mathfrak{a}$ runs over integral ideals. By comparing coefficients we are done. □

Define the **Stickelberger elements** for the extension M/F as

$$\alpha_n(M) = \sum_{\sigma \in \mathrm{Gal}(M/F)} \zeta_M(\sigma, -n) \sigma^{-1}.$$

Note that the Stickelberger elements belong to $\mathbb{Q}[\mathrm{Gal}(M/F)]$.

Theorem 3.7 (Stickelberger's theorem)**.** *If $F = \mathbb{Q}$, then $\alpha_0(M)\mathbb{Z}[Gal(M/F)] \cap \mathbb{Z}[Gal(M/F)]$ annihilates $Cl(M)$.*

Addendum 3.8. A common older definition of "the Stickelberger element" $\alpha_0(M)$ for extensions of $\mathbb{Q}$ was as

$$\frac{1}{m}\sum_{\mathrm{Gal}(a\in(\mathbb{Z}/m)^\times)} a\sigma_a^{-1}$$

for $M = \mathbb{Q}(\zeta_m)$, where $\sigma_a \in \mathrm{Gal}(M/\mathbb{Q})$ corresponded to the automorphism sending $\zeta_m \mapsto \zeta_m^a$, and then functorially extending to all abelian extensions by restriction. This differs only slightly from our definition above, as we can compute. The Fourier inversion formula from above tells us that

$$\zeta_m(\sigma_a, 0) = \frac{1}{\varphi(m)} \sum_{\chi\in\mathrm{Gal}(\widehat{\mathbb{Q}(\zeta_m)}/\mathbb{Q})} \overline{\chi}(\sigma_a)L(\chi, 0).$$

Calculating $L(\chi, 0)$ is a classical problem; one can compute $L(\chi, 1)$ from the Dirichlet series, using a Gauss sum identity and some analytic manipulations, then use the functional equation. Recall that for χ not the principal character, we have

$$L(\chi, 0) = -\frac{1}{m}\sum_{r=1}^{m}\chi(r)r.$$

Let us pretend for now that this formula holds even for χ principal; we will make the correction at the end. Combining our formulas, we find

$$\zeta_m(\sigma_a, 0) = -\frac{1}{m\varphi(m)}\sum_{\chi,r}\chi(a^{-1}r)r = -\frac{1}{m\varphi(m)}\sum_{r} r\sum_{\chi}\chi(a^{-1}r).$$

The inner sum will be zero unless $a^{-1}r = 1$, in which case it is φm. Hence our result is $-\frac{a\varphi(m)}{m\varphi(m)} = -\frac{a}{m}$. We make our correction now; we have an excess term corresponding to the wrong formula for $L(\chi_0, 0)$ (where χ_0 is the principal character):

$$-\frac{1}{m\varphi(m)}\sum_{r=1}^{m}\chi_0(r)r = -\frac{1}{m\varphi(m)}\frac{m\varphi(m)}{2} = -\frac{1}{2}$$

so that we finally get $-\frac{a}{m} + \frac{1}{2}$. Compare this to the coefficient $\frac{a}{m}$ in the "classical" definition.

The zeta definition is thus a "centralizing around zero" of the coefficients, and is also the definition which points to the generalization, revealing the relationship with the analytic world. A further advantage is how cleanly one case works:

Proposition 3.9. *Stickelberger's theorem is trivially true for totally real extensions of* $\mathbb{Q}$.

Proof. The partial zeta functions each evaluate to zero because σ_a and σ_{-a} become identified. □

We will not digress to prove Stickelberger's full theorem here, though the proof is not beyond our tools; it relies on an explicit computation with a canonical element of the Stickelberger ideal, which then extends without obstacle to the full ideal by Kummer theory. See section 3 of [5].

The strong generalization of the Kummer congruences, proven by Deligne and Ribet in [3], was first formulated by Coates in our reference [5]. We will state the congruences in Coates' formulation because it is most convenient, though this is at the expense of some naturality.

Define

$$\delta_n(\mathfrak{b},\mathfrak{c},\mathfrak{f}) = (\mathbb{N}c)^{n+1}\zeta_{\mathfrak{f}}(b,-n) - \zeta_{\mathfrak{f}}(\mathfrak{bc},-n).$$

for nonnegative integers n. Further, let $w_k(L)$ for a field L be the maximum integer n such that the exponent of the group $\mathrm{Gal}(L(\mu_n)/L)$ divides k.

Theorem 3.10 (Deligne-Ribet, Coates formulation). *We have the following two congruence relations:*

(A) For $n \geq 0$, if p does not divide $\mathbb{N}\mathfrak{c}$, then $\delta_n(\mathfrak{b},\mathfrak{c},\mathfrak{f}) \in \mathbb{Z}_p$

(B) For $n > 0$, if $p|\mathfrak{f}$, then for $n > 0$,

$$\delta_n(\mathfrak{b},\mathfrak{c},\mathfrak{f}) \cong (\mathbb{N}\mathfrak{bc})^n\delta_0(\mathfrak{b},\mathfrak{c},\mathfrak{f}) \pmod{w_n(M_{\mathfrak{f}})\mathbb{Z}_p}$$

where $M_{\mathfrak{f}}$ is the field associated to $C_{\mathfrak{f}}$.[20]

Proof. The heavily technical proof of [3] uses p-adic Hilbert modular forms, which certainly we cannot cover here. It should be noted that their result is formulated in a seemingly more general way, and is stated in terms of L-functions rather than partial zeta functions, but in fact Coates' congruences (A) and (B) are equivalent to Deligne and Ribet's statement. See [18] for details. □

This is a strict generalization of all our previous concrete Kummer congruences from this: (A) gives us the result, implicit in the Kummer congruences, that B_k/k is p-integral whenever $p-1$ does not divide k. Taking the linear combinations of the identities (B) over representatives $\mathfrak{b}$ of all classes inside $C_{\mathfrak{f}}$ for some $\mathfrak{f}$ to be determined, with weights $\chi(\mathfrak{bc})$, the expression we obtain is

$$((\mathbb{N}\mathfrak{c})^{n+1}\chi(\mathfrak{c})-1)L(\chi,-n) \equiv ((\mathbb{N}\mathfrak{c})^{n+1}\chi(\mathfrak{c})-1)L(\omega^n\chi,0) \pmod{w_n(M_{\mathfrak{f}})\mathbb{Z}_p}.$$

Notice that the norm map modulo p^n can be identified with ω. Take $F = \mathbb{Q}$. Setting χ to be the principal character modulo $\mathfrak{f} = (p)$ recovers the congruence

$$B_{1,\omega^n} \equiv \frac{B_{n+1}}{n+1} \pmod{p}$$

given that $p-1$ does not divide $n+1$, for example, since $w_1(M_{\mathfrak{f}})$ here is p, and we need $(\mathbb{N}\mathfrak{c})^{n+1} \cong \omega^{n+1}(\mathfrak{c})$ to not be the identity character. Setting $\mathfrak{f} = (p^k)$ recovers the classical Kummer congruences, because the order of ω with respect to $\mathfrak{f}$ is precisely $\varphi(p^k)$. The only potential problems arise when a power of p divides $(\mathbb{N}\mathfrak{c})^{n+1}-1$, but in fact this leads precise to the same "excess" power of p dividing $w_n(M_{\mathfrak{f}})$, as by class field theory the latter is also the greatest common divisor of $(\mathbb{N}\mathfrak{c})^n - 1$ across all ideals $\mathfrak{c}$ whose Artin symbol

[20] The use of the strange constants $w_n(-)$ here is probably the part of our convenient formulation which detracts most from comprehensibility, especially since we did not discuss the Stickelberger ideals in depth, where they play an important role. It may be enlightening to state that in fact, congruence (B) is equivalent to the apparently weaker statement that for for every $n \geq 0$, there exists an integer $m = m(n)$ such that if $p^m|\mathfrak{f}$, then the congruence holds modulo $p^n\mathbb{Z}_p$. There is a natural way to strengthen these congruences by breaking up the partial zeta functions into sums of partial zetas with respect to larger conductors; $w_n(-)$ can be thought of as precisely the strongest possible modulus one obtains via this technique.

under the modulus $\mathfrak{f}$ is the identity, so we simply have to use the change of character $\chi \to \chi\omega^{-1}$ and take $\mathfrak{c}$ an ideal in the kernel of the Artin symbol; one sees that this leads to the correct twisting as well.

We can also see the relation to the p-adic integration approach discussed above. Indeed, congruence (B) resembles the result of theorem 3.5. This is not a coincidence; it is in fact a special case, and we can sketch the more general application as follows: for $\mathfrak{c}$ an ideal of F prime to p, define the measures $\mu_{\mathfrak{c}}$ by

$$\mu_{k,\mathfrak{c}}(a + p^n\mathbb{Z}_p) = (\mathbb{N}\mathfrak{c})^{-k}\zeta_{p^n}(\mathfrak{c}a, 1-k) - \zeta_{p^n}(a, 1-k).$$

This is a direct generalization of Mazur's construction which we discussed earlier.

Theorem 3.11. *Each $\mu_{k,\mathfrak{c}}$ defines a measure on $\mathbb{Z}_p$, and $\mu_{k,\mathfrak{c}} = x^{k-1}\mu_{1,\mathfrak{c}}$ as measures.*

Proof. That the measures assign $\mathbb{Z}_p$ values to locally constant $\mathbb{Z}_p$-valued functions is congruence (A), and the identity is congruence (B). Notice that taking formal differentials of both sides in the identity results in the identity from the proof of theorem 3.5. $\mu_{1,\mathfrak{c}}$ is certainly a measure, so the result follows. □

With this, one can write down the formula

$$L_p(\chi, s) = \frac{1}{\chi(\mathfrak{c})\langle\kappa((\mathfrak{c}, F(\mu_{p^\infty})/F))\rangle^{s-1} - 1}\int_{\mathbb{Z}_p^\times} \chi(x)\langle x\rangle^{-s} d\mu_{1,\mathfrak{c}}$$

where $(-, L/K)$ denotes the Artin symbol and κ is the character giving the isomorphism of Γ with $\mathbb{Z}_p$. Once again, one can also twist by powers of $\omega(x)$ to get the "branches" of the L-function in the spirit of Tate's thesis, or equivalently an L-function with domain $\mathbb{Z}/(p-1) \times \mathbb{Z}_p$. See [18] for details.[21]

With this understanding of the congruences, we are ready for our main construction. We will follow [5] almost to the letter. Let M/F be an arbitrary finite abelian extension with Galois group G, and let χ be a faithful multiplicative character of G; we regard its values as living in $\mathbb{C}_p$. Set $M_0 = M(\mu_p)$ and let $q_n = p^n w_1(M_0)$; then take $M_0 \subset M_1 \subset \ldots \subset M_n \subset \ldots$ to be the cyclotomic $\mathbb{Z}_p$ extension $M_0 \subset M_0(\mu_{q_1}) \subset \ldots \subset M_0(\mu_{q_n}) \subset \ldots$. As before, let M_∞ be the total composite extension, and we set $G_n = \mathrm{Gal}(M_n/F)$, so that $G_\infty = \mathrm{Gal}(M_\infty/F)$. Denote $\xi_n = \alpha_0(M_n)$ for the zeroth Stickelberger element of M_n/F.

The gist of our approach is to projectively glue together the ξ_n along our tower of extensions, pushed down to Λ with a twist by the character χ; the presence of zeta in the definition of Stickelberger elements, along with the fact that they naturally live (up to some factor of integrality) in quotients of the associated Iwasawa algebra, is what suggests this approach. In fact it is not hard to see that the evaluation maps from the cyclotomic characters κ^n assemble the partial zeta summands of the Stickelberger elements in just the right way, and this is the essence of the proof that our construction works below.

Furthermore, Stickelberger's theorem for $\mathbb{Q}$ suggests to us that the Stickelberger elements are somehow related to the characteristic power series of X_∞, since they annihilate the class group.[22]

[21] Be warned that Ribet uses nontrivially different conventions than we have used here; we changed the conventions to match [17] so as to make the generalization clear.

[22] One might think that Stickelberger's theorem would follow from the main conjecture (see later) since the latter is such a strong

Write $\mathcal{O}_\chi$ for the ring of integers of $\mathbb{Q}_p(\chi)$; all our group rings will be with $\mathcal{O}_\chi$ coefficients, since our L-function will live in Λ_χ.

Let $F_0 \subset F_1 \subset \ldots$ be the unique $\mathbb{Z}_p$-extension of F contained its p-cyclotomic extension. Let $p^{e+1} = q_0$ so that we have a restriction map $r_n : G_n \to \mathrm{Gal}(F_{n+e}/F)$, and let $\rho_n : \mathcal{O}_\chi[G_n] \to \mathcal{O}_\chi[\mathrm{Gal}(F_{n+e}/F)]$ be the induced homomorphism extending linearly by sending $\sigma \in G_n$ to $\chi(\sigma) r_n(\sigma)$.

For $\mathfrak{c}$ a nontrivial ideal of F prime to p and $\mathfrak{f}(\chi)$, set $\nu_n(\mathfrak{c}) = \mathbb{N}\mathfrak{c} - (\mathfrak{c}, F_{n+e}/F)\chi(\mathfrak{c})$; this is the factor needed to make the Stickelberger coefficients integral so we can work in the integral group algebra, by congruence (A). Explicitly, we have the elements

$$\nu_n(\mathfrak{c})\rho_n(\xi_n) = \sum_{\sigma \in G_n} \delta_0(\sigma, \mathfrak{c}, M_n)\chi(\sigma)^{-1} r_n(\sigma)^{-1} \in \mathcal{O}_\chi[\mathrm{Gal}(F_{n+e}/F)]$$

which we will denote by $\eta_n(\mathfrak{c})$. Since the primes ramified in the extensions M_n/F are eventually stable, a final collection of these elements fit together in the inverse limit to yield $\eta \in \varprojlim \mathcal{O}_\chi[\mathrm{Gal}(F_{n+e}/F)] \cong \Lambda_\chi$. Denote by $f_{\mathfrak{c}}(T, \chi)$ the corresponding power series.

A final collection of the ν_n also yield an element ν of Λ_χ for the same reason, which corresponds to the formal power series

$$u_{\mathfrak{c}}(T, \chi) = \mathbb{N}\mathfrak{c} - \chi(\mathfrak{c})(1+T)^{\tau(\mathfrak{c})}.$$

where $\tau(\mathfrak{c})$ is defined as the additive character which satisfies $(\mathfrak{c}, F_\infty/F) = \gamma^{\tau(\mathfrak{c})}$. This allows us then to define the twisted pro-Stickelberger element, as we wanted:

$$G(T, \chi) := f_{\mathfrak{c}}(T, \chi)/u_{\mathfrak{c}}(T, \chi) \in Q(\Lambda_\chi).$$

We drop the $\mathfrak{c}$ from the subscript since by construction the result is independent of it.

Before taking the short step to defining the L-functions themselves, we analyze the poles of $G(T, \chi)$. Recall that ω (previously referred to as θ) denotes the cyclotomic character of the Galois action of $\mathrm{Gal}(M_0/F)$ on μ_p.

Theorem 3.12. *If $\chi\omega^{-1}$ is nontrivial on the subgroup of $Gal(M_0/F)$ fixing F_e, $G(T, \chi) \in \Lambda_\chi$. Otherwise, we have $(\mu_\chi(1+T) - u)G(T, \chi) \in \Lambda_\chi$, where μ_χ is the p^eth root of unity given by $\chi\omega^{-1}(\gamma)$.*

Proof. Let $h(T)$ be the greatest common divisor of $u_{\mathfrak{c}}(T, \chi)$ as $\mathfrak{c}$ varies; certainly any poles must be present in $h(T)$. Let us suppose that it is nontrivial, since otherwise everything lives in Λ_χ.

structural statement, but only small parts of it follow - for one thing, by nature, the main conjecture (and p-adic L-functions) can only "see" the negative part of the class group; for another, the p-adic L-functions are built from "stable" Stickelberger elements, in that they do not project down to the group rings corresponding to extensions with any lesser ramification. However, results in the reverse direction are considered by Coates in section 5.3 of [5]; in particular, the main conjecture over F would follow from a generalization of Stickelberger's theorem over F and a strong structural assumption on the class group. Little progress has been made in either direction, and with the Mazur-Tate and Rubin proofs, interest seems to have waned.

Let π be a local parameter for $\mathcal{O}_\chi$. Picking $\mathfrak{c}$ so that $\tau(\mathfrak{c})$ is a unit, we see that $h(T)$ cannot reduce to zero modulo π, so by the structure theory, we can assume that $h(T)$ is a distinguished polynomial. Say that it has a root $\alpha \in \mathbb{C}_p$; we find then that

$$\chi(\mathfrak{c})\omega^{-1}(\mathfrak{c}) = \left(\frac{u}{1+\alpha}\right)^{\tau(\mathfrak{c})}$$

for all $\mathfrak{c}$, where we have used the factorization $\mathbb{N}\mathfrak{c} = u^{\tau(\mathfrak{c})}\omega(\mathfrak{c})$ coming from considering the norm map as being defined on the Galois group of the ray class field of $F[\mathfrak{f}(\chi)p^\infty]$, then taking the projections onto Γ and its direct complement.

Since $|\alpha|_p < 1$, $\chi\omega^{-1}$ must be trivial on the subgroup fixing F_e, as the RHS is some power of a p-power root of unity divided by something $1 \pmod{\pi}$, while the image of $\mathfrak{c}$ corresponding to an element of $\mathrm{Gal}(M_0/F_e)$ necessarily has non-p power torsion.

Then taking $\mathfrak{c}$ in the preimage of γ under the Artin symbol, we obtain $\alpha = \mu_\chi^{-1}u - 1$. Since by taking a formal derivative $u_\mathfrak{c}(T,\chi)$ can only have simple roots, the result follows. □

We now finally define $L_p(\chi, s) := G(u^s - 1, \chi^{-1}\omega)$. Note that the potential poles of $G(T, \chi^{-1}\omega)$, which arise when ω is trivial on $\mathrm{Gal}(M_0/F_e)$, do not actually in general give rise to poles of $L_p(\chi, s)$: we would need $u^{s-1} = \mu_\chi$, which is impossible for $\mu_\chi \neq 1$, i.e. the case of χ principal and $s = 1$. Hence we recover the same meromorphicity/analyticity conditions for $L_p(\chi, s)$, as expected.[23]

Theorem 3.13. *This is a good definition of $L_p(\chi, s)$; i.e.*

$$G(u^{-k} - 1, \chi) = L_S(\chi^{-1}\omega^{-k}, -k)$$

for integers $k \geq 0$.

Proof. We need to work out that the evaluation map κ^{-k} combines the partial zetas in the elements $\eta_n(\mathfrak{c})$ in the correct way. Indeed, since κ also acts on each $\mathcal{O}_\chi[\mathrm{Gal}(F_n/F)]$, we have that $f_\mathfrak{c}(u^{-k} - 1, \chi) = \varprojlim \kappa_*^{-k}\eta_n(\mathfrak{c})$. We calculate

$$\kappa_*^{-k}\eta_n(\mathfrak{c}) \equiv \sum_{b \in C_{\mathfrak{f}(M_n/F)}} \delta_0(\mathfrak{b}, \mathfrak{c}, \mathfrak{f}(M_n/F))\chi^{-1}(\mathfrak{b})\langle \mathbb{N}\mathfrak{b}\rangle^k \pmod{q_n\mathcal{O}_\chi}.$$

By congruence (B), this is further congruent to

$$(\mathbb{N}\mathfrak{c})^{-s} \sum_{b \in C_{\mathfrak{f}(M_n/F)}} \delta_k(\mathfrak{b}, \mathfrak{c}, \mathfrak{f}(M_n/F))\chi^{-1}(\mathfrak{b})\omega^{-k}(\mathfrak{b}) \pmod{q_n\mathcal{O}_\chi}$$

which assemble in the inverse limit to give $(\mathbb{N}\mathfrak{c} - \chi^{-1}(\mathfrak{c})\omega^{-k}(\mathfrak{c}))(\mathbb{N}\mathfrak{c})^{-k}L_S(\chi^{-1}\omega^{-k}, -k)$. Similarly, the first two terms come from $u_\mathfrak{c}(u^{-k} - 1, \chi)$, so we are done. □

[23] The condition on a p-adic function f of having an analytic function φ for which $f(s) = \varphi(u^s - 1)$ is not unimportant, however, and is called **Iwasawa analyticity**. Thus $L_p(\chi, s)$ is Iwasawa analytic precisely when χ is nontrivial on $\mathrm{Gal}(M_0/F_e)$.

Recall Dirichlet's classical formula

$$L(\chi, 1) = -\frac{\tau(\chi)}{\mathfrak{f}(\chi)} \sum_{r=1}^{\mathfrak{f}(\chi)} \overline{\chi}(r) \log(1 - \zeta_{\mathfrak{f}(\chi)}^{-r}).$$

Leopoldt derived an exact analogue for the p-adic L-functions abelian over $\mathbb{Q}$:

Theorem 3.14 (Leopoldt's formula)**.** *Let χ be a nonprincipal even Dirichlet character. Then*

$$L_p(\chi, 1) = -\left(1 - \frac{\chi(p)}{p}\right) \frac{\tau(\chi)}{\mathfrak{f}(\chi)} \sum_{r=1}^{\mathfrak{f}(\chi)} \chi^{-1}(r) Log(1 - \zeta_{\mathfrak{f}(\chi)}^{-r}).$$

Proof. The proof is a long and technical manipulation of formal power series, based on the formula from the first analytic construction. See chapter 8 of [17]. □

From this, we can derive an equally precise analogue of the analytic class number formula for abelian extensions of $\mathbb{Q}$; this will wait until the next section.

4 Main conjecture

Finally, we can state the main conjecture of Iwasawa. Let F be a totally real number field, p ann odd number, and let ω be the cyclotomic character of the extension $F(\mu_p)/F$. Recall $G_i(T) \in \Lambda$ is the power series for $e_i A_\infty$, i.e. such that $G_i((1+T)^{-1} - 1)$ is the characteristic power series for $\hom(e_i A_\infty, \mathbb{Q}_p/\mathbb{Z}_p) \cong e_{1-i} X_\infty(-1)$, where here A_∞ and X_∞ are defined with respect to the cyclotomic $\mathbb{Z}_p$-extension starting at $F(\mu_p)$.

Theorem 4.1 ("Main conjecture")**.** *For odd integers $i \not\equiv 1 \pmod{[F(\mu_p) : F]}$, $(G(T, \omega^i)) = (G_i(T))$ as ideals of Λ. Else, $((1 + T - u)G(T, \omega)) = (G_i(T))$.*

Proof. This was first proven by Mazur and Wiles in a very technical way using modular forms; later, Rubin gave a simpler argument using Euler systems. □

Here, finally, we have interpreted zeta (or rather its "Iwasawa analytified" version $G(T, \omega^i)$) as the characteristic polynomial of γ acting on a Jacobian-like module, as Weil did in the geometric case, up to a unit in Λ.

4.1 Beginnings

We give some indications of further results which historically pointed towards the truth of the main conjecture, which also become important first steps in the proof of Rubin and the development of Euler systems by Kolyvagin.

First, as promised, there is a p-adic analytic class number formula. Compare the following theorem to theorem 2.4; no explanation of its suggestiveness is needed. We will use the notation $\zeta_p(K,s)$ for $L_p(\chi_0, s)$ where χ_0 is the principal character of the absolute Galois group of K.

Theorem 4.2 (p-adic analytic class number formula)**.** *Let $F/\mathbb{Q}$ be a totally real abelian extension of degree d. Then*

$$\lim_{s\to 1}(s-1)\zeta_p(K,s) = \frac{2^{d-1}h_F R_p}{\sqrt{\Delta_{F/\mathbb{Q}}}}\prod_{\mathfrak{p}|p}(1-(\mathbb{N}\mathfrak{p})^{-1})$$

Proof. Analogously to the archimedean case for p-power cyclotomic fields, the Frobenius determinant formula yields the following formula for the regulator of the 1-cyclotomic units:

$$R_p(C_1) = (-1)^{d-1}\prod_{\chi\neq 1}\frac{1}{2}\sum_{\sigma\neq 1}\chi^{-1}(\sigma)\mathrm{Log}(\xi_\sigma)$$

where characters and Galois elements range over $\widehat{\mathrm{Gal}(K/\mathbb{Q})}$ and $\mathrm{Gal}(K/\mathbb{Q})$ respectively, and

$$\xi_\sigma := \prod_{(a,K/\mathbb{Q})=\sigma}\xi_a := \prod_{(a,K/\mathbb{Q})=\sigma}\frac{1-\zeta^a_{\mathfrak{f}(\chi)}}{1-\zeta_{\mathfrak{f}(\chi)}}$$

is the cyclotomic unit basis element associated to $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, $\sigma \neq 1$; these project onto a basis of the 1-cyclotomic units $C_1 \subset E_1$. This works because the Iwasawa logarithm factors through the projector to E_1 in the unit group (i.e. non-p-power roots of unity are sent to zero).

The rest is much as we have seen before: we may write

$$\zeta_p(F,s) = \prod_{\chi\in\widehat{\mathrm{Gal}(F/\mathbb{Q})}} L_p(\chi,s),$$

and recalling that the residue of $\zeta_p(\mathbb{Q},s)$ at $s=1$ is $1-1/p$, and using Leopoldt's formula for the other L-terms, we see that the Euler factors cancel out and our statement is equivalent to

$$(-1)^{d-1}\prod_{\chi\neq\chi_0}\frac{\tau(\chi)}{\mathfrak{f}(\chi)}\sum_{r=1}^{\mathfrak{f}(\chi)}\chi^{-1}(r)\mathrm{Log}(1-\zeta^{-r}_{\mathfrak{f}(\chi)}) = \frac{2^{d-1}h_F R_p}{\sqrt{\Delta_{F/\mathbb{Q}}}}$$

which by the conductor-discriminant formula reduces to

$$(-1)^{d-1}\prod_{\chi\neq\chi_0}\sum_{r=1}^{\mathfrak{f}(\chi)}\chi^{-1}(r)\mathrm{Log}(1-\zeta^{-r}_{\mathfrak{f}(\chi)}) = 2^{d-1}h_F R_p.$$

We rewrite the LHS as

$$(-1)^{d-1}\prod_{\chi\neq\chi_0}\left(\sum_{r=1}^{\mathfrak{f}(\chi)}\chi^{-1}(r)\mathrm{Log}(\xi_r)+\sum_{r=1}^{\mathfrak{f}(\chi)}\chi^{-1}(r)\mathrm{Log}(1-\zeta_{\mathfrak{f}(\chi)})\right) = (-1)^{d-1}\prod_{\chi\neq\chi_0}\sum_{\sigma\neq 1}\chi^{-1}(\sigma)\mathrm{Log}(\xi_\sigma)$$

since the second term in each sum is a constant times character sum of a nonprincipal character, and then we collapse the residues modulo the conductor to their Artin symbols in the Galois group. But as above, this last expression is precisely $2^{d-1}R_p(C_1)$. Looking at volumes, $R_p(C_1) = [E_1 : C_1]R_p = h_K R_p$ by classical results on the Archimedean side, so this completes the proof.

For details of Leopoldt's original proof, see [15]. Note that this gives us yet another way to state Leopoldt's conjecture. □

The proof for any totally real base K instead of $\mathbb{Q}$ was resolved more recently by Colmez, in [2]; the proof would take us considerably far afield.

For convenience, denote by $H(T, \omega^i)$ either $G(T, \omega^i)$ or $(1 + T - u)G(T, \omega^i)$ as per the conditions in the theorem. An immediate corollary here is that $G_1(T)$ and $H(T, \omega)$ have the same λ and μ values. This relationship between numerical invariants is actually extremely important, and the following generalization is the foundation for the proof by method of Euler systems employed by Rubin:

Theorem 4.3. *If F is abelian over K totally real, then $\prod G_i(T)$ and $\prod H(T, \omega^i)$ have the same λ and μ values.*

Proof. This follows immediately by functoriality of p-adic L-functions and the more general p-adic analytic class number formula of Colmez (though in fact this result precedes that one). □

The results generated by the method of Euler systems involve showing a whole array of annihilation results, which in total prove that $G_i(T)$ divides $H(T, \omega^i)$ for each i. When combined with the above numerical result, this yields the main conjecture. Kolyvagin's article [13] introducing Euler systems discusses these two fragments of the proof as "$x_i \leq y_i$" and "$\sum x_i = \sum y_i$," respectively.

5 Acknowledgements

Many thanks to Prof. Peter May for putting in the tireless work to run the UChicago REU this year, as every year; this applies doubly so for dealing with my antics this time around. I would also like to thank Profs. Kazuya Kato and Matthew Emerton for sharing their expertise on number theory, without which I would have been even more lost than I was and am. Finally, much gratitude to Yun Cheng and John Halliday for putting up with my inane complaints during the process of writing this essay.

References

[1] Brumer, Armand. "On the units of algebraic number fields." *Mathematika, A Journal of Pure and Applied Mathematics* 14/2 (1967), p. 121—124.

[2] Colmez, Pierre. "Résidu en $s = 1$ des fonctions zêta p-adiques." *Inventiones mathematique* 91 (1988), p. 371—389.

[3] Deligne, Pierre and Ribet, Kenneth. "Values of abelian L-functions at negative integers over totally real fields." *Inventiones mathematicae* 59 (1980), p. 227–286.

[4] Ferrero, Bruce and Washington, Lawrence. "The Iwasawa invariant μ_p vanishes for abelian number fields." *Annals of Mathematics. Second Series* 109/2 (1979), p. 377—395.

[5] Coates, John. "p-adic zeta functions and the theory of Iwasawa." Collected in *Algebraic Number Fields*, ed. Albrecht Fröhlich. Academic Press (1977), p. 269—353.

[6] Hasse, Helmut. *Über die Klassenzahl abelscher Zahlkörper.* Erstauflage Akademie-Verlag (1952).

[7] Iwasawa, Kenkichi. "A note on class numbers of algebraic number fields." *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 20 (1956), p. 257—258.

[8] Iwasawa, Kenkichi. "On Γ-extensions of algebraic number fields." *Bulletin of the American Mathematical Society* 65 (1959), p. 183—226.

[9] Iwasawa, Kenkichi. "On some invariants of cyclotomic fields." *Bulletin of the American Mathematical Society* 80 (1958), p. 773—783.

[10] Iwasawa, Kenkichi. "Sheaves for algebraic number fields." *Annals of Mathematics* 69 (1959), p. 408—413.

[11] Kida, Yoshikata. "l-extensions of CM fields and cyclotomic invariants." *Journal of Number Theory* 12 (1980), p. 519—528.

[12] Koblitz, Neal. "p-adic analysis: a short course on recent work." *London Mathematical Societal Lecture Notes* 46 (1980).

[13] Kolyvagin, Victor. "Euler systems," in *The Grothendieck Festschrift (Vol. II)*, Pierre Cartier et al., eds. *Progress in Mathematics.* 87 (1990), p. 435—483.

[14] Kummer, Ernst. "Allgemeiner Beweis des Fermatschen Satzes, dass die Gleichung $x^\lambda + y^\lambda = z^\lambda$ durch ganze Zahlen unlösbar ist, für alle diejenigen Potenz-Exponenten λ, welche ungerade Primzahlen sind und in den Zählern der ersten $(\lambda - 3)/2$ Bernoullischen Zahlen als Factoren nicht vorkommen." *Journal für die reine und angewandte Mathematik* 40 (1850), p. 131—138.

[15] Leopoldt, Heinrich-Wolfgang. "Zur Arithmetik in abelschen Zahlkörpern." *Journal für die reine und angewandte Mathematik* 209 (1962), p. 54—71.

[16] Barry Mazur. "Analyse p-adique." Unpublished (1972).

[17] Murty, Maruti Ram. *Introduction to p-adic Analytic Number Theory.* International Press (2002).

[18] Ribet, Kenneth. "Report on p-adic L-functions over totally real fields." *Astérisque* 61 (1979), p. 177—192.

[19] Serre, Jean-Pierre. "Classes des corps cyclotomiques." *Séminaire N. Bourbaki* 174 (1958—1960), p. 83—93.

[20] Siegel, Carl. "Bernoullische Polynome und quadratische Zahlkorper." *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen* 2 (1968), p. 7—38.

[21] Siegel, Carl. "Über die Fouriersche Koeffizienten von Modulformen." *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen* 3 (1970), p. 15—56.

[22] Weil, André. Letter to Simone Weil. March 26, 1940. Collected in: Weil, Simone, ed. Robert Chenavier. *Oeuvres complétes, tome VII: correspondance.* Èditions Gallimard (2002), p. 535—552.