2005 Summer Research Journal

Bob Hough

Last updated: August 11, 2005

Contents

Chapter 0

Preface

Let χ be a multiplicative character on the finite field F_q and let ψ be an additive character. Then the Gauss sum $g(\chi, \psi)$ is defined by

$$g(\chi, \psi) = \sum_{t \in F_q} \chi(t)\psi(t).$$

The goal of this research journal is to document the properties, uses, and interpretations of such sums.

Gauss first introduced the Gauss sum in one of his proofs of quadratic reciprocity. Gauss's sum was in the form

$$g_p = \sum_{t \in F_p} \left(\frac{t}{p}\right) \zeta_p^t$$

where p is a prime, $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol and $\zeta_p = e^{\frac{2\pi i}{p}}$. We take Gauss's proof as a starting point and then consider Gauss sums of third and fourth order characters over F_p to give proofs of cubic and biquadratic reciprocity laws respectively. This is the substance of the first chapter of the journal.

Throughout chapter one an underlying theme is that the exact value of $g(\chi, \zeta_p)^l$ can be determined, where $l = 2, 3, 4$ is the order of χ. In fact, it is a simple computation to show that for any character χ, $|g(\chi, \zeta_p)| = \sqrt{p}$ and in the cases $l = 2, 3, 4$ we are able to determine the angle of $g(\chi, \zeta_p)$ up to an lth root of unity. Each of the proofs of reciprocity then works by employing congruences on Gauss sums in algebraic number fields in order to extract information about the characters in the sum.

Chapter two consists of a collection of more advanced results on Gauss sums. We begin by investigating the relationship between the value of the Gauss sum and the reciprocity laws of chapter one and are led to consider the theory of cyclotomic field extensions. Viewing the field $\mathbb{Q}(\zeta_l, \zeta_p)$ as a vector space over $\mathbb{Q}(\zeta_l)$ where l is the order of the character χ, we deduce that the value of the Gauss sum $g(\chi)$ contains sufficient information to determine the value of χ at each element of the finite field F_p. The Gauss sum is thus a "compressed representation of χ."

We next treat the Gauss sum as a discrete Fourier expansion for the character χ. Using Fourier expansion techniques we prove two identities, the first being the norm relation on Gauss sums mentioned above and the second an expression for $L(1,\chi)$ as a product of $g(\chi)$ and the sum

$$h(\chi) = \sum_{t \in F_p} \chi(t) \frac{1}{1-\zeta_p^t}.$$

This leads to a digression in which we give a classical determination of the quadratic Gauss sum due to Estermann and try to adapt his technique to give an elementary determination of the argument of $h(\chi)$, an attempt which ultimately fails because the modulus of the function $\frac{1}{1-\zeta}$ grows arbitrarily large as ζ approaches 1.

In the remainder of chapter two we consider Gauss sums over finite fields F_q with $q = p^r$. We discuss the Davenport-Hasse identity for products of Gauss sums and give an alternate formulation in terms of Jacobi sums, J, defined by

$$J(\chi_1, \chi_2, ..., \chi_n) = \sum_{t_1+t_2+...+t_n=1} \chi_1(t_1)\chi_2(t_2)...\chi_n(t_n).$$

This formulation suggests a combinatorial approach to proving the identity that would be more elementary, although significantly more computationally intensive, than the known proof. We also give part of a proof due to Yamamoto [8] affirming Hasse's conjecture that the Davenport-Hasse identity and the norm relation are the only multiplicative relations on Gauss sums when the sums are considered as ideals of the cyclotomic field $\mathbb{Q}(\zeta_l, \zeta_p)$. At the end of his paper containing the proof of Hasse's conjecture, Yamamoto gives an counter-example to show that the conjecture does not hold for Gauss sums when considered as numbers. We show here, by imitating calculations in a paper of Muskat and Whiteman [6], that Yamamoto's counter-example can be reconciled with Hasse's conjecture if we include the additional relations associated with the behavior of Gauss sums under automorphisms of $\mathbb{Q}(\zeta_l, \zeta_p)/\mathbb{Q}$.

We conclude chapter two with a proof of Eisenstein reciprocity law, which generalizes the reciprocity laws of chapter one to any prime-order character. The proof of Eisenstein reciprocity is analogous to our earlier reciprocity proofs in that it hinges on a computation regarding $g(\chi)^l$ where l is the order of χ. In the case of Eisenstein reciprocity, this computation is the Stickelberger identity, which gives a factorization of the principal ideal $(g(\chi)^l)$ in $\mathbb{Q}(\zeta_l, \zeta_p)$ into automorphisms of a prime ideal. The proof then follows by again working with congruences in an algebraic number field. As an application, we give a proof of Wieferich's theorem, which places a limit on solutions to the Fermat equation: $x^n + y^n + z^n = 0$, $n > 2$.

I have made an effort to ensure that this journal reflects my learning process over the past eight weeks and consequently very little background knowledge is assumed. Thus, chapter one contains substantial sections devoted to motivating multiplicative characters and describing the behavior of the rings of integers $\mathbb{Z}[\omega]$ ($\omega = \frac{-1+i\sqrt{3}}{2}$) and $\mathbb{Z}[i]$; the reader familiar with these topics will most likely find these discussions superfluous. In chapter two

I do not include all of the background material relating to the splitting of prime ideals under cyclotomic extension because the theory is fairly extensive. I also regret the omission of a section pertaining to the Stickelberger relation, which plays a central role in all modern computations regarding Gauss sums.

The theory in the journal is almost entirely condensed from my assorted readings. The reciprocity laws of chapter one, as well as the proof of Eisenstein's reciprocity law in chapter two are due to Ireland and Rosen [4]. The definition of multiplicative characters over F_q and the discussion of the Davenport-Hasse identity is from Berndt et. al. [2] while the reformulation of Davenport-Hasse in terms of Jacobi sums is derived using formulas for products of Gauss sums given in Ireland and Rosen. Estermann's determination of the argument of the quadratic Gauss sum can be found in [2] while the partial proof of Hasse's conjecture is from Yamamoto's paper itself [8]. The observations regarding the duality between χ and its Gauss sum, as well as the formulation of Fourier analysis on the finite field F_p were not taken from a text, although they are familiar in the literature. For example, Weil views the Gauss sum as a Fourier expansion in his excellent paper [7].

Most of my investigations for the summer have focused on three problems: determining the argument of the cubic Gauss sum, giving an elementary proof of the Davenport-Hasse identity, and giving a direct proof that $L(1,\chi) > 0$ in the case χ is a quadratic character with $\chi(-1) = -1$. Toward the first problem I have written C++ code to calculate cubic Gauss sums for primes that are not too large, but I have been unable as yet to find a pattern in the argument. The source code can be found in the appendix. For the second problem, the expression of the Davenport-Hasse identity in terms of Jacobi sums represents the point at which a combinatorial proof of that identity becomes dauntingly cumbersome. The only noteworthy original result in the journal is the computation of the alternate expression for $L(1,\chi)$. Our determination via Fourier expansion gives a simpler proof than the one provided by Beck et. al. in their recent paper [1], which utilizes contour integration. My attempts to prove that the sum $h(\chi)$ is positive imaginary, however, have thus far proved unsuccessful, so problem three remains incomplete.

The difficulty of applying classical techniques to, for instance, prove the Davenport-Hasse identity or to determine the argument of the cubic Gauss sum emphasizes the need for more modern techniques. For instance, some progress has been made on the cubic Gauss sum by introducing elliptic curves and the Weierstrass $\wp$-function (see Matthews [5]). Thus further investigation of the first two problems may require more background information. I still remain hopeful that I will be able to give a proof of the third problem. It would also be interesting to try to extend Yamamoto's proof of Hasse's conjecture for ideals to numbers by including properties of Gauss sums under automorphism.

Chapter 1

Gauss Sums and Reciprocity

Reciprocity laws hold the key to the question: For which residues a modulo prime p does $x^n \equiv a$ have a solution? The standard approach to this problem is to assign a character to the elements of $\mathbb{Z}/p\mathbb{Z}$ that distinguishes between nth-residues and non-residues. If the character is chosen to be multiplicative, then it is sufficient to determine the character of all primes $q < p$ in order to determine all characters in $\mathbb{Z}/p\mathbb{Z}$. Reciprocity laws simplify this computation by relating the character of q in $\mathbb{Z}/p\mathbb{Z}$ to the character of p in $\mathbb{Z}/q\mathbb{Z}$, thus reducing the modulus, a process which can be repeated until the modulus prime is sufficiently small that its residue characters are known.

The Gauss sum is an object that compresses information about the characters of the residues modulo p into a single complex number. First introduced by Gauss in his sixth proof of quadratic reciprocity [4], Gauss sums play an integral role in the general theory of reciprocity.

1.1 Quadratic Gauss sums and quadratic reciprocity

Gauss used a quadratic Gauss sum in his proof of quadratic reciprocity and the sum is most easily understood in this special case, although it is easily generalized to higher powers.

Definition 1.1 *The quadratic Gauss sum associated with residue a modulo p is given by*

$$g_a = \sum_t \left(\frac{t}{p}\right)\zeta^{at}$$

where t runs over all residue classes modulo p, $\left(\frac{\cdot}{p}\right)$ denotes the Legendre symbol, and $\zeta = e^{\frac{2\pi i}{p}}$.

Ireland and Rosen note that for non-zero residues a,

$$\left(\frac{a}{p}\right)g_a = \sum_t \left(\frac{a}{p}\right)\left(\frac{t}{p}\right)\zeta^{at} = \sum_t \left(\frac{at}{p}\right)\zeta^{at}.$$

Since at ranges over all residues mod p as t does, the final sum is equal to g_1 and $\left(\frac{a}{p}\right)g_a = g_1$. Equivalently, as $\left(\frac{a}{p}\right)^2 = 1$, $g_a = \left(\frac{a}{p}\right)g$ where we write g for g_1. Meanwhile, we have $g_0 = 0$ since $\sum\left(\frac{a}{p}\right) = 0$ follows from the familiar fact that there are an equal number of quadratic residues and non-residues modulo p.

Having made this observation, it is possible to compute g^2.

Proposition 1.1 *$g^2 = (-1)^{\frac{p-1}{2}}p$.*

Proof. (from [4], pp. 71-72) Consider the sum $S = \sum g_{-a}g_a$ ranging over residues a mod p. From our observation, $g_{-a}g_a = \left(\frac{-a}{p}\right)\left(\frac{a}{p}\right)g^2$ for $a \neq 0$ and is 0 otherwise. Then since $\left(\frac{-a}{p}\right)\left(\frac{a}{p}\right) = \left(\frac{-a^2}{p}\right) = \left(\frac{-1}{p}\right)$,

$$S = \sum_{a\neq 0}\left(\frac{-1}{p}\right)g^2 = (p-1)\left(\frac{-1}{p}\right)g^2.$$

On the other hand,

$$g_{-a}g_a = \sum_x \left(\frac{x}{p}\right)\zeta^{-ax}\sum_y\left(\frac{y}{p}\right)\zeta^{ay}$$

so

$$S = \sum_a\sum_{x,y}\left(\frac{x}{p}\right)\zeta^{-ax}\left(\frac{y}{p}\right)\zeta^{ay} = \sum_a\sum_{x,y}\left(\frac{xy}{p}\right)\zeta^{a(y-x)}.$$

Exchanging the order of summation,

$$S = \sum_{x,y}\left(\frac{xy}{p}\right)\sum_a \zeta^{a(y-x)}.$$

But $\sum_a \zeta^{az}$ is p if $z \equiv 0 \pmod{p}$ and 0 otherwise so

$$\sum_{x,y}\left(\frac{xy}{p}\right)\sum_a\zeta^{a(y-x)} = \sum_{x=y}\left(\frac{xy}{p}\right)p.$$

When $x = y$, $(\frac{xy}{p}) = (\frac{x^2}{p})$, which is 0 if $x = 0$ and 1 otherwise. Hence $S = \sum_{x\neq 0} p = (p-1)p$. Thus we have $S = (p-1)(\frac{-1}{p})g^2 = (p-1)p$. Canceling $p-1$ and multiplying each side by $(\frac{-1}{p})$ gives $g^2 = (\frac{-1}{p})p = (-1)^{\frac{p-1}{2}}p$, the desired result. □

Determining the sign of g requires elaborate computation that we postpone until later. Fortunately, knowing g^2 is sufficient to prove the Law of Quadratic Reciprocity.

1.1.1 Quadratic reciprocity and the algebraic integers

Because the value of g is not an integer, proving the Quadratic Reciprocity Law using Gauss Sums requires working with congruences in a slightly more general setting. This setting is the algebraic integers, Ω, defined to be the set of roots to monic polynomials, $P \in \mathbb{Z}[x]$. Ireland and Rosen prove that Ω, inheriting addition and multiplication from the complex numbers, is a ring. We have, trivially, that $\mathbb{Z} \subset \Omega$ since for $n \in \mathbb{Z}$, n is the root of $x - n = 0$. Moreover, observe that if $m, n \in \mathbb{Z}$, $g.c.d(m,n) = 1$ and m/n is a root to the polynomial $x^k + a_{k-1}x^{k-1} + ... + a_0 = 0$ with $a_i \in \mathbb{Z}$ then

$$\frac{m^k}{n^k} + a_{k-1}\frac{m^{k-1}}{n^{k-1}} + ... + a_0 = 0.$$

Subtracting the m^k/n^k term and multiplying each side of the equation by n^k leaves

$$na_{k-1}m^{k-1} + n^2 a_{k-2}m^{k-2} + ... + n^k a_0 = -m^k$$

so $n \mid m^k$ implying $n = 1$ and $m/n \in \mathbb{Z}$. Hence $\Omega \cap \mathbb{Q} \subset \mathbb{Z}$, but since $\mathbb{Z} \subset \Omega$ we have $\Omega \cap \mathbb{Q} = \mathbb{Z}$.

Generalizing the notion of congruence in the ordinary integers, we define congruence in the arithmetic integers so that for $a, b, c \in \Omega, c \neq 0$

$$a \equiv b \pmod{c} \Leftrightarrow a - b = cd$$

for some $d \in \Omega$. To see that this definition is actually just an extension of the notion of congruence in the integers, take $a, b, c \in \mathbb{Z}$ with $a \equiv b \ (c)$ in Ω. Then $a - b = cd$ for some $d \in \Omega$ implying that $d = (a-b)/c \in \mathbb{Q}$ (the division can be understood to take place in $\mathbb{C}$ which contains Ω). Hence $d \in \Omega \cap \mathbb{Q}$ so $d \in \mathbb{Z}$ and $a \equiv b \ (c)$ in $\mathbb{Z}$. With these facts about the algebraic integers we are now able to prove the Law of Quadratic Reciprocity.

Theorem 1.2 *(Quadratic Reciprocity) Given odd primes p, q the quadratic residue of p w.r.t. q and of q w.r.t. p are related by*

$$\left(\frac{p}{q}\right) = (-1)^{(\frac{p-1}{2})(\frac{q-1}{2})}\left(\frac{q}{p}\right).$$

Proof. (Due to [4] p.72). Let g be the quadratic Gauss sum modulo p and, for convenience of notation, write $g^2 = (-1)^{\frac{p-1}{2}}p = p^*$. Observe that $g = \sum_t \left(\frac{t}{p}\right)\zeta^t$ is a linear combination of pth roots of unity, hence an arithmetic integer, so we can work with g modulo q in Ω. The idea of the proof is to compute the residue of $g^q \ (q)$ in two ways.

Observe that

$$g^q = gg^{q-1} = g(g^2)^{\frac{q-1}{2}} = gp^{*(\frac{q-1}{2})}.$$

Now p^* is an ordinary integer so by Euler's criterion, $p^{*(\frac{q-1}{2})} \equiv \left(\frac{p^*}{q}\right) \mod q$ and

$$g^q \equiv g\Big(\frac{p^*}{q}\Big) \quad (q).$$

But we also have $g = \sum_t \big(\frac{t}{p}\big)\zeta^t$, so expanding g^q using the Multinomial Theorem gives

$$g^q = \Bigg(\sum_t \Big(\frac{t}{p}\Big)\zeta^t\Bigg)^q \equiv \sum_t \Bigg(\Big(\frac{t}{p}\Big)\zeta\Bigg)^q \quad (\text{mod } q)$$

by recalling that for prime q, $\binom{q}{a_1 a_2 \ldots a_q}_{\sum a_i = q} \equiv 0 \ (q)$ unless some a_i is q and all the rest are 0. Now $\big(\frac{\cdot}{p}\big)$ is an integer, so applying Fermat's Little Theorem, $\big(\frac{\cdot}{p}\big)^q \equiv \big(\frac{\cdot}{p}\big)$ (mod q). Hence,

$$g^q \equiv \sum_t \Bigg(\Big(\frac{t}{p}\Big)\zeta\Bigg)^q \equiv \sum_t \Big(\frac{t}{p}\Big)\zeta^{qt} \quad (\text{mod } q)$$

and $\sum_t \big(\frac{t}{p}\big)\zeta^{qt}$ (mod q) $= g_q = \big(\frac{q}{p}\big)g$. Combining our results we have $g^q \equiv g\big(\frac{p^*}{q}\big) \equiv \big(\frac{q}{p}\big)g$ (mod q), or, multiplying each side by g,

$$\Big(\frac{p^*}{q}\Big)p^* \equiv \Big(\frac{q}{p}\Big)p^* \quad (\text{mod } q).$$

Both $\big(\frac{\cdot}{p}\big)$ and p^* are integers and $\mathbb{Z}/q\mathbb{Z}$ is a field, so cancelation is permissible, leaving $\big(\frac{p^*}{q}\big) \equiv \big(\frac{q}{p}\big)$ (q). Each number is plus or minus one, so equality holds. To complete the proof, observe that $p^* = (-1)^{\frac{p-1}{2}}p$ so

$$\Big(\frac{p^*}{q}\Big) = \Bigg(\frac{(-1)^{\frac{p-1}{2}}}{q}\Bigg)\Big(\frac{p}{q}\Big) = \Big((-1)^{\frac{p-1}{2}}\Big)^{\frac{q-1}{2}}\Big(\frac{p}{q}\Big)$$

by applying Euler's criterion to $\Big(\frac{(-1)^{\frac{p-1}{2}}}{q}\Big)$. Hence,

$$\Big(\frac{q}{p}\Big) = (-1)^{(\frac{p-1}{2})(\frac{q-1}{2})}\Big(\frac{p}{q}\Big)$$

as desired. □

This proof of quadratic reciprocity is based upon two computations involving the value of the Gauss sum; g is used to determine both $\big(\frac{p^*}{q}\big)$ and $\big(\frac{q}{p}\big)$ modulo q. The first of these calculations was more direct: we took advantage of the fact that that Gauss sum has magnitude $\sqrt{p}$, then used a property of the Legendre symbol, namely Euler's criterion that $a^{\frac{q-1}{2}} \equiv \big(\frac{a}{q}\big)$ mod q. The second computation was more subtle, but perhaps less surprising since the Gauss sum contains $\big(\frac{q}{p}\big)$. In fact, the value of the Gauss sum explicitly describes $\big(\frac{a}{p}\big)$ for all residues a mod p, a result that we elaborate more fully in our later section on Fourier analysis in finite fields. In our proofs of cubic and biquadratic reciprocity, we will utilize the same strategy that we have here for quadratic reciprocity, first determining the value of a (generalized) Gauss sum up to a cubic or quartic root of unity, and then using the value to determine a relation between the character of primes.

1.2 General Gauss sums

1.2.1 The general character

Before we discuss higher reciprocity laws, however, we need a generalization of the Legendre symbol $\left(\frac{\cdot}{p}\right)$. This desired generalization is that of a multiplicative character, χ, mapping $(\mathbb{Z}/p\mathbb{Z})^\times$ to $\mathbb{C}^\times$ and satisfying

$$\chi(a)\chi(b) = \chi(ab)$$

for all $a, b \in (\mathbb{Z}/p\mathbb{Z})^\times$. The set of all such characters forms a group under the multiplication operation defined by $\chi\lambda(a) = \chi(a)\lambda(a)$ for all characters χ, λ and all $a \in (\mathbb{Z}/p\mathbb{Z})^\times$. The trivial character, ϵ, with $\epsilon(a) = 1$ for all $a \in (\mathbb{Z}/p\mathbb{Z})^\times$, is the group identity. The group inverse is defined by $\chi^{-1}(a) = \chi(a)^{-1}$. It follows from the definition that both the product of two characters and the inverse of a character preserve the multiplicative property and hence are characters. The character group is thus well defined. It is customary to extend characters on $(\mathbb{Z}/p\mathbb{Z})^\times$ to characters on $\mathbb{Z}/p\mathbb{Z}$ by setting $\chi(0) = 0$ if $\chi \neq \epsilon$ and $\epsilon(0) = 1$. This extension preserves the multiplicative definition of the character although it does not maintain the group structure (for instance, the inverse of a non-trivial character at 0 is no longer well defined).

The streamlined definition of χ belies its underlying complexity. Ireland and Rosen note that for all characters χ, $\chi(1) = 1$ and for all $a \in (\mathbb{Z}/p\mathbb{Z})^\times$, $\chi(a)$ is a $(p-1)$st root of unity. Moreover the group of characters is cyclic of order $p-1$. The first observation follows because $\chi(1) = \chi(1 \cdot 1) = \chi(1)\chi(1)$ which implies that $\chi(1)$ is either 0 or 1. Since χ maps to $\mathbb{C}^\times$, we have $\chi(1) = 1$. For general $a \in (\mathbb{Z}/p\mathbb{Z})^\times$, $a^{p-1} = 1$ so $\chi(a)^{p-1} = \chi(1) = 1$ from which we conclude that $\chi(a)$ is a $(p-1)$st root of unity. For the final observation let a generate $(\mathbb{Z}/p\mathbb{Z})^\times$. Then for any character χ, the values of χ are uniquely determined on $(\mathbb{Z}/p\mathbb{Z})^\times$ by $\chi(a)$. Thus there are at most $p-1$ characters because $\chi(a)$ must be a $(p-1)$st root of unity. Setting $\chi(a^t) = e^{\frac{2\pi i t}{p}}$ describes a well defined character. The characters $\chi^1, \chi^2, ...\chi^{p-1}$ are all distinct because $\chi^t(a) = e^{\frac{2\pi i t}{p}}$. This is a list of $p-1$ characters, hence all characters on $(\mathbb{Z}/p\mathbb{Z})^\times$. The group of characters is thus cyclic with generator χ and of order $p-1$.

At this point it is instructive to take a step back and examine why the multiplicative character is the natural generalization of the Legendre symbol for $n > 2$. Let f be the function $f(x) = x^n$ on $(\mathbb{Z}/p\mathbb{Z})^\times$. Because $(\mathbb{Z}/p\mathbb{Z})^\times$ is cyclic of order $p-1$, if n and $p-1$ are relatively prime then f maps generators to generators in $(\mathbb{Z}/p\mathbb{Z})^\times$. As a result, f is a permutation of the group and $f(x) = a$ has a solution for all $a \in (\mathbb{Z}/p\mathbb{Z})^\times$. Now write $n = dm$ where $d \mid p-1$ and $(m, p-1) = 1$. Let g be the function $g(x) = x^d$. If $g(x) = a$ has solution a_0 then, as m and $p-1$ are relatively prime, there is a_1 with $a_1^m = a_0$ so $a_1^n = a_1^{md} = a_0^d = a$. Meanwhile, if $f(x) = a$ has solution a_2 then $a_2^n = a$ or $a_2^{md} = a$ implying that a_2^m is a solution to $g(x) = a$. Hence $f(x) = a$ has solution if and only if $g(x) = a$ has a solution and it is sufficient to consider only those exponents n dividing $p - 1$.

Now take n with $dn = p - 1$ and let a generate $(\mathbb{Z}/p\mathbb{Z})^\times$. Suppose $x^n = a^m$ has solution b. Raising both sides of the equation to the power d gives $a^{md} = b^{p-1} = 1$ which is possible only if $p - 1$ divides md, or equivalently, if $n \mid m$. If we view $(\mathbb{Z}/p\mathbb{Z})^\times$ as the additive cyclic group $\mathbb{Z}/(p-1)\mathbb{Z}$ via the isomorphism $a^t \mapsto t$ then since a^t is an nth power iff $n \mid t$, the collapsing homomorphism $t\ (p-1) \mapsto t\ (n)$ is a natural way to understand the character of $(\mathbb{Z}/p\mathbb{Z})^\times$ with respect to $f(x) = x^n$. Now if $\mathbb{Z}/n\mathbb{Z}$ is taken to be the group of nth roots of unity under multiplication, then our "nth character" homomorphism is $a^t \mapsto \zeta^t$ where ζ is any primitive nth root of unity. We recognize the image group of the homomorphism as a character χ of order n in the group of characters on $(\mathbb{Z}/p\mathbb{Z})^\times$. In fact, mapping a in turn to each of the $\phi(n)$ primitive nth roots of unity, we derive all $\phi(n)$ order n characters in the group so that the order n characters on $(\mathbb{Z}/p\mathbb{Z})^\times$ correspond exactly to our understanding of the behavior of $f(x) = x^n$ on $(\mathbb{Z}/p\mathbb{Z})^\times$. Moreover, we retrieve the Legendre symbol as the lone character of order 2.

1.2.2 The general Gauss sum

Having generalized the Legendre symbol to higher order characters, it is now possible to define a general Gauss sum that plays a role in the theory of higher order reciprocity laws analogous to the role of the quadratic Gauss sum in quadratic reciprocity.

Definition 1.2 *The Gauss sum associated with the character χ at the residue a (mod p) is given by*

$$g_a(\chi) = \sum_t \chi(t)\zeta^{at}$$

where, as usual, t runs over residues modulo p and $\zeta = e^{\frac{2\pi i}{p}}$.

We note that the character χ is the extended character on all of $\mathbb{Z}/p\mathbb{Z}$ since t runs over all residues modulo p (the distinction is only important in the case $\chi = \epsilon$ since otherwise $\chi(0) = 0$). The general Gauss sum has analogs to many of the properties of the quadratic Gauss sum. In particular, Ireland and Rosen prove the following two propositions:

Proposition 1.3 *For $\chi \neq \epsilon$ if $a \neq 0$ then $g_a(\chi) = \chi(a^{-1})g_1(\chi)$ and if $a = 0, g_a(\chi) = 0$. Meanwhile, $g_0(\epsilon) = p$ and $g_a(\epsilon) = 0$ for all $a \neq 0$.*

Proof. For $\chi \neq \epsilon$ and $a \neq 0$, we have $g_a(\chi) = \sum_t \chi(a^{-1})\chi(a)\chi(t)\zeta^{at}$ since $\chi(a^{-1})\chi(a) = \chi(1) = 1$. But then

$$g_a(\chi) = \chi(a^{-1})\sum_t \chi(at)\zeta^{at} = \chi(a^{-1})g_1(\chi)$$

since at runs over all residues modulo p as t does. For $a = 0$, $g_0(\chi) = \sum_t \chi(t)$. Recalling that for $\chi \neq \epsilon, \chi(0) = 0$, we can find primitive root b mod p and rewrite the sum as

$$g_0(\chi) = \sum_{t=1}^{p-1} \chi(b^t) = \sum_{t=1}^{p-1} \chi(b)^t = \frac{\chi(b)^p - \chi(b)}{\chi(b) - 1}$$

Now $\chi(b)$ is a $(p-1)$st root of unity not equal to 1 since $\chi \neq \epsilon$. Thus $\chi(b) - 1 \neq 0$ and $\chi(b)^{p-1} = 1$ implying $\chi(b)^p - \chi(b) = 0$. Hence $g_0(\chi) = 0$ as desired.

Now for the case $\chi = \epsilon$,

$$g_a(\epsilon) = \sum_t \zeta^{at}$$

which is equal to p if $p \mid a$ and zero otherwise. This completes the proposition. □

As before, we will drop the subscript on $g_1(\chi)$ and write only $g(\chi)$.

Proposition 1.4 *For $\chi \neq \epsilon, |g(\chi)| = \sqrt{p}$.*

Proof. The proof follows from a similar approach to the one used to derive the absolute value of the quadratic Gauss sum. Summing $g_a(\chi)\overline{g_a(\chi)}$ over all residues a, we have, by Proposition 1.3,

$$\sum_a g_a(\chi)\overline{g_a(\chi)} = \sum_a \chi(a^{-1})\overline{\chi(a^{-1})}g(\chi)\overline{g(\chi)}$$

But $\chi(a^{-1}) = \chi(a)^{-1} = \overline{\chi(a)}$ since $|\chi(a)| = 1$ so the sum reduces to

$$g(\chi)\overline{g(\chi)} \sum_a \chi(a)\overline{\chi(a)} = |g(\chi)|^2(p-1).$$

On the other hand, expanding $g_a(\chi)\overline{g_a(\chi)}$ as a sum we have

$$\sum_a g_a(\chi)\overline{g_a(\chi)} = \sum_a \sum_x \chi(x)\zeta^{ax} \sum_y \overline{\chi(y)\zeta^{ay}} = \sum_{x,y} \chi(x)\overline{\chi(y)} \sum_a \zeta^{a(x-y)}$$

Now $\sum_a \zeta^{a(x-y)}$ is p if $x = y$ and zero otherwise so

$$\sum_a g_a(\chi)\overline{g_a(\chi)} = p \sum_x \chi(x)\overline{\chi(x)}.$$

Since $\chi \neq \epsilon$, $|\chi(x)|^2 = 1$ if $x \neq 0$ and zero otherwise. Hence $\sum_a g_a(\chi)\overline{g_a(\chi)} = p(p-1)$. Then with

$$|g(\chi)|^2(p-1) = p(p-1)$$

we have the desired result. □

1.2.3 Jacobi sums

We would like to use generalized Gauss sums to prove theorems of cubic and biquadratic reciprocity, but in order to do so it will be necessary to compute the exact values of $g(\chi)^3$ and $g(\lambda)^4$ where χ and λ are third and fourth order characters on $\mathbb{Z}/p\mathbb{Z}$ respectively (a similar computation of g^2 was central to our proof of quadratic reciprocity). In making these computations, we will find it useful to use a tool called the Jacobi sum, defined below.

Definition 1.3 *The Jacobi sum $J(\chi, \lambda)$ of characters χ and λ on $\mathbb{Z}/p\mathbb{Z}$ is given by*

$$\sum_{a+b=1} \chi(a)\lambda(b).$$

The Jacobi sum is motivated by consideration of the question: How many solutions exist to the equation $x^n + y^n = a$ in $\mathbb{Z}/p\mathbb{Z}$? and it has a well-developed theory in its own right, but we shall be primarily interested in its relations to Gauss sums. We prove two properties of the Jacobi sum now and will prove other necessary results as they are needed. Both proofs are due to [4].

Proposition 1.5 *Let χ and λ be non-trivial characters such that $\chi\lambda$ is also non-trivial. Then*

$$J(\chi, \lambda) = \frac{g(\chi)g(\lambda)}{g(\chi\lambda)}$$

Proof. Expanding $g(\chi)g(\lambda)$,

$$g(\chi)g(\lambda) = \sum_{x,y} \chi(x)\lambda(y)\zeta^{x+y}$$

Indexing according to $x + y$ we have, equivalently,

$$g(\chi)g(\lambda) = \sum_{t} \Big(\sum_{x+y=t} \chi(x)\lambda(y) \Big) \zeta^{t}$$

For $t = 0$, $\sum_{x+y=0} \chi(x)\lambda(y) = \sum_x \chi(x)\lambda(-x)$, but the latter is equal to $\lambda(-1)\sum_x \chi\lambda(x) = 0$ since we assumed that $\chi\lambda$ is non-trivial. For $t \neq 0$, we can make the substitution $x = tx', y = ty'$ in the sum $\sum_{x+y=t} \chi(x)\lambda(y)$ to arrive at $\sum_{x'+y'=1} \chi(tx')\lambda(ty') = \chi\lambda(t)J(\chi, \lambda)$. Thus

$$g(\chi)g(\lambda) = \sum_{t=1}^{p-1} J(\chi, \lambda)\chi\lambda(t)\zeta^{t} = J(\chi, \lambda)g(\chi\lambda).$$

Upon dividing both the left and right hand sides by $g(\chi\lambda)$ we have our desired result. □

Our next proposition follows by iterating the above result.

Proposition 1.6 *Suppose $n > 2$ and $n \mid p-1$ and let χ be a character of order n. Then*

$$g(\chi)^n = \chi(-1)pJ(\chi,\chi)J(\chi,\chi^2)...J(\chi,\chi^{n-2}).$$

Proof. By Proposition 1.5 $g(\chi)^2 = J(\chi,\chi)g(\chi^2)$. In general, $g(\chi)g(\chi^m) = J(\chi,\chi^m)g(\chi^{m+1})$ so, iterating,

$$\begin{aligned}
g(\chi)^3 &= J(\chi,\chi)J(\chi,\chi^2)g(\chi^3)\\
g(\chi)^4 &= J(\chi,\chi)J(\chi,\chi^2)J(\chi,\chi^3)g(\chi^4)\\
&\vdots\\
g(\chi)^{n-1} &= J(\chi,\chi)J(\chi,\chi^2)...J(\chi,\chi^{n-2})g(\chi^{n-1})
\end{aligned}$$

But $\chi^{n-1} = \chi^{-1}$ and $g(\chi^{-1}) = \sum_t \chi(t)^{-1}\zeta^t = \chi(-1)^{-1}\sum_t \chi(-t)^{-1}\zeta^t$. This last equation is equal to

$$\overline{\chi(-1)}\sum_t \overline{\chi(t)}\zeta^{-t} = \overline{\chi(-1)g(\chi)} = \chi(-1)\overline{g(\chi)}$$

since $\chi(-1)$ is plus or minus 1. Thus

$$g(\chi)^n = g(\chi)J(\chi,\chi)J(\chi,\chi^2)...J(\chi,\chi^{n-2})\chi(-1)\overline{g(\chi)}.$$

But by Proposition 1.4, $g(\chi)\overline{g(\chi)} = p$ so

$$g(\chi)^n = \chi(-1)pJ(\chi,\chi)J(\chi,\chi^2)...J(\chi,\chi^{n-2})$$

as desired. □

1.3 Cubic reciprocity

In Jacobi and generalized Gauss sums we have the tools necessary to prove cubic and biquadratic reciprocity. We do not yet, however, have the correct setting. In our proof of quadratic reciprocity, we made use of the fact that the quadratic character was a member of the ring in question (the integers) meaning that it could be included in congruence computations. We will find this necessary in the corresponding theories of cubic and biquadratic reciprocity, but, because the cubic and biquadratic characters are not restricted to the integers, we will need to expand our ring of interest in order formulate the two reciprocity laws. In the cubic case, the cubic character takes on values that are cube roots of unity so we need to extend $\mathbb{Z}$ to include the complex numbers $\omega = \frac{-1+\sqrt{-3}}{2}$ and $\omega^2 = \frac{-1-\sqrt{-3}}{2}$. The smallest ring that contains these two numbers as well as $\mathbb{Z}$ is the ring $\mathbb{Z}[\omega] = \{a + b\omega | a, b \in \mathbb{Z}\}$ and this will be the ring that we work in. In the biquadratic case we will work in the ring $\mathbb{Z}[i] = \{a + bi | a, b \in \mathbb{Z}\}$ which contains the fourth roots of unity.

1.3.1 The ring $\mathbb{Z}[\omega]$

The ring $\mathbb{Z}[\omega]$ is an Euclidean domain (and hence a unique factorization domain) under the norm N where $N(\alpha) = \alpha\overline{\alpha} = a^2 - ab + b^2$ for $\alpha = a + b\omega$. The units of $\mathbb{Z}[\omega]$ are the elements of norm 1, namely 1, -1, ω, $-\omega$, ω^2 and $-\omega^2$. The primes in $\mathbb{Z}[\omega]$ are not generally the same as the primes in $\mathbb{Z}$. Ireland and Rosen show that if $q \equiv 2 \ (3)$ is prime in $\mathbb{Z}$ then q is prime in $\mathbb{Z}[\omega]$ while if $p \equiv 1 \ (3)$ is prime in $\mathbb{Z}$ it splits as $\pi\overline{\pi}$ in $\mathbb{Z}[\omega]$ where π is prime in $\mathbb{Z}[\omega]$. Meanwhile, $3 = -\omega^2(1-\omega)^2$ and $1-\omega$ is prime in $\mathbb{Z}[\omega]$. Up to multiplication by units, these are all the primes in $\mathbb{Z}[\omega]$. In actuality, each prime in $\mathbb{Z}[\omega]$ is related to six other primes by units. As in $\mathbb{Z}$ where we do not consider negative primes, we would like to study only one out of each group of six associated primes. We do this by defining a *primary* prime to be a prime congruent to 2 mod 3. An integer prime congruent to 2 mod 3 is clearly primary, and Ireland and Rosen show that for each prime π with norm $p \equiv 1 \ (p)$, π has one and only one associate that is primary. There are no primary associates to $1 - \omega$ but this prime will not play an important role in the theory of cubic reciprocity.

Because $\mathbb{Z}[\omega]$ is an Euclidean domain, its prime and maximal ideals coincide so that if π is a prime in $\mathbb{Z}[\omega]$ then $\mathbb{Z}[\omega]/(\pi)$ is a field. (Here congruence is defined as usual so that $a \equiv b \ (\pi)$ means $\pi \mid (a-b)$ in $\mathbb{Z}[\omega]$). We claim that this field has $N(\pi)$ elements. There are three cases to consider: case 1 is when $q \equiv 2 \ (3)$ is a prime in $\mathbb{Z}$; case 2 is π such that $\pi\overline{\pi} = p$ where $p \equiv 1 \ (3)$ is a prime in $\mathbb{Z}$; case 3 is $1 - \omega$.

In case 1, $N(q) = q^2$. We claim that $S = \{a + b\omega | a, b \in \mathbb{Z}, 0 \leq a, b < q\}$ is a complete set of residues in $(\mathbb{Z}[\omega])/q(\mathbb{Z}[\omega])$. Clearly for any $a + b\omega$ in $\mathbb{Z}[\omega]$ there exist a' and b' such that $0 \leq a - qa', b - qb' < q$ by the division algorithm in $\mathbb{Z}$. This shows that for all α in $\mathbb{Z}[\omega]$, $\alpha \equiv \beta \ (q)$ for some $\beta \in S$. To see that each of the elements in S is distinct modulo q take $\alpha_0 = a_0 + b_0\omega$ and $\alpha_1 = a_1 + b_1\omega$ in S and suppose $\alpha_0 \equiv \alpha_1 \ (q)$. Then $q \mid \alpha_0 - \alpha_1$ which implies that q divides $a_0 - a_1$ and q divides $b_0 - b_1$ in $\mathbb{Z}$. These two conditions imply that $a_0 = a_1$ and $b_0 = b_1$ so $\alpha_0 = \alpha_1$. Hence the residues in S are distinct and the q^2 residues in S are a complete set of residues modulo q.

In case 2, $N(\pi) = \pi\overline{\pi} = p$. We claim that the set $S = \{0, 1, 2, ..., p-1\}$ forms a complete set of residues modulo π. Let $\pi = a + b\omega$. Observe then that $p = a^2 - ab + b^2$ so if p divides either a or b it divides the other, implying that $p^2 \mid p$, which is false. Hence p divides neither a nor b. Then for any $\alpha = c + d\omega$ there exists n so that $bn \equiv d \ (p)$. Hence $\alpha - n\pi \equiv c - na$ (p) so $\alpha - n\pi \equiv c - na \ (\pi)$. This shows that every element of $\mathbb{Z}[\omega]$ is congruent to an integer, modulo π. Then since every integer is congruent to an element of S modulo p, every element of $\mathbb{Z}[\omega]$ is congruent to an element of S modulo π. It only remains to show that the residues $\{0, 1, 2, ..., p-1\}$ are distinct modulo π. Let $s, s' \in S$ and suppose $\pi \mid s - s'$. Then $p \mid N(s - s') = (s - s')^2$. Since $s - s' \leq p - 1$ this is only possible if $s - s'$ are distinct. This implies that all the residues of S are distinct modulo π so S is a complete set of residues modulo π.

In case 3, $N(1-\omega) = 3$. We claim that $\{0, 1, 2\}$ is a complete set of residues modulo $1-\omega$. Since $1-\omega$ divides neither 1 nor 2 the three residues are distinct. For any $\alpha = a + b\omega \in \mathbb{Z}[\omega]$,

$\alpha + b(1-\omega) \in \mathbb{Z}$ so for all $\alpha \in \mathbb{Z}[\omega]$ there exists $n \in \mathbb{Z}$ with $\alpha \equiv n \ (1-\omega)$. Then since all $n \in \mathbb{Z}$ are congruent to one of $0, 1, 2 \ (3)$ (and hence also mod π) it follows that $\{0, 1, 2\}$ is a complete set of residues mod $1-\omega$.

1.3.2 The cubic character

Since $(\mathbb{Z}[\omega])/\pi(\mathbb{Z}[\omega])$ is a field with $N(\pi)$ elements, its multiplicative group has $N(\pi)-1$ elements so if π doesn't divide α then $\alpha^{N(\pi)-1} \equiv 1 \ (\pi)$. For π a prime with norm not equal to 3, then π is either an integer prime, q, congruent to 2 mod 3 or has norm equal to an integer prime, p, congruent to 1 mod 3. In either case, $N(\pi) \equiv 1 \ (3)$ so $\frac{N(\pi)-1}{3}$ is an integer. Now observe that the units $1, \omega$ and ω^2 are not congruent modulo π (or else $\pi \mid 1-\omega$) so since $x^3 - 1$ factors as $(x-1)(x-\omega)(x-\omega^2)$, for all $\alpha \in \mathbb{Z}[\omega]$ such that π doesn't divide α, $\alpha^{\frac{N(\pi)-1}{3}}$ is either $1, \omega$ or ω^2. Hence we can make the following definition for the cubic character modulo π on $\mathbb{Z}[\omega]$:

Definition 1.4 *Let π be a prime in $\mathbb{Z}[\omega]$ with $N(\pi) \neq 3$. The cubic character associated with π at the residue α, $\chi_\pi(\alpha)$, is 0 if $\pi \mid \alpha$ and otherwise is that member of $\{1, \omega, \omega^2\}$ such that*

$$\alpha^{\frac{N(\pi)-1}{3}} \equiv \chi_\pi(\alpha) \quad (\pi).$$

It follows from this definition that if $\alpha \equiv \beta \ (\pi)$ then $\chi_\pi(\alpha) = \chi_\pi(\beta)$ and that χ_π is multiplicative. Furthermore, it follows from the theory of finite fields that $\chi_\pi(\alpha) = 1$ iff $x^3 = \alpha$ has a solution in $(\mathbb{Z}[\omega])/\pi(\mathbb{Z}[\omega])$. (Note that in making this definition we have a generalization of the quadratic Euler criterion for cubic characters by definition.)

1.3.3 The Law of Cubic Reciprocity

With this definition of the cubic character, it is now possible to state and prove the Law of Cubic Reciprocity.

Theorem 1.7 *(Cubic Reciprocity) Let π_1 and π_2 be primary primes with norm not equal to 3 in $\mathbb{Z}[\omega]$. Then*

$$\chi_{\pi_1}(\pi_2) = \chi_{\pi_2}(\pi_1).$$

In the proof of the theorem, there are evidently three cases to consider: case 1 is when $\pi_1 = q_1, \pi_2 = q_2$ with q_1, q_2 integer primes congruent to 2 mod 3; case 2 is when $\pi_1 = q$ is an integer prime congruent to 2 mod 3 while $\pi_2 \mid p$ where p is an integer prime congruent to 1 mod 3; case 3 is when both π_1 and π_2 divide integer primes congruent to 1 mod 3. We can prove case 1 immediately. Before we prove cases 2 and 3, which are more difficult, we will introduce a preliminary lemma.

Proof. (Case 1) Given that $q_1, q_2 \equiv 2 \ (3)$, then $(3, q_i - 1) = 1$ so the function $f(x) = x^3$ is a permutation on both $(\mathbb{Z}/q_1\mathbb{Z})^\times$ and $(\mathbb{Z}/q_2\mathbb{Z})^\times$. Consequently both of the equations $x^3 \equiv q_2$ (q_1) and $x^3 \equiv q_1 \ (q_2)$ have a solution so $\chi_{q_1}(q_2) = \chi_{q_2}(q_1) = 1$. This completes case 1. □

Before considering cases 2 and 3, observe that if π is a non-integer prime with norm $p \equiv 1$ (3) then the finite field $(\mathbb{Z}[\omega])/\pi(\mathbb{Z}[\omega])$ contains p elements and hence is isomorphic to $\mathbb{Z}/p\mathbb{Z}$. The isomorphism is made precise by mapping the coset $a \bmod \pi$ to the coset $a \bmod p$ for $a = 0, 1, ..., p-1$. With this association in mind we can view the cubic character χ_π as a order three character on $\mathbb{Z}/p\mathbb{Z}$, and as a result the Gauss sum $g(\chi_\pi)$ and the Jacobian sum $J(\chi_\pi, \chi_\pi)$ are well defined. The lemma that we need to complete parts 2 and 3 of the proof of Cubic Reciprocity is as follows:

Lemma 1.8 *Let π be a primary prime in $\mathbb{Z}[\omega]$ with $N(\pi) = p \equiv 1 \ (3)$. Then $g(\chi_\pi)^3 = \pi p$.*

Proof. From Proposition 1.6, $g(\chi_\pi)^3 = p\chi_\pi(-1)J(\chi_\pi, \chi_\pi)$. Since -1 is always a perfect cube, $\chi_\pi(-1) = 1$ and our problem is reduced to showing that $J(\chi_\pi, \chi_\pi) = \pi$. We know from Proposition 1.5 that $J(\chi_\pi, \chi_\pi) = \frac{g(\chi_\pi)^2}{g(\chi_\pi^2)}$ so $|J(\chi_\pi, \chi_\pi)|^2 = p$. This implies that $J(\chi_\pi, \chi_\pi)$ is a prime. Now Ireland and Rosen prove (p. 96) that if χ is a cubic character and $J(\chi, \chi) = a + b\omega$ then $a \equiv 2 \ (3)$ and $b \equiv 0 \ (3)$. Thus $J(\chi_\pi, \chi_\pi)$ is primary. To complete the proof, observe

$$J(\chi_\pi, \chi_\pi) = \sum_x \chi_\pi(x)\chi_\pi(1-x) \equiv \sum_x x^{\frac{p-1}{3}}(1-x)^{\frac{p-1}{3}} \quad (\pi)$$

The latter sum, upon binomial expansion, has terms $a_k x^k$ where k ranges from $\frac{p-1}{3}$ to $\frac{2(p-1)}{3}$. But for $0 < k < p$, $\sum_x x^k \equiv 0 \ (p)$ since the sum can be rewritten as

$$\sum_{t=1}^{p-1} a^{tk} = \frac{a^{kp} - a^k}{a^k - 1}$$

and $a^{kp} = a^k$ by Fermat's Little Theorem while $a^k \neq 1$. Hence $p \mid \sum_x x^{\frac{p-1}{3}}(1-x)^{\frac{p-1}{3}}$ implying $\pi \mid \sum_x x^{\frac{p-1}{3}}(1-x)^{\frac{p-1}{3}}$ and $\pi \mid J(\chi_\pi, \chi_\pi)$. Then as $J(\chi_\pi, \chi_\pi)$ is a primary prime, the only possibility is that $J(\chi_\pi, \chi_\pi) = \pi$ completing the proof of the lemma. □

With this lemma we are now ready to prove cases 2 and three of Cubic Reciprocity. *Proof.* (Case 2) We know from the lemma that $g(\chi_\pi)^3 = p\pi$. Raising both sides of the equation to the $\frac{q^2-1}{3}$ power and working modulo q we have

$$(p\pi)^{\frac{q^2-1}{3}} \equiv \chi_q(p\pi) \pmod{q}.$$

From our argument in case 1 we know that $\chi_q(p) = 1$ so $g(\chi_\pi)^{q^2-1} \equiv \chi_q(\pi) \pmod{q}$. Hence

$$g(\chi_\pi)^{q^2} \equiv g(\chi_\pi)\chi_q(\pi) \pmod{q}.$$

On the other hand, expanding $g(\chi_\pi)^{q^2}$ by the multinomial theorem, we find that

$$\Big(\sum_t \chi_\pi(t)\zeta^t\Big)^{q^2} \equiv \sum_t (\chi_\pi(t)\zeta^t)^{q^2} \equiv \sum_t \chi_\pi(t)^{q^2}\zeta^{q^2 t} \pmod{q}.$$

Since $q \equiv 2 \ (3), q^2 \equiv 1 \ (3)$ and since $\chi_\pi(t)^3 = 1$, $\chi_\pi(t)^{q^2} = \chi_\pi(t)$. Hence,

$$g(\chi_\pi)^{q^2} \equiv g_{q^2}(\chi_\pi) \equiv \overline{\chi_\pi(q^2)} g(\chi_\pi) \quad (q).$$

Now $\chi_\pi(q^2)$ is a third root of unity so $\overline{\chi_\pi(q^2)} = \chi_\pi(q^2)^2 = \chi_\pi(q)^4 = \chi_\pi(q)$. Thus, collecting our results, we have

$$g(\chi_\pi)\chi_q(\pi) \equiv g(\chi_\pi)\chi_\pi(q) \pmod q$$

or multiplying each side by $g(\chi_\pi)^2$ and canceling $p\pi$ from each side, $\chi_q(\pi) \equiv \chi_\pi(q) \pmod q$ which implies

$$\chi_q(\pi) = \chi_\pi(q),$$

the desired result. □

The proof in case 3 follows from the same techniques as the proof in case 2, but with a little more bookkeeping.

Proof. We have π_1 and π_2 primary complex primes with $N(\pi_1) = p_1 \equiv 1 \ (3)$ and $N(\pi_2) = p_2 \equiv 1 \ (3)$. Let $\gamma_1 = \overline{\pi_1}$ and $\gamma_2 = \overline{\pi_2}$. Observe that both γ_1 and γ_2 are primary primes. Using the technique from case 2, we start with $g(\chi_{\gamma_1})^3 = \gamma_1 p_1$ and raise each side to the $\frac{N(\pi_2)-1}{3} = \frac{p_2-1}{3}$ power working mod π_2 to find $g(\chi_{\gamma_1})^{p_2-1} \equiv \chi_{\pi_2}(\gamma_1 p_1) \ (\pi_2)$, so

$$g(\chi_{\gamma_1})^{p_2} \equiv g(\chi_{\gamma_1})\chi_{\pi_2}(\gamma_1 p_1) \quad (\pi_2).$$

Expanding $g(\chi_{\gamma_1})^{p_2}$ using the multinomial theorem we have

$$\Big(\sum_t \chi_{\gamma_1}(t)\zeta^t\Big)^{p_2} \equiv \sum_t \chi_{\gamma_1}(t)^{p_2}\zeta^{p_2 t} \quad (\pi_2).$$

Since $p_2 \equiv 1 \ (3)$ and χ_{γ_1} is a third root of unity, $\chi_{\gamma_1}^{p_2} = \chi_{\gamma_1}$ so the latter sum becomes $g_{p_2}(\chi_{\gamma_1}) = \chi_{\gamma_1}(p_2)^{-1} g(\chi_{\gamma_1})$. Comparing this with the above result, we have $g(\chi_{\gamma_1})\chi_{\pi_2}(\gamma_1 p_1) \equiv g(\chi_{\gamma_1})\chi_{\gamma_1}(p_2)^{-1} \ (\pi_2)$. Upon multiplying each side by $\overline{g(\chi_{\gamma_1})}$ and canceling the resultingp_1 from each side,

$$\chi_{\pi_2}(\gamma_1 p_1) \equiv \chi_{\gamma_1}(p_2)^{-1} \quad (\pi_2)$$

so $\chi_{\pi_2}(\gamma_1 p_1) = \chi_{\gamma_1}(p_2)^{-1}$ since $1, \omega$ and ω^2 are distinct modulo π_2.

Using an identical approach but starting with the identity $g(\chi_{\pi_2})^3 = p_2\pi_2$ and working modulo π_1 we also have

$$\chi_{\pi_2}(p_1)^{-1} = \chi_{\pi_1}(\pi_2 p_2).$$

The two equations, $\chi_{\pi_2}(\gamma_1 p_1) = \chi_{\gamma_1}(p_2)^{-1}$ and $\chi_{\pi_2}(p_1)^{-1} = \chi_{\pi_1}(\pi_2 p_2)$ give us enough information to prove the result, but we need to know how to work with the inverse of characters. For $\chi_{\gamma_1}(p_2)^{-1} = \overline{\chi_{\gamma_1}(p_2)}$ observe that

$$\chi_{\gamma_1}(p_2) \equiv p_2^{\frac{p_1-1}{3}} \quad (\gamma_1)$$

so

$$\overline{\chi_{\gamma_1}(p_2)} \equiv \overline{p_2}^{\frac{p_1-1}{3}} \quad (\overline{\gamma_1}).$$

But $p_2 = \overline{p_2}$, $\overline{\gamma_1} = \pi_1$ and $N(\pi_1) = N(\gamma_1)$ so the above line reduces to

$$\overline{\chi_{\gamma_1}(p_2)} \equiv \chi_{\pi_1}(p_2) \quad (\pi_1),$$

which implies that equality actually holds. For $\chi_{\pi_2}(p_1)^{-1}$ observe that this is a cube root of unity so

$$\chi_{\pi_2}(p_1)^{-1} = \chi_{\pi_2}(p_1)^2 = \chi_{\pi_2}(p_1^2).$$

Now working with $\chi_{\pi_2}(\gamma_1 p_1) = \chi_{\pi_1}(p_2)$ and $\chi_{\pi_2}(p_1^2) = \chi_{\pi_1}(\pi_2 p_2)$, we have

$$\begin{aligned} \chi_{\pi_1}(\pi_2)\chi_{\pi_2}(p_1\gamma_1) &= \chi_{\pi_1}(\pi_2)\chi_{\pi_1}(p_2) \\ &= \chi_{\pi_1}(\pi_2 p_2) = \chi_{\pi_2}(p_1^2) \\ &= \chi_{\pi_2}(p_1\pi_1\gamma_1) = \chi_{\pi_2}(\pi_1)\chi_{\pi_2}(p_1\gamma_1) \end{aligned}$$

Canceling $\chi_{\pi_2}(p_1\gamma_1)$ from the beginning and the end of our equality, we have the desired result. □

Borrowing from Legendre, in common usage the cubic character symbol modulo prime π is often denoted $\left(\frac{\cdot}{\pi}\right)_3$. As an easy example of the use of the cubic reciprocity law, we determine $\left(\frac{7}{29}\right)_3$. Observe that 29 is primary and 7 is factored into primary primes as $(2+3\omega)(-1-3\omega)$. Hence

$$\left(\frac{7}{29}\right)_3 = \left(\frac{2+3\omega}{29}\right)_3\left(\frac{-1-3\omega}{29}\right)_3 = \left(\frac{29}{2+3\omega}\right)_3\left(\frac{29}{-1-3\omega}\right)_3$$

by the cubic reciprocity law. But $29 \equiv 1 \ (7)$ so $29 \equiv 1$ modulo both $2+3\omega$ and $-1-3\omega$. Since 1 is always a perfect cube, both $\left(\frac{29}{2+3\omega}\right)_3$ and $\left(\frac{29}{-1-3\omega}\right)_3$ are 1 and $\left(\frac{7}{29}\right)_3 = 1$. It requires rather intensive computation to verify, by checking the cube of each residue mod 29, that $16^3 = 4096 \equiv 7 \mod 29$ so 7 is indeed a perfect cube.

In general, applying cubic reciprocity to evaluate cubic characters is not quite so easy as in the above example. To evaluate $\left(\frac{\mu}{\pi}\right)_3$ for primary prime π and arbitrary element $\mu \in \mathbb{Z}[\omega]$, one first factors μ into powers of $-1, \omega, \gamma, \lambda_1, \lambda_2, ..., \lambda_t$ where γ is the prime $1-\omega$ and the λ_i are primary primes (the fact that every prime in $\mathbb{Z}[\omega]$ has a primary associate guarantees μ can be factored in this way). One applies cubic reciprocity to reduce each of the characters $\left(\frac{\lambda_i}{\pi}\right)_3$ to $\left(\frac{\pi}{\lambda_i}\right)_3$. For any primary prime λ one always has $\left(\frac{-1}{\lambda}\right)_3 = 1$. Writing $\lambda = a + b\omega$ with $a = 3m-1$, $b = 3n$ Eisenstein proved that $\left(\frac{\gamma}{\lambda}\right)_3 = \omega^{2m}$ (See [4] pp.114-115). Recalling the definition of the cubic character, one also has $\left(\frac{\omega}{\lambda}\right)_3 = \omega^{\frac{a^2+b^2-ab-1}{3}}$, which, upon substitution and reduction modulo 3 in the exponent, gives $\left(\frac{\omega}{\lambda}\right)_3 = \omega^{m+n}$. Armed with these facts, the cubic character is always computable in a manner still much easier than cubing each of the residues modulo π.

1.4 Biquadratic reciprocity

In almost every respect, the theory of biquadratic reciprocity mirrors that of cubic reciprocity so we will only outline the proof, indicating where the two approaches diverge (for a complete treatment, see [4], pp.119-127). As in the cubic case, we begin by making a field extension from $\mathbb{Q}$ to $\mathbb{Q}(i)$ so that the quadratic character is contained in the field in question. We then work in the associated ring of integers $\mathbb{Z}[i]$. Once again we define the norm N on $\mathbb{Z}[i]$ via multiplication by the complex conjugate: $N(\alpha) = \alpha\overline{\alpha}$. The norm is multiplicative and the units of $\mathbb{Z}[i]$ are the elements of norm 1, namely $1, -1, i, -i$. This norm makes $\mathbb{Z}[i]$ into a unique factorization domain.

As in other unique factorization domains, we do not wish to distinguish between associate elements of $\mathbb{Z}[i]$ that are related by units so again we introduce the notion of *primary* elements. The element $\alpha = a + bi \in \mathbb{Z}[i]$ is defined to be primary if and only if either $a \equiv 1$ $(4), b \equiv 0$ (4) or $a \equiv 3$ $(4), b \equiv 2$ (4). Provided that $(1 + i)$ does not divide α, exactly one of α's four associates is primary. Since in our discussion of biquadratic reciprocity, we will only be interested odd primes (primes not divisible by 2 and consequently not divisible by $(1 + i)$) this is a useful definition of primary elements.

In $\mathbb{Z}[i]$ as in $\mathbb{Z}[\omega]$ some rational primes split while others remain irreducible. 2 factors as $(1+i)(1-i)$ and odd primes $p \equiv 1 \mod 4$ split into irreducibles $\pi\overline{\pi}$ with $N(\pi) = N(\overline{\pi}) = p$. Primes $q \equiv 3 \mod 4$ remain irreducible. Up to multiplication by unit, these are all the irreducibles in $\mathbb{Z}[i]$.

Like $\mathbb{Z}[\omega]$, because $\mathbb{Z}[i]$ is a principle ideal domain, maximal and prime ideals coincide, so for irreducible $\pi \in \mathbb{Z}[i]$, $\mathbb{Z}[i]/(\pi)$ is a field. This resulting finite field has order $N(\pi)$ giving the analog to Fermat's Little Theorem: $\alpha^{N(\pi)-1} \equiv 1$ (π) for $\alpha \in (\mathbb{Z}[i]/(\pi))^{\times}$. Now for irreducible π, not an associate of $(1 + i)$ the four units $1, i, -1, -i$ are distinct and we either have $N(\pi) = p \equiv 1$ (4) or π is a rational prime $q \equiv 3$ (4) with norm $q^2 \equiv 1$ (4). In either case, $N(\pi) \equiv 1$ (4), enabling us to make the following definition of the quartic character,

Definition 1.5 *Let π be an irreducible in $\mathbb{Z}[i]$ not divisible by $(1 + i)$ and let $\alpha \in \mathbb{Z}[i]$. The quadratic character of α mod π, $\chi_\pi(\alpha)$, is zero if $\pi \mid \alpha$ and is i^j such that $\pi \mid \alpha^{\frac{N(\pi)-1}{4}} - i^j$ otherwise.*

Ireland and Rosen prove that this character has the expected properties, namely, $\chi_\pi(\alpha) = 1$ iff α is a fourth power mod π, the character is multiplicative, its value depends only upon the residue of α mod π, and π and λ have the same characters if they are associates. They also note that $\chi_\pi(-1) = (-1)^{\frac{a-1}{2}}$ if π is a primary irreducible equal to $a + bi$ (this follows from the definition of a primary element, working mod 4).

With this definition we are able to formulate a quadratic reciprocity law, but for the purposes of proving the law it is most convenient to first generalize the quartic character to all elements of $\mathbb{Z}[i]$ not divisible by $(1 + i)$ (this represents the only major digression from the proof of cubic reciprocity). This generalization is achieved by setting

$$\chi_\alpha(\beta) = \prod_i \chi_{\pi_i}(\beta)$$

where α is an arbitrary non-unit of $\mathbb{Z}[i]$ not divisible by $(1+i)$ with prime factorization $\prod \pi_i$. Having made this definition, we now state the law of biquadratic reciprocity.

Theorem 1.9 *(Biquadratic Reciprocity) Let α and β be primary elements of $\mathbb{Z}[i]$ that are relatively prime. Then*

$$\chi_\alpha(\beta) = \chi_\beta(\alpha)(-1)^{((N(\alpha)-1)/4)((N(\beta)-1)/4)}.$$

As was the case in both the proofs of quadratic and cubic reciprocity, at the heart of the proof of biquadratic reciprocity is the computation of the value of a Gauss sum; the remainder of the proof is built upon this computation. The Gauss sum is introduced by taking an irreducible, $\pi \in \mathbb{Z}[i]$ of norm $p \equiv 1 \ (4)$ and viewing the field $\mathbb{Z}/(\pi)$ containing p elements as $\mathbb{Z}/p\mathbb{Z}$. The quartic character, χ_π, then becomes an order 4 character on $\mathbb{Z}_p$ so we can consider its Gauss sum, $g(\chi_\pi)$. We will compute $g(\chi_\pi)^4$. To do this, recall that we proved in our section on Jacobi sums that if χ and λ are non-trivial characters with $\chi\lambda$ also non-trivial then

$$J(\chi, \lambda) = \frac{g(\chi)g(\lambda)}{g(\chi\lambda)}.$$

Applying this identity, we have

$$J(\chi_\pi, \chi_\pi)^2 = \frac{g(\chi_\pi)^4}{g(\chi_\pi^2)^2}.$$

But $p \equiv 1 \ (4)$ and χ_π is a quartic character so χ_π^2 is a quadratic character, i.e. the Legendre symbol, so $g(\chi_\pi^2)^2 = p$ and $g(\chi_\pi)^4 = pJ(\chi_\pi, \chi_\pi)$. We already know, from our results on Jacobi sums, that $N(J(\chi_\pi, \chi_\pi)) = p$. Using almost identical techniques to those used to determine $J(\chi, \chi)$ in the proof of cubic reciprocity, one first shows that $-\chi_\pi(-1)J(\chi_\pi, \chi_\pi)$ is primary and then concludes that $-\chi_\pi(-1)J(\chi_\pi, \chi_\pi) = \pi$. Since $-\chi_\pi(-1)$ is 1 or -1 this implies $J(\chi_\pi, \chi_\pi)^2 = \pi^2$ so

$$g(\chi_\pi)^4 = p\pi^2 = \pi^3\overline{\pi}.$$

We are now ready to prove the substance of biquadratic reciprocity, which comes in two lemmas.

Lemma 1.10 *(Part 1) Let $q > 0$, $q \equiv 3 \ (4)$ be a prime in $\mathbb{Z}$. Then*

$$\chi_\pi(-q) = \chi_q(\pi).$$

Proof. Working modulo q,

$$g(\chi_\pi)^q \equiv \sum_t \chi_\pi(t)^q \zeta^{qt} \quad (q).$$

Since $\chi_\pi(t)$ is a fourth root of unity and $q \equiv 3 \ (4)$, $\chi_\pi(t)^q = \chi_\pi(t)^3 = \overline{\chi}_\pi(t)$. Thus,

$$\sum_t \chi_\pi(t)^q \zeta^{qt} = g_q(\overline{\chi}_\pi) = \chi_\pi(q) g(\overline{\chi}_\pi).$$

Multiplying each side by $g(\chi_\pi)$ yields

$$g(\chi_\pi)^{q+1} = (g(\chi_\pi)^4)^{(q+1)/4} \equiv \chi_\pi(q) g(\chi_\pi) g(\overline{\chi}_\pi) \quad (q).$$

Now $g(\chi_\pi)^4 = \pi^3\overline{\pi}$ and $g(\overline{\chi}_\pi) = \chi_\pi(-1)\overline{g(\chi_\pi)}$ and $g(\chi_\pi)\overline{g(\chi_\pi)} = p = \pi\overline{\pi}$ so the above congruence reduces to

$$(\pi^3\overline{\pi})^{(q+1)/4} \equiv \chi_\pi(q)\chi_\pi(-1)\pi\overline{\pi} \quad (q).$$

But for a general element $\lambda \in \mathbb{Z}[i]$, $\lambda^q \equiv \overline{\lambda} \ (q)$, a fact that we will prove shortly. Assuming this congruence, we then have:

$$(\pi^{(q+3)})^{(q+1)/4} \equiv \chi_\pi(-q)\pi^{q+1} \quad (q)$$

or, canceling the π^{q+1} from each side,

$$\pi^{(q^2-1)/4} \equiv \chi_\pi(-q) \quad (q).$$

Finally, recalling the definition, that $\pi^{(q^2-1)/4} \equiv \chi_q(\pi) \ (q)$ we arrive at

$$\chi_q(\pi) \equiv \chi_\pi(-q) \quad (q).$$

Since each side is a unit, equality actually holds and this gives our desired result.

It remains to prove the fact that $\lambda^q \equiv \overline{\lambda} \ (q)$. For this, let $\lambda = a + bi$ and consider the two values $(\lambda + \overline{\lambda})^q$ and $(\lambda - \overline{\lambda})^q$ modulo q. Since we are working mod q, $(\lambda + \overline{\lambda})^q \equiv \lambda^q + \overline{\lambda}^q$ (q) and $(\lambda - \overline{\lambda})^q \equiv \lambda^q - \overline{\lambda}^q \ (q)$. But $\lambda + \overline{\lambda} = 2a$ is an integer, so Fermat's Little Theorem guarantees that $(\lambda + \overline{\lambda})^q \equiv \lambda + \overline{\lambda}$. Meanwhile, $\lambda - \overline{\lambda} = 2bi$ so $(\lambda - \overline{\lambda})^q \equiv (2b)^q i^q \ (q)$. Again by Fermat's Little Theorem, $(2b)^q \equiv 2b \ (q)$ while $i^q = -i$ since $q \equiv 3 \ (4)$. Thus $(\lambda - \overline{\lambda})^q \equiv -2bi = -(\lambda - \overline{\lambda}) \ (q)$. Combining results, we have $\lambda^q + \overline{\lambda}^q \equiv \lambda + \overline{\lambda} \ (q)$ while $\lambda^q - \overline{\lambda}^q \equiv -(\lambda - \overline{\lambda}) \ (q)$. Together these imply $\lambda^q \equiv \overline{\lambda} \ (q)$. This completes the proof of part 1. □

Lemma 1.11 *(Part 2) Let $q \equiv 1 \ (4)$ be a prime in $\mathbb{Z}$. Then*

$$\chi_\pi(q) = \chi_q(\pi).$$

Proof. We observe that each side of the equality is trivially zero if $\pi \mid q$ so we may assume that π and q are relatively prime. Once again working modulo q,

$$g(\chi_\pi)^q \equiv \sum_t \chi_\pi(t)^q \zeta^{qt} = \sum_t \chi_\pi(t)\zeta^{qt} = g_q(\chi_\pi) \quad (q),$$

where in the first equality we have made use of the fact that $\chi_\pi(t)$ is a fourth root of unity and $q \equiv 1 \ (4)$. Thus $g(\chi_\pi)^q \equiv \overline{\chi}_\pi(q)g(\chi_\pi) \ (q)$, or, multiplying each side by $g(\chi_\pi)^3$,

$$g(\chi_\pi)^{q+3} \equiv \overline{\chi}_\pi(q)g(\chi_\pi)^4 \quad (q).$$

Recalling that $g(\chi_\pi)^4 = \pi^3\overline{\pi}$ gives

$$(\pi^3\overline{\pi})^{(q+3)/4} \equiv \overline{\chi}_\pi(q)\pi^3\overline{\pi} \quad (q).$$

Now since π and q are relatively prime, we can cancel $\pi^3\overline{\pi}$ from each side, resulting in

$$(\pi^3)^{(q-1)/4}(\overline{\pi})^{(q-1)/4} \equiv \overline{\chi}_\pi(q) \quad (q).$$

If we write $q = \lambda\overline{\lambda}$ where λ is an irreducible in $\mathbb{Z}[i]$ then the above congruence implies that

$$\chi_\lambda(\pi^3)\chi_\lambda(\overline{\pi}) \equiv \overline{\chi_\pi(q)} \quad (\lambda).$$

Since both sides of the congruence are units, equality holds. Finally, capitalizing on the properties of the quartic character, $\chi_\lambda(\pi^3) = \chi_\lambda(\pi)^3 = \overline{\chi_\lambda(\pi)}$ and $\chi_\lambda(\overline{\pi}) = \overline{\chi_{\overline{\lambda}}(\pi)}$ so the equality can be rewritten as

$$\overline{\chi_\lambda(\pi)\chi_{\overline{\lambda}}(\pi)} = \overline{\chi}_\pi(q).$$

Conjugating each side then gives $\chi_q(\pi) = \chi_\pi(q)$, completing part 2. □

Having proven these two lemmas, only a handful of additional facts are needed to prove biquadratic reciprocity. One readily observes that if q is a prime $q \equiv 3 \ (4)$ and a is an integer not divisible by q then $\chi_q(a) = 1$. This holds since the function $f(x) = x^4$ is a permutation on $\mathbb{Z}_q$. This fact can then be extended to show that if a and α are relatively prime integers with a odd, then $\chi_a(\alpha) = 1$. One factors a into primes $a = \prod p_i \prod q_i$ with the $p_i \equiv 1 \ (4)$ and the $q_i \equiv 3 \ (4)$. Then $\chi_{q_i}(\alpha) = 1$ and $\chi_{p_i}(\alpha) = \chi_\pi(\alpha)\chi_{\overline{\pi}}(\alpha) = \chi_\pi(\alpha)\overline{\chi_\pi(\alpha)} = 1$. Ireland and Rosen compute that $\chi_n(i) = (-1)^{(n-1)/4}$ for integers $n \equiv 1 \ (4)$ by considering the two cases of $\chi_p(i)$ and $\chi_{-q}(i)$ for primes $p \equiv 1 \ (4)$ and $q \equiv 3 \ (4)$. Together with the two lemmas we proved, these facts imply biquadratic reciprocity in the case of an integer $a \equiv 1 \ (4)$ and a primary element λ. One then extends this to prove biquadratic reciprocity in the case of relatively prime primary elements $\pi = a+bi$ and $\lambda = c+di$ in which $(a, b) = 1$ and $(c, d) = 1$. This then implies the general statement. The proofs of these three cases are not particularly enlightening and are omitted here. Full details can be found in [4] pp. 126-127.

Chapter 2

Gauss Sums, Field Theory and Fourier Analysis

In our proofs of quadratic, cubic and biquadratic reciprocity in the previous chapter, the numerical value of the Gauss sum played an integral, albeit somewhat mysterious, role. In order to understand this role, as well as to prove the more general Eisenstein reciprocity law, field theory is needed.

2.1 Cyclotomic extensions and the duality between χ and $g(\chi)$

Thus far, we have considered Gauss sums in the following context: p is a prime and χ is a character of order $k \mid p-1$ mapping $(\mathbb{Z}/p\mathbb{Z})^{\times}$ to the kth roots of unity. The Gauss sum $g(\chi)$ is given by $\sum \chi(t)\zeta^t$ where $\zeta = e^{2\pi i/p}$. If we let $\zeta_k = e^{2\pi i/k}$ then χ takes on values in the field $\mathbb{Q}(\zeta_k)$, which is a degree $\phi(k)$ extension over $\mathbb{Q}$ (where ϕ denotes Euler's function). The value of the Gauss sum is thus an element of the field $\mathbb{Q}(\zeta_k, \zeta)$. Since ζ is a pth root of unity while ζ_k is a kth root of unity, and p and k are relatively prime, $\mathbb{Q}(\zeta_k, \zeta)$ is a degree $\phi(p) = p-1$ extension over $\mathbb{Q}(\zeta_k)$ and the field $\mathbb{Q}(\zeta_k, \zeta)$ can be viewed as a $(p-1)$-dimensional vector space over $\mathbb{Q}(\zeta_k)$ with basis $\{\zeta, \zeta^2, ..., \zeta^{p-1}\}$. The Gauss sum is a linear combination of basis elements in this vector space and we know that any two distinct linear combinations must take on different values. Thus the value of the Gauss sum uniquely determines the coefficients $\chi(1), \chi(2), ...\chi(p-1)$ in the sum $g(\chi) = \sum \chi(t)\zeta^t$. This justifies our claim in the discussion following the proof of quadratic reciprocity that the value of the quadratic Gauss sum contains all the necessary information to deduce $\left(\frac{a}{p}\right)$ for all residues $a \mod p$.

In our three reciprocity proofs, we did not determine the value of the Gauss sum, g, however, but rather g^k. Still, in doing so we did not actually limit our ability to determine χ. Knowing the value of g^k guarantees that the value of g is among k values, $g_1, g_2, ..., g_k$, that differ by kth roots of unity. The bijection between (p-1)-dimensional vectors over $\mathbb{Q}(\zeta_k)$ and the field $\mathbb{Q}(\zeta_k, \zeta)$ guarantees that each of these values corresponds to exactly one set of coefficients in $\mathbb{Q}(\zeta_k)$ for the basis elements $\zeta, \zeta^2, ..., \zeta^{p-1}$ and since g is among

$g_1, g_2, ..., g_k$ we know that $[\chi(1), \chi(2), ..., \chi(p-1)]$ is one of the sets of coefficients. Since each of the g_i is a kth root of unity times g, the other sets of coefficients must be given by $[\chi(1)\zeta_k^j, \chi(2)\zeta_k^j, ..., \chi(p-1)\zeta_k^j]_{j=1,2,...,k-1}$ and together with the values of χ, these are all sets of coefficients. Now $\chi(1) = 1$ and this last observation shows that exactly one set of coefficients for the g_i will have coefficient 1 on ζ, namely, the set for g. Thus, to determine $[\chi(1), \chi(2), ..., \chi(p-1)]$ from the value of g^k, one simply computes the set of coefficients for g_1, g_2, up to g_k and chooses the one that has coefficient 1 on ζ. Thus the value of $g(\chi)^k$ together with the knowledge that $\chi(1) = 1$ effectively contains all the information needed to determine χ. This makes the Gauss sum a natural tool for proving theorems of reciprocity.

2.2 Fourier analysis on $\mathbb{Z}_p$

While the determination of χ from the value of $g(\chi)$ is theoretically possible, e.g. via an exhaustive search, there is no practical known method for doing so (if there were, reciprocity laws would be dispensable!). In effect, what each reciprocity proof in chapter 1 accomplished was to extract some knowledge about a character from the value of the Gauss sum, by working modulo a prime. Another promising method of extracting information from a Gauss sum is via a discrete analog of Fourier Analysis.

2.2.1 Fourier coefficients and Fourier expansion

If we define the inner product

$$< f, g >= \frac{1}{p} \sum_{t=0}^{p-1} f(t)g(-t)$$

for complex valued functions f and g on $\mathbb{Z}_p$, then the set of functions $\{e^{2\pi int/p}\}_{n=0,1,2,...,p-1}$ on $\mathbb{Z}_p$ plays a role analogous to that of the set of functions $\{e^{2\pi inx}\}_{n=...-2,-1,0,1,2,...}$ on the real line in the sense that they form an orthonormal basis for complex valued functions on $\mathbb{Z}_p$. Specifically, if $f : \mathbb{Z}_p \to \mathbb{C}$ and $\zeta = e^{2\pi i/p}$, setting

$$\hat{f}_n = \frac{1}{p} \sum_{s=0}^{p-1} f(s)\zeta^{-ns},$$

then

$$\sum_{n=0}^{p-1} \hat{f}_n \zeta^{nt} = \frac{1}{p} \sum_{n,s} f(s)\zeta^{n(t-s)} = \frac{1}{p} \sum_{s} f(s) \sum_{n} \zeta^{n(t-s)}.$$

The inner sum is p if $t = s$ and 0 otherwise, so we retrieve:

$$\sum_{n=0}^{p-1} \hat{f}_n \zeta^{nt} = \frac{1}{p} \sum_{t} f(t)p = f(t).$$

The values $\hat{f}_n$ are the Fourier coefficients for the function f.

2.2.2 Two exercises in Fourier techniques

As an exercise in the use of discrete Fourier transforms, we take the following two results:

Proposition 2.1 *For any character, χ, on $\mathbb{Z}_p$,*

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

Proof. Let $f = \chi$. Then, computing the Fourier coefficients of f,

$$\hat{f}_0 = \frac{1}{p}\sum_t \chi(t) = 0,$$

and for $1 \leq n \leq p-1$

$$\hat{f}_n = \frac{1}{p}\sum_s \chi(s)\zeta^{-ns} = \frac{1}{p}\chi(-1)\sum \chi(-s)\zeta^{-ns} = \frac{1}{p}\chi(-1)g_n(\chi).$$

Now $g_n(\chi) = \overline{\chi}(n)g(\chi)$, so applying the Fourier transform,

$$f(t) = \sum_n \hat{f}_n\zeta^{nt} = \frac{\chi(-1)g(\chi)}{p}\sum_n \overline{\chi}(n)\zeta^{nt},$$

which reduces to

$$f(t) = \frac{\chi(-1)g(\chi)g_t(\overline{\chi})}{p}.$$

But, $g_t(\overline{\chi}) = \chi(t)g(\overline{\chi})$ so canceling $f(t) = \chi(t)$ from each side and multiplying each side by p, we arrive at

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

our desired result. □

This fact, which also follow from our earlier computations that $|g(\chi)| = \sqrt{p}$ and $g(\overline{\chi}) = \chi(-1)\overline{g(\chi)}$, is an easy consequence of the Fourier analysis, and is one of two known multiplicative identities on Gauss sums (the other being the Hasse-Davenport identity that we will introduce shortly). Our second result follows from similar techniques:

Proposition 2.2 *Let χ be a character on $\mathbb{Z}_p$ and $\zeta = e^{2\pi i/p}$. Then*

$$\sum_{n=0}^{p-1} n\chi(n) = g(\chi)\sum_{n=0}^{p-1}\frac{\overline{\chi}(n)}{1-\zeta^n}.$$

Proof. Let $f(t) = \sum_{j=0}^{t} \chi(t)$. Then the Fourier coefficients for f are given by

$$\hat{f}_0 = \frac{1}{p}\sum_{s=0}^{p-1} f(s) = \frac{1}{p}\sum_{s=0}^{p-1}\sum_{j=0}^{s}\chi(j) = \frac{1}{p}\sum_{j=0}^{p-1}(p-j)\chi(j).$$

And for $1 \leq n \leq p-1$

$$\hat{f}_n = \frac{1}{p}\sum_{s} f(s)\zeta^{-ns} = \frac{1}{p}\sum_{s=0}^{p-1}\sum_{j=0}^{s}\chi(j)\zeta^{-ns}.$$

Exchanging the order of summation, this becomes

$$\frac{1}{p}\sum_{j=0}^{p-1}\chi(j)\sum_{s=j}^{p-1}\zeta^{-ns},$$

or, recalling that $\sum_{a=0}^{p-1}\zeta^{-ns} = 0$, we may write this as

$$-\frac{1}{p}\sum_{j=0}^{p-1}\chi(j)\sum_{s=0}^{j-1}\zeta^{-ns}.$$

But

$$\sum_{s=0}^{j-1}\zeta^{-ns} = \frac{\zeta^{-nj}-1}{\zeta^{-n}-1}$$

so

$$\hat{f}_n = \frac{1}{p(1-\zeta^{-n})}\sum_{j}\chi(j)(\zeta^{-nj}-1).$$

Since $\sum \chi(j) = 0$ and $\sum \chi(j)\zeta^{-nj} = \chi(-1)\sum \chi(-j)\zeta^{-nj} = \overline{\chi}(-1)g_n(\chi)$ and $g_n(\chi) = \overline{\chi}(n)g(\chi)$, we finally have that

$$\hat{f}_n = \frac{\overline{\chi}(-n)g(\chi)}{p(1-\zeta^{-n})}$$

for $1 \leq n \leq p-1$.

Now analyzing the Fourier expansion $f(t) = \sum_n \hat{f}_n\zeta^{nt}$, we have

$$\begin{aligned}\sum_{j=0}^{t}\chi(j) &= \hat{f}_0 + \sum_{n=1}^{p-1}\hat{f}_n\zeta^{nt} \\ &= \frac{1}{p}\sum_{j=0}^{p-1}(p-j)\chi(j) + \frac{1}{p}\sum_{n=1}^{p-1}\frac{\overline{\chi}(-n)g(\chi)}{1-\zeta^{-n}}\zeta^{nt} \\ &= \frac{1}{p}\sum_{j=0}^{p-1}(p-j)\chi(j) + \frac{\overline{\chi}(-1)g(\chi)}{p}\sum_{n=1}^{p-1}\frac{\overline{\chi}(n)\zeta^{nt}}{1-\zeta^{-n}}\end{aligned}$$

But now $\frac{1}{1-\zeta^{-n}} = 1+\zeta^{-n}+\zeta^{-2n}+...+\zeta^{-tn}+\frac{\zeta^{-(t+1)n}}{1-\zeta^{-n}}$, so

$$\sum_{n=1}^{p-1}\frac{\overline{\chi}(n)\zeta^{nt}}{1-\zeta^{-n}} = \sum_{n=1}^{p-1}\overline{\chi}(n)\zeta^{nt}\left(1+\zeta^{-n}+\zeta^{-2n}+...+\zeta^{-tn}+\frac{\zeta^{-(t+1)n}}{1-\zeta^{-n}}\right)$$

or, by the distributive law,

$$= \sum_n \overline{\chi}(n)\zeta^{nt} + \sum_n \overline{\chi}(n)\zeta^{n(t-1)} + \sum_n \overline{\chi}(n)\zeta^{n(t-2)} + ... \sum_n \overline{\chi}(n)\zeta^0 + \sum_n \frac{\overline{\chi}(n)\zeta^{-n}}{1-\zeta^{-n}}$$

which gives, by collecting Gauss sums,

$$= g_t(\overline{\chi}) + g_{t-1}(\overline{\chi}) + g_{t-2}(\overline{\chi}) + ... + g_0(\overline{\chi}) + \sum_n \frac{\overline{\chi}(n)\zeta^{-n}}{1-\zeta^{-n}}$$

and finally

$$= g(\overline{\chi})\sum_{j=0}^{t}\chi(j) + \sum_n \frac{\overline{\chi}(n)\zeta^{-n}}{1-\zeta^{-n}}$$

from the identity $g_a(\overline{\chi}) = \chi(a)g(\overline{\chi})$.

Thus we have

$$\sum_{j=0}^{t}\chi(j) = \frac{1}{p}\sum_{j=0}^{p-1}(p-j)\chi(j) + \frac{\overline{\chi}(-1)g(\chi)}{p}\left(g(\overline{\chi})\sum_{j=0}^{t}\chi(j) + \sum_n \frac{\overline{\chi}(n)\zeta^{-n}}{1-\zeta^{-n}}\right).$$

Recalling that we proved in the previous proposition that $\chi(-1)g(\chi)g(\overline{\chi}) = p$, the right hand side collapses leaving

$$\sum_{j=0}^{t}\chi(j) = \frac{1}{p}\sum_{n=0}^{p-1}(p-n)\chi(n) + \sum_{j=0}^{t}\chi(j) + \frac{\overline{\chi}(-1)g(\chi)}{p}\left(\sum_{n=0}^{p-1}\frac{\overline{\chi}(n)\zeta^{-n}}{1-\zeta^{-n}}\right).$$

After canceling terms and multiplying each side by $\chi(-1)p$, we are left with

$$\sum_{n=0}^{p-1}(p-n)\chi(-n) = -g(\chi)\sum_{n=0}^{p-1}\frac{\overline{\chi}(n)\zeta^{-n}}{1-\zeta^{-n}}$$

which, after the change of variables $n \rightarrow p-n$ on the left and multiplying both top and bottom of the fraction on the right by ζ^n and switching sign, leaves

$$\sum_{n=0}^{p-1}n\chi(n) = g(\chi)\sum_{n=0}^{p-1}\frac{\overline{\chi}(n)}{1-\zeta^n},$$

the desired result. □

This last proposition is of interest because in the case that $\chi(-1) = -1$, the value of the Dirichlet L function at 1 is given by

$$L(1, \chi) = \frac{-i\pi}{g(\overline{\chi})p} \sum_{n=1}^{p-1} n\overline{\chi}(n)$$

(see [3], pp. 8-9) so our formula gives the alternate evaluation:

$$L(1, \chi) = \frac{-i\pi}{p} \sum_{n=1}^{p-1} \frac{\chi(n)}{1 - \zeta^n}.$$

Moreover, in the special case that χ is a quadratic character and $p \equiv 3 \ (4)$, we have

$$\sum_{n=0}^{p-1} n\Big(\frac{n}{p}\Big) = g \sum_{n=0}^{p-1} \Big(\frac{n}{p}\Big) \frac{1}{1 - \zeta^n}.$$

Davenport notes, ([3] p. 9), that this sum is known to be negative, although no elementary proof has been given.

2.3 Estermann's determination of $g(\chi_2)$

Motivated by the determination of the sign of $\sum n\big(\frac{n}{p}\big)$, we give a determination of the argument of the quadratic Gauss sum due to Estermann and attempt to adapt his method to determine the argument of $\sum \frac{\chi(n)}{1-\zeta^n}$.

Theorem 2.3 *Let g_p be the quadratic Gauss sum modulo odd prime p. Then the value of g_p is $\sqrt{p}$ if $p \equiv 1 \ (4)$ and $i\sqrt{p}$ if $p \equiv 3 \ (4)$.*

Remarks. In his proof, Estermann actually works with the sum

$$g(1, p) = \sum_{n=0}^{p-1} e^{2\pi n^2 i/p}$$

To see that this implies our result, observe that we can equivalently express the Gauss sum, g_p, as

$$g_p = \sum_n \Big(\frac{n}{p}\Big)\zeta^n = \sum_{n \in R} \zeta^n - \sum_{n \in N} \zeta^n$$

where R denotes the set of quadratic residues mod p and N denotes the set of non-residues. Since 0 is included in neither of these sets, we have:

$$\sum_{n\in R}\zeta^n+\sum_{n\in N}\zeta^n=\sum_{n=1}^{p-1}\zeta^n=-1$$

so

$$\sum_{n\in R}\zeta^n-\sum_{n\in N}\zeta^n=2\sum_{n\in R}\zeta^n+1.$$

Now the sum $\sum_{n=0}^{p-1}\zeta^{n^2}$ contains two contributions of ζ^j for each quadratic residue $j\ \ (p)$ and one contribution of $\zeta^0=1$. Thus

$$g_p=2\sum_{n\in R}\zeta^n+1=\sum_{n=0}^{p-1}\zeta^{n^2},$$

so our claim is equivalent to Estermann's statement.

Recall that in our discussion of quadratic reciprocity, we computed $g_p^2=\pm p$ according to whether $p\equiv 1$ or 3 mod 4. Thus $g_p=\pm\sqrt{p}$ if $p\equiv 1\ \ (4)$ and $g_p=\pm i\sqrt{p}$ if $p\equiv 3\ \ (4)$. This can be compactly restated as

$$\frac{1}{2}(1-i)(1+i^p)g_p=\pm\sqrt{p}.$$

Estermann shows that $\Re\big(\frac{1}{2}(1-i)(1+i^p)g(1,p)\big)>-\sqrt{p}$, implying the result.

Proof. (Estermann) The idea of the proof is to split the sum $\sum\zeta^{n^2}$ into the first few terms, which are close together in the first quadrant and hence give a positive sum, and a tail that is not too large. Estermann first manipulates the sum to divide the exponents on zeta by 4 (thus creating more terms in the first quadrant). He writes

$$g(1,p)-1=\sum_{n=1}^{p-1}e^{2\pi n^2i/p}=2\sum_{n=1}^{(p-1)/2}e^{2\pi n^2i/p}$$

so that

$$\frac{1}{2}(1+i^p)(g(1,p)-1)=(1+i^p)\sum_{n=1}^{(p-1)/2}e^{2\pi n^2i/p}.$$

Since $(\frac{p}{2}-n)^2=\frac{p^2}{4}+pn+n^2$ then $\frac{2\pi(\frac{p}{2}-n)^2i}{p}\equiv\frac{p\pi i}{2}+\frac{2\pi n^2i}{p}$ (mod 2π) and since $e^{p\pi i/2}=i^p$ we have

$$(1+i^p)\sum_{n=1}^{(p-1)/2}e^{2\pi n^2i/p}=\sum_{n=1}^{(p-1)/2}e^{\frac{2\pi n^2i}{p}}+e^{\frac{2\pi(p/2-n)^2i}{p}}.$$

which is equivalently rewritten as

$$(1+i^p)(g(1,p)-1) = \sum_{n=1}^{p-1} e^{\frac{2\pi(n/2)^2 i}{p}} = \sum_{n=1}^{p-1} e^{\pi n^2 i/(2p)}.$$

This last sum splits into leading part L and tail T with

$$L = \sum_{n=1}^{\lfloor\sqrt{p}\rfloor} e^{\pi n^2 i/(2p)}, T = \sum_{n=\lfloor\sqrt{p}\rfloor+1}^{p-1} e^{\pi n^2 i/(2p)}$$

and $L+T = \sum_n e^{\pi n^2 i/(2p)}$. We want a lower bound on $\Re\big((1-i)(L+T)\big)$. Beginning with L, we have

$$\Re\big((1-i)L\big) = \sum_{n=1}^{\lfloor\sqrt{p}\rfloor}\Big(\cos\frac{\pi n^2}{2p} + \sin\frac{\pi n^2}{2p}\Big)$$

and since $1 < \cos x + \sin x$ for $0 < x < \frac{\pi}{2}$, this sum is at least $\lfloor\sqrt{p}\rfloor > \frac{1}{2}\sqrt{p}$.

To get a bound on $|T|$ we create a telescoping sequence. Set $a_n = e^{\frac{\pi n(n+1)i}{2p}}$ and $b_n = \frac{1}{\sin\frac{\pi n}{2p}}$. Recalling that $\sin x = \frac{e^{ix}+e^{-ix}}{2i}$ we then have that

$$(a_n - a_{n-1})b_n = 2i\frac{e^{\frac{\pi n(n+1)i}{2p}} - e^{\frac{\pi n(n-1)i}{2p}}}{e^{\frac{\pi n i}{2p}} - e^{-\frac{\pi n i}{2p}}} = 2ie^{\pi n^2 i/(2p)}$$

so

$$2iT = \sum_{n=\lfloor\sqrt{p}\rfloor+1}^{p-1}(a_n - a_{n-1})b_n = a_{p-1}b_p - a_{\lfloor\sqrt{p}\rfloor}b_{\lfloor\sqrt{p}\rfloor+1}\sum_{n=\lfloor\sqrt{p}\rfloor+1}^{p-1}(b_n - b_{n+1})a_n.$$

Then by the triangle inequality, since for all n, $|a_n| = 1$, we have

$$|T| \leq \frac{1}{2}\Bigg(b_p + b_{\lfloor\sqrt{p}\rfloor+1} + \sum_{n=\lfloor\sqrt{p}\rfloor+1}^{p-1}|b_n - b_{n+1}|\Bigg).$$

But $b_n = \frac{1}{\sin\frac{\pi n}{2p}}$ so for $\lfloor\sqrt{p}\rfloor < n < p$, b_n is strictly decreasing. Thus we have

$$|T| \leq \frac{1}{2}\Bigg(b_p + b_{\lfloor\sqrt{p}\rfloor+1} + \sum_{n=\lfloor\sqrt{p}\rfloor+1}^{p-1}(b_n - b_{n+1})\Bigg) = b_{\lfloor\sqrt{p}\rfloor+1}.$$

Now using the approximation $\frac{\pi}{2}\sin x > x$ which is valid for $0 < x < \frac{\pi}{2}$,

$$b_{\lfloor\sqrt{p}+1\rfloor} = \frac{1}{\sin\frac{\pi(\lfloor\sqrt{p}\rfloor+1)}{2p}} < \frac{p}{\lfloor\sqrt{p}\rfloor+1} < \sqrt{p}$$

so $|T| < \sqrt{p}$.

Now recalling that $\Re\Big(\frac{1}{2}(1-i)(1+i^p)\Big)$ is non-negative, we know

$$\Re\bigg(\frac{1}{2}(1-i)(1+i^p)g(1,p)\bigg) \geq \Re\bigg(\frac{1}{2}(1-i)(1+i^p)\big(g(1,p)-1\big)\bigg) = \Re\big((1-i)(L+T)\big)$$

and

$$\Re\big((1-i)(L+T)\big) \geq (1-i)L - |1-i||T|.$$

Substituting our lower bound for L and our upper bound for $|T|$ into this expression, leaves

$$\Re\bigg(\frac{1}{2}(1-i)(1+i^p)g(1,p)\bigg) \geq \frac{1}{2}\sqrt{p} - \sqrt{2}\sqrt{p} = \bigg(\frac{1}{2} - \sqrt{2}\bigg)\sqrt{p} \geq -\sqrt{p}$$

as desired. This completes the determination of the value of the quadratic Gauss sum. □

2.3.1 A similar approach to a related sum

We would like to employ the method of Estermann to demonstrate that for primes $p \equiv 3$ (4) the sum $S = \sum_n \big(\frac{n}{p}\big)\frac{1}{1-\zeta^n}$ is positive imaginary. We first record several observations regarding the sum $\sum_n n\big(\frac{n}{p}\big)$.

Fact 2.4 *$\sum_n n\big(\frac{n}{p}\big)$ is odd.*

Proof. (Due to [3]) Since $p \equiv 3$ (4),

$$\sum_{n=1}^{p-1} n = \frac{p(p-1)}{2}$$

has both p and $\frac{p-1}{2}$ odd, hence is odd. Thus, with N the set of quadratic non-residues mod p,

$$\sum_n n\bigg(\frac{n}{p}\bigg) = \sum_{n=1} p - 1n - 2\sum_{n \in N} n,$$

implies that $\sum_n n\big(\frac{n}{p}\big)$ is also odd. □

Fact 2.5 *If $p > 3$, p divides $\sum_n n\big(\frac{n}{p}\big)$.*

Proof. Working modulo p,

$$\sum_n n\Big(\frac{n}{p}\Big) \equiv \sum_n n\big(n^{\frac{p-1}{2}}\big) \equiv \sum_n n^{\frac{p+1}{2}} \quad (p).$$

But for $p > 3$, $\frac{p+1}{2}$ is an integer between 1 and $p-1$ and we have, with a a generator for $\mathbb{Z}_p^\times$ for all $0 < k < p$

$$\sum_{n=1}^{p-1} n^k \equiv \sum_{n=1}^{p-1} a^{kn} \equiv \frac{a^p - a}{a-1} \equiv 0 \quad (p).$$

Thus $p \mid \sum_n n\big(\frac{n}{p}\big)$. $\square$

Together, these two facts imply that for $p > 3$,

$$gS = \sum_n n\Big(\frac{n}{p}\Big) = mp$$

for some odd integer m. Thus, as $g = i\sqrt{p}$ we have $S = -im\sqrt{p}$ and our task reduces to proving that $\Im(S) > -\sqrt{p}$.

The first step in emulating Estermann's proof is to convert the sum for S from a sum over n to a sum over n^2. To do this, write

$$\sum_n \Big(\frac{n}{p}\Big)\frac{1}{1-\zeta^n} = \sum_{n\in R}\frac{1}{1-\zeta^n} - \sum_{n\in N}\frac{1}{1-\zeta^n}$$

where, as before, R denotes the set of quadratic residues mod p and N denotes the set of non-residues. Now as $p \equiv 3\ (4)$, -1 is a non-residue mod p so for all n, $n \in R$ iff $-n \in N$. Thus we have

$$S = \sum_{n\in R}\frac{1}{1-\zeta^n} - \sum_{n\in R}\frac{1}{1-\zeta^{-n}} = \sum_{n\in R}\left(\frac{1}{1-\zeta^n} - \frac{\zeta^n}{\zeta^n - 1}\right)$$

and this last expression is equal to

$$\sum_{n\in R}\frac{1+\zeta^n}{1-\zeta^n} = \sum_{n\in R}\frac{\zeta^{\frac{n}{2}} + \zeta^{-\frac{n}{2}}}{\zeta^{\frac{n}{2}} - \zeta^{-\frac{n}{2}}} = i\sum_{n\in R}\cot\frac{\pi n}{p}$$

Now as n varies over $\mathbb{Z}_p^\times$, n^2 varies over R twice, so we have

$$S = i\sum_{n\in R}\cot\frac{\pi n}{p} = \frac{i}{2}\sum_{n=1}^{p-1}\cot\frac{\pi n^2}{p} = i\sum_{n=1}^{\frac{p-1}{2}}\cot\frac{\pi n^2}{p}.$$

Thus we wish to show that

$$\sum_{n=1}^{\frac{p-1}{2}} \cot \frac{\pi n^2}{p} > -\sqrt{p}.$$

If we are to follow the approach of Estermann, we should split the sum in question into a dominate, positive leading part and a tail part of smaller modulus. This suggests putting

$$L = \sum_{n=1}^{\lfloor \sqrt{p/2} \rfloor} \cot \frac{\pi n^2}{p}, T = \sum_{\lfloor \sqrt{p/2} \rfloor + 1}^{\frac{p-1}{2}} \cot \frac{\pi n^2}{p}$$

so that L contains the first group of positive terms in the sum. If we recall that cot is strictly decreasing on $(0, \pi)$ and that the gap between successive squares is also strictly increasing $((n+1)^2 - n^2 = 2n+1)$ we can get an upper bound for L by writing

$$\sum_{n=1}^{\lfloor \sqrt{p/2} \rfloor} \cot \frac{\pi n^2}{p} > \sum_{n=1}^{\lfloor \frac{\sqrt{p/2}}{2} \rfloor - 1} \cot \frac{\pi n^2}{p} + \sum_{\lfloor \frac{\sqrt{p/2}}{2} \rfloor}^{\lfloor \sqrt{p/2} \rfloor - 1} \cot \frac{\pi n^2}{p} \geq \sum_{n=1}^{\lfloor \frac{\sqrt{p/2}}{2} \rfloor - 1} \cot \frac{\pi n^2}{p} + \cot \left(\frac{\pi}{2} - \frac{\pi n^2}{p} \right)$$

where the last inequality is attained by pairing the kth term of the first sum with the kth-from-the-last from the second and observing that angles in question sum to less than $\frac{\pi}{2}$ (since the gaps between squares is increasing).

But $\cot x + \cot \left(\frac{\pi}{2} - x \right) = \frac{\cos x}{\sin x} + \frac{\sin x}{\cos x} = \frac{1}{\sin x \cos x} = \frac{2}{\sin 2x}$ so

$$\sum_{n=1}^{\lfloor \sqrt{p/2} \rfloor} \cot \frac{\pi n^2}{p} > 2 \sum_{n=1}^{\lfloor \frac{\sqrt{p/2}}{2} \rfloor - 1} \frac{1}{\sin \frac{2\pi n^2}{p}}.$$

Using the estimate $\sin x < x$ for $0 < x < \pi$ this gives

$$\sum_{n=1}^{\lfloor \sqrt{p/2} \rfloor} \cot \frac{\pi n^2}{p} > 2 \sum_{n=1}^{\lfloor \frac{\sqrt{p/2}}{2} \rfloor - 1} \frac{p}{2\pi n^2} = \frac{p}{\pi} \sum_{n=1}^{\lfloor \frac{\sqrt{p/2}}{2} \rfloor - 1} \frac{1}{n^2}.$$

Recalling that

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \qquad \sum_{n=k}^{\infty} \frac{1}{n^2} \leq \int_{k-1}^{\infty} \frac{dx}{x^2} = \frac{1}{k-1}$$

then

$$\sum_{n=1}^{\lfloor \frac{\sqrt{p/2}}{2} \rfloor - 1} \frac{1}{n^2} \geq \frac{\pi^2}{6} - \frac{1}{\lfloor \frac{\sqrt{p/2}}{2} \rfloor - 2}.$$

For suitably large p,

$$\lfloor\frac{\sqrt{p/2}}{2}\rfloor - 2 > \frac{\sqrt{p}}{\pi}$$

so in sum we have

$$\sum_{n=1}^{\lfloor\sqrt{p/2}\rfloor} \cot\frac{\pi n^2}{p} > \frac{p}{\pi}\Big(\frac{\pi^2}{6} - \frac{\pi}{\sqrt{p}}\Big) = \frac{p\pi}{6} - \sqrt{p}.$$

Thus we would be finished if we could show that $|T| < \frac{p\pi}{6}$. For p not overly large (e.g $p \sim 10^6$) this is true and on average we have $S_p \sim \frac{p\pi}{6}$. Using methods of complex analysis, however, one can show that S is $O(n\log n)$ so that L is not the dominant term in the sum. Thus this approach of Estermann's will not work to show that S is positive imaginary. The approach does, however, shed light on why the sum $\sum_n n\big(\frac{n}{p}\big)$ is negative. Since $\cot x$ is π-periodic, if we split the sum $\sum_{n=1}^{\frac{p-1}{2}} \cot\frac{\pi n^2}{p}$ into the terms between 0 and $\sqrt{p}$, $\sqrt{p}$ and $\sqrt{2p}$, ... then each of these successive sums falls in one period of cot. In each of these sums, the terms are more densely populated toward the beginning of the period than toward the end because the gap between terms is increasing. Since cot is big and positive at the beginning of its period and big and negative at the end, this would tend to imply that, at least while there are many points in the period, the positive terms outweigh the negative ones. It is difficult to make this intuition fruitful, however, because cot tends to plus or minus infinity at the ends of its period so that only a few terms are influential in each sum.

2.4 The Davenport-Hasse Identity

Earlier we remarked that the modulus identity $g(\chi)g(\overline{\chi}) = \chi(-1)p$ was one of two known multiplicative identities concerning Gauss sums; the second is the identity of Davenport and Hasse. Before we state this identity, however, we first generalize the notion of character and Gauss sum to general finite fields.

2.4.1 Gauss sums over F_q

The setting is as follows: let $q = p^r$ with p a prime and r a positive integer. F_q denotes the finite field of q elements and F_p denotes the finite field of p elements associated with residues $0, 1, 2, ..., p-1$. We define the trace map $tr : F_q \to F_p$ such that for $\alpha \in F_q$,

$$tr(\alpha) = \alpha + \alpha^p + \alpha^{p^2} + ... + \alpha^{p^{r-1}}.$$

Berndt et. al. prove the essential properties of the trace. They are:

Proposition 2.6 *Suppose $\alpha, \beta \in F_q$ and $a \in F_p$. Then*
1. $tr(\alpha) \in F_p$
2. $tr(\alpha + \beta) = tr(\alpha) + tr(\beta)$

3. $tr(a\alpha) = atr(\alpha)$
4. tr *maps* F_q *onto* F_p
5. $tr(\alpha^p) = tr(\alpha)$

We prove the first two items and refer the reader to Berndt (pp. 7, 8) for the other (similar) deductions.

Proof. For (1.) consider $tr(\alpha)^p$. By the multinomial theorem, and because F_q is a field of characteristic p,

$$tr(\alpha)^p = \left(\alpha + \alpha^p + ... + \alpha^{p^{r-1}}\right)^p = \alpha^p + \alpha^{p^2} + ... + \alpha^{p^r}.$$

But $\alpha^{p^r} = \alpha^q = \alpha$ so in fact we have $tr(\alpha)^p = tr(\alpha)$. But $x^p = x$ is a pth order polynomial with p roots in F_p. Thus these are all the roots so $tr(\alpha) \in F_p$.

For (2.) apply the binomial theorem:

$$tr(\alpha + \beta) = \sum_{j=0}^{r-1} (\alpha + \beta)^{p^j} = \sum_{j=0}^{r-1} \alpha^{p^j} + \beta^{p^j} = tr(\alpha) + tr(\beta).$$

□

With the trace, it is now possible to define the Gauss sum for a character χ on F_q.

Definition 2.1 *Let* χ *be a multiplicative character on* F_q *and let* $\beta \in F_q$*. The Gauss sum with respect to* β *and* χ *is given by*

$$\tau(\beta, \chi) = \sum_{\alpha \in F_q} \chi(\alpha) e^{\frac{2\pi tr(\alpha\beta)}{p}}$$

Using the same techniques used to derive the properties of the Gauss sum over F_p we have

Proposition 2.7 *With* χ *and* β *as in the definition and* l *an integer,*
1. $\tau(\beta, \chi) = \chi(\beta^{-1})\tau(1, \chi)$
2. $\tau(\beta, \chi)\tau(\beta, \overline{\chi}) = \chi(-1)p^r$
3. $\overline{\tau(\beta, \chi)} = \chi(-1)\tau(\beta, \chi)$
4. $|\tau(\beta, \chi)| = p^{\frac{r}{2}}$
5. $\tau(\beta, \chi^l) = \tau(\beta^l, \chi)$

We refer the reader to chapter 1 and to ([2] p. 10) for proofs. In view of (1.) it is customary to denote $\tau(1, \chi)$ by $\tau(\chi)$.

2.4.2 The character of a prime ideal

Thus far we have defined the Gauss sum of a character on a general finite field but we have suppressed the description of the character. In order to develop the kth order character χ on F_q observe that we must have $k \mid q-1 = p^r - 1$ and suppose for now that k does not divide $p^a - 1$ for $a < r$ so that r is the order of p mod k. Let $\zeta = e^{\frac{2\pi i}{p}}$ and $\zeta_k = e^{\frac{2\pi i}{k}}$. Define the field extensions $K = \mathbb{Q}(\zeta_k)$ and $M = K(\zeta) = \mathbb{Q}(\zeta_k, \zeta)$ and let these fields have associated rings of integers O_K and O_M respectively. Since $k \mid p^r - 1$, p splits in O_K. Let P be a prime ideal of O_K dividing pO_K. Then O_K/P is a finite field of p^r elements associated with F_q and we define the character χ_P on this field mapping to kth roots of unity by

$$\chi_P(\alpha + P) \equiv \alpha^{\frac{q-1}{k}} \pmod{P}.$$

This definition makes implicit use of the fact that, mod P, the first k powers of ζ_k are distinct. To see that this holds suppose $P \mid \zeta_k^s - \zeta_k^t$ so $P \mid 1 - \zeta_k^{t-s}$ and write $v = t - s$. Let $\sigma_j \in Gal(K/\mathbb{Z})$ denote the automorphism mapping ζ_k to ζ_k^j. Then with $P_j = \sigma_j(P)$ we have for each j relatively prime to k, $P_j \mid 1 - \zeta_k^{jv}$. If we let R be the group of reduced residues mod k and $T = R/\{1, p, p^2, ..., p^{r-1}\}$ then ([2] p. 343) shows that

$$pO_K = \prod_{j \in T} P_j$$

so

$$(p) \mid \prod_{j \in T} (1 - \zeta_k^{jv}).$$

But this is impossible because, writing

$$1 + x + x^2 + ... + x^{k-1} = (x - \zeta_k)(x - \zeta_k^2)...(x - \zeta_k^{k-1})$$

and substituting 1 for x we see that $\prod_{j \in T}(1 - \zeta_k^{jv})$ is a power of a factor of k, and p and k are relatively prime. Thus each of the ζ_k^j are distinct mod P.

Our definition of kth order character so far has depended upon the choice of prime ideal P. In fact we get another kth order character for each of the prime ideals P_j and these are all the kth order characters (since there are $\phi(k)$ of them). Although we have only defined the kth order character in the case that $ord_k(p) = r$, we know that the group of characters is a cyclic group of order $q - 1$ generated by the characters of order $q - 1$ and these characters are of the type defined. Thus effectively we have given a description of all characters on F_q.

2.4.3 The Davenport-Hasse identity in several formulations

With this discussion of Gauss sums and characters on F_q it is now possible to state the Davenport-Hasse product formula for Gauss sums on F_q.

Theorem 2.8 *(Davenport-Hasse) Let ψ be an lth order character, $l > 1$, on F_q and let χ be any character. Set $\tau(\epsilon) = -1$ for the trivial character ϵ. Then*

$$N(l, \chi) = \frac{\chi^l(l)\tau(\chi)}{\tau(\chi^l)} \prod_{i=1}^{l-1} \frac{\tau(\chi\psi^i)}{\tau(\psi^i)} = 1.$$

There is no known elementary proof of this identity. Davenport and Hasse's proof relies on Stickelberger's congruence for Gauss sums and is involved, so we will not give their proof here. We will make some observations, however. Berndt remarks that if χ is a power of ψ then the identity trivially holds since each of the Gauss sums cancel. If χ is not a power of ψ then none of the characters in the product is trivial. Then since we know that the norm of each Gauss sum is $p^{\frac{r}{2}}$ we observe that the product has norm 1 since there are an equal number of Gauss sums in the numerator and denominator. The problem reduces to computing the angle of the product.

There are many ways of expressing the identity. For instance, Berndt writes

$$N(l, \chi) = \chi^l(l) \prod_{i=1}^{l-1} \frac{J(\chi, \chi^i)}{J(\chi, \psi^i)}$$

where $J(\chi, \lambda)$ is the Jacobi sum $\sum_{a+b=1} \chi(a)\lambda(a)$ (the change in the product formula follow from $J(\chi, \lambda) = \frac{\tau(\chi)\tau(\lambda)}{\tau(\chi\lambda)}$). His formalism shows that if O_K is the ring of integers $\mathbb{Z}[\zeta_k]$ from above with $k = q-1$ then $N(l, \chi) \in O_K$ since each character has values in O_K. Alternatively, clearing the denominator from the product leaves

$$\prod_{i=0}^{l-1} \tau(\chi\psi^i) = \overline{\chi}^l(l)\tau(\chi^l) \prod_{i=1}^{l-1} \tau(\psi^i).$$

Then employing the identity

$$\prod_{i=1}^{t} \tau(\chi_i) = \tau(\chi_1\chi_2 ... \chi_t) J(\chi_1, \chi_2, ..., \chi_t)$$

which holds for non-trivial characters $\chi_1, \chi_2, ..., \chi_t$ and $\chi_1\chi_2...\chi_t$ non-trivial with J the t-order Jacobi sum

$$J(\chi_1, ..., \chi_t) = \sum_{s_1+s_2+...+s_t=1} \chi_1(s_1)\chi_2(s_2)...\chi_t(s_t)$$

(see [4], p. 100) the Davenport-Hasse identity reduces to:

$$\tau(\chi^l\psi^{\frac{l(l-1)}{2}})J(\chi, \chi\psi, \chi\psi^2, ..., \chi\psi^{l-1}) = \overline{\chi}^l(l)\tau(\chi^l\psi^{\frac{l(l-1)}{2}})J(\chi^l, \psi, \psi^2, ..., \psi^{l-1})$$

or simply

$$J(\chi, \chi\psi, ..., \chi\psi^{l-1}) = \overline{\chi}^l(l)J(\chi^l, \psi, ..., \psi^{l-1}).$$

Upon expansion, this yields:

$$\sum_{t_0+t_1+...+t_{l-1}=1} \chi(t_0)\chi\psi(t_1)...\chi\psi^{l-1}(t_{l-1}) = \overline{\chi}^l(l) \sum_{t_0+t_1+...+t_{l-1}=1} \chi^l(t_0)\psi(t_1)\psi^2(t_2)...\psi^{l-1}(t_{l-1})$$

or more compactly,

$$\sum_{t_0+t_1+...+t_{l-1}=1} \chi(t_0 t_1 ... t_{l-1})\psi(t_1^1 t_2^2 ... t_{l-1}^{l-1}) = \sum_{t_0+t_1+...+t_{l-1}=1} \chi\left(\frac{t_0^l}{l^l}\right)\psi(t_1^1 t_2^2 ... t_{l-1}^{l-1}).$$

It seems hopeful that a combinatorial analysis of these two sums might yield an elementary proof of the identity, for instance, via permuting $t_1, t_2, ..., t_{l-1}$.

2.5 A conjecture of Hasse

Hasse conjectured that the only two multiplicative relations on Gauss sums are the modulus relation and the Davenport-Hasse formula. Hasse's conjecture has been proven correct by Yamamoto in the case where Gauss sums are viewed as ideals of O_K [8]. In his proof, Yamamoto also makes use of the Stickelberger congruence while viewing Gauss sums as a discrete analog of the p-adic gamma function. We reproduce an elegant section of Yamamoto's proof that highlights his general approach and refer the interested reader to his paper for the rest of the argument.

2.5.1 Multiplicatively independent Gauss sums

The setting for Yamamoto's proof is as follows. Let p be a prime and χ an eth order character on $\mathbb{Z}/(p\mathbb{Z})$ where e is an even integer. With ζ_e a primitive eth root of unity, p splits in $\mathbb{Q}(\zeta_e)$ as

$$p = \prod_{\sigma \in G} \sigma(P)$$

where G denotes the Galois group of $\mathbb{Q}(\zeta_e)/\mathbb{Q}$. In the field $\mathbb{Q}(\zeta_e, \zeta_p)$ the prime ideals $\sigma(P)$ are completely ramified as

$$\sigma(P) = \sigma(\wp)^{p-1}$$

with $\wp$ a prime ideal of $\mathbb{Q}(\zeta_e, \zeta_p)$. The Gauss sum $\tau(\chi^a)$ is given by

$$\tau(\chi^a) = \sum_{t \bmod p} \chi^a(t)\zeta_p^t,$$

and with $< x >= x - \lfloor x \rfloor$ being the fractional part of x, Yamamoto invokes the Stickelberger congruence in the form

$$(\tau(\chi^a)) = \prod_{(t,e)=1} \sigma_{-t}^{-1}(\wp)^{(p-1)< \frac{at}{e} >}$$

where $(\tau(\chi^a))$ represents the principal ideal of $\tau(\chi^a)$ in the field $\mathbb{Q}(\zeta_e, \zeta_p)$.

He then considers the Davenport-Hasse identity in the form[1]

$$\tau(\chi^{la}) = \frac{\chi(l^{la})\psi(l)}{(\psi(2)\tau(\psi))^{l-1}} \prod_{j=0}^{l-1} \tau(\chi^{a+\frac{ej}{l}})$$

where ψ is the order 2 character and asserts:

Theorem 2.9 *(Yamamoto) If $e \geq 4$ is even then the number of multiplicatively independent Gauss sums $\tau(\chi^a)$ is $\frac{\phi(e)}{2} + 1$ where ϕ denotes Euler's function.*

To prove this theorem, Yamamoto introduces the notation

$$\rho_a = \frac{\tau(\chi^a)}{\tau(\psi)} \quad \text{and} \quad \{x\} =< x > -\frac{1}{2}.$$

With this notation, he expresses the norm relation, Davenport-Hasse identity and Stickelberger congruence in terms of ideals of $\mathbb{Q}(\zeta_e, \zeta_p)$. Specifically he proves

Lemma 2.10 *Let a be a residue mod p and $l \mid e$. Then we have*

1. $(\rho_a)(\rho_{-a}) = (1)$ *provided e does not divide a*
2. $(\rho_{la}) = \prod_{j=0}^{l-1} (\rho_{a+\frac{ej}{l}})$
3. $(\rho_a) = \prod_{(t,e)=1} \sigma_{-t}^{-1}(\wp)^{(p-1)\{\frac{ta}{e}\}}$.

To justify the lemma, (1.) and (2.) hold simply by dropping units from respectively the norm relation $\tau(\chi^a)\tau(\chi^{-a}) = \chi^a(-1)p = \pm\tau(\psi)^2$ and from Yamamoto's earlier formulation of the Davenport-Hasse identity. Comparing the right hand side of (3.) to the Stickelberger identity, we see that it differs by

$$\prod_{(t,e)=1} \sigma_{-t}^{-1}(\wp)^{\frac{p-1}{2}} = \prod_{\sigma \in G} \sigma(\wp)^{\frac{p-1}{2}} = \prod_{\sigma \in G} \sigma(P)^{\frac{1}{2}} = p^{\frac{1}{2}}$$

which is, up to a unit, precisely the difference between $\tau(\chi^a)$ and ρ_a since $\tau(\psi)$ is a unit times $\sqrt{p}$.

[1] This is accomplished by pairing Gauss sums of conjugate characters in the denominator of the original Davenport-Hasse identity.

In light of (3.) in the lemma, Yamamoto notes that his theorem is equivalently restated as $Rank(A) = 1 + \frac{\phi(e)}{2}$ where (A_{at}) is the $e \times \phi(e)$ matrix

$$A = \left(\left\{ \frac{at}{e} \right\} \right)_{\substack{a \in \mathbb{Z}_e \\ t \in \mathbb{Z}_e^\times}}$$

since multiplication on Gauss sums is the same as addition on vectors of exponents in the prime ideal factorization. Now $\{x\}$ has the property that if x is not an integer then $\{-x\} = -\{x\}$ and otherwise $\{x\} = \{-x\} = -\frac{1}{2}$. Thus the first row of A, with $a = 0$ has all entries equal to $-\frac{1}{2}$ and every other row has entries that sum to 0. Hence the first row is linearly independent of all the others and we need to show that the $(e-1) \times \phi(e)$ matrix

$$A_1 = \left(\left\{ \frac{at}{e} \right\} \right)_{\substack{a \in \mathbb{Z}_e, a \neq 0 \\ t \in \mathbb{Z}_e^\times}}$$

has rank $\frac{\phi(e)}{2}$. Now since none of $\frac{at}{e}$ is an integer for $1 \leq a < e$ and $(t, e) = 1$, the column $-t$ is -1 times the column t in A_1 so A_1 has rank at most $\frac{\phi(e)}{2}$. It only remains to show that $rank(A_1) \geq \frac{\phi(e)}{2}$. For this, suppose we have some non-trivial linear combination $\{c_t\}$ on the columns of A_1 satisfying

$$\sum_{(t,e)=1} c_t \left\{ \frac{at}{e} \right\} = 0 \quad \text{for } 1 \leq a < e$$

and the additional condition $c_{-t} = -c_t$ so that the linear combination is effectively only over the first $\frac{\phi(e)}{2}$ columns.

Now if we take s with $(s, e) = 1$ then certainly we have $sa \neq 0$ in $\mathbb{Z}_e$ for any $a \in \mathbb{Z}_e, a \neq 0$ so, in particular, we have

$$\sum_{(t,e)=1} c_t \left\{ \frac{ast}{e} \right\} = 0 \quad \text{for } 1 \leq a < e,$$

or, for any character θ on $\mathbb{Z}_e^\times$,

$$\sum_{(t,e)=1} c_t \overline{\theta}(t)\ \theta(st) \left\{ \frac{ast}{e} \right\} = 0 \quad \text{for } 1 \leq a < e$$

since $\theta(s)$ is a constant. But then summing over all s in $\mathbb{Z}_e^\times$,

$$\sum_{(t,e)=1} c_t \overline{\theta}(t) \sum_{(s,e)=1} \theta(st) \left\{ \frac{ast}{e} \right\} = \sum_{(t,e)=1} c_t \overline{\theta}(t) \sum_{(s,e)=1} \theta(s) \left\{ \frac{as}{e} \right\} = 0 \quad \text{for } 1 \leq a < e$$

where we have made the substitution $st \mapsto s$ on $\mathbb{Z}_e^\times$. But now these are independent sums, and choosing $c_{t'} \neq 0$ we have the Fourier expansion for $c_{t'}$,

$$c_{t'} = \sum_{(n,e)=1} \hat{c}_n \theta_n(t')$$

where the sum runs over all characters on $\mathbb{Z}_e^\times$ and

$$\hat{c}_n = \frac{1}{\phi(e)} \sum_{(t,e)=1} c_t \overline{\theta_n}(t).$$

Hence as $c_{t'} \neq 0$ we cannot have all $\hat{c}_n = 0$, so we can choose θ with

$$\sum_{(t,e)=1} c_t \overline{\theta}(t) \neq 0.$$

Then since $c_{-t} = -c_t$ we know that $\theta(-1) = -1$ and

$$\sum_{(s,e)=1} \theta(s) \left\{ \frac{as}{e} \right\} = 0 \quad \text{for } 1 \leq a < e.$$

Now let θ_1 be a primitive character on $\mathbb{Z}_{e_1}^\times$ satisfying $\theta(t) = \theta_1(s)$ if $t \equiv s \ (e_1), t \in \mathbb{Z}_e^\times, s \in \mathbb{Z}_{e_1}^\times$ (in this case we say that e_1 is the *conductor* for θ) and observe that $\left\{ \frac{t}{e_1} \right\} = \left\{ \frac{s}{e_1} \right\}$. There are $\frac{\phi(e)}{\phi(e_1)}$ such $t \in \mathbb{Z}_e^\times$ congruent to s mod e_1 for each $s \in \mathbb{Z}_{e_1}^\times$ so, choosing $a = \frac{e}{e_1}$,

$$\sum_{(t,e)=1} \theta(t) \left\{ \frac{at}{e} \right\} = \sum_{(t,e)=1} \theta(t) \left\{ \frac{t}{e_1} \right\} = \frac{\phi(e)}{\phi(e_1)} \sum_{(s,e_1)=1} \theta_1(s) \left\{ \frac{s}{e_1} \right\} = 0.$$

But $\left\{ \frac{s}{e_1} \right\} = \frac{s}{e_1} - \frac{1}{2}$ so we have that

$$\sum_{(s,e)=1} \theta_1(s) \left(\frac{s}{e_1} - \frac{1}{2} \right) = \frac{1}{e_1} \sum_{(s,e)=1} s\theta_1(s) = 0$$

where we have made use of the fact that the sum of a non-trivial character over its domain of definition is zero. But we may not have $\sum_{(s,e)=1} s\theta_1(s) = 0$ since we recall from our section on Fourier series that in the case that $\theta_1(-1) = -1$ the sum in question is a factor of $L(1, \theta_1) \neq 0$ where L denotes the the Dirichelet L-function. Thus we have reached a contradiction so our original choice of the c_t must have been impossible, and we must have that the first $\phi(e)$ columns of A_1 are linearly independent. Hence A_1 has rank at least $\frac{\phi(e)}{2}$ and A has rank at least $1 + \frac{\phi(e)}{2}$, proving our result. □

This proof of Yamamoto's is powerful because it employs the Stickelberger identity to convert a question about the multiplicative relations among Gauss sums into a problem of linear algebra. The introduction of the character θ and the sum $\sum s\theta(s)$ is particularly clever. In the proof that we elaborated, Yamamoto used only property (3.) of his lemma (the Stickelberg relation in ideals). In the remainder of his paper he shows that the first two properties, namely the norm relation and Davenport-Hasse identity, account for all $1+\frac{\phi(e)}{2}$ linearly independent columns in the matrix A. The approach is similar, so we do not repeat it.

2.5.2 A counter-example when Gauss sums are numbers

At the end of his paper ([8] p. 489), however, Yamamoto does make the interesting observation that Hasse's conjecture does not hold for Gauss sums when viewed as numbers and he offers the following counterexample:

Example 2.1 *Let χ be an order 12 character and write $\tau(n)$ for $\tau(\chi^n)$. Then the Davenport-Hasse formula together with the norm relation imply that*

$$\tau(2)^2\tau(5)^2 = \chi(-4)\overline{\chi}(-27)\tau(3)^2\tau(4)^2,$$

from which we conclude that $\tau(2)\tau(5) = \eta\tau(3)\tau(4)$ where η is a 12th root of unity. However, it is possible to show via other methods that $\eta^4 = \chi(-4)$ and $\eta^6 = \overline{\chi}(-27)$. Thus there is more data available than can be derived from the norm relation and the Davenport-Hasse formula, so there exist other multiplicative relations when Gauss sums are considered as numbers.

In addition to Yamamoto, Berndt references the computations of Muskat and Whiteman [6] concerning the values of Jacobi sums as other counterexamples to Hasse's conjecture for Gauss sums as numbers. In fact, we will show that Yamamoto's evaluation of η^4 and η^6 can be derived from the norm relation and the Davenport-Hasse formula, together with a fact concerning the behavior of Gauss sums under automorphism. The same is true for the other counterexamples. Specifically we show

Proposition 2.11 *Define η by $\tau(2)\tau(5) = \eta\tau(3)\tau(4)$ where $\tau(n) = \tau(\chi^n)$ for the order 12 character χ. Then the following three identities imply that $\eta^4 = \chi(-4)$ and $\eta^6 = \overline{\chi}(-27)$:*

1. $\tau(n)\tau(12-n) = \chi^n(-1)p$

2. If $mn = 12$ then $\chi^m(m)\prod_{j=0}^{m-1}\tau(nj+x) = \tau(mx)\prod_{j=1}^{m-1}\tau(nj)$

3. Suppose $(j, 12) = 1$ and let σ_j be the automorphism of $\mathbb{Q}(\zeta_{12})$ that maps ζ_{12} to ζ_{12}^j. Then $\sigma_j\tau(n) = \tau(nj)$.

Proof. For ease of computation we will work with Jacobi sums rather than with Gauss sums; we do not gain any information by doing this. Let $J(m,n)$ denote the Jacobi sum $J(\chi^m,\chi^n)$ so that we have $J(m,n) = \frac{\tau(m)\tau(n)}{\tau(m+n)}$. From this expansion, we recognize that

$$J(m,n)J(m+n,r) = J(m,r)J(m+r,n).$$

Also,

$$J(m,n) = \frac{\tau(m)\tau(n)}{\tau(m+n)} = \frac{\tau(m)\tau(-m)\tau(n)}{\tau(m+n)\tau(-m-n)}\frac{\tau(-m-n)}{\tau(-m)}$$

and employing (1.),

$$J(m,n) = \chi^n(-1)\frac{\tau(-m-n)\tau(n)}{\tau(-m)} = \chi^n(-1)J(-m-n,n).$$

Meanwhile, interpreting (3.) in terms of Jacobi sums give

$$\sigma_j J(m,n) = \sigma_j\left(\frac{\tau(m)\tau(n)}{\tau(m+n)}\right) = \frac{\tau(jm)\tau(jn)}{\tau(jm+jn)} = J(jm,jn).$$

Now writing our information about η in terms of Jacobi sums we have $J(2,5) = \eta J(3,4)$. Applying σ_5 to each side gives

$$J(1,10) = \eta^5 J(3,8).$$

To compute η^4, write

$$J(3,4)J(1,10) = \eta^4 J(3,8)J(2,5).$$

By applying $J(m,n) = \chi^n(-1)J(-m-n,n)$ to $J(3,4)$ with $m = 3, n = 4$, $J(1,10)$ with $m = 10, n = 1$, $J(3,8)$ with $m = 3, n = 8$ and to $J(2,5)$ with $m = 2, n = 5$ this gives

$$J(4,5)J(1,1) = \eta^4 J(1,8)J(5,5)$$

or expanding in terms of Gauss sums,

$$\frac{\tau(4)\tau(5)\tau(1)^2}{\tau(9)\tau(2)} = \eta^4\frac{\tau(1)\tau(8)\tau(5)^2}{\tau(9)\tau(10)}$$

which leaves, after collecting terms,

$$\tau(1)\tau(4)\tau(10) = \eta^4\tau(2)\tau(5)\tau(8).$$

If we multiply each side by $\tau(7)$ we know $\tau(1)\tau(4)\tau(7)\tau(10) = \eta^4\tau(2)\tau(5)\tau(7)\tau(8)$. But now by (2.) we have $\tau(2)\tau(8) = \chi^{-4}(2)\tau(4)\tau(6)$ and $\tau(1)\tau(7) = \chi^{-2}(2)\tau(2)\tau(6)$ so making these substitutions

$$\chi^{-2}(2)\tau(2)\tau(4)\tau(6)\tau(10) = \eta^4\chi^{-4}(2)\tau(4)\tau(5)\tau(6)\tau(7).$$

Canceling like terms, this reduces to

$$\chi^2(2)\tau(2)\tau(10) = \eta^4\tau(5)\tau(7)$$

and now applying (1.) to each side we finally have

$$\chi(4)\chi^2(-1)p = \eta^4\chi^5(-1)p$$

or, recalling that $\chi(-1)$ is plus or minus 1,

$$\chi(4)\chi(-1) = \chi(-4) = \eta^4$$

as desired.

To compute η^6 recall that $J(2,5) = \eta J(3,4)$ and $J(1,10) = \eta^5 J(3,8)$ so

$$J(1,10)J(2,5) = \eta^6 J(3,4)J(3,8).$$

Applying $J(m,n) = \chi^n(-1)J(-m-n,n)$ to $J(2,5)$ with $m = 2, n = 5$, $J(3,4)$ with $m = 4, n = 3$ and $J(3,8)$ with $m = 3, n = 8$ gives

$$J(1,10)J(5,5) = \eta^6 J(3,5)J(1,8).$$

Then applying $J(m,n)J(m+n,r) = J(m,r)J(m+r,n)$ to each side of this equation, we are left with

$$J(1,5)J(5,6) = \eta^6 J(1,5)J(3,6)$$

so $J(5,6) = \eta^6 J(3,6)$. Expanding this in terms of Gauss sums

$$\frac{\tau(5)\tau(6)}{\tau(11)} = \eta^6 \frac{\tau(3)\tau(6)}{\tau(9)}$$

or $\tau(5)\tau(9) = \eta^6\tau(3)\tau(11)$. Then multiplying each side of this equation by $\tau(3)\tau(7)$ and applying (1.) twice to the left side,

$$p^2 = \eta^6\tau(3)^2\tau(7)\tau(11).$$

But then by (2.), $\tau(3)\tau(7)\tau(11) = \chi^{-9}(3)\tau(9)\tau(4)\tau(8)$ so making this substitution, with $\chi^{-9} = \chi^3$,

$$p^2 = \eta^6\chi^3(3)\tau(3)\tau(4)\tau(8)\tau(9)$$

and applying (1.) twice to the right side leaves

$$p^2 = \eta^6\chi^3(3)p^2\chi^3(-1)\chi^4(-1)$$

or $1 = \eta^6\chi(27)\chi(-1)$ so

$$\eta^6 = \overline{\chi}(-27)$$

completing the computation. □

In view of the fact that all of the known contradictions to Hasse's conjecture arise from including properties of Gauss sums under automorphism, it is possible that Yamamoto's proof for ideals might be adapted to a proof for numbers by including the automorphism transformation as a third multiplicative property of Gauss sums on numbers.

2.6 Eisenstein Reciprocity

We conclude this chapter with a theory of reciprocity due to Eisenstein that generalizes the theories of quadratic, cubic and biquadratic reciprocity from chapter 1. Recall that in our discussion of the Davenport-Hasse formula we considered the ring of integers $\mathbb{Z}[\zeta_k] = O_K$ ($\zeta_k = e^{\frac{2\pi i}{k}}$) and defined the kth order character χ_P for prime ideal P of O_K by

$$\chi_P(\alpha) = \zeta_k^j : P \mid \left(\alpha^{\frac{q-1}{k}} - \zeta_k^j\right)$$

where q is the cardinality of the finite field O_K/P and $\alpha \in O_K/P$, $\alpha \neq 0$. For $\alpha = 0$ we set $\chi_P(\alpha) = 0$. In our discussion of Eisenstein reciprocity we introduce the familiar notation

$$\left(\frac{\alpha}{P}\right)_k = \chi_P(\alpha)$$

and extend this character to non-prime ideals A, relatively prime to k in O_K, by

$$\left(\frac{\alpha}{A}\right)_k = \prod_i \left(\frac{\alpha}{P_i}\right)_k$$

where $\prod_i P_i$ is the prime ideal decomposition of A in O_K. It is important to note that in calculating $\left(\frac{\cdot}{P_i}\right)_l$ there is an implicit dependence on the order of P_i in O_l and to emphasize this dependence we introduce the norm $N(A)$ of an ideal A in O_K to be the order of O_K/A. Thus

$$\left(\frac{\alpha}{P_i}\right) \equiv \alpha^{\frac{N(P_i)-1}{l}} \pmod{P_i}.$$

In our later discussion, we will make use of the following two facts about N, which we state without proof (for proofs see [4] pp. 203-4).

Fact 2.12 *Let A and B be ideals of O_K. Then $N(AB) = N(A)N(B)$.*

Fact 2.13 *Let G be the Galois group of the field extension $\mathbb{Q}(\zeta_k)/\mathbb{Q}$ and let A be an ideal of O_K. Then*

$$\prod_{\sigma \in G} \sigma(A) = (N(A)).$$

As in our previous reciprocity proofs, we find it necessary to restrict to a set of primary elements in order to state Eisenstein's reciprocity law. To this end, let l be an odd prime, $\lambda = 1 - \zeta_l$ in O_l and define

Definition 2.2 *For non-zero, non-unit $\alpha \in O_l$, α relatively prime to l, α is primary if $\alpha \equiv n \pmod{\lambda^2}$ for some $n \in \mathbb{Z}$.*

Having made this definition, we can state the law of Eisenstein reciprocity.

2.6.1 The Eisenstein Law

Theorem 2.14 *Let l be an odd prime and suppose that $a \in \mathbb{Z}$ and $\alpha \in O_l$ are relatively prime with a prime to l and α primary. Then*

$$\left(\frac{\alpha}{a}\right)_l = \left(\frac{a}{\alpha}\right)_l.$$

Here $\left(\frac{\cdot}{a}\right)_l$ and $\left(\frac{\cdot}{\alpha}\right)_l$ refer to the characters of the principal ideals generated by a and α respectively.

Once again we will find that the Gauss sum is a powerful tool for proving Eisenstein reciprocity since it stores the information about characters that we are interested in. In our earlier proofs, we made use of the fact that we could determine the value of $g(\chi)$ up to a power: for quadratic reciprocity we calculated $g(\chi_p)^2 = (-1)^{\frac{p-1}{2}}p$, for cubic reciprocity, we had $g(\chi_\pi)^3 = \pi^2\overline{\pi}$, for biquadratic reciprocity we evaluated $g(\chi_\pi)^4 = \pi^3\overline{\pi}$, where π was a prime in $\mathbb{Z}[\omega]$ or $\mathbb{Z}[i]$ respectively. This suggests that to prove Eisenstein reciprocity, we should look for a relation regarding $g(\chi)^l$.

2.6.2 Breaking down $g(\chi)^l$

Ultimately our desired result on $g(\chi)^l$ will be the Stickelberger relation, but we first prove a simpler fact. Recall that we defined the Gauss sum $g(\chi_P) = g(P)$ over the finite field F_q where $q = p^r$ by

$$g(P) = \sum_{t \in F_q} \chi_P(t)\zeta_p^{tr(t)}$$

where tr was a map from F_q to $F_p = \mathbb{Z}/(p\mathbb{Z})$ satisfying $tr(t) = t + t^p + t^{p^2} + ... + t^{p^{r-1}}$. If χ_P is an l order character, then χ_P maps F_q into $\mathbb{Q}(\zeta_l)$ and, since tr maps into F_p, $\zeta_p^{tr(t)}$ is an element of $\mathbb{Q}(\zeta_p)$. Hence we have $g(P) \in \mathbb{Q}(\zeta_l, \zeta_p) = \mathbb{Q}(\zeta_{lp})$.

Now introduce the function Φ mapping prime ideals of O_l into $\mathbb{Q}(\zeta_{lm})$ and satisfying

$$\Phi(P) = g(\overline{\chi_P})^l.$$

The reason for choosing $\overline{\chi_P}$ instead of χ_P in the definition of Φ is for convenience in stating the Stickelberger relation, but first we show that $im(\Phi) \subset \mathbb{Q}(\zeta_l)$. For this, let σ_a be the automorphism of $\mathbb{Q}(\zeta_{lm})$ that sends ζ_{lm} to ζ_{lm}^a. We have $(a, lm) = 1$ and σ_a fixes $\mathbb{Q}(\zeta_l)$ iff $a \equiv 1$ (l) and similarly for $\mathbb{Q}(\zeta_p)$. To show that $\Phi(P) \in \mathbb{Q}(\zeta_l)$ it is sufficient to show that $\Phi(P)^{\sigma_a} = \Phi(P)$ for all $a \equiv 1$ (l), where we have adopted exponential notation for the action of the automorphism. Take such an a. Then we have $\overline{\chi_P}(t)^{\sigma_a} = \overline{\chi_P}(t)$ and $(\zeta_p^{tr(t)})^{\sigma_a} = \zeta_p^{atr(t)} = \zeta_p^{tr(at)}$. Hence

$$g(\overline{\chi_P})^{\sigma_a} = \sum_{t \in F_q} \overline{\chi}(t)\zeta_p^{tr(at)} = \chi_p(a)g(\overline{\chi_P}).$$

Thus $\Phi(P)^{\sigma_a} = \big(\chi_P(a)g(\overline{\chi_P})\big)^l = g(\overline{\chi_P})^l = \Phi(P)$ as desired.

The Stickelberger relation takes this result much deeper, giving a prime ideal factorization of $(\Phi(P))$. It states

Theorem 2.15 *Let O_m be the ring of integers of $\mathbb{Q}(\zeta_m)$ and let P be a prime ideal of O_m not containing m. Then*

$$(\Phi(P)) = \prod_{\substack{1 \le t < m \\ (t,m)=1}} P^{t\sigma_t^{-1}}.$$

As we have already observed in our discussion of the Davenport-Hasse identity, the Stickelberger relation is fundamental to understanding the behavior of generalized Gauss sums. It's proof, however, is long, and we defer it for the time being in order to streamline our treatment of Eisenstein reciprocity.

2.6.3 Two lemmas on cyclotomic fields

Before proceeding with the primary argument of Eisenstein reciprocity, Ireland and Rosen prove two preliminary lemmas. They are:

Lemma 2.16 *The roots of unity in $\mathbb{Q}(\zeta_m)$ are given by $\pm\zeta_m^i$ $i = 1, 2, ..., m$.*

and

Lemma 2.17 *Let $\sigma_1, \sigma_2, ..., \sigma_{\phi(k)}$ be the $\phi(k)$ automorphisms in the Galois group of $\mathbb{Q}(\zeta_k)/\mathbb{Q}$. If $\alpha \in \mathbb{Q}(\zeta_k)$ with $\alpha^{\sigma_i}| \le 1$ for all $i = 1, 2, ..., \phi(k)$ then α is a root of unity.*

To prove the first lemma, one argues that if there existed other roots of unity in $\mathbb{Q}(\zeta_m)$, the degree of the field extension would be larger than $\phi(m)$. The proof of the second lemma given by [4] is clever, so we give the argument in detail.

Proof. (2nd lemma) We define the function

$$f_\alpha(x) = \prod_{i=1}^{\phi(k)} (x - \alpha^{\sigma_i})$$

and since f_α is a power of the minimal polynomial for α, we have $f_\alpha(x) \in \mathbb{Z}[x]$. The bound on the modulus of α^{σ_i} implies that the coefficient on x^n in f is bounded by $\binom{\phi(k)}{n}$, hence only finitely many degree $\phi(k)$ polynomials in $\mathbb{Z}[x]$ can be created in from elements of $\mathbb{Q}(\zeta_k)$ with the given modulus property. But every power of α has the given property. Hence among all the polynomials generated by the powers of α, one must occur infinitely often. Then since all these polynomials' roots coincide, we must have for $i \ne j$, $\alpha^{i\sigma_m} = \alpha^{j\sigma_m}$ which implies that $\alpha^i = \alpha^j$ and α is a root of unity. □

With these two facts, we are now ready to attack Eisenstein reciprocity.

2.6.4 Proof preliminaries

We begin by extending Φ to non-prime ideals by setting

$$\Phi(A) = \prod_i \Phi(P_i)$$

where $\prod_i P_i$ is the prime ideal decomposition of A. Since, by this definition Φ is multiplicative, we recover $\Phi(AB) = \Phi(A)\Phi(B)$. Then, since $|\Phi(P)|^2 = |g(P)|^{2m} = p^{km}$, where k is the order of the prime ideal P, we have $|\Phi(P)|^2 = (N(P))^m$, and then, because both the norm and Φ are multiplicative, we have in general $|\Phi(A)|^2 = (N(A))^m$.

Because both sides of the Stickelberger relation are multiplicative, we have

$$(\Phi(A)) = \prod_{\substack{1 \le t < m \\ (t,m)=1}} A^{t\sigma_t^{-1}}.$$

Thus if α is any element of $\mathbb{Q}(\zeta_m)$, then we have, from the equality of ideals,

$$\Phi((\alpha)) = \epsilon(\alpha) \prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}}.$$

where $\epsilon(\alpha)$ is some unit. Henceforth we write $\Phi(\alpha)$ for $\Phi((\alpha))$.

We would like to determine the value of $\epsilon(\alpha)$ more precisely. To do this we calculate:

Lemma 2.18 *For $\alpha \in O_m$,*

$$\left| \prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}} \right|^2 = |N\alpha|^m.$$

Proof. As Ireland and Rosen note, σ_{-1} is complex conjugation in $\mathbb{Q}(\zeta_m)$ so

$$\left| \prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}} \right|^2 = \Bigg(\prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}} \Bigg) \sigma_{-1} \Bigg(\prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}} \Bigg).$$

But

$$\sigma_{-1} \Bigg(\prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}} \Bigg) = \prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_{-t}^{-1}} = \prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{(m-t)\sigma_t^{-1}}$$

where in the last equality we have made the substitution $t \mapsto m - t$. Thus

$$\left| \prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}} \right|^2 = \Bigg(\prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{t\sigma_t^{-1}} \Bigg) \Bigg(\prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{(m-t)\sigma_t^{-1}} \Bigg) = \prod_{\substack{1 \le t < m \\ (t,m)=1}} \alpha^{m\sigma_t^{-1}}$$

But

$$N\alpha = \prod_{\substack{1\le t<m\\(t,m)=1}} \alpha^{\sigma_t^{-1}}$$

and this completes the proof. □

From this last lemma we have

$$\left| \prod_{\substack{1\le t<m\\(t,m)=1}} \alpha^{t\sigma_t^{-1}} \right|^2 = |N\alpha|^m = |\Phi(\alpha)|^2$$

so since

$$\Phi(\alpha) = \epsilon(\alpha) \prod_{\substack{1\le t<m\\(t,m)=1}} \alpha^{t\sigma_t^{-1}}$$

we know $|\epsilon(\alpha)| = 1$. But if we expand Φ as a power of Gauss sums and apply an automorphism of $\mathbb{Q}(\zeta_m,\zeta_p)/\mathbb{Q}$ fixing ζ_p, we conclude that $\Phi(A)^\sigma = \Phi(A^\sigma)$ for ideals A of O_m and automorphisms σ of $\mathbb{Q}(\zeta_m)/\mathbb{Q}$. Thus $|\epsilon(\alpha)^\sigma| = 1$ for all automorphisms σ of $\mathbb{Q}(\zeta_m)/\mathbb{Q}$ and hence by our two preliminary lemmas, $\epsilon(\alpha)$ is a root of unity in $\mathbb{Q}(\zeta_m)$ so $\epsilon(\alpha) = \pm\zeta_m^i$ for some $i \in \{1,2,...,m\}$.

2.6.5 The congruence argument

We now begin the familiar character computations that will lead to Eisenstein reciprocity.

Proposition 2.19 *Let A, B be ideals of O_m relatively prime to m and suppose that $N(A)$ and $N(B)$ are relatively prime. Then*

$$\left(\frac{\Phi(A)}{B}\right)_m = \left(\frac{N(B)}{A}\right)_m$$

Proof. Since the norm N, Φ, and the character $\left(\frac{\cdot}{*}\right)_m$ are multiplicative, it is sufficient to consider prime ideals P and P'. Let $N(P') = q' = p'^{f'}$. Then working modulo p' in O_m, we have by the multinomial theorem

$$g(\overline{\chi_P})^{q'} \equiv \sum_{t\in O_m/P} \overline{\chi_P}(t)^{q'} \zeta_p^{q' tr(t)} \quad (p').$$

Now $q' \equiv 1 \ (m)$ so $\overline{\chi_P}^{q'} = \overline{\chi_P}$. Thus

$$g(\overline{\chi_P})^{q'} \equiv \sum_{t\in O_m/P} \overline{\chi_P}(t)\zeta_p^{tr(q't)} \equiv \left(\frac{q'}{P}\right)_m g(\overline{\chi(P)}) \quad (p').$$

But working mod P'

$$g(\overline{\chi_P})^{q'-1} = \Phi(P)^{\frac{q'-1}{m}} \equiv \left(\frac{\Phi(P)}{P'}\right)_m \quad (P').$$

Now $P' \mid (p')$ so we have

$$\left(\frac{q'}{P}\right)_m = \left(\frac{N(P')}{P}\right)_m \equiv \left(\frac{\Phi(P)}{P'}\right)_m \quad (P').$$

Then since P' does not divide m, the mth roots of unity are distinct mod P' and equality holds, completing the proof. □

Now in the case that A is the principal ideal (α) we have that

$$\left(\frac{\Phi(\alpha)}{B}\right)_m = \left(\frac{\epsilon(\alpha)}{B}\right)_m \left(\frac{\prod_t \alpha^{t\sigma_t^{-1}}}{B}\right)_m$$

or, by the multiplicativity of the character,

$$\left(\frac{\Phi(\alpha)}{B}\right)_m = \left(\frac{\epsilon(\alpha)}{B}\right)_m \prod_{(m,t)=1} \left(\frac{\alpha^{t\sigma_t^{-t}}}{B}\right)_m.$$

But

$$\left(\frac{\alpha^{t\sigma_t^{-1}}}{B}\right)_m = \left(\frac{\alpha^{\sigma_t^{-1}}}{B}\right)_m^t = \left(\frac{\alpha^{\sigma_t^{-1}}}{B}\right)_m^{\sigma_t} = \left(\frac{\alpha}{B^{\sigma_t}}\right)_m$$

where in the last equality we used the fact that $\left(\frac{X}{Y}\right)^{\sigma} = \left(\frac{X^{\sigma}}{Y^{\sigma}}\right)$. Thus

$$\prod_{(m,t)=1} \left(\frac{\alpha^{t\sigma_t^{-t}}}{B}\right)_m = \left(\frac{\alpha}{\prod_t B^{\sigma_t}}\right)_m = \left(\frac{\alpha}{N(B)}\right)_m,$$

and

$$\left(\frac{\Phi(\alpha)}{B}\right)_m = \left(\frac{\epsilon(\alpha)}{B}\right)_m \left(\frac{\alpha}{N(B)}\right)_m = \left(\frac{N(B)}{A}\right)_m.$$

We need two more lemmas before we conclude our proof of Eisenstein reciprocity.

Lemma 2.20 *Let l be an odd prime and A be an ideal of O_l relatively prime to l. Then $\Phi(A) \equiv \pm 1\ (l)$.*

Proof. By the multiplicativity of Φ it is sufficient to show that $\Phi(P) \equiv -1\ (l)$ for prime ideals P not dividing l. But $\Phi(P) = g(\overline{\chi_P})^l$ and

$$g(\overline{\chi_P})^l \equiv \sum_t \overline{\chi_P}(t)^l \zeta_p^{tr(lt)} \quad (l).$$

Now $\overline{\chi_P}(t)^l$ is 1 if $t \neq 0$ and 0 otherwise so this last sum is equal to $\sum_{t\neq 0} \zeta_p^{tr(lt)} = -1$ since $\sum_t \zeta_p^{tr(lt)} = 0$. Thus $\Phi(P) \equiv -1 \;\; (l)$. □

Lemma 2.21 *If $\alpha \in O_l$ is primary, then $\epsilon(\alpha) = \pm 1$.*

Proof. Since α is primary it is prime to l and we can apply the above lemma:

$$\Phi(\alpha) = \epsilon(\alpha) \prod_{(t,l)=1} \alpha^{t\sigma_t^{-1}} \equiv \pm 1 \quad (l).$$

Now $l = \prod_{i=1}^{l-1} (1-\zeta_l)^{\sigma_t}$ and for each t, $(1-\zeta_l)^{\sigma_t} = 1 - \zeta_l^t$ is a unit times $1-\zeta_l$ since, with s the multiplicative inverse of t mod l, $1 - \zeta_l = \big(1-\zeta_l^t\big)\big(1 + \zeta_l^t + \zeta_l^{2t} + ... + \zeta_l^{(s-1)t}\big)$ and $1 - \zeta_l^t = \big(1-\zeta_l\big)\big(1 + \zeta_l + ... + \zeta_l^{t-1}\big)$. Thus $(1-\zeta_l)^2 \mid l$ and

$$\epsilon(\alpha) \prod_{t=1}^{l-1} \alpha^{t\sigma_t^{-1}} \equiv \pm 1 \quad ((1-\zeta_l)^2).$$

But α primary also implies that $\alpha \equiv n \;\; ((1-\zeta_l)^2)$ for some $n \in \mathbb{Z}$. Hence

$$\prod_{t=1}^{l-1} \alpha^{t\sigma_t^{-1}} \equiv \prod_{t=1}^{l-1} n^{t\sigma_t^{-1}} \equiv \prod_{t=1}^{l-1} n^t \equiv n^{\frac{l(l-1)}{2}} \quad ((1-\zeta_l)^2),$$

and by Euler's criterion, $n^{\frac{l-1}{2}} \equiv \pm 1 \;\; (l)$ so $n^{\frac{l(l-1)}{2}} \equiv \pm 1 \;\; ((1-\zeta_l)^2)$. Thus $\epsilon(\alpha) \equiv \pm 1 \;\; ((1-\zeta_l)^2)$.

But we know that $\epsilon(\alpha) = \pm \zeta_l^j$ for some $j \in \{1, 2, ..., l\}$. With $\zeta_l = 1 - (1-\zeta_l)$ this implies that $\epsilon(\alpha) \equiv \pm(1 - j(1-\zeta_l)) \;\; ((1-\zeta_l)^2)$ by binomial expansion. Hence, $1 - j(1-\zeta_l) \equiv \pm 1$ $((1-\zeta_l)^2)$. In the case of -1, we have $(1-\zeta_l) \mid 2$, a contradiction, so we have $1 - j(1-\zeta_l) \equiv 1$ $((1-\zeta_l)^2)$. But, subtracting 1 from both sides, this implies that $(1-\zeta_l) \mid j$ so $j = l$ (for instance, by taking norms). Thus $\epsilon(\alpha) = \pm 1$ as desired. □

With this last result, we can strengthen our previous proposition. If $\alpha \in O_l$ is primary and B is an ideal that is prime to l and has $N(B)$ prime to $N(\alpha)$ then

$$\left(\frac{\alpha}{N(B)}\right)_l = \left(\frac{N(B)}{\alpha}\right)_l.$$

This holds because $\epsilon(\alpha) = \pm 1$ and l is odd so

$$\left(\frac{\epsilon(\alpha)}{N(B)}\right)_l = 1.$$

We are now finally ready to prove Eisenstein reciprocity.

Proof. (Eisenstein Reciprocity) We have α which is primary, so take a prime $p \in \mathbb{Z}$ other than l and let P be a prime ideal factor of p in O_l with $N(P) = p^f$ for some $f \mid l-1$. Then we know that

$$\left(\frac{\alpha}{N(P)}\right)_l = \left(\frac{N(P)}{\alpha}\right)_l$$

so

$$\left(\frac{\alpha}{p}\right)_l^f = \left(\frac{p}{\alpha}\right)_l^f.$$

Since the characters are powers of ζ_l and $f \mid l-1$ is prime to l, we have

$$\left(\frac{\alpha}{p}\right)_l = \left(\frac{p}{\alpha}\right)_l.$$

Then by multiplicativity, we have

$$\left(\frac{\alpha}{a}\right)_l = \left(\frac{a}{\alpha}\right)_l$$

for all $a \in \mathbb{Z}$ such that a is prime to both l and α. □

Remarks: We have given a proof of Eisenstein reciprocity due to Ireland and Rosen. Berndt et. al. give a more general formulation that relates kth order characters for $k = l^n$, $(k \neq 2, 4)$ where l is any prime and n is a positive integer (p. 474), but the proof of Ireland and Rosen has the advantage that it mimics the proofs of the reciprocity laws that it generalizes: for the l-order character χ_α Ireland and Rosen compute $g(\chi_\alpha)^l$ up to a sign when α is primary, then work with congruences modulo a prime ideal to extract information about the value of characters from this determination. The proof of Berndt, while more general, lacks this intuitive structure. Primary integers are defined implicitly from a ratio of Gauss sums rather than being set congruent to an integer modulo the square of a prime ideal, and a more general version of the Stickelberger congruence is employed that obscures the tidy factorization of $(g(\chi_P)^l)$ into automorphisms of P in Ireland and Rosen.

2.7 Wieferich's Theorem

As an application of Eisenstein Reciprocity we give a proof of Wieferich's theorem which places a severe limit on the potential solutions to Fermat's equation: $x^n + y^n = z^n, n > 2$.

Theorem 2.22 *Let l be an odd prime and suppose x, y, and z are integers not divisible by l that satisfy $(*)$ $x^l + y^l + z^l = 0$. Then $2^{l-1} \equiv 1$ (l^2).*

Proof. (from [2] pp. 490-2) Observe that if any two of x, y, z have common factor k then the third must be divisible by k as well so that $(x_1, y_1, z_1) = (\frac{x}{k}, \frac{y}{k}, \frac{z}{k})$ is another solution to $(*)$ satisfying the given conditions with $gcd(x_1, y_1, z_1) < gcd(x, y, z)$. Thus there is no loss of generality in assuming that x, y and z are pairwise relatively prime. Next note that exactly two of x, y, z must be odd and the other even so, again without loss of generality suppose that x and z are odd and y is even.

Now write $(*)$ as $x^l + y^l = -z^l$ and with $\zeta_l = e^{\frac{2\pi i}{l}}$ we have the cyclotomic factorization:

$$x^l + y^l = (x+y)(x+\zeta_l y)(x+\zeta_l^2 y^2)...(x+\zeta_l^{l-1} y^{l-1}) = (-z)^l$$

where we have used the fact that l odd implies $-z^l = (-z)^l$. Now let I_j be the principal ideal $(x+\zeta_l^j y)O_l$ for $j = 0, 1, ..., l-1$ and O_l the ring of integers of $\mathbb{Q}(\zeta_l)$ and let $J = zO_l$. Then our above factorization translates to

$$I_0 I_1 ... I_{l-1} = J^l$$

in terms of ideals.

We show that the I_j are pairwise relatively prime and conclude that each is an lth power. For this, suppose I_i and I_j have common prime ideal factor P. Then we have $x + \zeta_l^i \in P$ and $x + \zeta_l^j \in P$ from which we conclude via linear combination that $(\zeta_l^i - \zeta_l^j)x, (\zeta_l^i - \zeta_l^j)y \in P$. Since x and y are coprime in $\mathbb{Z}$ and hence also in O_l these two conditions together imply that $\zeta_l^i - \zeta_l^j \in P$. But then $1 - \zeta_l^{j-i} \in P$, and with $k = j - i$ and k^{-1} the multiplicative inverse of k in $\mathbb{Z}/(l\mathbb{Z})$ we have $(1-\zeta_l^k)(1+\zeta_l^k+\zeta_l^{2k}+...+\zeta_l^{(k^{-1}-1)k}) = 1-\zeta_l \in P$. But $\lambda = 1 - \zeta_l$ is a prime since it has norm l so we must have $P = \lambda O_l$. But then $\lambda O_l \mid zO_l$ implying, taking norms, that $l \mid z$, a contradiction. Hence the I_j are pairwise relatively prime and each is an lth power. Write $I_j = J_j^l$.

Now put $\alpha = (x+y)^{l-2}(x+\zeta_l y)$. The idea of the proof is to relate the value of the characters $\left(\frac{\alpha}{2}\right)_l$ and $\left(\frac{2}{\alpha}\right)_l$ in two ways in order to extract information about $\left(\frac{\zeta_l}{2}\right)_l$. Observe that we have

$$\alpha O_l = I_0^{l-1} I_1 = J_0^{l(l-1)} J_1^l$$

so that the principal ideal generated by α is a perfect lth power. Hence, as $\left(\frac{\cdot}{*}\right)_l$ is an lth root of unity,

$$\left(\frac{2}{\alpha}\right)_l = 1.$$

Since $x \equiv 1 \ (2), y \equiv 0 \ (2)$, $\alpha \equiv 1 \ (2)$ so

$$\left(\frac{\alpha}{2}\right)_l = \left(\frac{1}{2}\right)_l = 1.$$

Now set $t = (x+y)^{l-2}y (\in \mathbb{Z})$ and observe $\alpha = (x+y)^{l-1} - (x+y)^{l-2}(1-\zeta_l)y = (x+y)^{l-1} - \lambda t$. Since $x^l + y^l = (-z)^l$ we have $x + y \equiv -z$ from Fermat's Little Theorem so l doesn't divide $x+y$ and $(x+y)^{l-1} \equiv 1 \ (l)$. Moreover

$$l = (1-\zeta_l)(1-\zeta_l^2)...(1-\zeta_l)^{l-1}$$

and each $1-\zeta_l^i$ is a unit times λ (we have already shown that $1-\zeta_l^k \mid 1-\zeta_l$ but $1-\zeta_l^k = (1-\zeta_l)(1+\zeta_l+...+\zeta_l^{k-1})$ showing the two are related by a unit) so since $l \geq 3$ we have $\lambda^2 \mid l$ and $(x+y)^{l-1} \equiv 1 \ (\lambda^2)$. Hence

$$\alpha \equiv 1 - \lambda t \quad (\lambda^2).$$

Now by the binomial theorem,

$$\zeta_l^{-t} = (1-\lambda)^{-t} \equiv 1 + \lambda t \quad (\lambda^2)$$

so $\alpha\zeta_l^{-t} \equiv 1 \ (\lambda^2)$ and in the sense of Eisenstein, $\alpha\zeta_l^{-t}$ is primary. We know ζ_l^{-t} is a unit so it follows that

$$\left(\frac{2}{\alpha}\right)_l = \left(\frac{2}{\alpha\zeta^{-t}}\right)_l = \left(\frac{\alpha\zeta_l^{-t}}{2}\right)_l = \left(\frac{\zeta_l^{-t}}{2}\right)_l \left(\frac{\alpha}{2}\right)_l$$

and hence

$$\left(\frac{\zeta_l^{-t}}{2}\right)_l = 1 = \left(\frac{\zeta_l}{2}\right)_l^t.$$

But in O_l we have the prime factorization

$$2O_l = P_1P_2...P_{(l-1)/s}$$

where s is the order of 2 mod l and each P_i has norm 2^s (see, for instance, [2] pp. 342-3 or [4] pp. 196-8). Hence

$$\left(\frac{\zeta_l}{2}\right)_l = \prod_{i=1}^{(l-1)/s} \left(\frac{\zeta_l}{P_i}\right)_l = \prod_{i=1}^{(l-1)/s} \zeta_l^{\frac{2^s-1}{l}} = \zeta_l^{\frac{(2^s-1)(l-1)}{sl}}$$

so

$$\left(\frac{\zeta_l}{2}\right)_l^t = 1$$

implies that

$$l \mid \frac{(2^s-1)}{l}\frac{(l-1)}{s}t.$$

But l divides neither t nor $l-1$ so we must have that $l \mid \frac{2^s-1}{l}$ or $l^2 \mid 2^s-1$. But then $s \mid l-1$ so $2^{l-1} \equiv 1 \ (l^2)$ as desired. □

Remarks: This theorem of Wieferich represents a very strong elementary result on Fermat's Last Theorem: Berndt notes that the only primes less than $4 \cdot 10^{12}$ satisfying $2^{l-1} \equiv 1 \ (l^2)$ are 1093 and 3511. Ireland and Rosen prove a somewhat stronger result due to Furtwängler that replaces 2 with any prime p dividing y. The approach is almost entirely identical.

The cyclotomic factorization of x^l+y^l is a common trick used in elementary approaches to Fermat's Last Theorem; for instance, Ireland and Rosen make use of the same factorization to show that $x^l + y^l = z^l$ does not have solution in positive integers x, y, z prime to l if l is a prime not dividing the class number of $\mathbb{Q}(\zeta_l)$. After each of the principal ideals $x + \zeta_l^i y$ is shown to be an lth power, most of the work of the theorem is done. The choice of computing the character of ζ_l is clever because it allows the congruence in the definition of χ_P to be replaced with equality. After this choice, the proof is essentially careful bookkeeping.

Unfortunately, it is unclear how to generalize this approach to other high-order Diophantine equations. The cyclotomic factorization, for instance, is relatively specific to Fermat's equation since it is applicable only in the case of a prime exponent and two variables. The trick of using Eisenstein reciprocity to relate two characters that differ by a unit seems more amenable to generalization, but Eisenstein reciprocity suggests the case of a homogenous, high order Diophantine equation, and among these, the Fermat equation is the simplest!

Chapter 3

Concluding remarks and future considerations

Our classical approach to understanding Gauss sums is frustrated by an apparent randomness in the sums' multiplicative behavior - a notion made precise in Hasse's conjecture - that underscores the need for more modern and powerful techniques. A thorough understanding of the Stickelberger relation, which underlies most work done on Gauss sums during the 20th century, is a natural starting point for further investigation.

During our present discussion, as suggested by Weil [7], we have already seen that an interpretation of the Gauss sum as a Fourier expansion is natural, and leads to useful results. Meanwhile, Yokoyama [] considers the Gauss Jacobi sums as discrete analogues of the p-adic gamma and beta functions. A search for similar analogues to special p-adic functions might prove fruitful.

In his 1979 paper [5] Matthews gives an expression for the cubic Gauss sum in terms of the Weierstrass $\wp$-function satisfying $\wp'^2 = 4\wp^3 - 1$ by considering the related elliptic curve over a finite field. In particular, Matthews proves

$$g_3(\chi_\pi) = p^{\frac{1}{3}} \pi \alpha(S)^{-1} \prod_{s \in S} \wp\left(\frac{s\theta}{\pi}\right)$$

where π is the principal prime factor of p in $\mathbb{Z}[\omega]$, S is a set of $\frac{p-1}{3}$ residues modulo π such that $S \cup \pi S \cup \pi^2 S$ contains all the non-zero residues modulo π and $\alpha(S) \equiv \prod_{s\in S} s \ (\pi)$. This represents an unlooked-for connection between the study of elliptic curves and Gauss sums, and suggests that an entire field of study in modern number theory might be brought to bear upon the determination of Gauss sums.

While the approaches of Weil, Yamamoto and Matthews may seem haphazard and distinct, if there is common ground among contemporary results on Gauss sums it lies in drawing parallels between the sums and other number-theoretic formalisms. Thus, in our ongoing study of Gauss sums we should remain open to ideas from disparate areas that might bring structure to a field which lacks readily-apparent tools of its own to work with.

Chapter 4

Appendix: Source code for computing Gauss sums

The following is C++ source code for calculating Gauss sums of various orders mod a prime.

```
#include <stdlib.h>
#include <math.h>
#include <stdio.h>
#include <iostream>
#include <fstream>
using namespace std;

#define maxPrimeSize 10
#define pi 3.14159265
#define filename "29.txt"
#define constPrime 73

bool generates(int n, int prime, int arr[]);
int getprimroot(int prime);
void getprime(/*fstream fout*/);
void gauss(int prime, int order, int power, int root,
double & real, double & imag);
int gcd(int a, int b);
void getGaussSums(int prime /*, fstream fout*/);

/* To print to file, set filename above and delete commented out sections */

int main(int argc, char **argv)
{
/*fstream fout;*/
```

```
// must pass fout to either getprime or getGaussSums if writing to file
/*fout.open(filename.c_str(), fstream::out);*/
getprime(/*fout*/);
// will find all primes less than maxPrimeSize (set above) and
// compute all Gauss sums for these primes
// getGaussSums(constPrime /*,fout*/);
// computes all Gauss sums for prime set in constPrime above
/*fout.close;*/
}


//prime sieve that calls getGaussSums on every prime that it finds
void getprime(/*fstream fout*/)
{
int primearray[maxPrimeSize+1];
for(int i = 0; i<maxPrimeSize+1; i++)
    {
    primearray[i]=1;
    }
for(int i = 2; i<maxPrimeSize+1; i++)
    {
    if(primearray[i] == 1)
        {
        getGaussSums(i /*,fout*/);       // pass fout if writing to file
        for(int j = i*i; j<maxPrimeSize+1; j+=i)
                {
                primearray[j] = 0;
                }
        }
    }
}

// first calls getprimroot to find a primitive root mod the prime.
// then calls gauss for all the powers 1 to p-2.  this computes the gauss
// sum of the order p-1 character to the power 1, 2, ..., p-2
void getGaussSums(int prime /*, fstream fout*/)
{
int primroot = getprimroot(prime);
double real, imag;
cout<<"Prime "<<prime<<" has Gauss sums:" <<endl; // fout to write to file
for(int k = 1; k < prime-1; k++)
```

```
    {
    gauss(prime, prime-1, k, primroot, real, imag);
    cout<< "g(X^" << k << ") = " << real << "  +  " << imag << " i" <<endl;
// fout to write to file
    }
}


// primitive root sieve.  cycles through the residues mod the prime,
// calling generates if the flag for the element is set to 1
// (see generates) in checkarray.
int getprimroot(int prime)
{
int *checkarray = new int[prime];
for(int i = 0; i< prime; i++)
    {
    checkarray[i] = 1;
    }
for(int j = 2; j<prime; j++)
    {
    if(checkarray[j] == 1)
        {
        if(generates(j, prime, checkarray))
                {
                delete[] checkarray;
                return j;
                }
        }
    }
delete checkarray;
return -1;
}

// determines if the input element, n, is a generator mod the prime.
// does this by verifying that the first (p-1)/2 powers of n are not 1
// mod p.  as it works, sets each of these residues to 0 in the checkarray.
bool generates(int n, int prime, int *arr)
{
int res = 1;
for(int i = 1; i< (prime+1)/2; i++)
    {
```

```
    res = (res*n)%prime;
    arr[res] = 0;
    if(res == 1)
        {
        return false;
        }
    }
return true;
}

// takes in a prime, the order of the character, the power of the character,
// a primitive root mod the prime and stores the gauss sum in the
// referenced variables real and imag
void gauss(int prime, int order, int power, int root,
double & real, double & imag)
{
real = imag = 0;
int prodcount=1;
for(int count = 1; count<prime; count++)
    {
    prodcount*= root;
    prodcount = prodcount%prime;
    real += cos(2 * pi *(double(count) * power / order +
double(prodcount) / prime));
    imag += sin(2 * pi *(double(count) * power / order +
double(prodcount) / prime));
    }
}


// a helper function not currently being used (although it’s handy,
// for instance, in checking the order of a power of a character).
// harnesses the euclidean algorithm to return the gcd of two numbers.
int gcd(int a, int b)
{
int temp;
if(b == 0) return a;
if(a == 0) return b;
if(a%b == 0) return b;
else return gcd(b, a%b);
}
```

Bibliography

[1] M. Beck, B. Berndt, O. Chan, A. Zaharescu. *Determination of analogues of Gauss sums and other trigonometric sums* avail. online: arXiv:math.NT/0504160v2, 2005.

[2] B. Berndt, R. Evans, K. Williams. *Gauss and Jacobi Sums*, volume 21 of *Canadian Mathematics Society Series of Monographs and Advanced Texts.* Wiley-Interscience, 1998.

[3] H. Davenport. *Multiplicative Number Theory*, volume 74 of *Graduate Texts in Mathematics.* Springer, New York, third edition, 2000.

[4] K. Ireland and M. Rosen. *A Classical Introduction to Modern Number Theory*, volume 84 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1990.

[5] C. Matthews. Gauss Sums and Elliptic Functions. *Inv. Math.*, 52:163–185, 1979.

[6] J. Muskat. On Jacobi sums of certain composite orders. *Trans. Amer. Math. Soc.*, 134:483–502, 1969.

[7] A. Weil. Numbers of solutions of equations in finite fields. *Bull. Amer. Math. Soc.*, 55:497–508, 1949.

[8] K. Yamamoto. On a conjecture of Hasse concerning multiplicative relations of Gauss sums. *J. Comb. Theory*, 476–489, 1966.