On Some Special Cases of Fermat's Last Theorem

Aniruddh V. *

December 15, 2024

Contents

1 Introduction

Fermat's Last Theorem asserts that the equation $x^n + y^n = z^n$ has no solutions $(x, y, z) \in \mathbf{Z}_+$ whenever $n \in \mathbf{Z}$ is an integer greater than 2. Although it was stated by Fermat in the 17th century, a full proof of Fermat's Last Theorem was only provided in the late 20th century by Wiles, involving high-powered machinery from modern number theory.

1.1 Some History

Fermat first proposed his last theorem in 1637. Since then, many have tried and failed to provide a complete proof of Fermat's Last Theorem. One particularly important example is Lamé's attempted proof in 1844. Lamé noticed that one could factor $x^n + y^n = z^n$ as

$$\prod_{i=1}^{n}(x + \zeta_n^i y) = z^n$$

*Reviewers: Andrew C. and David Z.

and thus began studying the ring of cyclotomic integers $\mathbf{Z}[\zeta_n]$. This marks the inception of the use of algebraic number theoretic techniques to prove Fermat's Last Theorem. Lamé's proof used algebraic properties of the ring $\mathbf{Z}[\zeta_n]$ to construct an infinite descent argument, thus arriving at a contradiction. When Lamé presented his solution, Liouville was quick to notice that Lamé implicitly assumed unique factorization of elements in $\mathbf{Z}[\zeta_n]$. We now know this is not the case, for example in the case of $\mathbf{Z}[\zeta_{23}]$ - but this fact was known back then as well, as Kummer had shown a few years prior.

Nonetheless, Kummer began work trying to modify Lamé's proof - and had mild success. He proved Fermat's Last Theorem for the so-called regular primes.

Theorem 1.1 (Kummer)**.** *Let p be a regular prime, that is an odd prime such that p does not divide the class number of $\mathbf{Q}(\zeta_p)$. Then $x^p + y^p = z^p, (p, xyz) = 1$ has no solutions in the rational integers.*

The condition $(xyz, p) = 1$ is often referred to as the first case of Fermat's Last Theorem. In this case, the arguments are often much easier. One can use similar methods to give a proof of the second case (where p divides one of x, y, or z). Whether or not a given prime is regular seems a-priori difficult to determine - if all one knows is how to compute the class group via Minkowski'sn theorem and related facts from the geometry of numbers, computing the class number of $\mathbf{Q}(\zeta_n)$ amounts to calculating the class group, which becomes more and more difficult as n grows larger. Fortunately, Kummer also came up with an elegant criterion for determining when a given prime p is regular.

Theorem 1.2 (Kummer)**.** *A prime p is regular if and only if p does not divide the numerator of the Bernoulli numbers B_k, $k = 2, 4, 6, \dots, p-3$.*

For example, $B_{12} = -\frac{691}{2730}$, so $p = 691$ is not regular.
Kummer's work made apparent the importance of the class number of $\mathbf{Q}(\zeta_n)$, and in particular its relation to solutions of Fermat's Last Theorem. Many have given other proofs of more general cases given slightly weaker assumptions on the class number. We also present one such example towards the end of this paper.

Theorem 1.3. *Suppose p is prime, and the index of regularity of p satisfies $i(p) < \sqrt{p} - 2$. Then $x^p + y^p = z^p, (p, xyz) = 1$ has no solutions in the rational integers.*

1.2 Structure of the Paper

The paper is structured as follows: In Section 2, we present a proof of Theorem 1.1, which essentially comes out of an understanding of the ring $\mathbf{Z}[\zeta_p]$. In section 3, introduce p-adic L-functions. Using this theory, we present a proof of Theorem 1.2. Finally, in Section 4 we return to more algebraic notions by proving Stickelberger's theorem and Herbrand's theorem. Using these, we prove Theorem 1.3. In the appendices, we review basic facts about Dirichlet characters and L-series.

2 Fermat's Last Theorem

The goal of this section is to prove the following case of Fermat's Last Theorem, due to Kummer:

Theorem 2.1 (Theorem 1.1)**.** *Let p be an odd prime such that p does not divide the class number of $\mathbf{Q}(\zeta_p)$. Then $x^p + y^p = z^p, (p, xyz) = 1$ has no solutions in the rational integers.*

We can factor the above equation as

$$\prod_{i=0}^{p-1}(x + \zeta_p^i y) = z^p$$

where ζ_p is a primitive pth root of unity. Thus we are led to consider the ring $\mathbf{Z}[\zeta_p]$. For the remainder of this section, fix some prime p and let $\zeta := \zeta_p$. The following proposition helps to understand the ring $\mathbf{Z}[\zeta]$.

Proposition 2.2. *The ring $\mathbf{Z}[\zeta]$ is the ring of integers of $\mathbf{Q}(\zeta_p)$*

From this, we immediately deduce that $\mathbf{Z}[\zeta_p]$ is a Dedekind domain ([8], pp 9 Proposition 9).

Proof. If $\mathcal{O}$ is the ring of integers in $\mathbf{Q}[\zeta]$, then certainly $\mathbf{Z}[\zeta] \subset \mathcal{O}$. We need to show the reverse inclusion. Before giving the proof of Proposition 2.2, we need the following auxiliary results.

Lemma 2.3. *Let $r, s \in \mathbf{Z}$ be such that $(p, rs) = 1$. Then $\frac{\zeta^r - 1}{\zeta^s - 1}$ is a unit in $\mathbf{Z}[\zeta]$.*

Proof. Writing $r \equiv st (\mathrm{mod}\ p)$, we calculate

$$\frac{\zeta_p^r - 1}{\zeta_p^s - 1} = \frac{\zeta_p^{st} - 1}{\zeta_p^s - 1} = \zeta_p^{s(t-1)} + \zeta_p^{s(t-2)} + \cdots + \zeta_p^s + 1 \in \mathbf{Z}[\zeta_p]$$

Similarly, if $s \equiv ru (\mathrm{mod}\ p)$, we have

$$\frac{\zeta_p^s - 1}{\zeta_p^r - 1} = \frac{\zeta_p^{ru} - 1}{\zeta_p^r - 1} = \zeta_p^{r(u-1)} + \zeta_p^{r(u-2)} + \cdots + \zeta_p^r + 1 \in \mathbf{Z}[\zeta_p]$$

□

Lemma 2.4. *Let $\mathcal{O}$ denote the ring of integers of $\mathbf{Q}[\zeta]$. Then $(1 - \zeta)$ is prime in $\mathcal{O}$, and is totally ramified, so that $(1 - \zeta)^{p-1} = (p)$.*

Proof. Using the cyclotomic polynomial $\Phi_p(x) = \prod_{i=1}^{p-1}(x - \zeta^i) = x^{p-1} + x^{p-2} + ... + x + 1$, plugging in $x = 1$ gives $p = \prod_i (1 - \zeta^i)$, and $(p) = (1 - \zeta)$, where we have equality of ideals $(1 - \zeta) = (1 - \zeta^i)$ by 2.3. The ideal $(1 - \zeta)$ is prime since (p) can split into at most $p - 1 = [\mathbf{Q}(\zeta) : \mathbf{Q}]$ prime factors. □

We can now finish the proof of Proposition 2.2. Let υ denote the valuation corresponding to the ideal $(1 - \zeta)$. Let $\{1, 1 - \zeta, (1 - \zeta)^2, ..., (1 - \zeta)^{p-2}\}$ be a basis for $\mathbf{Q}(\zeta)$ over $\mathbf{Q}$. Any element $\alpha \in \mathcal{O}$ can be written uniquely as

$$\alpha = \sum_{i=0}^{p-2} c_i (1 - \zeta)^i$$

for $c_i \in \mathbf{Q}$. To show $\mathbf{Z}[\zeta] = \mathcal{O}$, we need to show $c_i \in \mathbf{Z}$. To see this, we first reduce to the case where p does not divide the denominator of c_i. This is possible since the numbers $\upsilon(c_i(1 - \zeta)^i), c_i \neq 0$ are distinct, so $\upsilon(\alpha) = min(\upsilon(c_i(1 - \zeta)^i))$. We know $\upsilon(\alpha) \geq 0$ and $\upsilon((1 - \zeta)^i) < p - 1$, so $\upsilon(c_i) \geq 0$. This shows that p is not in the denominator of c_i. So rearranging, we may write

$$\alpha = c_0 + c_1\zeta + \cdots + \cdot c_{p-2}\zeta^{p-2}$$

Then $\zeta^{-1}\alpha \in \mathcal{O}$, and thus $\mathrm{Tr}(\zeta^{-1}\alpha) \in \mathbf{Z}$. Since the minimal polynomial of ζ^j is $x^{p-1} + x^{p-2} + \cdots + x + 1$ whenever $(j, p) = 1$, we have

$$pc_i - \sum_{j=0}^{p-2} b_j = (p - 1)b_i \sum_{i \neq j} b_j = \mathrm{Tr}(\zeta^{-1}\alpha)$$

Applying this for two distinct indices, say $i = 0$ and $i = i$, gives $p(c_0 - c_i) \in \mathbf{Z}$, and thus $c_0 - c_i \in \mathbf{Z}$. We claim $c_0 \in \mathbf{Z}$; from this it will follow that $c_i \in \mathbf{Z}$. To see this, note we can write

$$\alpha = c_0(1 + \zeta + \cdots + \zeta^{p-2}) + \underbrace{(c_1 - c_0)\zeta + \cdots + (c_{p-2} - c_0)\zeta^{p-2}}_{\in \mathcal{O}}$$

Then

$$-\zeta^{p-1}c_0 = c_0(1 + \zeta + \cdots + \zeta^{p-2}) \in \mathcal{O}$$

and thus $c_0 \in \mathcal{O} \cap \mathbf{Q} = \mathbf{Z}$. □

We can now begin with the proof of Theorem 2.1 We can treat the case $p = 3$ first in a simple manner. If $3 \nmid x$, then $x \equiv \pm 1 \bmod 9$, and similarly for y and z. Thus $x^3 + y^3 \equiv 0, \pm 2 \bmod 9$, and so $x^3 + y^3 \not\equiv z^3 \bmod 9$, and thus $x^3 + y^3 \neq z^3$. Now, assume $p \geq 5$ and $x^p + y^p = z^p$ with $p \nmid xyz$. We may assume $x \not\equiv y \bmod p$, since if we did have $x \equiv y \equiv -z \bmod p$, then $-2z^p \equiv z^p$, which is impossible since $p \nmid 3z$. Finally assume $(x, y, z) = 1$, by dividing each by the greatest common divisor if necessary.

Before proceeding further, we need the following lemma.

Lemma 2.5. *Under the above assumptions, the ideals $(x + \zeta^i y), 0 \leq i \leq p - 1$ are pairwise relatively prime.*

Proof. Let $0 \leq i \neq j \leq p - 1$, be two distinct integers, and suppose there is some prime $\mathfrak{p}$ of $\mathbf{Z}[\zeta]$ such that $\mathfrak{p} \mid (x + \zeta^i y)$ and $\mathfrak{p} \mid (x + \zeta^j y)$. Then $\mathfrak{p} \mid (\zeta^i y - \zeta^j y)$, but as ideals we have $(\zeta^i y - \zeta^j y) = (1 - \zeta)y$, and so either $\mathfrak{p} = (1 - \zeta)$ or $\mathfrak{p} \mid (y)$. Similarly $\mathfrak{p}$ divides $\zeta^j(x + \zeta^i y) - \zeta^i(x + \zeta^j y) = (1 - \zeta)x$, so $\mathfrak{p} = (1 - \zeta)$ or $\mathfrak{p} \mid (x)$. So we must have $\mathfrak{p} = (1 - \zeta)$, otherwise $\mathfrak{p} \mid (x)$ and $\mathfrak{p} \mid (y)$, contradicting $(x, y) = 1$. Now $x + y \equiv x + \zeta^i y \equiv \bmod \mathfrak{p}$, and so $x + y \equiv 0 \bmod p$. But $x + y \equiv x^p + y^p = z^p \equiv 0 \bmod p$, so $p \mid z$, contradicting our initial assumption. □

Returning back to the proof, consider the equation

$$\prod_{i=0}^{p-1}(x + \zeta^i y) = (z)^p$$

as an equality of ideals. The ideals $(x + \zeta^i y), 0 \leq i \leq p - 1$ are relatively prime by the above lemma, and so each is the pth power of some ideal, say

$$(x + \zeta^i y) = A_i^p$$

Since the class number of $\mathbf{Q}(\zeta)$ is assumed to not be divisible by p, each A_i is also a principal ideal, say $A_i = (\alpha_i)$. It follows that $x + \zeta^i y = u_i \alpha_i^p$, where u is some unit in $\mathbf{Z}[\zeta]$. Now, fix $i = 1$, and omit subscripts so $x + \zeta y = u\alpha^p$. Again, we need a lemma.

Lemma 2.6. *Let u be a unit of $\mathbf{Z}[\zeta]$. Then there are $u_1 \in \mathbf{Q}(\zeta + \zeta^{-1})$ and $r \in \mathbf{Z}$ such that $u = \zeta^r u_1$.*

Proof. Let $\beta = u/\bar{u}$. Since u is a unit, $\beta \in \mathbf{Z}[\zeta]$. Also, since complex conjugation commutes with every other element of the Galois group, all conjugates of β have absolute value 1, and so β is a root of unity. So $u/\bar{u} = \pm\zeta^a$. We claim $u/\bar{u} = \zeta^a$ for some a. Indeed, suppose for the sake of contradiction that $u/\bar{u} = -\zeta^a$. Writing $u = b_0 + b_1\zeta + \cdots + b_{p-2}\zeta^{p-2}$, we see $u \equiv b_0 + b_1 + \cdots + b_{p-2} \bmod (1 - \zeta)$. Similarly, we have $\bar{u} = b_0 + b_1\zeta^{-1} + \cdots$, and we have $\bar{u} \equiv b_0 + b_1 + \cdots b_{p-2} \equiv u = -\zeta^a\bar{u} \equiv -\bar{u} \bmod (1 - \zeta)$. Adding these two congruences, we have $2\bar{u} \equiv 0 \bmod (1-\zeta)$. Since $2 \notin (1-\zeta)$, we must have $\bar{u} \in (1-\zeta)$, which is a contradiction since $\bar{u}$ is a unit. Therefore $u/\bar{u} = \zeta^a$ for some a. Let r be so that $a \equiv 2r (\bmod\, p)$, and let $u_1 := \zeta^{-r}u$. Then $u = \zeta^r u_1$, with $\bar{u_1} = u_1$. □

Returning to the proof, let $x + \zeta y = u\alpha^p$. The above lemma implies that $u = \zeta^r u_1$ for $r \in \mathbf{Z}$ and $u_1 = \bar{u_1}$. There is some $a \in \mathbf{Z}$ so that $\alpha^p \equiv a (\bmod\, p)$. Therefore we have

$$x + \zeta y = \zeta^r u_1 \alpha^p \equiv \zeta^r u_1 a (\bmod\, p)$$

and

$$x + \zeta^{-1} y = \zeta^{-r} u_1 \bar{\alpha}^p \equiv \zeta^{-r} u_1 a (\bmod\, p)$$

Combining these gives

$$x + \zeta y - \zeta^{2r} x - \zeta^{2r-1} y \equiv 0 (\bmod\, p) \tag{2.7}$$

We are ready to complete the proof. We need two final lemmas.

Lemma 2.8. *Let $\alpha \in \mathbf{Z}[\zeta]$. Then α^p is congruent to a rational integer* mod p.

Proof. Let $\alpha = b_0 + b_1\zeta + \cdots + b_{p-1}\zeta^{p-2}$. Then $\alpha^p \equiv b_0^p + (b_1\zeta)^p + \cdots + (b_{p-2}\zeta^{p-2})^p = b_0^p + b_1^p + \cdots + b_{p-2}^p (\bmod\, p)$. □

Lemma 2.9. *Let $\alpha = a_0 + a_1\zeta + \cdots + a_{p-1}\zeta^{p-1}$ with $a_i \in \mathbf{Z}$, and at least one $a_i = 0$. If $n \in \mathbf{Z}$ and n divides α, then n divides each α_j. Similarly, if all the $a_i \in \mathbf{Z}_p$ and at least one $a_i = 0$, if p divides α_i then p divides a_j.*

Proof. We have $1+\zeta+\zeta^2+\cdots+\zeta^{p-1}=0$, so any subset of $\{1,\zeta,\dots,\zeta^{p-1}\}$ with $p-1$ elements gives a $\mathbf{Z}$ basis for the abelian group $\mathbf{Z}[\zeta]$. Since at least one $a_i=0$, the other a_j give coefficients with respect to the basis. The first result follows. The second result follows from the exact same argument, replacing $\mathbf{Z}$ with $\mathbf{Z}_p$. □

If $1,\zeta,\zeta^{2r},\zeta^{2r-1}$ are distinct, then the above lemma implies p divides x and y, which contradicts our original assumptions. Thus they cannot be distinct, and since $1\neq\zeta$ and $\zeta^{2r}\neq\zeta^{2r-1}$, we have three cases.

- If $\zeta^2 r=1$, then 2.7 gives $x+\zeta y-x-\zeta^{-1}y\equiv 0(\text{mod } p)$, and so $\zeta y-\zeta^{p-1}y\equiv 0(\text{mod } p)$. Applying Lemma 2.9, we see $y\equiv 0(\text{mod } p)$, a contradiction.
- If $\zeta^{2r-1}=1$, then $\zeta^2 r=\zeta$. Then 2.7 becomes

  $$(x-y)-(x-y)\zeta\equiv 0(\text{mod } p)$$

  Applying Lemma 2.9, we have $(x-y)\equiv 0(\text{mod } p)$, contradicting the choice of x and y made at the beginning of the proof.
- If $\zeta=\zeta^{2r-1}$, then 2.7 becomes $x-\zeta^2 x\equiv 0(\text{mod } p)$, so $x\equiv 0(\text{mod } p)$, which is a contradiction.

This completes the proof of Theorem 2.1.

The statement and proof of Theorem 2.1 of the previous section raise the two following natural questions:

Question 2.10. Is there a simpler way (as compared to brute force computation) to determine whether or not a given prime p divides the class number of $\mathbf{Q}(\zeta_p)$?

Question 2.11. Can other conditions on the class number of $\mathbf{Q}(\zeta_p)$ be given to ensure that Fermat's last theorem holds?

3 p-adic L-functions and Bernoulli numbers

The goal of this section is to prove certain congruence relations for (generalized) Bernoulli numbers. To do this, we need to introduce p-adic L-functions and study some of their basic properties. We omit most proofs, and refer to [10] and [7]. We routinely refer to (generalized) Bernoulli numbers and Bernoulli polynomials, which we discuss in Appendix B.

3.1 p-adic Functions

Before discussing p-adic L-functions, we discuss the basic theory of p-adic functions. Throughout, fix some prime p, let $\mathbf{Q}_p$ denote the p-adic rationals, $\overline{\mathbf{Q}}_p$ its algebraic closure (which is not complete), and $\mathbf{C}_p$ the completion of $\overline{\mathbf{Q}}_p$ (which is algebraically closed).

We first define the p-adic exponential and logarithmic functions.

Definition 3.1. Define the p-adic exponential by

$$\exp(X)=\sum_{n=0}^{\infty}\frac{X^n}{n!}$$

Definition 3.2. Define the p-adic logarithm by

$$\log_p(1+X)=\sum_{n=1}^{\infty}\frac{(-1)^{n+1}X^n}{n}$$

Proposition 3.3. *There is a unique extension of* $\log_p$ *to all of* $\mathbf{C}_p^\times$ *such that* $\log_p(p)=0$ *and* $\log_p(xy)=\log_p(x)+\log_p(y)$ *for all* $x,y\in\mathbf{C}_p$.

Proof. See [10, Proposition 5.4]. □

3.2 p-adic L-functions

We can now work with p-adic L-functions. To begin, let

$$H(s,a,F) = \sum_{m \equiv a (\mathrm{mod}\, F)} m^{-s} = \sum_{n=0}^{\infty} \frac{1}{(a+nF)^s}$$

where s is a complex variable and $0 < a < F$ are integers. Then we have

$$H(1-n,a,F) = -\frac{F^{n-1}B_n(a/F)}{n} \in \mathbf{Q}$$

for $n \geq 1$, and H has a simple pole at $s = 1$ with residue $1/F$.

Throughout the rest of this section, we let $q := 4$ if $p = 2$, and $q = p$ otherwise for our fixed prime p.

Theorem 3.4. *Suppose $q|F$ and $p \nmid a$. Then there is a p-adic meromorphic function $H_p(s,a,F)$ defined on*

$$\{s \in \mathbf{C}_p \,:\, |s| < qp^{-1/(p-1)} > 1\}$$

such that

$$H_p(1-n,a,F) = \omega^{-n}(a)H(1-n,a,F)$$

for $n \geq 1$, where $\omega(a)$ is the $\phi(q)$-th root of unity such that $a \equiv \omega(a) (\mathrm{mod}\, q)$. In particular, when $n \equiv 0 (\mathrm{mod}\, p-1)$, or $\mathrm{mod}\, 2$ is $p = 2$, then

$$H_p(1-n,a,F) = H(1-n,a,F)$$

Furthermore, H_p is analytic except for a simple pole at $s = 1$, with residue $1/F$. Here

Proof. Let

$$H_p(s,a,F) = \frac{1}{s-1}\frac{1}{F}\langle a \rangle^{1-s} \sum_{j=0}^{\infty} \binom{1-s}{j} \left(\frac{F}{a}\right)^j B_j$$

where $\langle a \rangle = \omega(a)^{-1}a$, so that $\langle a \rangle \equiv 1 (\mathrm{mod}\, q)$ and $\log_p a = \log_p \langle a \rangle$. For now, ignore convergence issues - we refer to [10, Theorem 5.10]. Then

$$\begin{aligned} H_p(1-n,a,F) &= \frac{-1}{nF}\langle a \rangle^n \sum_{j=0}^{\infty} \binom{n}{j} \left(\frac{F}{a}\right)^j B_j \\ &= -\frac{F^{n-1}\omega^{-n}(a)}{n} B_n\left(\frac{a}{F}\right) \\ &= \omega^{-n}(a)H(1-n,a,F) \end{aligned}$$

as desired. At $s = 1$, we have the residue

$$\frac{1}{F}\langle a \rangle^0 \sum_{j=0}^{\infty} \binom{0}{j} \left(\frac{F}{a}\right)^j B_j = \frac{1}{F}$$

□

We are now ready to construct p-adic L-functions. Once and for all, fix an embedding of $\overline{\mathbf{Q}}$ into $\mathbf{C}_p$. We may therefore regard a Dirichlet character as having values in $\mathbf{C}_p$.

It will be useful to regard $\omega(a)$ as a p-adic character - while it may be treated as a complex character, in this case the choice is noncanonical and depends on a choice of embedding of $\mathbf{Q}(\zeta_{p-1})$ into $\mathbf{Q}_p$. Note that ω generates the group of Dirichlet characters defined mod q.

Theorem 3.5. *Let χ be a Dirichlet character of conductor f, and let F be any multiple of q and f. Then there is a p-adic meromorphic (analytic if $\chi \neq 1$) function $L_p(s,\chi)$ on $\{s \in \mathbf{C}_p : |s| < qp^{-1/(p-1)}\}$ such that*

$$L_p(1-n,\chi) = -(1-\chi\omega^{-n}(p)p^{n-1})\frac{B_{n,\chi\omega^{-n}}}{n}$$

for $n \geq 1$. Furthermore, if $\chi = 1$, then $L_p(s,1)$ is analytic except for a pole at $s=1$ with residue $1-1/p$. We have the formula

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

Proof. We show that the given formula works. Analyticity immediately follows from the fact that

$$L_p(s,\chi) = \sum_{a=1,p\nmid a}^{F}\chi(a) = H_p(s,a,F)$$

At $s=1$, the residue of $L_p(s,\chi)$ is

$$\sum_{a=1,p\nmid a}^{F}\chi(a)\frac{1}{F}$$

If $\chi = 1$, this sum is $1-1/p$. Otherwise, we may split the sum as

$$\frac{1}{F}\sum_{a=1}^{F}\chi(a) - \frac{1}{F}\sum_{b=1}^{F/p}\chi(pb)$$

The first sum is always 0. If $p|f$, then $\chi(pb)=0$ for all b. Otherwise, if $p\nmid f$, then $f|(F/p)$, so the second sum is 0. Therefore $L_p(s,\chi)$ has no pole at $s=1$ if $\chi\neq 1$. If $n\geq 1$, then we see

$$\begin{aligned}
L_p(1-n,\chi) &= \sum_{a=1,p\nmid a}^{F}\chi(a)H_p(1-n,a,F)\\
&= -\frac{1}{n}F^{n-1}\sum_{a=1,p\nmid a}^{F}\chi\omega^{-n}(a)B_n\left(\frac{a}{F}\right)\\
&= -\frac{1}{n}F^{n-1}\sum_{a=1}^{F}\chi\omega^{-n}(a)B_n\left(\frac{a}{F}\right) + \frac{1}{n}p^{n-1}\left(\frac{F}{p}\right)^{n-1}\sum_{b=1}^{F/p}\chi\omega^{-n}(pb)B_n\left(\frac{bp}{F}\right)
\end{aligned}$$

If $p|f_{\chi\omega^{-n}}$, then $\chi\omega^{-n}(pb)=0$. Otherwise, $f_{\chi\omega^{-n}}|(F/p)$. By Theorem B.6, we have

$$\begin{aligned}
L_p(1-n,\chi) &= -\frac{1}{n}\left(B_{n,\chi\omega^{-n}} - \chi\omega^{-n}(p)P^{n-1}B_{n,\chi\omega^{-n}}\right)\\
&= -\frac{1}{n}\left(1-\chi\omega^{-n}(p)p^{n-1}\right)B_{n,\chi\omega^{-n}}
\end{aligned}$$

This completes the proof. □

3.3 p-adic Congruences

In this section, we develop formulas for p-adic congruences. Our main tool is the following theorem, which we will use without proof.

Theorem 3.6. *Let χ be a nontrivial Dirichlet character, and suppose $pq\nmid f_\chi$. Then*

$$L_p(s,\chi) = a_0 + a_1(s-1) + a_2(s-1)^2 + \cdots$$

with $|a_0|\leq 1$ and $p|a_i$ for all $i\geq 1$.

Proof. See [10, Theorem 5.12]. □

From this, we easily deduce the following corollaries.

Corollary 3.7. *Suppose $\chi \neq 1$, and $pq \nmid f$. Let $m, n \in \mathbf{Z}$. Then*

$$L_p(m, \chi) \equiv L_p(n, \chi) (\mathrm{mod}\ p)$$

and both are p-integral.

Proof. Both $L_p(m, \chi)$ and $L_p(n, \chi)$ are congruent to a_0 in the notation of Theorem 3.6. □

Here is our main result, regarding Bernoulli numbers.

Corollary 3.8 (Kummer's Congruences)**.** *Suppose $m \equiv n \not\equiv 0 (\mathrm{mod}\ p-1)$ are positive even integers. Then*

$$\frac{B_m}{m} \equiv \frac{B_n}{n} (\mathrm{mod}\ p)$$

More generally, if m and n are positive even integers with $m \equiv n (\mathrm{mod}(p-1)p^a)$ and $n \not\equiv 0 (\mathrm{mod}\ p-1)$, then

$$(1 - p^{m-1})\frac{B_m}{m} \equiv (1 - p^{n-1})\frac{B_n}{n} (\mathrm{mod}\ p^{a+1})$$

Proof. Note that $L_p(s, \omega^m) = L_p(s, \omega^n)$. Then

$$L_p(1 - m, \omega^m) = -1(1 - p^{m-1})\frac{B_m}{m}$$

and similarly

$$L_p(1 - n, \omega^n) = -1(1 - p^{n-1})\frac{B_n}{n}$$

Thus we have

$$\begin{aligned} L_p(1 - m, \omega^m) &= a_0 + a_1(-m) + a_2(-m)^2 + \cdots \\ &\equiv a_0 + a_1(-n) + a_2(-n)^2 + \cdots (\mathrm{mod}\ p^{a+1}) \\ &= L_p(1 - n, \omega^n) \end{aligned}$$

□

Corollary 3.9. *Suppose n is odd and $n \not\equiv -1 (\mathrm{mod}\ p-1)$. Then*

$$B_{1,\omega^n} \equiv \frac{B_n}{n+1} (\mathrm{mod}\ p)$$

and both sides are p integral.

Proof. Since $n \not\equiv -1$, the character ω^{n+1} is not trivial, and thus $\omega^n(p) = 0$. Then, by Corollary 3.7

$$\begin{aligned} B_{1,\omega^n} &= (1 - \omega^n(p))B_{1,\omega^n} \\ &= -L_p(0, \omega^{n+1}) \\ &\equiv -L_p(2 - n, \omega^{n+1}) (\mathrm{mod}\ p) \\ &= (1 - p^n)\frac{B_n}{n+1} (\mathrm{mod}\ p) \\ &\equiv \frac{B_n}{n+1} (\mathrm{mod}\ p) \end{aligned}$$

The p-integrality also follows from Corollary 3.7 □

Using this machinery of p-adic L-functions, we may partially answer one of the questions posed at the end of Section 2. Here is our result.

Theorem 3.10. *Let p be an odd prime, and let $h_p^- := h_p/h_p^+$ be the relative class number of $\mathbf{Q}(\zeta_p)$. Then p divides the relative class number if and only if p divides the numerator of B_j for some $j = 2, 4, \dots, p-3$.*

Later, we will show that p divides the relative class number if and only if p divides the class number, which, combined with this theorem, gives a beautiful answer to Question 2.10.

Proof. The odd characters of $\mathbf{Q}(\zeta_p)$ are $\omega, \omega^3, \dots, \omega^{p-2}$. Therefore, by Theorem B.21

$$h_p^- = 2p \prod_{j=1, j \text{ odd}}^{p-2} (-\frac{1}{2} B_{1,\omega^j})$$

where we take $Q = 1$ by Corollary B.16 and $w = 2p$. First, we have

$$\begin{aligned} B_{1,\omega^{p-2}} &= B_{1,\omega^{-1}} \\ &= \frac{1}{p} \sum_{a=1}^{p-1} a\omega^{-1}(a) \\ &\equiv \frac{p-1}{p} \bmod \mathbf{Z}_p \end{aligned}$$

Therefore $(2p)(-\frac{1}{2}B_{1,\omega^{p-2}}) \equiv 1 (\text{mod } p)$. Using Corollary 3.9, we may write the above as

$$h_p^- = \prod_{j=1, j \text{ odd}}^{p-4} \left(-\frac{1}{2} \frac{B_{j+1}}{j+1} \right) (\text{mod } p)$$

and the result follows. □

3.4 The Class Number Formula

We conclude this section by proving the implication made above, namely that p divides the relative class number if and only if p divides the class number. Before this, we need some preliminaries.

Definition 3.11. Let K be a number field. Fix an embedding of $\mathbf{C}_p$ into $\mathbf{C}$, so any embedding of K into $\mathbf{C}_p$ can be thought of as an embedding into $\mathbf{C}$, and thus as real or complex (depending on the embedding into $\mathbf{C}$). Let $r = r_1 + r_2 - 1$, where r_1 and r_2 denote the real and complex embeddings of K. Enumerate the embeddings of K into $\mathbf{C}_p$ by $\sigma_1, \dots, \sigma_{r_1}, \sigma_{r_1+1}, \bar{\sigma}_{r_1+1}, \dots, \sigma_{r_1+r_2}, \bar{\sigma}_{r_1+r_2}$, where the σ_i, $1 \le i \le r_1$ are real in the above sense, and the other embeddings are complex. Let $\epsilon_1, \dots, \epsilon_r$ be independent units of K. Define

$$R_{K,p}(\epsilon_1, \dots, \epsilon_e) = \det(\delta_i \log_p(\sigma_i \epsilon_j))_{1 \le i,j \le r}$$

If $\{\epsilon_1, \dots, \epsilon_r\}$ form a basis of units of K modulo roots of unity, then $R_p(K) = R_{K,p}(\epsilon_1, \dots, \epsilon_r)$ is called the **p-adic regulator** of K. In general, the p-adic regulator is determined only up to sign.

We state the following result without proof. For details, see [10, 151-153]

Theorem 3.12 (The Class Number Formula)**.** *Let K be a totally real abelian number field of degree n corresponding to a group of Dirichlet characters X. Let $h(K)$ be the class number of K. Then*

$$\frac{2^{n-1} h(K) R_p(K)}{\sqrt{d(K)}} = \prod_{\chi \in X, \chi \neq 1} \left(1 - \frac{\chi(p)}{p} \right)^{-1} L_p(1, \chi)$$

From the above, we can deduce our desired results on class numbers. We need one final proposition.

Proposition 3.13. *Let K be a totally real Galois number field. If there is only one prime of K above p, and the ramification index of p is at most $p-1$, then* $\left|\frac{[K:\mathbf{Q}]R_p(K)}{\sqrt{d(K)}}\right|_p \leq 1$.

Proof. Let K_p denote the completion of K at the prime p, let $\mathcal{O}_p$ be the ring of integers of K_p. We have $\deg(K_p/\mathbf{Q}_p) = \deg(K/\mathbf{Q})$, and also $\mathrm{Gal}(K_p/\mathbf{Q}_p) \cong \mathrm{Gal}(K/\mathbf{Q})$. If $x \in K_p$ and $|x| < 1$, then $|x| \leq p^{-1/(p-1)}$. Thus, we have

$$\log_p(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \in \mathcal{O}_p$$

since all terms of the sum are in $\mathcal{O}_p$. By extending $\log_p$, we see that $\log_p \epsilon \in \mathcal{O}_p$ for all $\epsilon \in K_p^\times$.
Let $\epsilon_1, \dots, \epsilon_{n-1}$ be a basis for the units of K modulo $\{\pm 1\}$, where $n = [K : \mathbf{Q}]$ is the extension degree. Let $\beta_i = \log_p \epsilon_i$ for $1 \leq i \leq n-1$, and set $\beta_n = 1$. Let $\alpha_1, \dots, \alpha_n$ be a basis for $\mathcal{O}_p$ as a $\mathbf{Z}_p$ module. Then we can write

$$\beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j$$

with $\alpha_{ij} \in \mathbf{Z}_p$. Let $\sigma \in \mathrm{Gal}(K_p/\mathbf{Q}_p)$. Since $B_i^\sigma = \sum a_{ij}\alpha_j^\sigma$, we have $\det(\beta_i^\sigma)_{i,\sigma} = \det(a_{ij})_{i,j} \det(\alpha_i^\sigma)_{i,\sigma}$. Since there is only one prime above p, the p-part of the discriminant of K is the discriminant of $K_p/\mathbf{Q}_p$, which gives

$$\left|\sqrt{d(K)}\right|_p = \left|\sqrt{d(K_p)}\right|_p = |\det(\alpha_i^\sigma)|_p$$

We also have a formula for $\det(\beta_i^\sigma)$ given by

$$\det(\beta_i^\sigma) = \begin{pmatrix} \cdots & \log_p(\epsilon_i^\sigma) & \cdots \\ \cdots & 1 & \cdots \end{pmatrix}$$

Since $\sum_\sigma \log_p(\epsilon^\sigma) = 0$, adding all of the columns onto the last gives $\det(\beta_i^\sigma) = nR_p(K)$. This gives

$$\left|\frac{[K:\mathbf{Q}]R_p(K)}{\sqrt{d(K)}}\right|_p = \left|\frac{\det(\beta_i^\sigma)}{\det(\alpha_i^\sigma)}\right|_P = |\det(a_{ij})|_p \leq 1$$

since all of the a_{ij} are in $\mathbf{Z}_p$. □

Theorem 3.14. *Let $h^+(\mathbf{Q}(\zeta_p))$ be the class number of the totally real field, and $h^-(\mathbf{Q}(\zeta_p))$ be the relative class number (i.e h/h^+). If p divides $h^+(\mathbf{Q}(\zeta_p))$, then p divides $h^-(\mathbf{Q}(\zeta_p))$.*

Proof. The characters corresponding to $\mathbf{Q}(\zeta_p)^+$ are $1, \omega, \dots \omega^{p-3}$. Let $n = \frac{p-1}{2}$. The class number formula (Theorem 3.12) gives

$$\frac{2^{2-1}h^+R_p^+}{\sqrt{d^+}} = \prod_{j=2, j \text{ even}}^{p-3} L_p(1, \omega^j)$$

Since $\mathbf{Q}(\zeta_p)^+$ satisfies the hypotheses of Proposition 3.13, we have $|R_p^+/\sqrt{d^+}| \leq 1$. If $p \nmid h^+$, then $p | L_p(1, \omega^j)$ for some $j = 2, 4, \dots, p-3$. By Corollary 3.7, we have

$$\begin{aligned} 0 \equiv L_p(1, \omega^j) &\equiv L_p(0, \omega^j) (\mathrm{mod}\ p) \\ &= -(1 - \omega^{j-1}(p))B_{1,\omega^{j-1}} (\mathrm{mod}\ p) \\ &= -B_{1,\omega^{j-1}} (\mathrm{mod}\ p) \end{aligned}$$

Since

$$h^- \equiv \prod_{i=1, i \text{ odd}}^{p-4} -\frac{1}{2} B_{1,\omega^i} (\mathrm{mod}\ p)$$

(see the proof of Theorem 3.10) and all the B_{1,ω^i} are p-integral by Corollary 3.9, we have $p | h^-$ as desired. □

Corollary 3.15. *Let h be the class number of $\mathbf{Q}(\zeta_p)$. Then p divides h if and only if p divides the numerator of B_j for some $j = 2, 4, \cdots p-3$.*

Proof. Follows immediately from Theorem 3.10 and the above theorem. □

4 Stickelberger and Herbrand's Theorems

In this section, we state and prove Stickelberger's and Herbrand's theorems, with the goal of proving a slightly more general form of Theorem 2.1. We begin with a brief discussion of Gauss sums.

4.1 Gauss Sums

Gauss sums will be of great importance to us. We now briefly recall some basic facts about them. Let $q = p^r$ be some prime power, and $\mathbf{F} = \mathbf{F}_q$ be the finite field of order q. Let ζ_p be a fixed primitive p-th root of unity, and let $\mathrm{Tr} : \mathbf{F} \to \mathbf{Z}/p\mathbf{Z}$ be the trace. Define a map $\psi : \mathbf{F} \to \mathbf{C}^\times$ by $\psi(x) = \zeta_p^{\mathrm{Tr}(x)}$. This is a well defined, surjective, character of the additive group of $\mathbf{F}$. Let $\chi : \mathbf{F}^\times \to \mathbf{C}^\times$ be a multiplicative character of $\mathbf{F}^\times$, extended to all of $\mathbf{F}$ by setting $\chi(0) = 0$. Since $\chi^{q-1} = 0$ is the trivial character, we note that the order of χ is prime to p.

Definition 4.1. Define the **Gauss sum**

$$g(\chi) := -\sum_{a \in \mathbf{F}} \chi(a)\psi(a)$$

We immedaitely see the following properties of Gauss sums :

Proposition 4.2. *Let g be the Gauss sum as above. Then*

(1) $g(\overline{\chi}) = \chi(-1)\overline{g(\chi)}$.

(2) *If $\chi \neq 1$, then $g(\chi)g(\overline{\chi}) = \chi(-1)q$.*

(3) *If $\chi \neq 1$, then $g(\chi)\overline{g(\chi)} = q$.*

Proof. (1) is immediate from the definitions. (2) follows from (1) and (3). To see (3), note that we have

$$\begin{aligned} g(\chi)\overline{g(\chi)} &= \sum_{a,b \neq 0} \chi(ab^{-1})\psi(a-b) \\ &= \sum_{b,c \neq 0} \chi(c)\psi(bc-b) \\ &= \sum_{b \neq 0} \chi(1)\psi(0) + \sum_{c \neq 0,1} \chi(c) \sum_{b \neq 0} \psi(b(c-1)) \\ &= (q-1) + \sum_{c \neq 0,1} \chi(c)(-1) = q \end{aligned}$$

□

This finishes the proof. It will often be fruitful to consider how Gauss sums interact with compositions of characters. The first such example is the Jacobi sum.

Definition 4.3. Let χ_1 and χ_2 be two multiplicative characters. We define the **Jacobi sum** of χ_1 and χ_2 by

$$J(\chi_1, \chi_2) := -\sum_{a \in \mathbf{F}} \chi_1(a)\chi_2(1-a)$$

Proposition 4.4. *We have the following basic properties of Jacobi sums:*

(1) $J(1,1) = 2 - q$.

(2) *If* $\chi \neq 1$, *then* $J(1,\chi) = J(\chi,1) = 1$.

(3) *If* $\chi \neq 1$, *then* $J(\chi,\overline{\chi}) = \chi(-1)$.

(4) *If* $\chi_1, \chi_2, \chi_1\chi_2 \neq 1$, *then* $J(\chi_1,\chi_2) = g(\chi_1)g(\chi_2)/g(\chi_1\chi_2)$.

Proof. (1) and (2) follow immediately from the definition. We compute

$$\begin{aligned} g(\chi_1)g(\chi_2) &= \sum_{a,b} \chi_1(a)\chi_2(b)\psi(a+b) \\ &= \sum_{a,b} \chi_1(a)\chi_2(b-a)\psi(b) \\ &= \sum_{a,b;b\neq 0} \chi_a(a)\chi_2(b-a)\psi(b) + \sum_a \chi_1(a)\chi_2(-a) \end{aligned}$$

If $\chi_1\chi_2 \neq 1$, then the second sum vanishes. If $\chi_1\chi_2 = 1$, then it is equal to $\chi_1(-1)(q-1)$. Letting $a = bc$, the first sum becomes

$$\sum_{b,c;b\neq 0} \chi_1(b)\chi_2(b)\chi_1(c)\chi_2(1-c)\psi(b) = g(\chi_1,\chi_2)J(\chi_1,\chi_2)$$

If $\chi_1\chi_2 \neq 1$, we get (4). If $\chi_1\chi_2 = 1$, using Proposition 4.2(2) along with $g(1) = 1$ gives (3). This completes the proof. □

Immediately, we have the following corollary:

Corollary 4.5. *If* χ_1, χ_2 *are characters of orders dividing* m, *then*

$$\frac{g(\chi_1)g(\chi_2)}{g(\chi_1\chi_2)}$$

is an algebraic integer in $\mathbf{Q}(\zeta_m)$.

Let m be an integer with $(m,p) = 1$. Then the fields $\mathbf{Q}(\zeta_p)$ and $\mathbf{Q}(\zeta_m)$ are disjoint - we have $\mathbf{Q}(\zeta_p) \cap \mathbf{Q}(\zeta_m) = \mathbf{Q}$. Thus, for $(b,m) = 1$, we define $\sigma_b \in \mathrm{Gal}(\mathbf{Q}(\zeta_m,\zeta_p)/\mathbf{Q})$ to be the element mapping $\zeta_p \mapsto \zeta_p$ and $\zeta_m \mapsto \zeta_m^b$.

Lemma 4.6. *Asusme* χ^m *is trivial. Then*

$$\frac{g(\chi)^n}{g(\chi)^{\sigma_b}} = g(\chi)^{b-\sigma b} \in \mathbf{Q}(\zeta_m)$$

and $g(\chi)^m \in \mathbf{Q}(\zeta_m)$.

Proof. We have

$$g(\chi)^{\sigma_b} = -\sum \chi(a)^b\psi(a) = g(\chi^b)$$

Let $\tau \in \mathrm{Gal}(\mathbf{Q}(\zeta_{mp})/\mathbf{Q}(\zeta))$ be an element such that $\tau(\zeta_p) = \zeta_p^c$ and $\tau(\zeta_m) = \zeta_m$, for some c such that $(c,p) = 1$. Then

$$\begin{aligned} g(\chi)^\tau &= -\sum \chi(a)\psi(ca) \\ &= -\chi(c)^{-1}\sum \chi(a)\psi(a) \\ &= \psi(c)^{-1}g(\chi) \end{aligned}$$

Repeating this calculation, we find $g(\chi^b)^\tau = \chi(c)^{-b}g(\chi^b)$. Thus τ fixes $g(\chi)^{b-\sigma_b}$. This shows the first claim. Taking $b = 1+m$ proves the second claim. □

We conclude this section with one final result.

Lemma 4.7. $g(\chi^p) = g(\chi)$

Proof. We have

$$\begin{aligned} g(\chi^p) &= -\sum \chi(a^p)\zeta_p^{\mathrm{Tr}(a)} \\ &= -\sum \chi(a^p)\zeta_p^{\mathrm{Tr}(a^p)} \\ &= g(\chi) \end{aligned}$$

□

4.2 Stickelberger's Theorem

The goal of this section is to state and prove Stickelberger's Theorem on annihilation of class groups.

Fix some finite abelian extension $M/\mathbf{Q}$, and embed M into $\mathbf{Q}(\zeta_m)$ (by Kronecker-Weber [10, Theorem 14.1]) for minimal m. Then the Galois group $G \coloneqq \mathrm{Gal}(M/\mathbf{Q})$ as a quotient of the multiplicative group $(\mathbf{Z}/p\mathbf{Z})^\times$. Given a such that $(a,m)=1$, let σ_a denote the element $\zeta_m \mapsto \zeta_m^a$ in $\mathrm{Gal}(Q(\zeta_m)/\mathbf{Q})$ and its reduction to M. We let $\{x\}$ denote the *fractional* part of x, so $x - \{x\} \in \mathbf{Z}$ and $0 \le \{x\} < 1$.

Definition 4.8. With notation as above, define the **Strickelberger element** by

$$\theta \coloneqq \theta(M) = \sum_{a(\mathrm{mod}\, m),(a,m)=1} \left\{\frac{a}{m}\right\}\sigma_a^{-1} \in \mathbf{Q}[G]$$

The **Strickelberger ideal** $I(M)$ is defined to be $\mathbf{Z}[G] \cap \theta\mathbf{Z}[G]$, the ideal of $\mathbf{Z}[G]$ mutiples of θ which have integral coefficients.

Lemma 4.9. *Suppose $M = \mathbf{Q}(\zeta_m)$. Let I' be the ideal of $\mathbf{Z}[G]$ generated by elements of the form $c - \sigma_c$, with $(c,m)=1$. Let $\beta \in \mathbf{Z}[G]$. Then $\beta \in I'$ if and only if $\beta\theta \in \mathbf{Z}[G]$. Thus $I = I'\theta$.*

Proof. We have

$$(c-\sigma_c)\theta = \sum_a \left(c\left\{\frac{a}{m}\right\} - \left\{\frac{ac}{m}\right\}\right)\sigma_a^{-1} \in \mathbf{Z}[G]$$

which gives one direction. Conversely, suppose $(\sigma_a x_a \sigma_a)\theta \in \mathbf{Z}[G]$, where $x_a \in \mathbf{Z}$. Then we have that

$$\left(\sum_a x_a\sigma_a\right)\left(\sum_c \left\{\frac{c}{m}\right\}\sigma_c^{-1}\right) = \sum_b \left(\sum_a x_a \left\{\frac{ab}{m}\right\}\right)\sigma_b^{-1}$$

Examining the coefficient of σ^{-1}, we see that $m \mid \sum x_a a$. Thus, since $m = (1+m) - \sigma_{1+m} \in I'$, so is $\sum x_a a$, and

$$\sum x_a\sigma_a = \sum x_a(\sigma_a - a) + \sum x_a a \in I'$$

□

The main goal of this section is to prove Strickelberger's Theorem:

Theorem 4.10 (Stickelberger's Theorem)**.** *Let A be a fractional ideal of M, $\beta \in \mathbf{Z}[G]$, and suppose $\beta\theta \in \mathbf{Z}[G]$. Then $A^{\beta\theta}$ is principal. Thus, the Strickelberger ideal annihilates the ideal class group of M.*

Our general setup for the proof is as follows: Let p be a prime number, and $q = p^f$ a prime power. Let $\mathfrak{p}$ be a prime of $\mathbf{Q}(\zeta_{q-1})$ lying above p. Then we have an isomorphism

$$\omega : \mathbf{F}_q^\times \to \{(q-1)\text{st roots of unity}\}$$

since $\mathbf{Z}[\zeta_{q-1}]$ reduced mod $\mathfrak{p}$ is a finite field of order q and the $(q-1)$st roots of unity are distinct modulo $\mathfrak{p}$. Moreover, this isomorphism satisfies $\omega(a) \bmod \mathfrak{p} \equiv a$.

Now, let $\mathfrak{q}$ be the prime lying above $\mathfrak{p}$ in $\mathbf{Q}(\zeta_{q-1}, \zeta_p)$. For every integer $\alpha \in \mathbf{Z}$, define $s(\alpha) \coloneqq v_\mathfrak{q}(g(\omega^{-a}))$, where g is the Gauss sum defined earlier, and $v_\mathfrak{q}$ is the $\mathfrak{q}$-adic valuation. We now collect some properties of $s(\alpha)$:

Proposition 4.11. (1) $s(0) = 0$.

(2) $0 \le s(\alpha+\beta) \le s(\alpha)+s(\beta)$.

(3) $s(\alpha+\beta) \equiv s(\alpha)+s(\beta) \pmod{p-1}$.

(4) $s(p\alpha) = s(\alpha)$.

(5) $\sum_{\alpha=1}^{q-2} s(\alpha) = f(q-2)(p-1)/2$.

Proof. (1) follows immediately from the definition. (2) follows directly from Corollary 4.5, and (4) follows directly Lemma 4.7. Since $\mathfrak{q}^{p-1} = \mathfrak{p}$, the values of $v_{\mathfrak{q}}$ on $\mathbf{Q}(\zeta_{q-1})$ are all divisible by $p-1$, so (3) also follows from Corollary 4.5. Finally, we have $g(\omega^{-a})g(\omega^a) = \pm q = \pm p^f$, we have $s(\alpha)+s(q-1-\alpha) = v_{\mathfrak{q}}(p^f) = (p-1)f$. Pairing up the terms in the sum gives (5). □

Proposition 4.12. *If $\alpha \not\equiv 0 \pmod{q-1}$, then $s(\alpha) > 0$, and $s(1) = 1$.*

Proof. Let $\pi := \zeta_p - 1$. Then $\pi \in \mathfrak{q}$, so we have

$$g(\omega^{-a}) = -\sum \omega^{-a}(a)\zeta_p^{\mathrm{Tr}(a)} = -\sum \omega^{-a}(a) \equiv 0 \pmod{\mathfrak{q}}$$

and thus $s(\alpha) > 0$. We also have

$$\begin{aligned} g(\omega^{-1}) &= -\sum \omega^{-1}(a)\zeta_p^{\mathrm{Tr}(a)} \\ &= -\sum \omega^{-1}(a)(1+\pi)^{\mathrm{Tr}(a)} \\ &\equiv -\sum \omega^{-1}(a)(1+\pi\,\mathrm{Tr}(a)) \pmod{\mathfrak{q}^2} \\ &\equiv -\pi \sum \omega^{-1}(a)\,\mathrm{Tr}(a) \end{aligned}$$

Since $\mathbf{F}_q \cong \mathbf{Z}[\zeta_{q-1}] \bmod \mathfrak{p}$, we have

$$\mathrm{Tr}(a) = a + a^p + \cdots + a^{p^{f-1}} \pmod{\mathfrak{p}}$$

and

$$\sum \omega^{-1}(a)\,\mathrm{Tr}(a) \equiv \sum_{a \not\equiv 0, a \bmod \mathfrak{p}} a^{-1}(a + a^p + \cdots + a^{p^{f-1}}) \pmod{\mathfrak{p}}$$

Now, if $0 < b < f$, then $\sum_{a \not\equiv 0} a^{p^b-1} \equiv 0 \pmod{\mathfrak{p}}$, so the sum becomes $\sum_{a \not\equiv 0} 1 = q-1 \equiv -1$. This gives $g(\omega^{-1}) \equiv \pi \pmod{\mathfrak{q}^2}$, and thus $s(1) = v_{\mathfrak{q}}(\pi) = 1$ □

Proposition 4.13. *Let $0 \le \alpha < q-1$ and let $\alpha = a_0 + a_1 p + \cdots + a^{f-1}p^{f-1}$ be the standard p-adic expansion of α. Then*

$$s(\alpha) = a_0 + a_1 + \cdots + a_{f-1}$$

Then $s(\alpha) = a_0 + a_1 + \cdots + a_{f-1}$.

Proof. From Proposition 4.11 (1), (2), (3) and Proposition 4.12 we have $s(\alpha) = \alpha$ for $0 \le \alpha \le p-2$. So if $q = p$, there is nothing to prove. Otherwise $s(p-1) > 0$ and thus $s(p-1) = p-1$. From Proposition 4.11 (2) and (4), we have $s(\alpha) \le a_0 + \cdots + a_{f-1}$. For each $0 \le \alpha \le q-1$, each coefficient of the p-adic expansion takes on each of the values from 0 to $p-1$ exactly p^{f-1} times, so

$$\sum_{\alpha=0}^{q-1}(a_0 + \cdots + a_{f-1}) = \frac{fp(p-1)}{2}p^{f-1} = \frac{p-1}{2}fq$$

Omitting $\alpha = q-1$, we have

$$\sum_{\alpha=0}^{q-2}(a_0 + \cdots + a_{f-1}) = \frac{p-1}{2}fq - (p-1)f = \sum_{\alpha=0}^{q-2} s(\alpha)$$

by Proposition 4.11 (5). This completes the proof. □

Lemma 4.14. *Let $0 \leq h \leq q-1$. Then*

$$s(h) = (p-1)\sum_{i=0}^{f-1}\left\{\frac{p^i h}{q-1}\right\}$$

Proof. Let $h = a_0 + a_1 p + \cdots + a_{f-1}p^{f-1}$ be the p-adic expansion of h. Then

$$p^i h \equiv a_0 p^i + a_1 p^{i+1} + \cdots + a_{f-1}p^{-1} (\mathrm{mod}\, q-1)$$

and it follows that

$$\left\{\frac{p^i h}{q-1}\right\} = \frac{1}{q-1}(a_0 p^i + \cdots a_{f-1}p^{i-1})$$

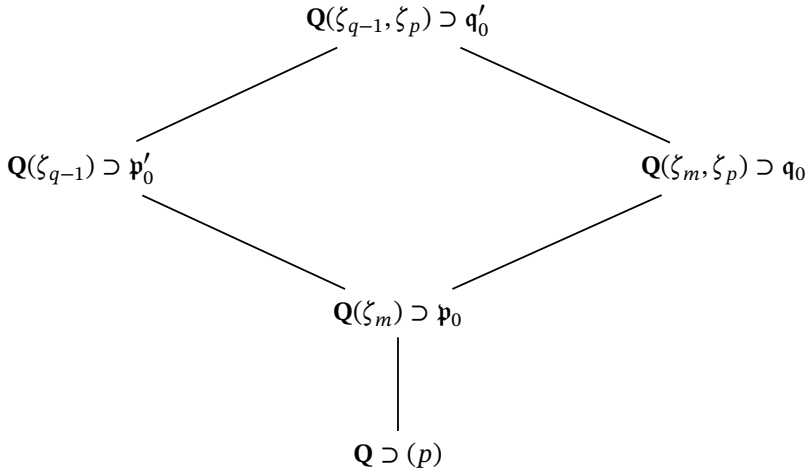
Summing over i, we see that

$$s(h) = (p-1)\sum_{i=0}^{f-1}\left\{\frac{p^i h}{q-1}\right\}$$

as desired. □

We now return to the proof of Strickelberger's theorem.

Proof. Fix some positive integer m. Let p be a prime such that $(p,m) = 1$, and let f be the order of $p (\mathrm{mod}\, m)$. In particular, we have that m divides $p^f - 1 := q-1$. Fix some prime $\mathfrak{p}_0$ of $\mathbf{Q}(\zeta_m)$ lying above p, and let $\mathfrak{q}_0$ be the unique prime of $\mathbf{Q}(\zeta_m, \zeta_p)$, so $\mathfrak{q}_0^{p-1} = \mathfrak{p}_0$. Let $\mathfrak{p}_0'$ be a prime of $\mathbf{Q}(\zeta_{q-1})$ lying above $\mathfrak{p}_0$, and let $\mathfrak{q}_0'$ be the unique prime of $\mathbf{Q}(\zeta_{q-1}, \zeta_p)$ lying above $\mathfrak{p}_0'$ and $\mathfrak{q}_0$, as in the following diagram:



Let $\omega := \omega_{\mathfrak{p}'}$ be as above, and let $\chi = \omega^{-d}$ where $d = (q-1)/m$. Then $\chi^m = 1$ so $g(\chi) \in \mathbf{Q}(\zeta_m, \zeta_p)$. Since $\chi(x)\overline{\chi(x)} = q = p^f$, the factorization of $g(\chi)$ depends only on the primes of $\mathbf{Q}(\zeta_m, \zeta_p)$ above p, which are just conjugates of $\mathfrak{q}_0$ over $\mathbf{Q}$. Let $(a,m) = 1$ and $\sigma_a \in \mathrm{Gal}(\mathbf{Q}(\zeta_m)/\mathbf{Q})$ be the corresponding element of the Galois group. For each such a, fix an extension σ_a to $\mathbf{Q}(\zeta_{q-1}, \zeta_p)$ such that $\zeta_p^{\sigma_a} = \zeta_p$.

The decomposition group for p in $(\mathbf{Z}/m\mathbf{Z})^\times$ is generated by $p (\mathrm{mod}\, m)$ - this follows from the so-called law of cyclotomic reciprocity, see [10, Theorem 2.13]. Let R denote a set of representatives for $(\mathbf{Z}/m\mathbf{Z})^\times$ modulo this decomposition group. Then the set $\{\mathfrak{p}_0^{\sigma_a^{-1}} : a \in R\}$ is the set of conjutgates of $\mathfrak{p}_0$. Since $\mathfrak{q}_0$ is the unique prime above $\mathfrak{p}_0$, all conjugates of $\mathfrak{q}_0$ are also of the form $\mathfrak{q}_0^{\sigma_a^{-1}}$. Let $\mathfrak{q}$ be one such conjugate. Then we have

$$v_{\mathfrak{q}}(g(\chi)) = v_{\mathfrak{q}_0}(g(\chi)^{\sigma_a}) = v_{\mathfrak{q}_0}(g(\chi^a)) = v_{\mathfrak{q}_0'}(g(\chi^a)) = s(ad)$$

where we have used the fact that $v_{\mathfrak{q}_0} = v_{\mathfrak{q}_0'}$, since $\mathfrak{q}_0'/\mathfrak{q}_0$ is unramified. Therefore, we conclude

$$(g(\chi)) = \mathfrak{p}_0^t$$

where $t := \sum_R s(ad)\sigma_a^{-1}$.

By Lemma 4.14, we have $s(ad) = \sum_{i=0}^{f-1}\{p^i a/m\}$, so

$$\sum_R s(ad)\sigma_a^{-1} = (p-1)\sum_{i=0}^{f-1}\sum_R \left\{\frac{p^i a}{m}\right\}\sigma_a^{-1}$$

Since $\mathfrak{q}_0^{p-1} = \mathfrak{p}_0$, and $\sigma_{p^i}(\mathfrak{p}) = \mathfrak{p}$, we have

$$(g(\chi)^m) = \mathfrak{p}_0^{m\sigma\sigma\{p^i a/m\}\sigma_{ap^i}^{-1}} = \mathfrak{p}_0^{m\theta}$$

where

$$\theta := \sum_{(b,m)=1}\{b/m\}\sigma_b^{-1}$$

is the Stickelberger element. This gives a partial result - if $\mathfrak{p}_0$ is a prime of $\mathbf{Q}(\zeta_m)$, with $\mathfrak{p}_0$ not dividing m, then $\mathfrak{p}_0^{m\theta}$ is principal in $\mathbf{Q}(\zeta_m, \zeta_p)$. We now want to work down to $\mathbf{Q}(\zeta_m)$, and eventually to M.

To this end, let A be an ideal of $M \subset \mathbf{Q}(\zeta_m)$, with $(A, m) = 1$. We may factor A into a product of prime ideals $A = \prod \mathfrak{p}_i$ in $\mathbf{Q}(\zeta_m)$. Then we also have a factorization of $A^{m\theta}$ given by

$$A^{m\theta} = \left(\prod g(\chi_{\mathfrak{p}_i})^m\right)$$

Let $G = \mathrm{Gal}(M/\mathbf{Q})$, and suppose $\beta \in \mathbf{Z}[G]$ and $\beta\theta \in \mathbf{Z}[G]$. By extending the elements of G, we may regard $\beta\theta$ as an element of $\mathbf{Z}[G']$, where $G' = \mathrm{Gal}(\zeta_{mp}/\mathbf{Q})$. Then, letting $P = \prod p_i$ of the product of all the primes divisible by the $\mathfrak{p}_i$, and $\gamma := \prod g(\chi_{\mathfrak{p}_i}) \in \mathbf{Q}(\zeta_{Pm})$, we have

$$A^{m\beta\theta} = (\gamma^{\beta m})$$

Since $\gamma^{m\beta} \in \mathbf{Q}(\zeta_m)$ by Lemma 4.6, and it is the mth power of an ideal of $\mathbf{Q}(\zeta_m)$, namely $A^{\beta\theta}$. From this, it follows that the extension $\mathbf{Q}(\zeta_m, \gamma^\beta)/\mathbf{Q}(\zeta_m)$ can only be ramified at primes dividing m. Indeed, locally $A^{\beta\theta}$ is principal, so we are adjoining the mth root of a local unit. But we have the following chain of inclusions:

$$\mathbf{Q}(\zeta_m) \subseteq \mathbf{Q}(\zeta_m, \gamma^\beta) \subseteq \mathbf{Q}(\zeta, \zeta_P) \tag{4.15}$$

So ramification can only occur at the p_i's. Since $(P, m) = 1$, the extension must be unramified.

Before proceeding further, we recall the following useful lemma:

Lemma 4.16. *Given a chain of inclusions $\mathbf{Q}(\zeta_m) \subseteq K \subseteq \mathbf{Q}(\zeta_n)$ with $K/\mathbf{Q}(\zeta_m)$ unramified at all primes, then $K = \mathbf{Q}(\zeta_m)$.*

Proof. Suppose $K \neq \mathbf{Q}(\zeta_m)$. Then there is a character χ for K of conductor not dividing m. By Theorem A.11, $K/\mathbf{Q}(\zeta_m)$ must be ramified at some prime, which is a contradiction. □

Applying the above lemma to the chain of inclusions in 4.15, we find that $\gamma^\beta \in \mathbf{Q}(\zeta_m)$. Therefore $A^{\beta\theta} = (\gamma^\beta)$ is principal as an ideal of $\mathbf{Q}(\zeta_m)$. This however, does not immediately imply that it is principal as an ideal of M. To show this, it suffices to show $\gamma^\beta \in M$, since if two ideals of M are equal in $\mathbf{Q}(\zeta_m)$, they must have been equal originally by unique factorization. To this end, let $\mathfrak{p}'$ be a prime in $\mathbf{Q}(\zeta_{q-1})$ lying over one of the prime factors $\mathfrak{p}_i$ of A. Fix some $\sigma \in \mathrm{Gal}(\mathbf{Q}(\zeta_{q-1})/M)$. Then σ defines an isomorphism

$$\sigma : \mathbf{Z}[\zeta_{q-1}] \bmod \mathfrak{p}' \to \mathbf{Z}[\zeta_{q-1}] \bmod (\mathfrak{p}')^\sigma$$

and so we see that if $\chi_{\mathfrak{p}'}(a) = \zeta$, then $\chi_{(\mathfrak{p}')^\sigma}(a) = \zeta^\sigma$. Therefore $\chi_{\mathfrak{p}'}^\sigma = \chi_{(\mathfrak{p}')^\sigma}$. Since $\chi_{\mathfrak{p}'}^m = 1$, we see $\chi_{\mathfrak{p}'}^\sigma = \chi_{(\mathfrak{p}')^\sigma}$ for all $\sigma \in \mathrm{Gal}(\mathbf{Q}(\zeta_{q-1})/\mathbf{Q}(\zeta_m))$, and $\chi_{\mathfrak{p}'}$ depends only on the $\mathfrak{p}_i$, so we may write $\chi_{\mathfrak{p}_i}$. Then the above argument shows $\chi_{\mathfrak{p}_i}^\sigma = \chi_{\mathfrak{p}_i^\sigma}$ for all $\sigma \in \mathrm{Gal}(\mathbf{Q}(\zeta_m)/M)$. Extending σ by letting $\sigma(\zeta_p) = \zeta_p$, we have $g(\chi_{\mathfrak{p}_i})^\sigma = g(\chi_{\mathfrak{p}_i}^\sigma) = g(\chi_{\mathfrak{p}_i^\sigma})$. Since $A^\sigma = A$ for all $\sigma \in \mathrm{Gal}(\mathbf{Q}(\zeta_m)/M)$, we see that σ permutes the $\mathfrak{p}_i$'s. Therefore, we have

$$\gamma^{\beta\sigma} = \prod g(\chi_{\mathfrak{p}_i})^{\beta\sigma} = \prod g(\chi_{\mathfrak{p}_i^\sigma})^\beta = \gamma^\beta$$

But γ^β is already in $\mathbf{Q}(\zeta_m)$; hence we conclude $\gamma^\beta \in M$. Thus $A^{\beta\theta}$ is principal in M. If A is an arbitrary ideal of M, then we may factor A as $A = (a)A_1$, where $a \in M$ and $(A_1, m) = 1$. Then $A^{\beta\theta} = (a^{\beta\theta})A_1^{\beta\theta}$, which is principal. This completes the proof. □

4.3 Herbrand's Theorem

The goal of this section is to prove Herbrand's theorem. We begin by recalling the definition of orthogonal idempotents.

Definition 4.17. Let G be a finite abelian group, and $\hat{G}$ be its character group. Let $\chi \in \hat{G}$, and define

$$\epsilon_\chi := \frac{1}{|G|} \sum_{\sigma \in G} \chi(\sigma)\sigma^{-1} \in \overline{\mathbf{Q}}[G]$$

where $\overline{\mathbf{Q}}$ is the algebraic closure of $\mathbf{Q}$. The ϵ_χ satisfy:

(1) $\epsilon_\chi^2 = \epsilon_\chi$.

(2) $\epsilon_\chi \epsilon_\psi = 0$ if $\chi \neq \psi$.

(3) $\sum_{\chi \in \hat{G}} \epsilon_\chi = 1$.

(4) $\epsilon_\chi \sigma = \chi(\sigma)\epsilon_\chi$.

The elements ϵ_χ are called the orthogonal idempotents of the group ring $\overline{\mathbf{Q}}[G]$.

If M is a $\overline{\mathbf{Q}}[G]$ module, then we may decompose M as

$$M \cong \bigoplus_\chi M_\chi$$

where $M_\chi = \epsilon_\chi M$. In greater generality, the above construction works when $\overline{\mathbf{Q}}$ is replaced with any (commutative) ring R that contains the values of χ for all $\chi \in \hat{G}$, and in which $|G|$ is invertible. In particular, we will work with the group ring $\mathbf{Z}_p[G]$, where $G = \mathrm{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q}) \cong (\mathbf{Z}/p\mathbf{Z})^\times$, and $\hat{G} = \{\omega^i : 0 \leq i \leq p-2\}$. The orthogonal idempotents are thus

$$\epsilon_i = \frac{1}{p-1} \sum_{a=1}^{p-1} \omega^i(a)\sigma_a^{-1}$$

Let

$$\theta := \frac{1}{p} \sum_{a=1}^{p-1} a\sigma_a^{-1}$$

be the Stickelberger element. Using property (4) of orthogonal idempotents, we find

$$\epsilon_i \theta = \frac{1}{p} \sum_{a=1}^{p-1} a\omega^{-1}(a)\epsilon_i = B_{1,\omega^{-i}}\epsilon_i$$

and

$$\epsilon_i(c - \sigma_c)\theta = (c - \omega^i(c))B_{1,\omega^{-i}}\epsilon_i$$

Now, let A be the p Sylow subgroup of the ideal class group of $\mathbf{Q}(\zeta_p)$. Since $p^n A = 0$ for some sufficiently large enough n, we can view A as a $\mathbf{Z}_p$ module by defining multiplication as

$$\left(\sum_{j=0}^{\infty} b_j p^j \right) a = \sum_{j=0}^{\infty} b_j p^j a$$

since the latter sum is finite. Since G also acts on A, we can regard A as a $\mathbf{Z}_p[G]$ module. Let

$$A = \bigoplus_{i=0}^{p-2} A_i$$

be the decomposition, as in the paragraph following Definition 4.17. By Stickelberger's theorem, (Theorem 4.10), we see that $(c - \sigma_c)\theta$ annihilates A, and hence each A_i. We have shown:

Lemma 4.18. *Let $c \in \mathbf{Z}$ be such that $(c, p) = 1$. Then $(c - \omega^i(c))B_{1,\omega^{-i}}$ annihilates A_i.*

Now, suppose $i \neq 0$ is even. Then $B_{1,\omega^{-i}} = 0$, so the above lemma gives us nothing. If $i = 0$, then $(c-1)/2$ annihilates A_0, so $A_0 = 0$. But this is immediate from the fact that $\epsilon_0 = (\text{Norm})/(p-1)$. Therefore, our main interest will be in the case when i is odd. We first consider the case where $i = 1$. Let $c = 1 + p$. Then we have

$$\begin{aligned}(c - \omega(c))B_{1,\omega^{-1}} &= pB_{1,\omega^{-1}} \\ &= \sum_{a=1}^{p-1} a\omega^{-1}(a) \\ &\equiv p - 1 (\text{mod } p)\end{aligned}$$

In particular, $p - 1 \not\equiv 0 (\text{mod } p)$, so since A_1 is a p-group, we must have $A_1 = 0$. If $i \neq 1$ is odd, we may choose an integer c such that $c \not\equiv c^i \equiv \omega^i(c) \bmod p$, and may consequently ignorethe factor $c - \omega^i(c)$. This gives the following result:

Proposition 4.19. *$A_0 = A_1 = 0$. For $i = 3, 5, \dots, p-2$, $B_{1,\omega^{-i}}$ annihilates A_i.*

In fact, we can get something even stronger - Herbrand's theorem, which was promised at the beginning of this section.

Theorem 4.20. *Let i be odd, with $3 \leq i \leq p-2$. If $A_i \neq 0$, then $p | B_{p-i}$.*

Proof. Suppose $A_i \neq 0$. Then we must have $B_{1,\omega^{-i}} \equiv 0 (\text{mod } p)$. But by Corollary Corollary 3.9, we know

$$B_{1,\omega^{-i}} \equiv \frac{B_{p-i}}{p-i} (\text{mod } p)$$

So $p | B_{p-i}$. This completes the proof. □

4.4 Fermat's Last Theorem, Revisited

Using the machinery developed in this chapter, we prove a slight strengthening of Theorem 2.1.

Theorem 4.21. *Suppose p is a prime, and the index of regularity (= the number of Bernoulli numbers divisible by p) satisfies $i(p) < \sqrt{p} - 2$. Then $x^p + y^p = z^p, (xyz, p) = 1$ has no integer solutions.*

Proof. Let $\zeta := \zeta_p$. As in the proof of Theorem 2.1, we assume a solution exists and obtain

$$(x + \zeta^i y) = C_i^p$$

for $0 \leq i \leq p-1$, and C_i an ideal of $\mathbf{Q}(\zeta_p)$. Let C be the subgroup of the ideal class group generated by $C_1, \dots, C_{p-1}$. Then C is an elementary p-group, so $\mathbf{Z}_p[G]$ acts naturally on C, where $G = \text{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q})$. Since $C_1^{\sigma_a} = C_a$, we see that C_1 generates C over the group ring $\mathbf{Z}_p[G]$, and obtain a decomposition

$$C = \bigoplus_i \langle \epsilon_i C_1 \rangle$$

, where ϵ_i is the orthogonal idempotent, and $\langle x \rangle$ denotes the cyclic subgroup generated by x. Define C^- by

$$C^- := \bigoplus_{i \text{ odd}} \langle \epsilon_i C_1 \rangle$$

Therefore, we have

$$\begin{aligned}p - \text{rank} C^i &= \#\{\epsilon_i C_1 \neq 0, i \text{ odd}\} \\ &\leq \#\{A_i \neq 0, i \text{ odd}\} \\ &\leq i(p) \text{ (by Herbrand's theorem)} \\ &\leq \sqrt{p} - 2 \text{ (by assumption)}\end{aligned}$$

We may assume $p > 3$. Let $r := [\sqrt{p}] - 1$. Consider the set of all products $C_1^{b_1} \cdots C_r^{b^r}$ with $0 \leq b_i \leq p-1$. The number of such products is $p^r > p^{\text{rank}C^-} = |C^-|$. Therefore, at least two must agree on their C^- components, so we divide to obtain

$$\prod_{i=1}^{r} C_i^{a_i} \in C^+$$

with $-p < a_i < p$ and some a_i nonzero. Therefore, we may write

$$\prod C_i^a = (\rho)S$$

with $\rho \in \mathbf{Q}(\zeta_p)$ and S an ideal satisfying $\bar{S} = S$. This gives

$$\left(\prod_{i=1}^{r}(x+\zeta^i y)^{a_i}\right) = (\rho^p)S^p$$

Since all the C_i's are prime to p, we may assume ρ and S are also prime to p. This implies S^p is principal in $\mathbf{Q}(\zeta_p)$. By Theorem B.17, the class group of $\mathbf{Q}(\zeta_p)^+$ injects into the class group of $\mathbf{Q}(\zeta_p)$, and thus S^p is principal in $\mathbf{Q}(\zeta_p)^+$. We conclude $S^p = (\alpha)$ for some α satisfying $\alpha = \bar{\alpha}$. Since any unit of $\mathbf{Q}(\zeta_p)$ is a root of unity times a real unit, we obtain

$$\prod_{i=1}^{r}(x+\zeta^i y)^{a_i} = \zeta^\mu \epsilon \alpha \rho^p$$

where $\mu \in \mathbf{Z}$ is an integer and ϵ is a real unit. This immedaitely gives

$$\prod_{i=1}^{r}(x+\zeta^{-i} y)^{a_i} = \zeta^{-\mu} \epsilon \alpha \bar{\rho}^p$$

By Lemma 2.8 we see $\rho^p \equiv \bar{\rho}^p \equiv$ rational integer (mod p). This gives

$$\prod_{i=1}^{r}\left(\frac{x+\zeta^i y}{x+\zeta^{-i} y}\right)^{a_i} \equiv \zeta^{2\mu} (\text{mod } p) \tag{4.22}$$

and

$$\prod_{i=1}^{r}\left(\frac{x+\zeta^i y}{y+\zeta^{i} x}\right)^{a_i} \equiv \zeta^{\nu} (\text{mod } p) \tag{4.23}$$

where $\nu := \sum_i i a_i (\text{mod } p)$. Define x_i and y_i by

$$x_i = \begin{cases} y & a_i < 0 \\ x & a_i \geq 0 \end{cases}$$

and

$$y_i = \begin{cases} x & a_i < 0 \\ y & a_i \geq 0 \end{cases}$$

Define polynomials F and G by $F(T) = \prod_i (x_i + T^i y_i)^{|a_i|}$ and $G(T) = \prod_i (y_i + T^i x_i)^{|a_i|}$. By construction, F and G yield the numerator and denominator respectively of 4.23, so we have

$$F(\zeta) \equiv \zeta^\nu G(\zeta) (\text{mod } p)$$

It follows that we may write $F(\zeta) = \zeta^\nu G(\zeta) + pK(\zeta)$ for some $K(T) \in \mathbf{Z}[T]$, and more generally, we can write $F(T) = T^\nu G(T) + pK(T) + (1 + T + \cdots + T^{p-1})H(T)$ for some $H(T) \in \mathbf{Q}[T]$. But since all of the coefficients are integral, we see that $H(T) \in \mathbf{Z}[T]$. The following process is meant to replicate the process of taking a logarithmic

derivative, formally. First, we multiply by $(1-T)$, then differentiate with respect to T, set $T=\zeta$, and reduce mod p. This gives

$$(1-\zeta)F'(\zeta)-F(\zeta)\equiv(1-\zeta)\zeta^{\nu}G'(\zeta)-\zeta^{\nu}G(\zeta)+\nu(1-\zeta)\zeta^{\nu-1}G(\zeta)\bmod p$$

Dividing by $F(\zeta)\equiv\zeta^{\nu}(\bmod p)$, we find

$$(1-\zeta)\frac{F'(\zeta)}{F(\zeta)}-1\equiv(1-\zeta)\frac{G'(\zeta)}{G(\zeta)}-1+\nu(1-\zeta)\zeta^{-1}$$

As mentioned previously, this is precisely what we would have gotten if we could take the logarithmic derivative of $F(\zeta)\equiv\zeta^{\nu}G(\zeta)$ with respect to ζ. We may rewrite the above as

$$(1-\zeta)\sum_{i=1}^{r}ia_i\zeta^i\left(\frac{y}{x+\zeta^i y}-\frac{x}{y+\zeta^i x}\right)\equiv\nu(1-\zeta)\bmod p$$

To conclude the proof, multiply by $\prod_{i=1}^{r}(x+\zeta^i y)(y+\zeta^i x)$, which is a polynomial of degree r^2+r in ζ. Let i_0 be the index of the first non-zero a_i. The left side becomes a polynomial in ζ of degree $1+i_0+r^2+r-2i_0=r^2+r-i_0+1$, with leading coefficient $i_0a_{i_0}(x^2-y^2)x^ry^r$. The right hand side becomes a polynomial in ζ of degree $1+r^2+r$ with leading coefficient $-x^ry^r\nu$. Note that we have

$$1+r^2+r<1+(\sqrt{p}-1)\sqrt{p}<p-1$$

so the right hand side becomes a polynomial of degree less than $p-1$. By Lemma 2.9, the corresponding coefficients are congruent mod p. Therefore $\nu\equiv 0(\bmod p)$, so the right hand side vanishes. The leading coefficient of the left hand side must also vanish mod p, so $x^2\equiv y^2$, and $x\equiv\pm y(\bmod p)$. We may repeat the same argument replacing y with z to obtain $x\equiv\pm z(\bmod p)$, which gives

$$\pm x^p\pm x^p\equiv\pm x^p(\bmod p)$$

which is impossible for $p>3$. This completes the proof. □

A Dirichlet Characters

In this appendix we briefly recall some facts about Dirichlet characters.

Definition A.1. A **Dirichlet character** is a multiplicative homomorphism $\chi:(\mathbf{Z}/n\mathbf{Z})^{\times}\to\mathbf{C}^{\times}$. If $n|m$, then this induces a multiplicative homomorphism $(\mathbf{Z}/m\mathbf{Z})^{\times}\to\mathbf{C}^{\times}$ by composition with the canonical map $(\mathbf{Z}/m\mathbf{Z})^{\times}\to(\mathbf{Z}/n\mathbf{Z})^{\times}$. Therefore, we can regard χ as being defined mod m or mod n. We will choose n minimal, and call it the **conductor** of χ, denoted f_χ or f.
Note that $\chi(-1)=\pm1$. If $\chi(-1)=1$, we say χ is **odd**, and if $\chi(-1)=1$, we say χ is even.

Example A.2. Let $\chi:(\mathbf{Z}/8\mathbf{Z})^{\times}\to\mathbf{C}^{\times}$ be defined by $\chi(1)=1$, $\chi(3)=-1$, $\chi(5)=1$ and $\chi(7)=-1$. Then $\chi(a+4)=\chi(a)$, so χ can be defined mod 4 by $\chi(1)=1$ and $\chi(3)=-1$. Since 4 is minimal, we have $f_\chi=4$. This is an odd character.

Given a Dirichlet character χ, we often consider χ as a homomorphism $\mathbf{Z}\to\mathbf{C}$ by defining $\chi(a)=0$ when $(a,f_\chi)=1$. When this is the case, to avoid ambiguity, we always assume χ is defined modulo its conductor. We call such characters **primitive**. Taking this convention has many advantages - first, χ is now periodic of period f_χ, and secondly, taking this convention makes it so that $\chi(a)=0$ happens as little as possible.

However, in the following discussion, when speaking of characters of $(\mathbf{Z}/n\mathbf{Z})^{\times}$, we include characters of conductor dividing n, including the trivial character of conductor 1.

Given two characters χ,ψ of conductors f_χ and f_ψ respectively, we define the product $\chi\psi$ as follows: Consider the homomorphism

$$\gamma:(\mathbf{Z}/\mathrm{lcm}(f_\chi,f_\psi)\mathbf{Z})^{\times}\to\mathbf{C}^{\times}$$

given by $\gamma(a)=\chi(a)\psi(a)$. We take $\chi\psi$ to be the primitive character associated to γ.

Example A.3. Define $\chi \bmod 12$ by $\chi(1) = 1, \chi(5) = -1, \chi(7) = -1$ and $\chi(11) = 1$, and $\psi \bmod 3$ by $\psi(1) = 1, \psi(2) = -1$. Then the character $\chi\psi$ on $(\mathbf{Z}/12\mathbf{Z})^\times$ has values $\chi\psi(1) = 1, \chi\psi(5) = 1, \chi\psi(7) = -1$, and $\chi\psi(11) = -1$. Therefore $\chi\psi$ has conductor 4 and satisfies $\chi\psi = 1$ and $\chi\psi(3) = -1$.

Example A.4. Let χ be any character, and $\psi := \bar{\chi}$ the the complex conjugate. Then $\psi(a) = \chi(a)^{-1}$ if $(a, f_\chi) = 1$, so we see that $\chi\bar{\chi} = 1$, the trivial character.

For any cyclotomic field $\mathbf{Q}(\zeta_n)$, we have $\mathrm{Gal}(\mathbf{Q}(\zeta_n)/\mathbf{Q}) \cong (\mathbf{Z}/n\mathbf{Z})^\times$. Making this identification, a Dirichlet character mod n is called a **Galois character**. Using Galois characters, we may re-interpret Example A.2 above as follows:

Example A.5. Let χ be as in Example A.2. The kernel of χ is $1 (\mathrm{mod}\, 8)$ and $5 (\mathrm{mod}\, 8)$. In the Galois group, these form $\mathrm{Gal}(\mathbf{Q}(\zeta_8)/\mathbf{Q}(\zeta_4))$, so we find χ is a character of the quotient of $\mathrm{Gal}(\mathbf{Q}(\zeta_8))$ by this subgroup, which of course is just $\mathrm{Gal}(\mathbf{Q}(\zeta_4)/\mathbf{Q}) \cong (\mathbf{Z}/4\mathbf{Z})^\times$.

In general, let χ be a character mod n, interpreted like above as a character of $\mathrm{Gal}(\mathbf{Q}(\zeta_n)/\mathbf{Q})$. Let K be the fixed field of the kernel of χ. Then $K \subseteq \mathbf{Q}(\zeta_n)$, and if n is minimal we have $n = f_\chi$. The field K depends only on χ, and we refer to it as the field **belonging to** χ. Generalizing this, let X be a finite group of Dirichlet characters, and let n be the least common multiple of the conductors of the characters in X. It follows that X is a subgroup of the characters of $\mathrm{Gal}(\mathbf{Q}(\zeta_n)/\mathbf{Q})$. Let H be the intersection of the kernels of these characters, and K the fixed field of H. Then X can be identified with $\mathrm{Hom}(\mathrm{Gal}(K/\mathbf{Q}), \mathbf{C}^\times)$. The field K is also called the field belonging to X, and $\deg(K/\mathbf{Q}) = |X|$.

Example A.6. If X is the group of characters of $(\mathbf{Z}/n\mathbf{Z})^\times$ satisfying $\chi(-1) = 1$, then complex conjugation is in the kernel of each $\chi \in X$. The field associated to X is the maximal real subfield $\mathbf{Q}(\zeta_n + \zeta_n^{-1})$ of $QQ(\zeta_n)$.

Before proceeding further, we recall some basic theory of characters of arbitrary finite groups. Let G be a finite group, and let $\hat{G}$ denote the group of multiplicative homomorphisms $G \to \mathbf{C}^\times$.

Theorem A.7. *If G is a finite abelian group, then $G \cong \hat{G}$. This isomorphism need not be natural.*

Proof. By the structure theorem for finite abelian groups, we may write G as the direct sum of groups of the form $\mathbf{Z}/m\mathbf{Z}$. Therefore, it suffices to show the theorem for $\mathbf{Z}/m\mathbf{Z}$. But if χ is a character of $\mathbf{Z}/m\mathbf{Z}$, it is uniquely determined by $\chi(1)$. Since $\chi(1)$ can be any mth root of unity, the theorem holds for $\mathbf{Z}/m\mathbf{Z}$, and hence for G. □

Corollary A.8. *$\hat{\hat{G}} \cong G$ for a finite group G. This isomorphism is canonical.*

Proof. Let $g \in G$. Then g defines a map $\hat{G} \to \mathbf{C}^\times$ given by $\chi \mapsto \chi(g)$. Suppose $\chi(g) = 1$ for all $\chi \in \hat{G}$. Let $H \subseteq G$ be the subgroup generated by G. Then $\hat{G}$ acts as a set of distinct characters on G/H. By the above theorem, there are at most $\#(G/H)$ such characters. Therefore $H = 1$ is the trivial group, so $g = 1$. We see that G injects into $\hat{\hat{G}}$, and since $\#G = \#\hat{G} = \#\hat{\hat{G}}$, the result follows. □

Since the above isomorphism is natural, we often equate $G = \hat{\hat{G}}$. This gives a nondegenerate pairing $G \times \hat{G} \to \mathbf{C}^\times$ given by sending the pair $(g, \chi) \mapsto \chi(g)$.
Now, let $H \subset G$ be a subgroup. Define

$$H^\perp := \{\chi \in \hat{G} : \chi(h) = 1 \text{ for all } h \in H\}$$

There is a natural isomorphism $H^\perp \cong \widehat{G/H}$

Proposition A.9. *We have an isomorphism $\hat{H} \cong \hat{G}/H^\perp$.*

Proof. Restriction gives a map $\hat{G} \to \hat{G}$, with kernel precisely $H^\perp$. So it remains to show surjectivity. But we have $\#(H^\times) = \#(\widehat{G/H}) = \#(G/H) = \#(G)/\#(H)$. Thus $\#(\hat{H}) = \#(H) = \#(G)/\#(H^\perp) = \#(\hat{G})/\#(H^\perp)$. This shows surjectivity, and the proposition follows. □

Proposition A.10. *Equating $\hat{\hat{G}} = G$, we have $(H^\perp)^\perp = H$.*

Proof. Both groups have the same order. If $h \in H$, then $h : \chi \mapsto \chi(h)$ maps $H^\perp$ to 1. Therefore $H \subset (H^\perp)^\perp$, and the groups are equal. □

Returning to Dirichlet characters, let X be the group of Dirichlet characters associated to a Galois number field K. This gives a pairing $\mathrm{Gal}(K/\mathbf{Q}) \times X \to \mathbf{C}^\times$. Let $L \subset K$ be a subfield, and let

$$Y = \{\chi \in X : \chi(g) = 1 \text{ for all } g \in \mathrm{Gal}(K/L)\}$$

Then we have

$$Y = \mathrm{Gal}(K/L)^\perp = (\mathrm{Gal}(K/\mathbf{Q})/\mathrm{Gal}(K/L))^\times = \mathrm{Gal}(\hat{L}/\mathbf{Q})$$

Conversely, if we start with a subgroup $Y \subset X$ and let L be the fixed field of $Y^\perp$. Then by Galois theory, we see $Y^\perp = \mathrm{Gal}(K/L)$. It follows that there is a bijective correspondence between subgroups of X and subfields of K given by

$$\begin{aligned} \mathrm{Gal}(K/L)^\times &\leftrightarrow L \\ Y &\leftrightarrow \text{ fixed field of } Y^\perp \end{aligned}$$

We conclude this section by showing how ramification may be detected using Dirichlet characters. Let $n = \prod p^a$. Then, corresponding to the decomposition

$$(\mathbf{Z}/n\mathbf{Z})^\times \cong \prod (\mathbf{Z}/p^a\mathbf{Z})^\times$$

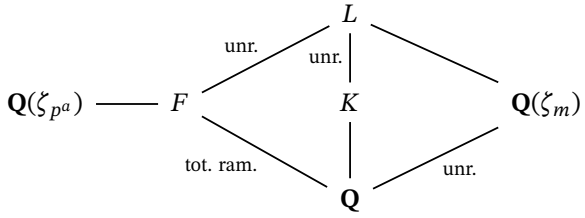
we may decompose any Dirichlet character defined mod n as

$$\chi = \prod \chi_p$$

where χ_p is a character defined mod p^a. If X is a group of Dirichlet characters, we define $X_p = \{\chi_p : \chi \in X\}$. For example, in Example A.3, we may write the character $\chi = \chi_2\chi_3$, where χ_2 is the character $\chi\psi$ of conductor 4 from that example, and $\chi_3 = \psi$.

Theorem A.11. *Let X be a group of Dirichlet characters and K the associated field. Let p be a prime number with ramification number e in K. Then $e = \#(X_p)$.*

Proof. Let n be the least common multiple of the conductors of the characters in X, so $K \subset \mathbf{Q}(\zeta_n)$. Let $n = p^a m$ with $p \nmid m$. Let $L = K(\zeta_m) = K \cdot \mathbf{Q}(\zeta_m)$ be the composite field, as below. Then the group of characters of L is generated by X and the characters of $(\mathbf{Z}/n\mathbf{Z})^\times$ with conductor prime to p, so it is the product of X_p and the characters of $\mathbf{Q}(\zeta_m)$. Thus L is the composite of $\mathbf{Q}(\zeta_m)$ and the field $F \subset \mathbf{Q}(\zeta_{p^a})$ belonging to X_p. Since p is unramified in $\mathbf{Q}(\zeta_m)$, the ramification index of p in K is the same as the ramification index of p in L. Since L/F is unramified for p, the ramification index is the same as that for F, which is just $\deg(F/\mathbf{Q}) = \#(X_p)$. This completes the proof.



□

A useful application of Dirichlet characters is the so-called conductor-discriminant formula:

Theorem A.12 (Conductor-Discriminant Formula)**.** *Let X be a group of Dirichlet characters, and K the associated number field. Then the discriminant of K can be computed by*

$$d(K) = (-1)^{r_2} \prod_{\chi \in X} f_\chi$$

B Dirichlet L-Series

In this section, we introduce Dirichlet L-Series, and recall some basic results about them. We omit most proofs, especially those which are mostly analytic in nature. We will freqeuently refer to Dirichlet *characters*, of which we present a brief intoduction in Appendix A.

Definition B.1. Let χ be a Dirichlet character of conductor f. The L-series attached to χ is defined by

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

for $\mathrm{Re}(s) > 1$.

We will want to give an explicit formula for $L(1-n,\chi)$. To do this, we need the generalized Bernoulli numbers. First, we recall the definition of the ordinary Bernoulli numbers.

Definition B.2. The **Bernoulli numbers** B_n are defined by

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!}$$

The first few Bernoulli numbers are $B_0 = 1, B_1 = -1/2, B_2 = 1/6, B_3 = 0, \ldots$. In fact, $B_{2k+1} = 0$ for $k \geq 1$.

Definition B.3. Let χ be a Dirichlet character of conductor f. Define the **generalized Bernoulli numbers** by

$$\sum_{a=1}^{f} \frac{\chi(a) t e^{at}}{e^{ft} - 1} = \sum_{n=0}^{\infty} B_{n,\chi} \frac{t^n}{n!}$$

When $\chi = 1$, we have

$$\sum_{n=0}^{\infty} B_{n,1} \frac{t^n}{n!} = \frac{te^t}{e^t - 1} = \frac{t}{e^t - 1} + t$$

Therefore, $B_{n,1} = B_n$, unless when $n = 1$, in this case we have $B_{1,1} = 1/2$ and $B_1 = -1/2$. Finally, if $\chi \neq 1$, then $B_{0,\chi} = 0$, since $\sum_{a=1}^{f} \chi(a) = 0$.

We will also need the Bernoulli polynomials.

Definition B.4. For a non-negative integer n. The n-th Bernoulli polynomial is defined by

$$\frac{te^{Xt}}{e^t - 1} = \sum_{n=1}^{\infty} B_n(X) \frac{t^n}{n!}$$

We immediately see that $B_n(1-X) = (-1)^n B_n(X)$. Since the generating function is the product of $\sum B_n t^n/n!$ and $\sum X^n t^n/n!$, we find

$$B_n(X) = \sum_{i=0}^{n} \binom{n}{i} B_i X^{n-1}$$

Proposition B.5. *Let F be any multiple of f (which is the conductor of χ). Then*

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

Proof. We have

$$\sum_{n=0}^{\infty} F^{n-1} \sum_{a=1}^{F} \chi(a) B_n\left(\frac{a}{F}\right) \frac{t^n}{n!} = \sum_{a=1}^{F} \chi(a) \frac{te^{(a/F)Ft}}{e^{Ft} - 1}$$

Let $g = F/f$ and $a = b + cf$. Then we have

$$\sum_{b=1}^{f}\sum_{c=0}^{g-1}\chi(b)\frac{te^{(b+cf)t}}{e^{fgt}-1} = \sum_{b=1}^{f}\chi(b)\frac{te^{bt}}{e^{ft}-1} = \sum_{n=0}^{\infty}B_{n,\chi}\frac{t^n}{n!}$$

The result follows. □

In particular, since $B_1(X) = X - \frac{1}{2}$, we have $B_{1,\chi} = \frac{1}{f}\sum_{a=1}^{f}\chi(a)a$ whenever $\chi \neq 1$. It is no coincidence that we introduced L-functions and generalized Bernoulli numbers immediately after each other.

Theorem B.6. *$L(1-n,\chi) = -B_{n,\chi}/n$ for $n \geq 1$.*

Proof. See [10, Theorem 4.2] □

The value at $s = 1$ is of particular interest to us. Our next goal is to show that $L(1,\chi) \neq 0$. We will derive this fact from the following theorem, which we will not prove.

Theorem B.7. *Let X be a group of Dirichlet characters, K the associated field, and $\zeta_K(s)$ the Dedekind zeta function of K. Then*

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

Corollary B.8. *$L(1,\chi) \neq 0$.*

Proof. Let K be the field belonging to χ. The Dedekind zeta function $\zeta_K(s)$ has a simple pole at $s = 1$. Let b be the order of χ. Then

$$\zeta_K(s) = \prod_{a=0}^{b-1} L(x,\chi^a) = \zeta(s)\prod_{a=1}^{b-1} L(s,\chi^a)$$

Since $\chi(s)$ only has a simple pole at $s = 1$, none of the factors $L(s,\chi^a)$ can vanish at $s = 1$. This completes the proof. □

Our next goal is to evaluate $L(1,\chi)$. When χ is odd, this is easily accomplished via the functional equation

$$L(1,\chi) = \frac{\tau(\chi)}{2i}\frac{2i}{f}L(0,\bar{\chi}) = \frac{\pi i\tau(x)}{f}B_{1,\chi}$$

where $\tau(\chi) = \sum_{a=1}^{f}\chi(a)e^{2\pi ia/f}$ and f is the conductor of χ. The even case requires a bit more work. To proceed, we need a few lemmas.

Lemma B.9. *For every integer b,*

$$\sum_{a=1}^{f}\bar{\chi}(a)e^{2\pi iab/f} = \chi(b)\tau(\bar{\chi})$$

In particular,

$$\overline{\tau(\chi)} = \chi(-1)\tau(\bar{\chi})$$

Proof. If $(b,f) = 1$, then change variables so that $c \equiv ab \pmod f$. Since everything only depends on the residue classes mod f, the result is immediate in this special case. If $(b,f) = d > 1$, we claim the result is true, since both sides vanish. The right hand side is clearly zero. We now show the left side is also zero. Note that if $\chi(y) = 1$ for all $y \equiv 1 \pmod{f/d}$ with $(y,f) = 1$, then χ would be defined mod f/d, and could not have conductor f. Therefore, there is some $y \equiv 1 \pmod{f/d}$ with $(f,y) = 1$ such that $\chi(y) \neq 1$. Since $dy \equiv d \pmod f$, we have $by \equiv b \pmod f$, and

$$\sum_{a=1}^{f}\chi(\bar{a})e^{2\pi iab/f} = \sum_{a=1}^{f}\chi(\bar{a})e^{2\pi aby/f} = \chi(y)\sum_{a=1}^{f}\bar{\chi}(a)e^{2\pi iab/f}$$

Since $\chi(y) \neq 1$, the sum is zero. This shows the first assertion. The second assertion follows from the first by taking $b = -1$. □

Lemma B.10. $|\tau(\chi)| = \sqrt{f_\chi}$.

Proof. Let ϕ denote Euler's ϕ function. Then we have

$$\begin{aligned}
\phi(f)|\tau(\chi)|^2 &= \sum_{b=1}^{f} |\chi(b)\tau(\chi)|^2 \\
&= \sum_{b=1}^{f}\sum_{a=1}^{f} \chi(a)e^{2\pi iab/f} \sum_{c=1}^{f} \bar{\chi}(c)e^{-2\pi ibc/f} \\
&= \sum_{a}\sum_{c} \chi(a)\bar{\chi}(c) \sum_{b} e^{2\pi ib(a-c)/f} \\
&= \sum_{a} \chi(a)\bar{\chi}(a)f \\
&= f\phi(f)
\end{aligned}$$

The second equality holds from Lemma B.9, and the fourth equality holds because the sum over b is zero unless $a = c$. The final equality holds since $\chi(a)\bar{\chi}(a) = 1$ if $(a, f) = 1$, and is 0 otherwise. This completes the proof. □

We now compute $L(1, \chi)$.

$$\begin{aligned}
L(1,\chi) &= \sum_{n=1}^{\infty} \frac{\chi(n)}{n} = \sum_{n=1}^{\infty} \frac{1}{n} \frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{f} \chi(a)e^{2\pi ian/f} \\
&= \frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{f} \bar{\chi}(a) \sum_{n=1}^{\infty} \frac{1}{n} e^{2\pi ian/f} \\
&= -\frac{1}{\tau(\bar{\chi})} \sum_{a=1}^{f} \bar{\chi}(a) \log(1-\zeta_f^a)
\end{aligned}$$

where $\zeta_f = e^{2\pi i/f}$. Since $\tau(\bar{\chi}) = \chi(-1)\overline{\tau(\chi)} = \chi(-1)f/\tau(\chi)$, we have

$$L(1,\chi) = -\frac{\chi(-1)\tau(\chi)}{f} \sum_{a=1}^{f} \bar{\chi}(a) \log(1-\zeta_a^f)$$

Note that $\log(1-\zeta_f^a) + \log(1-\zeta_f^{-a}) = 2\log|1-\zeta_a^f|$. Therefore, if χ is even, and $\chi(a) = \chi(-a)$, we have

$$L_{1,\chi} = -\frac{\tau(\chi)}{f} \sum_{a=1}^{f} \bar{\chi}(a) \log|1-\zeta_a^f|$$

We have thus shown the following:

Theorem B.11.

$$L_{1,\chi} = \begin{cases} \pi i \frac{\tau(\chi)}{f} B_{1,\bar{\chi}} & \text{if } \chi(-1) = -1 \\ -\frac{\tau(\chi)}{f} \sum_{a=1}^{f} \bar{\chi}(a) \log|1-\zeta_f^a| & \text{if } \chi(-1) = 1 \text{ and } \chi \neq 1 \end{cases}$$

We now shift gears to study class groups of certain types of number fields.

Definition B.12. Let K be a number field. We say K is **totally real** if all complex embeddings $K \to \mathbf{C}$ lie in $\mathbf{R}$, and **totally complex** if none of its embeddings lie in $\mathbf{R}$. A **CM-field** is a totally imaginary quadratic extension of a totally real number field.

The cyclotomic fields $\mathbf{Q}(\zeta_n)$ are all CM-fields. Their maximal real subfields are of the form $\mathbf{Q}(\zeta_n + \zeta_n^{-1})$.

Theorem B.13. *Let K be a CM-field, and K^+ be its maximal real subfield. Let h and h^+ be the respective class numbers. Then $h^+|h$. The quotient h/h^+ is called the* ***relative class number****.*

To prove this, we need the following result from class field theory:

Proposition B.14. *Let K/L be a finite extension of number fields such that there is no nontrivial unramified subextension F/L with $\mathrm{Gal}(F/L)$ abelian. Then the class number of L divides the class number of K.*

Proof. Let H be the Hilbert class field of L, that is H is the maximal unramified abelian extension of L. By class field theory, the Galois group $\mathrm{Gal}(H/L)$ is isomorphic to the class group of L. Our assumptions on the extension K/L imply that $H \cap K = L$, and hence $[KH : K] = [H : L]$. But KH/K is unramified and abelian, so it is contained in the maximal unramified abelian extension of K. Therefore, the class number of L, which is equal to $[H : L] = [KH : K]$ divides the class number of K. □

We now prove Theorem B.13.

Proof. Since K^+/K is totally ramified at the archimidean primes, the above proposition applies. This completes the proof. □

Theorem B.15. *Let K be a CM-field, and let E be its unit group. Let E^+ be the unit group of K^+, and let W be the group of roots of unity in K. Define $Q := [E : WE^+]$. Then $Q = 1$ or $Q = 2$.*

Proof. Define $\phi : E \to W$ by $\phi(\epsilon) = \epsilon/\bar{\epsilon}$. Since K is CM, $\bar{\epsilon^\sigma} = (\bar{\epsilon})^\sigma$ for all embeddings σ, and we have $|\phi(\epsilon)^\sigma| = 1$. Therefore $\phi(\epsilon) \in W$. Let $\psi : E \to W/W^2$ be the map induced by ϕ. Suppose $\epsilon = \zeta\epsilon_1$, where $\zeta \in W$ and $\epsilon_1 \in E^+$. Then $\phi(E) = \zeta^2 \in W^2$, so $\epsilon \in \ker\psi$. Conversely, if $\phi(\epsilon) = \zeta^2 \in W^2$, then $\epsilon_1 = \zeta^{-1}\epsilon$ is real, and we see that $\ker\psi = WE^+$. Since $|W/W^2| = 2$, we are done. In particular, if $\phi(E) = W$, then $Q = 2$, and if $\phi(E) = W^2$, then $Q = 1$. □

Corollary B.16. *Let $K = \mathbf{Q}(\zeta_n)$. Then $Q = 1$ if n is a prime power, and $Q = 2$ otherwise.*

Proof. Suppose ϵ is a unit in $\mathbf{Q}(\zeta_{2m})$ such that $\epsilon/\bar{\epsilon} \notin W^2$. Then $\epsilon/\bar{\epsilon} := \zeta$ is a primitive 2^mth root of unity. Let Norm denote the norm map from $\mathbf{Q}(\zeta_{2m})$ to $\mathbf{Q}(i)$. Then $\mathrm{Norm}(\zeta) = \zeta^a$, where a is given by

$$\begin{aligned} a &= \sum_{0<b<2^m; b\equiv 1 (\mathrm{mod}\ 4)} b \\ &= \sum_{j=0}^{2^{m-2}-1} (1+4j) \\ &= 2^{m-2} + 2^{m-1}(2^{m-2}-1) \\ &\equiv 2^{m-2} (\mathrm{mod}\ 2^{m-1}) \end{aligned}$$

Therefore ζ^a is a primitive 4th root of unity - $\zeta^a = \pm i$. It follows that $\mathrm{Norm}(\epsilon)/\overline{\mathrm{Norm}(\epsilon)} = \pm i$. But $\mathrm{Norm}(\epsilon)$ is a unit of $\mathbf{Q}(i)$, and is therefore ± 1 or $\pm i$. None of these possibilities work, so we have a contradiction. Therefore $Q = 1$ for $\mathbf{Q}(\zeta_{2^m})$. The case of an odd prime power is the same as in Lemma 2.6.
Now assume n is not a prime power. Then $1 - \zeta_n$ is a unit, and we have $(1-\zeta_n)(1-\bar{\zeta_n}) = -\zeta_n$. Suppose $-\zeta_n \in W^2$. Then $-\zeta_n = \zeta_n^{2r}$, so $-1 = \zeta_n^{2r-1}$. Clearly n must be even, so $n \equiv 0 (\mathrm{mod}\ 4)$. Since $-1 = \zeta_n^{n/2}$, we have $n/2 \equiv 2r - 1 (\mathrm{mod}\ n)$, and so $n/2 \equiv -1 (\mathrm{mod}\ 2)$, which is a contradiction. It follows that $-\zeta_n \notin W^2$, so $Q = 2$. □

When $K = \mathbf{Q}(\zeta_n)$, we can prove a stronger version of Theorem B.15.

Theorem B.17. *Let C be the class group of $\mathbf{Q}(\zeta_n)$, and C^+ the class group of the real subfield $\mathbf{Q}(\zeta_n)^+$. Then C^+ injects into C via the natural map.*

Proof. Suppose I is an ideal of $\mathbf{Q}(\zeta_n)^+$ which becomes principal when lifted to $\mathbf{Q}(\zeta_n)$. We need to show I was principal to begin with. Let $I = (\alpha)$ for some $\alpha \in \mathbf{Q}(\zeta_n)$. Then $(\bar{\alpha}/\alpha) = \bar{I}/I = (1)$, since I is real. Therefore $\bar{\alpha}/\alpha$ is a unit and has absolute value 1, so it is a root of unity.
If n is not a prime power, then $Q = 2$, and the proof of Theorem B.15 shows there is a unit ϵ such that $\bar{\epsilon}/\epsilon = \bar{\alpha}/\alpha$. Then $\alpha\epsilon$ is real, and $I = (\alpha) = (\alpha\epsilon)$. By unique factorization of ideals, $I = (\alpha\epsilon)$ in $\mathbf{Q}(\zeta_n)^+$, so I was originally

principal.
Now suppose $n = p^m$. Let $\pi = \zeta_{p^m} - 1$. We have $\pi/\bar{\pi} = -\zeta_{p^m}$, which generates the roots of unity in $\mathbf{Q}(\zeta_{p^m})$. Therefore $\bar{\alpha}/\alpha = (\pi/\bar{\pi})^d$ for some d. Since the π-adic valuation takes on only even values on $\mathbf{Q}(\zeta_{p^m})^+$, and $a\pi^d$ and I are real, we see

$$d = v_\pi(a\pi^d) - v_\pi(a) = v_\pi(a\pi^d) - v_\pi(I)$$

is even. Thus $\bar{\alpha}/\alpha = (-\zeta_{p^m})^d \in W^2$. In particular, $\bar{\alpha}/\alpha = \zeta/\bar{\zeta}$ for some root of unity ζ, and $\alpha\zeta$ is real. Therefore, $I = (\alpha\zeta)$, and thus I was originally principal. □

As a final application of Theorem B.15, we give a relationship between the *regulators* of K and K^+.

Definition B.18. Let L be a number field, and $r := r_1 + r_2 - 1$, where r_1 and r_2 are the number of real and complex embeddings of K. Let $\epsilon_1, \dots, \epsilon_r$ be an independent set of units of L. Write the embeddings of L into $\mathbf{C}$ as $\sigma_1, \dots, \sigma_{r_1+1}, \sigma_{r+1}, \bar{\sigma}_{r_1+1}, \dots, \bar{\sigma}_{r+1}$, where σ_j, $1 \le j \le r_1$ are real, and $\bar{\sigma}_k$, $r_1 + 1 \le k \le r + 1$, are pairs of complex embeddings. Finally, let $\delta_j = 1$ if σ_j is real and $\delta_j = 2$ ig σ is complex. The **regulator** is defined to be

$$R_L(\epsilon_1, \dots, \epsilon_r) = |\det(\delta_i \log|\epsilon_j^{\sigma_i}|)_{1\le i,j\le r}|$$

If $\epsilon_1, \dots, \epsilon_r$ is a basis for the group of units of L modulo the roots of unity, we say $R_L := R_L(\epsilon_1, \dots, \epsilon_r)$ is the **regulator of** L.

Let $\epsilon_1, \dots, \epsilon_r$ be a basis for the units of K^+ modulo $\{\pm 1\}$. Then $\epsilon_1, \dots, \epsilon_r$ form a basis for a subgroup of index $Q(= 1 \text{ or } 2)$ in the units of K modulo roots of unity. But each $\delta_i = 1$ for K^+ and $\delta_i = 2$ for K. This gives $R_K = 2^r R_{K^+}$.

We need the following result. We refer to [**Lemma 4.15**; 10] for the proof.

Lemma B.19. *Let $\epsilon_1, \dots, \epsilon_r$ be independent units of a number field K which generate a subgroup A of the units of K modulo roots of unity, and let $\eta_1, \dots, \eta_r$ generate a subgroup B. If $A \subset B$ is of finite index, then*

$$[B : A] = \frac{R_K(\epsilon_1, \dots, \epsilon_r)}{R_K(\eta_1, \dots, \eta_r)}$$

By the above lemma, we see

Proposition B.20. *Let K be a CM-field, and K^+ its maximal real subfield. Then*

$$\frac{R_K}{R_{K^+}} = \frac{2^r}{Q}$$

where $r = \frac{1}{2}[K : \mathbf{Q}] - 1$.

We conclude this section by giving another class number formula. Let X be a group of Dirichlet characters, and K the associated field. Assume K is totally complex, so half of the characters in X are even and the other half are odd. Let $n = [K : \mathbf{Q}]$. Then we have

$$\frac{2^{n/2} h(K^+) R_{K^+}}{2\sqrt{|d(K^+)|}} = \prod_{\chi \in X, \chi \text{ even}, \chi \neq 1} L(1, \chi)$$

and

$$\frac{(2\pi)^{n/2} h(K) R_K}{w\sqrt{|d(K)|}} = \prod_{\chi \in X, \chi \text{ odd}} L(1, \chi)$$

Dividing through, we find

$$\frac{\pi^{n/2} h^-(K) 2^{n/2}}{Qw\sqrt{|d(K)/d(K^+)|}} = \prod_{\chi \text{ odd}} L(1, \chi)$$

For odd χ, our earlier computation shows $L(1, \chi) = (\pi i \tau(\chi)/f_\chi) B_{1,\bar{\chi}}$, and the conductor-discriminant formula (Theorem A.12) gives $\sqrt{|d(K)/d(K^+)|} = (\prod_{\chi \in X} f_\chi)^{1/2}$. Putting it all together, we have the following result.

Theorem B.21. $h^-(K) = Qw \prod_{\chi \text{ odd}} -\frac{1}{2} B_{1,\chi}$.

References

[1] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969, pp. ix+128.

[2] D. Eisenbud, *Commutative algebra* (Graduate Texts in Mathematics). Springer-Verlag, New York, 1995, vol. 150, pp. xvi+785, With a view toward algebraic geometry, ISBN: 0-387-94268-8; 0-387-94269-6. DOI: 10.1007/978-1-4612-5350-1.

[3] S. Lang, *Algebra* (Graduate Texts in Mathematics), third. Springer-Verlag, New York, 2002, vol. 211, pp. xvi+914, ISBN: 0-387-95385-X. DOI: 10.1007/978-1-4613-0041-0.

[4] H. Matsumura, *Commutative ring theory* (Cambridge Studies in Advanced Mathematics). Cambridge University Press, Cambridge, 1986, vol. 8, pp. xiv+320, Translated from the Japanese by M. Reid, ISBN: 0-521-25916-9.

[5] J. S. Milne, *Algebraic number theory (v3.08)*, Available at www.jmilne.org/math/, 2020.

[6] J. Neukirch, *Algebraic number theory* (Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]). Springer-Verlag, Berlin, 1999, vol. 322, pp. xviii+571, Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder, ISBN: 3-540-65399-6. DOI: 10.1007/978-3-662-03983-0.

[7] A. M. Robert, *A course in p-adic analysis* (Graduate Texts in Mathematics). Springer-Verlag, New York, 2000, vol. 198, pp. xvi+437, ISBN: 0-387-98669-3. DOI: 10.1007/978-1-4757-3254-2.

[8] J.-P. Serre, *Local fields* (Graduate Texts in Mathematics). Springer-Verlag, New York-Berlin, 1979, vol. 67, pp. viii+241, Translated from the French by Marvin Jay Greenberg, ISBN: 0-387-90424-7.

[9] J. H. Silverman, *The arithmetic of elliptic curves* (Graduate Texts in Mathematics). Springer-Verlag, New York, 1986, vol. 106, pp. xii+400, ISBN: 0-387-96203-4. DOI: 10.1007/978-1-4757-1920-8.

[10] L. C. Washington, *Introduction to cyclotomic fields* (Graduate Texts in Mathematics), Second. Springer-Verlag, New York, 1997, vol. 83, pp. xiv+487, ISBN: 0-387-94762-0. DOI: 10.1007/978-1-4612-1934-7.