A History of Stickelberger's Theorem

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation *with research distinction* in Mathematics in the undergraduate colleges of The Ohio State University

by

Robert Denomme

The Ohio State University
June 8, 2009

Project Advisor: Professor Warren Sinnott, Department of Mathematics

Contents

Introduction

The late Professor Arnold Ross was well known for his challenge to young students,

"Think deeply of simple things."

This attitude applies to no story better than the one on which we are about to embark. This is the century long story of the generalizations of a single idea which first occurred to the 19 year old prodigy, Gauss, and which he was able to write down in no less than 4 pages. The questions that the young genius raised by offering the idea in those 4 pages, however, would torment the greatest minds in all the of the 19^{th} century. Just as Dr. Ross would advise, the answers that the successful minds eventually reached were all found by offering a slightly more elegant treatment of the known material, and then pushing that treatment in a very natural way to a more general setting. It was in discovering which setting the treatment needed to be considered that this single idea created an entirely new branch of mathematics which has become fundamental in the study of astonishingly many objects occurring throughout mathematics.

The particular goal of this paper is to examine the historical development of the area of 19^{th} century mathematics surrounding Gaussian sums and their generalizations leading to the proof of Stickelberger's theorem on ideal class annihilators. What the reader should take away from this presentation is the way in which the entire development of algebraic number theory in this period was inspired simply by Gauss' sixth proof of quadratic reciprocity, a short proof demonstrating essentially only a connection between cyclotomy and reciprocity laws. We stress again and again how the fundamental developments of the subject were due directly to the effort of generalizing this proof of quadratic reciprocity to higher reciprocity laws. The historical treatment allows a nearly self contained presentation of the material, and is meant to serve as a second introduction to the subject of reciprocity laws from the viewpoint that they were originally discovered in, ending at the following theorem due to Stickelberger.

Theorem. *Let m be a positive integer and set $K = \mathbb{Q}(\zeta)$ where ζ is a primitive m^{th} root of unity, $G = Gal(K/\mathbb{Q})$ where $\sigma_t \in G$ acts via $\sigma_t(\zeta) = \zeta^t$. Let*

$$\theta = \frac{1}{m} \sum_{\substack{(t,m)=1 \\ 0<t<m}} t\sigma_k^{-1} \in \mathbb{Q}[G]$$

be the Stickelberger element. For any $\beta \in \mathbb{Z}[G]$ which satisfies $\beta\theta \in \mathbb{Z}[G]$, then under the natural action of $\mathbb{Z}[G]$ on the class group of K as described in Section 4, $\beta\theta$ annihilates the class group of K. Moreover if $\mathfrak{p}$ is a prime ideal in the ring of integers of K which does not contain m then we have the explicit computation

$$\mathfrak{p}^{m\theta} = (G(\mathfrak{p})^m)$$

where $G(\mathfrak{p})^m \in K$ is the Gauss sum defined in beginning of Section 4.1.

Algebraic number theory in this period is perhaps too large a subject to focus on, as is seen in Hilbert's famous Zahlbericht, the fundamental text on the techniques of 19^{th} century algebraic number theory, which excludes completely Stickelberger's result.

Also the topic of Reciprocity laws, even when limiting oneself to the laws from Euler to Eisenstein, is quite broad, *cf.* [23]. We choose therefore to focus on the specific historical development of Stickelberger's theorem. This appears ideal looking through history as it stems from such a natural and highly motivated goal: generalize Gauss' sixth proof of quadratic reciprocity to obtain higher reciprocity laws. It also highlights in surprising ways the crucial developments of algebra that took place during this area, including nearly the entire area of commutative ring theory, and how they were motivated by this single goal. Finally the still active question of determining the structure of the ideal class group of a number field which was studied in response to this goal is also in some part answered by the study of Gaussian sums themselves which appear crucially in Stickelberger's theorem. In fact the answer that the theorem gives to the problem of determining the class group is of the most explicit answers to the question, and it plays a key role in proofs that are not yet even a decade old.

We pick the most influential papers of the topic and present them in as unfiltered a manner as possible, following a track that leads from Fermat's discovery of the first reciprocity laws all the way to Stickelberger's 1890 paper to a short discussion of material that is still active research today. Despite the possible clashes with modern notation, the original notation is often preserved, but as will be seen Kummer's use of the polynomial notation, $f(\alpha)$, to denote a polynomial in the λ^{th} root of unity, α, was a key tool in his explanation of ideal numbers. It is also surprising how simple the proofs of some familiar theorems are in their original notation, and how many interesting results are presented in these classical papers that might not appear in more modern treatments. One therefore can use this presentation as an assistant to reading these classic papers in a quest to study the masters.

The organization of the material is as follows. Section 1 attempts to cover the material up to Gauss' sixth proof of quadratic reciprocity. In particular it first covers the cyclotomy side, solving polynomial equations using roots of unity and then it covers the reciprocity law side, what the techniques of proving quadratic reciprocity were that existed leading up to and including the sixth. The next section covers Cauchy's, Jacobi's and Eisenstein's attempts to generalize the one proof to cubic and bi-quadratic reciprocity, though less effort is spent on biquadratic, as it is not that essentially different from the cubic case. This work contains the call to define algebraic integers and describe their properties, in particular to prove the laws of unique factorization in more exotic domains. The material up to this point may be understood by any student which has had a first course in number theory, most importantly a familiarity with primitive roots and to a lesser extent unique factorization domains would be a prerequisite. Section 3 covers some main parts of Kummer's influence on reciprocity laws. Much of the focus is spent on the paper [17] in which he both describes for the first time the properties of ideal numbers and proves a prime decomposition theorem for Gaussian sums, essentially the largest step towards proving the whole Stickelberger theorem. Somewhat deviating from the plan of this presentation the original proof of Eisenstein reciprocity is then given in an attempt to measure the amount of influence that Kummer's work had on the world of number theory in such a short amount of time. Finally, in Section 4 we come to Stickelberger's paper and other proofs of his theorem. The last

section is devoted to discussing the mathematics in the 20^{th} and 21^{st} century pertaining to Stickelberger's theorem, and is meant as a very brief survey of a few of the topics of modern number theory that Stickelberger's theorem applies to.

Acknowledgements. I would like to thank Dr. Warren Sinnott for all his help with this thesis. I thank Dr. Daniel Shapiro for encouraging my interest in number theory and for continuing the Ross mathematics program which started my interest in reciprocity laws. I would like also to thank Dr. Vitaly Bergelson for creating such a wonderful honors math program at Ohio State, and for running his number theory through history course. I thank Dr. G. Savin and Dr. J. Cogdell as they have been instrumental in learning the academic and social aspects of mathematics. I also would like to thank all of the Ohio State honors mathematics students, and of course, my family and friends for helping me be able to pursue the the further study of mathematics. Finally, I thank everyone I have met that has encouraged me to become a good writer.

1. Gauss's Cyclotomy and Quadratic Reciprocity

The two problems that have had the most influence on the discovery of classical Gauss sums were the solution of polynomial equations of higher degree and quadratic reciprocity. This history takes place almost entirely in the 18^{th} century, beginning with Leonhard Euler, who as we will see started both of these topics from our point of view. The relationship between these two problems was discovered by Gauss when he applied the techniques of the former to the solution of the latter and almost single handedly sparked the discovery algebraic number theory.

1.1. **Solution of the General Equation.**

1.1.1. *(Leonhard Euler & Étienne Bézout).* The 18th Century witnessed a great explosion of advancements in pure mathematics as Euler took on the challenges of Fermat and many topics motivated by only their own beauty. Of these challenges was the quest to solve a polynomial equation in one variable algebraically. Before the 18th century the classical quadratic formula would yield the solution to the degree two case, while the work of Cardano and Ferrari had given a formula for both the third and fourth degree cases. Newton also devised Newton's method for solving polynomial equations numerically, though this was an analytic answer to the question. A new, very simple method to solve equations of small degree was discovered in the 1760's by both Euler and Bézout which became known as Bézout's method.

Given a polynomial $p(x)$ of degree n, Bezout's method works by cleverly choosing coefficients $a_0, a_1 \dots a_{n-1}$ so that $p(x)$ is in fact equal to the polynomial,

$$R_n(x) = \prod_{\omega}(x - (a_0 + a_1\omega + \cdots + a_{n-1}\omega^{n-1}))$$

where ω runs over the n distinct nth roots of unity. Thus the roots of the equation $p(x) = R_n(x) = 0$ will be given by the values $a_0 + a_1\omega + \cdots + a_{n-1}\omega^{n-1}$ for the different roots of unity ω. For a given $p(x)$ there is no guarantee that such coefficients will have a

simple expression, and a sharp student of Galois theory will immediately see that it will be impossible for some $p(x)$ to choose the a_i to be in some solvable extension of $\mathbb{Q}$. Amazingly, though, this method gives explicit roots of a general polynomial of degree 2, 3 or 4, after, perhaps, a change of variables. For instance, to find the roots of $p(x) = x^3 + px + q$ we write out,

$$R_3(x) = (x - a_0)^3 - 3a_1a_2(x - a_0) - (a_1^3 + a_2^3).$$

Comparing with $p(x) = x^3 + px + q$ we set $a_0 = 0$ and find that $a_2 = -\frac{p}{3a_1}$. Finally we see that a_1 satisfies the equation,

$$a_1^6 + qa_1^3 - \left(\frac{p}{3}\right)^3 = 0$$

which is simply a quadratic polynomial in a_1^3. We may thus explicitly solve for a_1, a_2, a_3 and give the roots of $p(x)$ as $a_0 + a_1\omega + a_2\omega^2$ where ω ranges over the cubic roots of unity. Thus we have determined the a_i and hence the roots x_i in terms of radicals, sums, quotients and products of rational numbers and cubic roots of unity. Using similar observations one may arrive at the solution of the reduced quartic polynomial $p(x) = x^4 + px^2 + qx + r$ as well, though the calculations are more difficult. Because the cubic and quartic roots of unity also all have an expression in terms of radicals, sums quotients and products of rational numbers we can thus give a formula for the solutions of the equation $p(x) = 0$ in terms of the coefficients and the regular operations of arithmetic. In general we say that the polynomial $p(x) = 0$ is solvable in radicals when we may write each solution x as $a + bi$ for some real a, b which have an expression in terms of rational numbers and only the arithmetic operations on the real numbers of sums, products, quotients and root extractions. As an example one may write the 5^{th} root of unity $\zeta_5 = e^{\frac{2\pi i}{5}}$ as follows

$$\zeta_5 = \frac{-1 + \sqrt{5}}{4} + \sqrt{\frac{5 + \sqrt{5}}{8}}i$$

This definition is geometrically motivated by the complex plane, each complex number is given by its real part and its imaginary part.

At this time the solution of $Y^n = 1$ in radicals had been carried out for n less than 11 (*cf.* [7]), thus Euler and Bezout had explicit formulas for the cubic and quartic equations.

1.1.2. *(Joseph-Louis Lagrange & Alexandre Vandermonde).* The next major step towards the solutions of such equations came from Lagrange in his great work [21] of 1771. Lagrange hoped to demonstrate exactly why these methods would fail for general equations of degree greater than four. He considers the roots $x_1, x_2 \dots x_n$ of $R_n(x)$ again in the form,

$$\begin{aligned}
x_1 =& a_0 + a_1 + a_2 + \cdots + a_{n-1} \\
x_2 =& a_0 + a_1\zeta + a_2\zeta^2 + \cdots + a_{n-1}\zeta^{n-1} \\
x_3 =& a_0 + a_1\zeta^2 + a_2\zeta^4 + \cdots + a_{n-1}\zeta^{2(n-1)} \\
& \cdots \\
x_n =& a_0 + a_1\zeta^{n-1} + a_2\zeta^{2(n-1)} + \cdots + a_{n-1}\zeta^{(n-1)^2},
\end{aligned}$$

where ζ is a fixed primitive nth root of unity. He then notes that by multiplying each equation by an appropriate power of ζ so that the coefficient of a_k is 1, summing the equations and using the relation $1 + \zeta + \zeta^2 + \cdots + \zeta^{n-1} = 0$ we may obtain the relations

$$a_k = \frac{1}{n}\sum_{i=1}^{n} \zeta^{-(i-1)k} x_i$$

With very clever observations Lagrange then determines that each a_k may be expressed as a root of an equation of degree $(n-2)!$ when n is a prime, thus for the degree five case one still needs to solve a degree six polynomial. Lagrange thus concludes that one will probably not have luck looking at the polynomial $R_n(x)$ to help solve the higher degree cases.

Because of this original use of the nth roots of unity as a tool, Lagrange may be considered one of the fathers of our modern view of cyclotomy. In the later work [22] he recounts the properties of sums of the form

$$y = x_1 + x_2\zeta + x_3\zeta^2 + \cdots + x_n\zeta^{n-1} \tag{1}$$

and because of their ability to resolve the solutions of equations, names them *"resolvents"*, for which we have the name today, *Lagrange resolvents*. It is also noticed there that y^n is invariant under cyclic permutations of the x_i, a key factor that will carry over to Gauss sums as we will see. It is historically worth mentioning that in this paper Lagrange also discovers the fundamental theorem of symmetric functions and while considering groups of permutations concludes his famous theorem in group theory that the order of a subgroup divides the order of the group.

Almost concurrent with Lagrange's work was the work of Vandermonde. In the paper [29] he is led to introduce notation for permutations, worked with what were essentially Lagrange resolvents and went even further by successfully examining the specific case of resolving $Y^{11} = 1$ into radicals. In his work one sees the first sparks of Galois theory in that to study the structure of a polynomial one only needs consider the permutations of its roots which preserve the relations between the roots. This impressed the great Kronecker some 100 years later when he states in the preface to his German translation of [29],

> *With Vandermondes memoir on the resolution of equations, presented in 1770 to the Parisian Academy, began a new blossoming of algebra; the*

profundity of the view which is expressed in such clear words in this work arouses nothing less than our astonishment

Vandermonde's observations were initially much less influential than Lagrange's as his paper was held nearly two years longer than Lagrange's in the publication process in addition to the fact that the latter work was more comprehensive.

1.1.3. *(Carl Friedrich Gauss).* Gauss' contributions to the theory of cyclotomic fields were nothing short of monumental. Moreover, his largest contributions to the subject he completed before the age of 19 and were published in his famous Disquisitiones Arithmeticae, [9]. It is here that he defines the *periods* of a cyclotomic field. To define and work with these periods a few number theoretical remarks were in order. Foremost Gauss shows that for a prime p there exists a *primitive root* of p, that is an integer g such that the $p-1$ powers $g^0(=1), g^1, g^2 \dots g^{p-2}$ represent the $p-1$ distinct non-zero residues modulo p. Let f be a divisor of $p-1$ so that $p-1=ef$ and set $h=g^e$. For any λ prime to p the set of residues of $\lambda, \lambda h, \lambda h^2, \dots \lambda h^{f-1}$ does not depend on the choice of the primitive root g, changing g only permutes the order of the terms. Fixing a p^{th} root of unity ζ the summation

$$\zeta^{\lambda} + \zeta^{\lambda h} + \zeta^{\lambda h^2} + \cdots + \zeta^{\lambda h^{f-1}} \tag{2}$$

is denoted (f, λ) and the set of ζ^{α} for α equal to $\lambda, \lambda h, \lambda h^2, \dots \lambda h^{f-1}$ is called the period of (f, λ). One should immediately note that equation (2) is a special case of a Lagrange resolvent from equation (1) with each x_i either zero or one. It is then proven that for λ_1, λ_2 not divisible by p there exists a relation

$$(f, \lambda_1) = a_0 + a_1(f, \lambda_2) + a_2(f, \lambda_2)^2 + \cdots + a_{e-1}(f, \lambda_2)^{e-1}$$

where the a_i are determined rational numbers, thus if the sum of one of the periods for f is resolvable then any other such sum for f is. Next it is shown that if f' is a divisor of f then (f', λ_1) is a root of a determined equation of degree $\frac{f}{f'}$ whose coefficients are in the field generated over $\mathbb{Q}$ by (f, λ_2). Observing that $(1,1) = \zeta$ we may conclude that there is a bijective correspondence, $f \leftrightarrow \mathbb{Q}((f,1))$ between the divisors of $p-1$ and the subfields of $\mathbb{Q}(\zeta)$ where division on one side implies containment on the other. Thus Gauss has described explicitly the Galois theory of the cyclotomic extensions.

The importance of this observation alone set Gauss' place in the history of the great geometers by observing the example $p=17$. We may factor $17-1 = 2\cdot 2\cdot 2\cdot 2$. Thus $(2,1)$ is the root of a quadratic equation over $\mathbb{Q}$. $(4,1)$ is the root of a quadratic equation over $\mathbb{Q}((2,1))$. $(8,1)$ is the root of a quadratic equation over $\mathbb{Q}((4,1))$ and $(16,1) = \zeta$ is the root of a quadratic equation over $\mathbb{Q}((8,1))$. Thus it is possible to write each 17th root of unity by successively solving four quadratic equations. This can be shown to be equivalent to the statement that one may construct the regular 17-gon using ruler and compass.

For a short explanation of these results that utilizes Galois theory consult the aging standard text [30]

Gauss goes on further to show that each n^{th} root of unity is in fact resolvable into radicals. The importance to the theorem of Stickelberger is that Gauss was manipulating

with the periods of the cyclotomic field. As we will see he goes on to use them in what is one of the most important proofs of the law of quadratic reciprocity.

For a historical treatment of the progress leading into Galois theory and Algebraic equations see the wonderful text [28]

1.2. **Proof of Quadratic Reciprocity.** It would be easy to get distracted with the various and plentiful laws of reciprocity discovered in the 18th and 19th centuries. For a very comprehensive book on reciprocity laws between Euler and Eisenstein consult [23]. We shall instead maintain historical focus on the few proofs that seem to be the motivation to study the prime decomposition of certain Gauss sums.

1.2.1. *(Leonhard Euler & Adrien-Marie Legendre).* The history of quadratic reciprocity may be traced back to the work of Euler[1]. He states his four theorems equivalent to the law of quadratic reciprocity (posthumously) in 1783 in [8]. Lagrange in 1788 then states the theorem in a more familiar form in eight theorems, of which he was able to prove four. To state these laws, Let $p, q \equiv 1 \mod 4$ be primes, and let $b, B \equiv 3 \mod 4$. Then the eight theorems are

(1) If $b^{\frac{p-1}{2}} \equiv +1 \mod p$ then $p^{\frac{b-1}{2}} \equiv +1 \mod b$
(2) If $p^{\frac{b-1}{2}} \equiv -1 \mod b$ then $b^{\frac{p-1}{2}} \equiv -1 \mod p$
(3) If $p^{\frac{q-1}{2}} \equiv +1 \mod q$ then $q^{\frac{p-1}{2}} \equiv +1 \mod p$
(4) If $p^{\frac{q-1}{2}} \equiv -1 \mod p$ then $q^{\frac{p-1}{2}} \equiv -1 \mod b$
(5) If $p^{\frac{b-1}{2}} \equiv +1 \mod b$ then $b^{\frac{p-1}{2}} \equiv +1 \mod p$
(6) If $p^{\frac{q-1}{2}} \equiv -1 \mod p$ then $q^{\frac{p-1}{2}} \equiv -1 \mod b$
(7) If $b^{\frac{B-1}{2}} \equiv +1 \mod B$ then $B^{\frac{b-1}{2}} \equiv -1 \mod b$
(8) If $b^{\frac{B-1}{2}} \equiv -1 \mod B$ then $B^{\frac{b-1}{2}} \equiv +1 \mod b$

It was theorems (3)–(6) that Legendre was unable to prove completely, but he had an idea if he could prove the following theorem,

Theorem. *Let a, m be relatively prime positive integers. There exist infinitely many primes in the progression $a, a+m, a+2m \ldots$*

This theorem was not proven until 1837 when Dirichlet defined his L-functions and proved this analytically. According to [23] pp. 19, Ernst Kummer is quoted,

> *"In order to remedy this deficiency in Legendre's proof, Mr. Dirichlet later proved this property of arithmetic progressions rigorously... These celebrated papers of Mr. Dirichlet may therefore also be said to owe their existence to the occupation with reciprocity laws."*

This discussion is far from the topic of factoring of Gauss sums, but it should indicate the surprising connections that research into this general area has illuminated. We will see many more examples of this as we go along.

---

[1] or possibly with Fermat if one includes the supplementary law that $(\frac{-1}{p}) = (-1)^{\frac{p-1}{2}}$.

1.2.2. *(Carl Friedrich Gauss).* Gauss was the first author to publish a correct proof of the quadratic reciprocity law. In fact he gave two proofs in his publication of the *Disquisitiones Arithmeticae* and went on to publish a total of six, while two extra proofs he kept for himself in a private journal. Our main interest is in the fourth proof [11], and the sixth proof [12] as they deal directly with classical Gauss sums.

Gauss' fourth proof is derived by calculating the exact value of what is now called the quadratic Gauss sum. For t with $(t,p)=1$ let $(\frac{t}{p})$ denote the Legendre symbol, which takes the value 1 if t is a square modulo p, and -1 if t is a non-square modulo p. The quadratic Gauss sum is defined as $G=\sum_{t=1}^{p-1}(\frac{t}{p})\zeta^t$, where ζ is a fixed p^{th} root of unity. We may see the connection to the periods of the cyclotomic equations by setting $f=\frac{p-1}{2}$ and $\lambda=1$ in (2) to get $2(f,\lambda)=\sum_t \zeta^{t^2}=\sum_t(1+(\frac{t}{p}))\zeta^t=G$. It is determined explicitly in [11] that

$$G=\begin{cases}\sqrt{p} & \text{if } p\equiv 1 \mod 4\\ i\sqrt{p} & \text{if } p\equiv 3 \mod 4\end{cases} \tag{3}$$

Gauss derives from (3) the quadratic reciprocity law by some rather serious considerations. For a condensed account of this and the next proof of quadratic reciprocity consult the text [24], Chapter 6. The connection with Stickelberger's theorem is already clear, the only prime ideal divisors of the Gauss sum must also divide p. Thus we have started the road to discovering the complete decomposition of these sums.

To illustrate the way in which Gauss was thinking about the problem we will present elements of his paper [12]. In particular it should be noted that in this paper, unlike the previous one he did not work explicitly with the transcendental formula for the p^{th} roots of unity $\zeta=\cos(\frac{2\pi i}{p})+i\sin(\frac{2\pi i}{p})$. Thus Gauss' idea of working purely algebraically means without the use of irrationality as would be introduced by the transcendental formula. We will see what this means in the proof of the following,

Theorem (Gauss)**.** *For p,q distinct odd primes one has $(\frac{p}{q})(\frac{q}{p})=(-1)^{\frac{p-1}{2}\frac{q-1}{2}}$*

Proof. Gauss' sixth proof proceeds by fixing an odd prime p and a primitive root α of p, then defining the polynomial function

$$f(x)=x+x^{\alpha}+x^{\alpha^2}+\cdots+x^{\alpha^{p-2}}+1,$$

where x from here on out will be an indeterminate. Gauss desires to treat the unknown x as a p^{th} root of unity, so he manipulates with the function $f(x)$ modulo $\Phi_p(x)=1+x+x^2+\cdots+x^{p-1}$ (actually he never worked in the quotient rings of polynomial rings as no sufficient theory had been developed, thus all his statements assert that $\Phi_p(x)$ divides another polynomial, whcih to us is equivalent to this modular arithmetic), thus when he proves that $f(x^m)$ is divisible by this polynomial if $(p,m)=1$, he has shown that $f(\zeta)$ is zero whenever ζ is a primitive p^{th} root of unity. Gauss then sets the value

$$\xi=\xi(x)=x-x^{\alpha}+x^{\alpha^2}-x^{\alpha^3}+x^{\alpha^4}-\cdots-x^{\alpha^{p-2}} \tag{4}$$

The proof of quadratic reciprocity continues by summing the following expression horizontally and vertically,

$$
\begin{array}{lllll}
+x\xi & -x^2 & +x^{\alpha+1} & -x^{\alpha^2+1} & +\ldots+x^{\alpha^{p-2}+1} \\
-x^{\alpha}\xi & -x^{2\alpha} & +x^{\alpha^2+\alpha} & -x^{\alpha^3+\alpha} & +\ldots+x^{\alpha^{p-1}+\alpha} \\
+x^{\alpha^2}\xi & -x^{2\alpha^2} & +x^{\alpha^3+\alpha^2} & -x^{\alpha^4+\alpha^2} & +\ldots+x^{\alpha^{p}+\alpha^2} \\
-x^{\alpha^3}\xi & -x^{2\alpha^3} & +x^{\alpha^4+\alpha^3} & -x^{\alpha^5+\alpha^2} & +\ldots+x^{\alpha^{p+1}+\alpha^3} \\
\ldots & & & & \\
-x^{\alpha^{p-2}}\xi & -x^{2\alpha^{p-2}} & +x^{\alpha^{p-1}+\alpha^{p-2}} & -x^{\alpha^{p}+\alpha^{p-2}} & +\ldots+x^{\alpha^{2p-4}+\alpha^{p-2}}
\end{array} \tag{5}
$$

The entire expression is zero as each row in this sum is zero which is clear from (4). Summing the columns first yields

$$= \xi^2 - (f(x^2)-1) + (f(x^{\alpha+1})-1) - (f(x^{\alpha^2+1})-1) + (f(x^{\alpha^3+1})-1) - \cdots + (f(x^{\alpha^{p-2}+1})-1)$$
$$= \xi^2 - f(x^2) + f(x^{\alpha+1}) - f(x^{\alpha^2+1}) + f(x^{\alpha^3+1}) - \cdots + f(x^{\alpha^{p-2}+1})$$

Now as $1, \alpha, \alpha^2 \ldots \alpha^{p-2}$ runs over the set $\{1, 2, 3 \ldots p-1\}$ modulo p we see that $2, \alpha+1, \alpha^2+1 \ldots \alpha^{p-2}+1$ runs over the set $\{2, 3 \ldots p-1, 0\}$ modulo p. Thus all but one of the terms $f(x^{\alpha^k+1})$ are divisible by $\Phi_p(x)$ while the remaining term will be congruent to $\pm f(x^{pm}) \equiv \pm f(1) \equiv \pm p \mod \Phi_p(x)$. To determine the sign of this term we simply see that $\alpha^{\frac{p-1}{2}} + 1 \equiv 0 \mod p$ so that the sign is determined, and (5) becomes congruent to

$$\xi^2 - (-1)^{\frac{p-1}{2}} f(x^{\alpha^{\frac{p-1}{2}}+1}) \mod \Phi_p(x)$$

We will recognize ξ as the quadratic gauss sum upon setting $x = \zeta$, a primitive p^{th} root of unity, so to translate what Gauss has proven, if $G = \sum_{t=1}^{p-1} (\frac{t}{p})\zeta^t$ then

$$G^2 = (-1)^{\frac{p-1}{2}} p \tag{6}$$

The proof of the reciprocity law almost falls out of this theorem alone and will be easier to carry out in the latter notation, though Gauss certainly carried it out with divisibility properties of the polynomials we are working with. Consider the relation

$$
\begin{aligned}
G^q &= \left(\sum_{t=1}^{p-1} \left(\frac{t}{p}\right) \zeta^t\right)^q \\
&= \sum_{t=1}^{p-1} \left(\frac{t}{p}\right)^q \zeta^{tq} + q(A_0 + A_1\zeta + \cdots + A_{p-2}\zeta^{p-2}) \\
&= \left(\frac{q}{p}\right) G + q(A_0 + A_1\zeta + \cdots + A_{p-2}\zeta^{p-2})
\end{aligned} \tag{7}
$$

for some Integers $A_0, A_1, \ldots, A_{p-2}$. Using (6) we see the difference

$$
\begin{aligned}
G^q - \left(\frac{q}{p}\right) G &= \left((G^2)^{\frac{q-1}{2}} - \left(\frac{q}{p}\right)\right) G \\
&= \left((-1)^{\frac{p-1}{2}\frac{q-1}{2}} p^{\frac{q-1}{2}} - \left(\frac{q}{p}\right)\right) G
\end{aligned}
$$

is divisible by q in the integer ring $\mathbb{Z}[\zeta]$. The tricky part of this proof is to show that this implies $(-1)^{\frac{p-1}{2}\frac{q-1}{2}}p^{\frac{q-1}{2}} \equiv (\frac{q}{p}) \mod q$. Euler's theorem yields $p^{\frac{q-1}{2}} \equiv (\frac{p}{q}) \mod q$, and so because q is odd the result would follow,

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

To prove that in fact q divides $(-1)^{\frac{p-1}{2}\frac{q-1}{2}}p^{\frac{q-1}{2}} - (\frac{q}{p})$ Gauss shows that $A_0 + \cdots + A_{p-2}\zeta^{p-2} = GM$ where in fact $M \in \mathbb{Z}$. Dividing by the complex value G gives the result. □

In the introduction to this paper Gauss states that he has finally discovered a method that would generalize to the biquadratic and cubic case of reciprocity, though he never produced a full proof of either. Returning to this goal in, "Theory der biquadratischen Reste I, II" [10] Gauss declares (*cf.* [14] Chapter 9) that,

> *...the previously accepted principles of arithmetic are in no way sufficient for the foundations of a general theory, that rather such a theory necessarily demands that to a certain extent the domain of higher arithmetic needs to be endlessly enlarged...*

Gauss is calling for a way to introduce irrationalities to the integers and study arithmetic there, which will essentially become the theory of algebraic integers. In this paper in which Gauss states the law of biquadratic reciprocity Gauss is led to introduce arithmetic in the ring $\mathbb{Z}[i]$ where $i^2 = -1$, which now carries the name *Gaussian Integers*. Similarly the theory of Eisenstein integers $\mathbb{Z}[\varrho]$ with $\varrho^3 = 1$ was developed by Eisenstein in his proof of the law of cubic reciprocity. Thus the algebraic integers may well *owe their existence to the occupation with reciprocity laws.*

Thankfully from Gauss' sixth proof it is clear that the key to these higher reciprocity laws is in studying generalizations of the quadratic Gauss sums, and that it should not take much work the derive the laws after studying the structure of the sums. Thus the motivation of studying the sums is set in the very attractive problem of the time, reciprocity laws.

2. Jacobi's Congruence and Cubic Reciprocity

2.1. **Jacobi Sums.**

2.1.1. *(Augustin-Louis Cauchy & Carl Jacobi).* Crucial to the investigation of the decomposition of primes in cyclotomic fields was the study of Jacobi sums. These sums, studied simultaneously by Cauchy and Gauss were introduced in Jacobi's work in his letters to Gauss in 1827 [15], one year after his thesis defense, by the following procedure. Let ζ be a primitive p^{th} root of unity, g a primitive root modulo p and let l be a factor of $p-1$. Set $\xi(r)$, where r is any primitive l^{th} root of unity, to be

$$\xi(r) = \zeta + r\zeta^g + r^2\zeta^{g^2} + \cdots + r^{p-2}\zeta^{g^{p-2}} \tag{8}$$

We see $\xi(r)$ is a Gauss sum $G(\chi)$ as defined by [14] where χ is a character on $(\mathbb{Z}/(p))^*$ of order l satisfying $\chi(g) = r$. Jacobi then notes that

$$\frac{\xi(r)\xi(r^m)}{\xi(r^{m+1})} = A_0 + A_1 r + A_2 r^2 + \cdots + A_{l-1} r^{l-1} \tag{9}$$

where the A_i are integers. One has the relation $\xi(r)\xi(r^{-1}) = p$, which already appeared in the Disquisitiones. For $l = 3, m = 1$ these also lead to the following relations which we will see are the first steps to factoring non-quadratic Gauss sums

$$\xi(r)^2 = \frac{\xi(r)\xi(r)}{\xi(r^2)}\xi(r^2) \tag{10}$$

$$\xi(r)^3 = p\frac{\xi(r)\xi(r)}{\xi(r^2)} \tag{11}$$

where $\frac{\xi(r)\xi(r)}{\xi(r^2)}$ is shown to be of the form $\frac{a+b\sqrt{-3}}{2}$ with $a^2 + 3b^2 = 4p$.

Jacobi proceeds in the important work [16] to define what would become known as the Jacobi sum,

$$\psi_{n,m}(r) = \frac{\xi(r^{-n})\xi(r^{-m})}{\xi(r^{-n-m})} \tag{12}$$

When none of r^{-n}, r^{-m}, r^{-n-m} are 1 this is of the same form as (9). Jacobi seldom provided proofs for us to understand how he thought of these identities, which is why we lay down so many relations without immediate proof. It is worth noting that this hesitance led to a bitter dispute between himself, Eisenstein, and Cauchy over who had the right to claim the first proofs of cubic and biquadratic reciprocity. Essentially the results for the cubic case are due to the combined efforts of these three, as well as Legendre and Gauss, who very well might have had proofs of these theorems as early as 1807. An account of the dispute is given in the Notes section of chapter 8 in [23]. That said, proofs of the previous claims following Eisenstein's papers [4] and [5] will be laid out here.

Firstly, because g is a fixed primitive root of p we may define a *log* function for integers not divisible by p. **To avoid confusion we use** $Ind(x)$ **instead of** $log(x)$ to denote the unique integer z with $0 \leq z < p-1$ satisfying $g^z \equiv x \mod p$. We should note that if $r^{p-1} = 1$ then $r^{Ind(kk')} = r^{Ind(k)+Ind(k')}$. Presumably Eisenstein uses Ind to refer the index z of the term $r^z x^{g^z} = r^z x^k$. Carrying on we find

$$\begin{aligned}
\xi(r^{-n})\xi(r^{-m}) &= \left(\sum_{k=1}^{p-1} r^{-nInd(k)}\zeta^k\right)\left(\sum_{k'=1}^{p-1} r^{-mInd(k')}\zeta^{k'}\right) \\
&= \sum_{k'=1}^{p-1}\sum_{k=1}^{p-1} r^{-nInd(k)-mInd(k')}\zeta^{k+k'}
\end{aligned}$$

We now replace k with $k'\sigma \mod p$ for $\sigma \in \mathbb{Z}$ and notice as k ranges over $1, 2, \ldots p-1$ that σ does the same. Thus we are left with,

$$\begin{aligned} &= \sum_{k'=1}^{p-1}\sum_{\sigma=1}^{p-1} r^{-nInd(k')-nInd(\sigma)-mInd(k')}\zeta^{(\sigma+1)k'} \\ &= \sum_{\sigma=1}^{p-1} r^{-nInd(\sigma)}\sum_{k'=1}^{p-1} r^{(-n-m)Ind(k')}\zeta^{(\sigma+1)k'} \end{aligned}$$

When $\sigma = p-1$ the inner sum becomes

$$\sum_{k'=1}^{p-1} r^{(-n-m)Ind(k')} = 0$$

as $r^{-n-m} \neq 1$. For $\sigma \neq p-1$ the inner sum turns into

$$r^{(n+m)Ind(\sigma+1)}\sum_{k'=1}^{p-1} r^{(-n-m)Ind((\sigma+1)k')}\zeta^{(\sigma+1)k'}$$

The summation is simply $\xi(r^{-n-m})$. Thus the entire argument reveals the following,

$$\psi_{n,m}(r) = \frac{\xi(r^{-n})\xi(r^{-m})}{\xi(r^{-n-m})} = \sum_{\sigma=1}^{p-2} r^{-nInd(\sigma)+(n+m)Ind(\sigma+1)} \tag{13}$$

For $l = 3$ Jacobi notes that setting $r = \frac{1+\sqrt{-3}}{2}$ we may decompose any prime $p \equiv 1 \mod 3$ in the ring $\mathbb{Z}[r]$ by the formula $r^i\psi_{n,m}(r)\psi_{n,m}(r^{-1}) = p$ for some $i = 0, 1, 2$. The fact that $\psi_{n,m}(r)$ is in the ring $\mathbb{Z}[r]$ comes from the amazing fact that (13) has rid us of the p^{th} root of unity x!

Jacobi's next big insight was that the expression (13) may be considered as a *function* of r and modulo p one may replace the irrationality r with the primitive root g, which satisfies the same equation $g^l \equiv 1 \mod p$ when $l = p-1$ to obtain

$$\begin{aligned} \psi_{n,m}(g) &= \sum_{\sigma=1}^{p-2} g^{-nInd(\sigma)+(n+m)Ind(\sigma+1)} \\ &= \sum_{\sigma=1}^{p-2} \sigma^{-n}(\sigma+1)^{n+m} \\ &= \sum_{\sigma=1}^{p-2}\sum_{k=0}^{n+m} \binom{n+m}{k}\sigma^{m-k} \\ &= \sum_{k=0}^{n+m} \binom{n+m}{k}\sum_{\sigma=1}^{p-2}\sigma^{m-k} \end{aligned}$$

We may evaluate the inner sum in two cases. When $p-1$ does not divide $m-k$ then $\sum_{\sigma=1}^{p-1}\sigma^{m-k}\equiv 0 \mod p$ thus we are left with $\sum_{\sigma=1}^{p-2}\sigma^{m-k}\equiv -(p-1)^{m-k}\equiv -(-1)^{m-k} \mod p$. If $p-1$ does divide $m-k$ then each term $\sigma^{m-k}\equiv 1 \mod p$ so we are left with $p-2$ such terms and $\sum_{\sigma=1}^{p-2}\sigma^{m-k}\equiv -2 \mod p$. Thus the sum becomes

$$-\sum_{\substack{k=0\\k\neq m}}^{n+m}(-1)^{m-k}\binom{n+m}{k}+(-2)\binom{m+n}{m}$$

Of course $\sum_{k=0}^{n+m}(-1)^{m-k}\binom{n+m}{k}=0$ so that the left sum becomes the $k=m$ term, $(-1)^{m-m}\binom{m+n}{m}$ leaving the following incredible congruence between integers

$$\psi_{n,m}(g)\equiv -\frac{(n+m)!}{n!m!} \mod p \tag{14}$$

This congruence is the key to the way that Jacobi and Cauchy were able to factor Gauss sums without having ideals to work with. We will show how this leads to the factorization of a cubic Gaussian sum by looking at the Eisenstein integers.

2.1.2. *(Gotthold Eisenstein).* In [5] Eisenstein defines the complex integers of the form $a+b\varrho$ with a,b rational integers and $\varrho=e^{\frac{2\pi i}{3}}$ a primitive cube root of unity. He discusses the now familiar properties of these integers, divisibility, congruences, the norm map and in particular that one only needs to modify the proofs that Gauss gave that $\mathbb{Z}[i]$ is a Euclidean domain to obtain the same result for $\mathbb{Z}[\varrho]$. For these efforts this ring now bears his name, the Eisenstein integers. He shows that for a rational prime $p\equiv 2 \mod 3$ that p may not be decomposed as the product of two Eisenstein integers $(a+b\varrho)(a+b\varrho^2)$ by examining the relation $N(a+b\varrho)=a^2-ab+b^2\equiv 2 \mod 3$. So these integers are still primes in $\mathbb{Z}[\varrho]$. For $p\equiv 1 \mod 3$ we may fix a primitive root g of p and define the Jacobi sum

$$\psi_{2,2}(\varrho)=\frac{\xi(\varrho)\xi(\varrho)}{\xi(\varrho^2)}=\sum_{\sigma=1}^{p-2}\varrho^{Ind(\sigma)-2Ind(\sigma+1)}$$

We see that $\psi_{2,2}(\varrho)\psi_{2,2}(\varrho^2)=\xi(\varrho)\xi(\varrho^2)=p$ and so we have provided $\pi_1=a+b\varrho=\psi_{2,2}(\varrho)$ and $\pi_2=a+b\varrho^2=\psi_{2,2}(\varrho)$ with $\pi_1\pi_2=a^2-ab+b^2=p$. This analysis already appeared in [15] in 1827 and is a wonderful result on the representation of primes by quadratic forms. It follows that π_1 and π_2 are primes in $\mathbb{Z}[\varrho]$, which a priori may not be distinct, though we will show that this is the case. Now that we know what the primes in this ring are we should get right to the problem of investigating the prime decomposition of Gaussian sums in this ring. As we will see, for reciprocity there is no need for such sums over the finite field obtained by taking $\mathbb{Z}[\varrho]$ modulo a prime $p\equiv 2$, i.e. for fields with p^2 elements, and in fact these kinds of sums were not considered until much later. Thus we start with the problem of defining the cubic residue character χ_π on the field $\mathbb{Z}[\varrho]$ modulo a given prime π where $N(\pi)=\pi\overline{\pi}=p\neq 3$.

It was shown that $\mathbb{Z}[\varrho]$ modulo the prime π is a finite field with p elements, and so we have the relation for any Eisenstein integer α not divisible by π,

$$(\alpha^{\frac{p-1}{3}})^3 \equiv 1 \mod \pi.$$

This of course implies that π divides $(\alpha^{\frac{p-1}{3}} - 1)(\alpha^{\frac{p-1}{3}} - \varrho)(\alpha^{\frac{p-1}{3}} - \varrho^2)$. Thus $\alpha^{\frac{p-1}{3}}$ is congruent to 1, ϱ or ϱ^2 modulo π. The cubic residue symbol is defined then as

$$\left[\frac{\alpha}{\pi}\right] = 1, \varrho, \text{ or } \varrho^2$$

subject to the condition

$$\left[\frac{\alpha}{\pi}\right] \equiv \alpha^{\frac{p-1}{3}} \mod \pi.$$

In particular we see for rational integer k

$$k^{\frac{p-1}{3}} \equiv g^{\frac{p-1}{3}Ind(k)} \equiv \left[\frac{g}{\pi}\right]^{Ind(k)} \mod \pi$$

and so $[\frac{k}{\pi}] = [\frac{g}{\pi}]^{Ind(k)}$. Thus if we define the cubic residue character $\chi_\pi(t) = [\frac{t}{\pi}]$, then the Gaussian sum that appears in Stickelberger's Theorem is,

$$\begin{aligned} G((\pi)) &= \sum_{t=1}^{p-1} \chi_\pi(t)^{-1}\zeta^t \\ &= \sum_{t=1}^{p-1} \left[\frac{g}{\pi}\right]^{-Ind(t)} \zeta^t \\ &= \xi\left(\left[\frac{g}{\pi}\right]^{-1}\right) \end{aligned}$$

To determine the value of $[\frac{g}{\pi}]$ we must do some work. Recall that we may set $\pi_1 = \psi_{2,2}(\varrho)$ and $\pi_2 = \psi_{2,2}(\varrho^2)$ to get the prime factorization $p = \pi_1\pi_2$. By the theorem of unique factorization π must be associate to one of these primes, that is equal up to a multiplicative factor of $\pm\varrho^m$ for some $m = 0, 1, 2$. Continuing, we assume that $[\frac{g}{\pi}] = \varrho$. It follows

$$\psi_{2,2}(\varrho) \equiv \psi_{2,2}(g^{\frac{p-1}{3}}) \mod \pi$$

By definition $\psi_{n,m}(g^{\frac{p-1}{3}})$ is equal to $\psi_{n,m}(g)$ with m, n multiplied by a factor of $\frac{p-1}{3}$. By (14) we see that if $[\frac{g}{\pi}] = \varrho$ then

$$\pi_1 = \psi_{2,2}(\varrho) \equiv \psi_{2,2}(g^{\frac{p-1}{3}}) = \psi_{2\frac{p-1}{3},2\frac{p-1}{3}}(g) \equiv -\frac{(4\frac{p-1}{3})!}{(2\frac{p-1}{3})!(2\frac{p-1}{3})!} \mod \pi$$

We see that $4\frac{p-1}{3} > p$, but $2\frac{p-1}{3} < p$ and so p divides the term on the right. From this it follows that π divides π_1 and they are associate primes. A similar analysis shows that if $[\frac{g}{\pi}] = \varrho^2$ then π and π_2 are associate. It is interesting to note that one may similarly

use (14) to show that $\pi_1 \not\equiv 0 \mod \pi_2$, and as we mentioned before these primes are not associate. We have succeeded in factoring the Jacobi sum into its prime factors,

$$\psi_{2,2}([\frac{g}{\pi}]) = \frac{\xi(\varrho')\xi(\varrho')}{\xi(\varrho'^2)} = \varepsilon\pi$$

for some unit $\varepsilon = \pm\varrho^m$, $m = 0, 1, 2$.

It seems, perhaps for the first time, that we are now well on our way to Stickelberger's theorem! To finish analyzing the Gaussian sums that occur in Stickelberger's theorem we only require a few more details. Firstly, our goal is to analyze the decomposition of $G(\chi)^3 = \xi([\frac{g}{\pi}])^3$, and so we must prove that this is indeed an element of the ring $\mathbb{Z}[\varrho]$. Of course if we set $[\frac{g}{\pi}] = r$ then we see by the equations (10) and (11) the factorization

$$\xi\left(\left[\frac{g}{\pi}\right]\right)^3 = p\psi_{2,2}\left(\left[\frac{g}{\pi}\right]\right) = \varepsilon p\pi \tag{15}$$

and our analysis for these types of primes is complete! Historically it wasn't until much later that Gaussian sums over the finite residue fields with p^2 elements were defined. Thus we won't cover the Gaussian sums for these cases in this section. Besides this point the ideal class group of the Eisenstein integers is trivial and so there is nothing enlightening to say about its annihilators. It would, however, be enlightening to see how quickly the law of cubic reciprocity follows from our discussion.

2.2. **Proof of Cubic Reciprocity.** The first comment to make is that we must define the cubic residue symbol for a rational prime $p \equiv 2 \mod 3$. Because the multiplicative group of $\mathbb{Z}[\varrho]$ modulo the prime p is cyclic with $p^2 - 1$ elements and so the correct cubic character should be defined by the following

$$\left[\frac{\alpha}{p}\right] = 1, \varrho, or\, \varrho^2$$

subject to the condition

$$\left[\frac{\alpha}{p}\right] \equiv \alpha^{\frac{p^2-1}{3}} \mod p.$$

Before the statement and proof of the law of cubic reciprocity we need to discuss two last points. The first is that to gain a full understanding of the reciprocity law we must extend the definition of the the cubic residue symbol $[\frac{\alpha}{\pi}]$ to composite modulus π, and α not necessarily relatively prime to π. Unfortunately this point is only a distraction for us as our main goal is to study Stickelberger's theorem and so we opt not to cover these supplementary cases of the laws of cubic reciprocity.

The second point is that we need a notion of primary primes to determine exactly what the unit ε is in (15). It turns out for this cubic case that the only definition of primary prime needed is the following.

Definition. *A prime Eisenstein integer π is called* **primary** *if it satisfies,*

$$\pi \equiv 2 \mod 3$$

Along with this definition we find the following lemma which shows that this definition is not too restrictive.

Lemma. *Among the six associates of a given prime π which is not divisible by $1-\varrho$ exactly one is primary. Furthermore if π is primary and has norm p for some prime p with primitive root g then we have*

$$\xi\left(\left[\frac{g}{\pi}\right]\right)^3 = p\pi$$

Proof. The first claim takes nothing more than calculating that the multiplicative group modulo 3 has 6 elements which are given by the distinct classes of $1, \varrho, \varrho^2, -1, -\varrho, -\varrho^2$, only one of which is congruent to 2 modulo 3.

The second claim will follow if we can show that the cube of the Gauss sum given in (15) is primary. Similar to (7) we find that upon setting $[\frac{g}{\pi}] = \varrho'$

$$\begin{aligned}
\xi(\varrho')^3 &= \left(\sum_{k=1}^{p-1} \varrho'^{Ind(k)}\zeta^k\right)^3 \\
&= \sum_{k=1}^{p-1} \varrho'^{3Ind(k)}\zeta^{3k} + 3(A_0 + A_1\zeta + \cdots + A_{p-2}\zeta^{p-2}) \\
&= \sum_{k=1}^{p-1} \zeta^k + 3(A_0 + A_1\zeta + \cdots + A_{p-2}\zeta^{p-2}) \\
&= -1 + 3(A_0 + A_1\zeta + \cdots + A_{p-2}\zeta^{p-2})
\end{aligned}$$

for some integers $A_0, \ldots, A_{p-2}$. As before because both $\xi(\varrho')^3$ and -1 are Eisenstein integers we must have that the expression $A_0 + A_1\zeta + \cdots + A_{p-2}\zeta^{p-2}$ is also an Eisenstein integer. The lemma follows. □

We now state and prove the law of cubic reciprocity. One can see that the proof has cases and so might take some space, but every case here is handled exactly as in Gauss's sixth proof.

Theorem (Eisenstein)**.** *Let π_1 and π_2 be distinct primary primes. Then the following relationship exists*

$$\left[\frac{\pi_1}{\pi_2}\right] = \left[\frac{\pi_2}{\pi_1}\right]$$

Proof. Firstly let us consider the case where $\pi_1 = p$ and $\pi_2 = q$ are rational primes, which thus must be congruent to 2 modulo 3. By Fermat's little theorem we have the congruence $p^{q-1} \equiv 1 \mod q$. Thus because $q^2 - 1 = (q+1)(q-1)$ we have $p^{\frac{q^2-1}{3}} = (p^{q-1})^{\frac{q+1}{3}} \equiv 1 \mod q$. It is a similar story for the reverse situation, so we find

$$\left[\frac{p}{q}\right] = \left[\frac{q}{p}\right] = 1$$

Now suppose $p \equiv 1 \mod 3$ and $q \equiv 2 \mod 3$ are rational prime numbers where π is an Eisenstein prime with $N(\pi) = p$. As in (7) we analyze

$$\begin{aligned}
\xi\left(\left[\frac{g}{\pi}\right]\right)^{q^2} &= \left(\sum_{k=1}^{p-1}\left[\frac{k}{\pi}\right]^{q^2}\zeta^{kq^2}\right) + q(A_0 + \cdots + A_{p-2}\zeta^{p-2}) \\
&= \left[\frac{q^2}{\pi}\right]^{-1}\xi\left(\left[\frac{g}{\pi}\right]\right) + q(A_0 + \cdots + A_{p-2}\zeta^{p-2}) \\
&= \left[\frac{q}{\pi}\right]\xi\left(\left[\frac{g}{\pi}\right]\right) + q(A_0 + \cdots + A_{p-2}\zeta^{p-2})
\end{aligned}$$

This allows us to say that the following congruence is valid,

$$\xi\left(\left[\frac{g}{\pi}\right]\right)^{q^2-1} - \left[\frac{q}{\pi}\right] = (p\pi)^{\frac{q^2-1}{3}} - \left[\frac{q}{\pi}\right] \equiv 0 \mod q$$

From this congruence follows

$$\left[\frac{p\pi}{q}\right] = \left[\frac{p}{q}\right]\left[\frac{\pi}{q}\right] = \left[\frac{q}{\pi}\right]$$

it is clear however that $[\frac{p}{q}] = 1$ and so we are done with this case.

Lastly consider the case of two primary primes π_1 and π_2 with norms $p \equiv q \equiv 1 \mod 3$ for which g_1 and g_2 are primitive roots. Again as in (7) we see

$$\begin{aligned}
\xi\left(\left[\frac{g_1}{\pi_1}\right]\right)^{q} &= \left(\sum_{k=1}^{p-1}\left[\frac{k}{\pi_1}\right]^{q}\zeta^{kq}\right) + q(A_0 + \cdots + A_{p-2}\zeta^{p-2}) \\
&= \left[\frac{q}{\pi_1}\right]^{-1}\xi\left(\left[\frac{g}{\pi_1}\right]\right) + q(A_0 + \cdots + A_{p-2}\zeta^{p-2})
\end{aligned}$$

So similarly we conclude the following results

$$\left[\frac{p\pi_1}{\pi_2}\right] = \left[\frac{q}{\pi_1}\right]^{-1}$$

which becomes

$$\left[\frac{\pi_1}{\pi_2}\right]\left[\frac{\overline{\pi}_1}{\pi_2}\right]^2 = \left[\frac{\pi_2}{\pi_1}\right]\left[\frac{\overline{\pi}_2}{\pi_1}\right] \tag{16}$$

A similar analysis of $\xi([\frac{g_2}{\pi_2}])^p$ reveals the analogous equation

$$\left[\frac{\pi_2}{\pi_1}\right]\left[\frac{\overline{\pi}_2}{\pi_1}\right]^2 = \left[\frac{\pi_1}{\pi_2}\right]\left[\frac{\overline{\pi}_1}{\pi_2}\right] \tag{17}$$

Multiplying the left side of (16) with the right side of (17) and vice versa we see that in fact

$$\left[\frac{\pi_1}{\pi_2}\right] = \left[\frac{\pi_2}{\pi_1}\right]$$

These cover all the cases and we are done. □

3. Kummer's Unique Factorization and Eisenstein Reciprocity

As we will see Kummer's influence on reciprocity laws goes far beyond his contribution of the ideal numbers. In fact in the paper in which Kummer discovers the properties of ideal numbers he finds the prime ideal factorization of a Gauss sum for a prime p which splits completely as a product of prime numbers in $\mathbb{Z}[\zeta]$. It appears as though Kummer was determining the law of unique factorization in order to actually find reciprocity laws. We shall thus postpone defining an ideal number until we have seen proper motivation to do so, and see again that the reciprocity laws generated some of the most important math of the 19th century.

3.1. **Ideal Numbers.**

3.1.1. *(Ernst Kummer).* We first give a short explanation of Kummer's notation in [17], which was typical of his writing. Kummer was always working in the ring of integers $\mathbb{Z}[\alpha]$ where α is a primitive root of the equation $\alpha^\lambda - 1 = 0$ and λ is a prime. He writes a general element of the ring $\mathbb{Z}[\alpha]$ as a function of α

$$f(\alpha) = a + a_1\alpha + a_2\alpha^2 + \cdots + a_{\lambda-1}\alpha^{\lambda-1}$$

where the coefficients a_i are rational integers. In the year 1844 Galois' influence was yet to be felt, and this notation provided Kummer with techniques to deal with conjugates without such a theory by writing $f(\alpha^k)$ for the conjugate of $f(\alpha)$ which takes α to α^k. We will see that this notation has another very useful feature in a moment. He also writes the norm

$$Nf(\alpha) = f(\alpha)f(\alpha^2)f(\alpha^3)\dots f(\alpha^{\lambda-1})$$

and records the relation $N[f(\alpha)g(\alpha)] = Nf(\alpha)\cdot Ng(\alpha)$. It is then shown by comparing coefficients that if

$$\begin{aligned} f(\alpha) &= a + a_1\alpha + a_2\alpha^2 + \cdots + a_{\lambda-1}\alpha^{\lambda-1} \\ \varphi(\alpha) &= b + b_1\alpha + b_2\alpha^2 + \cdots + b_{\lambda-1}\alpha^{\lambda-1} \\ \psi(\alpha) &= c + c_1\alpha + c_2\alpha^2 + \cdots + c_{\lambda-1}\alpha^{\lambda-1} \end{aligned}$$

and $f(\alpha)\cdot\varphi(\alpha) = \psi(\alpha)$ then

$$\begin{aligned} (a + a_1 + a_2 + \cdots + a_{\lambda-1})(b + b_1 + b_2 + \cdots + b_{\lambda-1}) \\ \equiv c + c_1 + c_2 + \cdots + c_{\lambda-1} \mod \lambda \end{aligned}$$

This is used to show that for any $f(\alpha)$ we must have

$$Nf(\alpha) \equiv (a + a_1 + a_2 + \cdots + a_{\lambda-1})^{\lambda-1} \equiv 1 \mod \lambda$$

From this it immediately follows that if p is a rational prime distinct from λ and is the norm of an element $f(\alpha)$ then $p \equiv 1 \mod \lambda$. It is also easily seen that $f(\alpha)$ must be a prime as if $f(\alpha) = \varphi(\alpha)\cdot\psi(\alpha)$ then $p = Nf(\alpha) = N\varphi(\alpha)\cdot N\psi(\alpha)$ and so one of $\varphi(\alpha)$ or $\psi(\alpha)$ must be a unit.

The next few sections of [17] are devoted to proving a unique factorization theorem for primes p which decompose as the norm of an element. The theorem states that if $p = Nf(\alpha) = N\psi(\alpha)$ then for some unique k modulo λ, $\psi(\alpha^k) \equiv 0 \mod f(\alpha)$. The technique of the proof turns out to be substantially more important than the actual statement as it seems that this is what led Kummer to define his ideal numbers, and if a student has not seen these techniques it is a fascinatingly original, yet elementary exploration in number theory.

In order to work modulo $f(\alpha)$ it is important to find a complete set of residues modulo $f(\alpha)$. To do this Kummer amazingly constructs a rational integer $0 < \xi \leq p-1$ which satisfies $\xi^\lambda \equiv 1 \mod p$ for which $\alpha \equiv \xi \mod f(\alpha)$, and thus has constructed a homomorphism $\mathbb{Z}[\alpha] \to \mathbb{Z}/(p) \cong \mathbb{Z}[\alpha]/(f(\alpha))$ given by sending $g(\alpha) \to g(\xi)$. To accomplish this takes a little work. First write

$$F(\alpha) = f(\alpha^2)f(\alpha^3)\dots f(\alpha^{\lambda-1})$$

from which one immediately sees the relation $p = f(\alpha)F(\alpha)$. Now write

$$\begin{aligned}
F(\alpha) &= A + A_1\alpha + A_2\alpha^2 + \cdots + A_{\lambda-1}\alpha^{\lambda-1} \\
F(\alpha^2) &= A + A_1\alpha^2 + A_2\alpha^4 + \cdots + A_{\lambda-1}\alpha^{2\lambda-2} \\
F(\alpha^3) &= A + A_1\alpha^3 + A_2\alpha^6 + \cdots + A_{\lambda-1}\alpha^{3\lambda-3} \\
&\cdots \\
F(\alpha^{\lambda-1}) &= A + A_1\alpha^{\lambda-1} + A_2\alpha^{2\lambda-2} + \cdots + A_{\lambda-1}\alpha^{(\lambda-1)(\lambda-1)}
\end{aligned}$$

Next see that

$$\begin{aligned}
&\alpha^{-n}F(\alpha) + \alpha^{-2n}F(\alpha^2) + \cdots + \alpha^{-(\lambda-1)n}F(\alpha^{\lambda-1}) \\
&\quad = \lambda A_n - (A + A_1 + A_2 + \cdots + A_{\lambda-1})
\end{aligned}$$

to get the relation

$$\begin{aligned}
&\alpha^{-n}(1-\alpha)F(\alpha) + \alpha^{-2n}(1-\alpha^2)F(\alpha^2) + \cdots + \alpha^{-(\lambda-1)n}(1-\alpha^{\lambda-1})F(\alpha^{\lambda-1}) \\
&\quad = \lambda(A_n - A_{n-1})
\end{aligned}$$

With incredible insight Kummer replaces n with $n+1$ and squares the formula to obtain the relation

$$\lambda^2(A_{n+1} - A_n)^2 = \lambda^2(A_{n+2} - A_{n+1})(A_n - A_{n-1})$$

Fixing an integer ξ for which $A_{n+1} - A_n \equiv \xi(A_{n+2} - A_{n+1}) \mod p$ we observe the congruence

$$\xi \equiv \frac{A_{n+1} - A_n}{A_{n+2} - A_{n+1}} \equiv \frac{A_n - A_{n-1}}{A_{n+1} - A_n} \mod p$$

and see that the choice of ξ is independent of n. Some care is of course taken to show that each $A_{n+1} - A_n \not\equiv 0 \mod p$. We also observe the relations

$$\begin{aligned} A_{\lambda-1} - A &\equiv (A - A_1)\xi \\ A_{\lambda-2} - A_{\lambda-1} &\equiv (A - A_1)\xi^2 \\ A_{\lambda-3} - A_{\lambda-2} &\equiv (A - A_1)\xi^3 \\ &\cdots \\ A_1 - A_2 &\equiv (A - A_1)\xi^{\lambda-1} \end{aligned}$$

which yield

$$1 + \xi + \xi^2 + \cdots + \xi^{\lambda-1} \equiv 0 \mod p$$

and thus ξ is a primitive root of the equation $x^\lambda - 1 \equiv 0 \mod p$. More importantly ξ satisfies

$$(\xi - \alpha)F(\alpha) \equiv 0 \mod p.$$

Dividing this by $F(\alpha)$ finally yields the desired conclusion,

$$\xi \equiv \alpha \mod f(\alpha).$$

Now the real beauty of the work begins. If $f(\alpha)$ divides $f(\alpha^k)$ for some $k \not\equiv 1 \mod \lambda$ then $f(\alpha)^2$ will divide p. Because $\alpha^p = 1$ we have

$$f(\alpha)^p \equiv f(\alpha) \mod p$$

and thus $f(\alpha)^2$ would divide $f(\alpha)$, which is a contradiction. Thus if $k \not\equiv 1 \mod \lambda$ then $f(\alpha^k) \equiv f(\xi^k) \not\equiv 0 \mod f(\alpha)$. Because $f(\xi^k)$ is an integer then it is divisible by $f(\alpha)$ if and only if it is divisible by p, and so it follows that ξ is the unique λ^{th} root of unity modulo p for which

$$f(\xi) \equiv 0 \mod p.$$

It also follows that **an element $\varphi(\alpha)$ is divisible by $f(\alpha)$ if and only if $\varphi(\xi) \equiv 0 \mod p$.** The advantage of this statement is that one does not need the actual element '$f(\alpha)$' to talk about the prime divisor of $\varphi(\alpha)$ associated to ξ.

In modern language Kummer has shown that there are $\lambda - 1$ homomorphisms of $\mathbb{Z}[\alpha] \to \mathbb{Z}/(p)$ given by by sending $\alpha \to \xi$ for the different primitive λ^{th} roots of unity ξ modulo p, and that $\varphi(\alpha)$ is divisible by an ideal corresponding to ξ if it is in the kernel of corresponding homomorphism $\alpha \to \xi \mod p$. Of course the kernels of these homomorphisms are just the prime ideals above p.

Given these observations the unique factorization theorem mentioned earlier follows easily.

Theorem. *Suppose p is a rational prime $p \equiv 1 \mod \lambda$ and that there are elements $f(\alpha)$ and $\psi(\alpha)$ for which $p = Nf(\alpha) = N\psi(\alpha)$. Then for some $k \not\equiv 0 \mod \lambda$, $f(\alpha)$ divides $\psi(\alpha^k)$.*

Proof. By the previous observations there must be some integer ξ' which is a primitive λ^{th} root of 1 modulo p for which $\psi(\xi') \equiv 0 \mod p$. The group of λ^{th} roots of unity modulo p form a cyclic group and so there is a $k \not\equiv 0 \mod \lambda$ for which $\xi' \equiv \xi^k \mod p$, from which it follows that $\psi(\xi^k) \equiv 0 \mod p$ and so $f(\alpha)$ divides $\psi(\alpha^k)$. □

The final chapter in [17] is devoted to determining the prime factorization of the Gaussian sum formed by letting ζ be a primitive p^{th} root of unity and setting,

$$(\alpha, \zeta) = \zeta + \alpha\zeta^g + \alpha^2\zeta^{g^2} + \cdots + \alpha^{p-2}\zeta^{g^{p-2}}$$

and

$$\psi(\alpha) = \frac{(\alpha^n, \zeta)(\alpha^m, \zeta)}{(\alpha^{n+m}, \zeta)}$$

similar to (8) and (12). Now we have the well established relations

$$\psi(\alpha)\psi(\alpha^{-1}) = p$$

and define the Jacobi sums $\psi_i(\alpha)$ by the following

$$\begin{aligned}
(\alpha, \zeta)(\alpha, \zeta) &= \psi_1(\alpha)(\alpha^2, \zeta) \\
(\alpha, \zeta)(\alpha^2, \zeta) &= \psi_2(\alpha)(\alpha^3, \zeta) \\
(\alpha, \zeta)(\alpha^3, \zeta) &= \psi_3(\alpha)(\alpha^4, \zeta) \\
&\cdots \\
(\alpha, \zeta)(\alpha^{\lambda-2}, \zeta) &= \psi_{\lambda-2}(\alpha)(\alpha^{\lambda-1}, \zeta)
\end{aligned}$$

The formula $(\alpha, \zeta)(\alpha^{-1}, \zeta) = p$ yields the important decomposition,

$$(\alpha, \zeta)^\lambda = p\psi_1(\alpha)\psi_2(\alpha)\ldots\psi_{\lambda-1}(\alpha)$$

which demonstrates that $(\alpha, \zeta)^\lambda$ is an element of $\mathbb{Z}[\alpha]$. Furthermore if p decomposes as the norm of an element $p = Nf(\alpha)$ then we have by the unique factorization theorem that

$$(\alpha, \zeta)^\lambda = \epsilon(\alpha)f(\alpha)^{m_1}f(\alpha^2)^{m_2}\ldots f(\alpha^{\lambda-1})^{m_{\lambda-1}}$$

for some unit $\epsilon(\alpha)$, which the must satisfy $\epsilon(\alpha)\epsilon(\alpha^{-1}) = 1$, and each $m_i \geq 1$. Of course if one writes

$$\epsilon(\alpha) = a + a_1\alpha + a_2\alpha^2 + \cdots + a_{\lambda-1}\alpha^{\lambda-1}$$

then

$$\epsilon(\alpha)\epsilon(\alpha^{-1}) = A + A_1\alpha + A_2\alpha^2 + \cdots + A_{\lambda-1}\alpha^{\lambda-1}$$

where

$$\begin{aligned}
A &= a^2 + a_1^2 + a_2^2 + \cdots + a_{\lambda-1}^2 \\
A_1 &= aa_1 + a_1a_2 + a_2a_3 + \cdots + a_{\lambda-1}a \\
A_2 &= aa_2 + a_1a_3 + a_2a_4 + \cdots + a_{\lambda-1}a_1 \\
&\cdots
\end{aligned}$$

Summing these equations we see then that

$$A + A_1 + A_2 + \cdots + A_{\lambda-1} = (a + a_1 + a_2 + \ldots a_{\lambda-1})^2$$

and

$$A + A_1\alpha + A_2\alpha^2 + \dots A_{\lambda-1}\alpha^{\lambda-1} = 1$$

from which it follows

$$A_1 = A_2 = A_3 = \cdots = A_{\lambda-1} = m$$

and

$$A = m + 1$$

for some integer m. Thus $m\lambda + 1 = (a + a_1 + a_2 + \cdots + a_{\lambda-1})^2$ and so

$$a + a_1 + a_2 + \cdots + a_{\lambda-1} = \pm 1 + k'\lambda$$

Without changing $\epsilon(\alpha)$ we may replace a_i with $a_i + k'$ to get the relation, $A = 1$, from which it follows that

$$1 = a^2 + a_1^2 + \dots a_{\lambda-1}^2$$

and so $\epsilon(\alpha) = \pm\alpha^\chi$ for some $0 \leq \chi < \lambda$. It follows that one really has

$$(\alpha, x)^\lambda = \pm\alpha^\chi f(\alpha)^{m_1} f(\alpha^2)^{m_2} \dots f(\alpha^{\lambda-1})^{m_{\lambda-1}}$$

The relation $(\alpha, x)(\alpha^{-1}, x) = p = f(\alpha)f(\alpha^2)\dots f(\alpha^{\lambda-1})$ implies that

$$m_i + m_{\lambda-i} = \lambda \tag{18}$$

This technique is used in [31]. The relation

$$\psi_r(\alpha) = \frac{(\alpha, \zeta)(\alpha^r, \zeta)}{(\alpha^{r+1}, \zeta)}$$

gives

$$\psi_r(\alpha)^\lambda = \frac{f(\alpha)^{m_1} f(\alpha^2)^{m_2} \dots f(\alpha^{\lambda-1})^{m_{\lambda-1}} \cdot f(\alpha^r)^{m_1} f(\alpha^{2r})^{m_2} \dots f(\alpha^{(\lambda-1)r})^{m_{\lambda-1}}}{f(\alpha^{r+1})^{m_1} f(\alpha^{2(r+1)})^{m_2} \dots f(\alpha^{(\lambda-1)(r+1)})^{m_{\lambda-1}}}$$

which yield the congruences

$$m_k + m_\mu - m_\nu \equiv 0, \text{ when } \mu \equiv \frac{k}{r}, \nu \equiv \frac{k}{r+1} \mod \lambda$$

To simplify these congruences set $n_k \equiv km_k \mod \lambda$ so that for $k = 1$ the above equation becomes

$$n_1 + rn_\mu - (r+1)n_\nu \equiv 0, \text{ when } \mu \equiv \frac{1}{r} \text{ and } \nu \equiv \frac{1}{r+1} \mod \lambda$$

For $r = 1$ this implies that $n_1 \equiv n_{2^{-1}} \mod \lambda$, where k^{-1} denotes a multiplicative inverse of k modulo λ. For $r = 2$ this implies that $n_1 + 2n_{2^{-1}} \equiv 3n_{3^{-1}}$, which simplifies to $n_1 \equiv n_{3^{-1}} \mod \lambda$. Continuing on we see that $n = n_i$ is a constant modulo λ, and so

$$m_k \equiv \frac{n}{k} \mod \lambda$$

where n depends on $f(\alpha)$. This is essentially Stickelberger's theorem for primes $p \equiv 1 \mod \lambda$ which split as the norm of an element $f(\alpha)$. We can already see the surprising result that **the m_k do not depend of he prime p!**

Kummer extrapolates the ideas in this paper into the three papers [18], [19], [20], all of which were published around 1847, in which he defines Ideal numbers for any prime p in

a cyclotomic field generated by a λ^{th} root of unity, and proves the decomposition law of Gauss sums for the primes $p \equiv 1 \mod \lambda$. The most troubling problem one might expect Kummer to have encountered is how to determine the prime ideals above primes $q \not\equiv 1 \mod \lambda$ as it is not clear what the congruence relation $\xi \equiv \alpha \mod f(\alpha)$ or $f(\xi) \equiv 0 \mod p$ should be replaced with. We have developed enough of the theory here already that it would take us too far afield to discuss this topic further, so we leave the interested reader to explore the papers of Kummer that followed [17], and simply note that Kummer derived the following decomposition theorem in $\mathbb{Z}[\alpha]$, given a prime q coprime to λ, let ν be the least positive integer for which $q^\nu \equiv 1 \mod \lambda$. Then q splits into $\frac{\lambda-1}{\nu}$ distinct ideal primes. If $\phi(\alpha)$ is one of the ideal primes dividing q then a complete set of residues modulo $\phi(\alpha)$ forms a finite field with q^ν elements.

We now turn our attention instead to the application of these concepts to the burning problem of the day, reciprocity laws.

3.2. **Proof of Eisenstein Reciprocity.** We now turn to one of the most general reciprocity laws which can be proven using these techniques, the law of Eisenstein reciprocity. We will follow Eisenstein's paper [6] which quite conveniently uses Kummer's new notation for the integers $\mathbb{Z}[\alpha]$ and uses ideal numbers as Kummer has described them thus far. We begin by briefly defining the relevant notation. Let λ be an odd prime and α a primitive λ^{th} root of unity. Let p be an odd prime of the form $p = \lambda\pi + 1$ and fix g a primitive root modulo p. Finally fix ζ to be a primitive p^{th} root of unity and define as before,

$$\begin{aligned}(\alpha, \zeta) =& \alpha\zeta^g + \alpha^2\zeta^{g^2} + \cdots + \alpha^{p-1}\zeta^{g^{p-1}} \\ =& \sum_{k=1}^{p-1} \alpha^{Ind(k)}\zeta^k\end{aligned}$$

This is consistent with the previous subsections notation. By the previous subsection and the rest of Kummer's related papers we have the prime decomposition

$$(\alpha, \zeta)^\lambda = \pm\alpha^r f(\alpha)^{m_1} f(\alpha^2)^{m_2} \dots f(\alpha^{\lambda-1})^{m_{\lambda-1}}$$

where $f(\alpha)$ is the ideal number, whether it is an actual number (think principal ideal) or not, for which $p = f(\alpha)f(\alpha^2)\dots f(\alpha^{\lambda-1})$ and for which the congruence

$$g^\pi \equiv \alpha^{-1} \mod f(\alpha)$$

holds. By Kummer the numbers m_k satisfy $0 < m_k < \lambda$ and $km_k \equiv 1 \mod \lambda$. When $f(\alpha)$ is not an actual number the meaning of the above equations can cause confusion as it is not an equality between numbers. For the first part of the proof though we will assume that $f(\alpha)$ is an actual complex number, that is, assume that p is the norm of an element of $\mathbb{Z}[\alpha]$. Let q be an odd prime distinct from p and λ and set ν to be the least positive integer for which $q^\nu \equiv 1 \mod \lambda$, and set $q^\nu = \lambda\varrho + 1$. To define the λ^{th} power residue symbol modulo any ideal prime $\phi(\alpha)$ dividing q we note the fact, which Eisenstein attributes to Gauss, that the finite field which $\phi(\alpha)$ defines has a primitive solution to the equation $x^{q^{\nu-1}} \equiv 1 \mod \phi(\alpha)$, that is, a solution γ, of which every other solution is a power. Because every non-zero element $g(\alpha)$ of the residue system is a solution of this equation we see that in fact

$g(\alpha)^{q^\nu-1} = (g(\alpha)^\varrho)^\lambda \equiv 1 \mod \phi(\alpha)$. Now because the powers α^k for $1 \leq k < \lambda$ are distinct modulo $\phi(\alpha)$ they must form a set of representatives of all the solutions to the equation $x^\lambda \equiv 1 \mod \phi(\alpha)$, and so we find for some integer k we have $g(\alpha)^\varrho \equiv \alpha^k \mod \phi(\alpha)$, and thus we define the residue symbol to be this power. That is,

$$\left(\frac{g(\alpha)}{\phi(\alpha)}\right) = \alpha^k \text{ where } k \text{ is such that } g(\alpha)^\varrho \equiv \alpha^k \mod \phi(\alpha).$$

Following Gauss we immediately find the relations

$$(\alpha,\zeta)^{\lambda\varrho+\lambda} = (\alpha,\zeta)^{q^\nu} \cdot (\alpha,\zeta)^{\lambda-1}. \tag{19}$$

By the fact that q divides the binomial coefficients $\binom{q}{k}$ for $0 < k < q$ we have $(\alpha,\zeta)^{q^\nu} \equiv (\alpha^{q^\nu},\zeta^{q^\nu}) \mod q$ where $\alpha^{q^\nu} = \alpha$ and $(\alpha,\zeta^{q^\nu}) = \alpha^{-Ind(q^\nu)}(\alpha,\zeta)$.

Setting $\epsilon = \pm\alpha^r$ and $V = f(\alpha)^{m_1} f(\alpha^2)^{m_2} \dots f(\alpha^{\lambda-1})^{m_{\lambda-1}}$ we find also $(\alpha,\zeta)^{\lambda\varrho+\lambda} = (\epsilon V)^{\varrho+1}$ from which we derive the relation

$$(\epsilon V)^\varrho \equiv \alpha^{-\nu Ind(q)} \mod q.$$

Now $\alpha^{-\nu Ind(q)} \equiv g^{\pi\nu Ind(q)} \equiv q^{\nu\pi} \mod f(\alpha)$ so that in fact by the definition of the λ^{th} power residue symbols,

$$\left(\frac{q}{f(\alpha)}\right)^\nu = \left(\frac{\epsilon V}{\phi(\alpha)}\right). \tag{20}$$

Using the binomial coefficient trick again we find that $(\alpha,\zeta)^\lambda \equiv (\alpha^\lambda,\zeta^\lambda) \equiv -1 \mod \lambda$. From this it follows that $\epsilon V \equiv -1 \mod \lambda$.

We now begin a discussion of the unit ϵ. We define the number $\eta = 1-\alpha$, so that p and η^{p-1} are associate numbers, $p = u\eta^{p-1}$ for a unit u. First suppose $\psi(\alpha) = a_0 + a_1\alpha + a_2\alpha^2 + \dots$ so that $\psi(\alpha) \equiv a_0 + a_1 + a_2 + \dots - (a_1 + 2a_2 + 3a_3 + \dots)\eta \mod \eta^2$. If we formally define the derivative operator, $d\phi(x) = a_1 + 2a_2x + 3a_3x^2 + \dots$ we find that $\phi(\alpha) \equiv \phi(1) - d\phi(1)\eta \mod \eta^2$. We simplify this with the notion $d\log f(1) = \frac{f(1)}{df(1)}$ and find more generally $f(\alpha^k) \equiv f(1)(1 - kd\log f(1)\eta) \mod \eta^2$. Finally we stitch this information together along with $(1 - kd\log f(1)\eta)^{m_k} \equiv 1 - d\log f(1)\eta \mod \eta^2$ to obtain

$$\begin{aligned} V = \prod_{k=1}^{\lambda-1} f(\alpha^k)^{m_k} \equiv & f(1)^{\sum m_k}(1 - d\log f(1)\eta)^{\lambda-1} \mod \eta^2 \\ \equiv & f(1)^{\lambda\frac{\lambda-1}{2}}(1 + d\log f(1)\eta) \mod \eta^2 \\ \equiv & f(1)^{\frac{\lambda-1}{2}}(1 + d\log f(1)\eta) \mod \eta^2 \end{aligned}$$

We now combine this with $\epsilon V \equiv -1 \mod \eta^2$ to get $f(1)r \equiv df(1) \mod \lambda$. If we choose $f(\alpha)$ beforehand so that for some rational integer m, $f(\alpha) \equiv m \mod \eta^2$ then we would find that $r \equiv 0 \mod \lambda$, and in fact $\epsilon = 1$. Multiplying $f(\alpha)$ by an appropriate power of α we may in fact satisfy the relation $f(\alpha) \equiv m \mod \eta^2$, and if $f(\alpha)$ is of this form we call $f(\alpha)$ **primary**, this condition is related to, but not the same as the previous definition of

a primary number in the Eisenstein integers. We can then under the assumption that $f(\alpha)$ is primary make the claim that,

$$\left(\frac{q}{f(\alpha)}\right)^{\nu} = \left(\frac{V}{\phi(\alpha)}\right) = \prod_{k=1}^{\lambda-1}\left(\frac{f(\alpha^k)}{\phi(\alpha)}\right)^{m_k}$$

By exchanging α for α^{m_k} in the definition of the power residue symbol we find the relation

$$\left(\frac{f(\alpha^k)}{\phi(\alpha)}\right)^{m_k} = \left(\frac{f(\alpha)}{\phi(\alpha^{m_k})}\right).$$

and by extending the power residue symbol multiplicatively in the denominator to form the Jacobi symbol we find that in fact

$$\prod_{k=1}^{\lambda-1}\left(\frac{f(\alpha^k)}{\phi(\alpha)}\right)^{m_k} = \prod_{k=1}^{\lambda-1}\left(\frac{f(\alpha)}{\phi(\alpha^{m_k})}\right) = \left(\frac{f(\alpha)}{q^{\nu}}\right) = \left(\frac{f(\alpha)}{q}\right)^{\nu}$$

Now p and ν are coprime from which the reciprocity law follows,

$$\left(\frac{q}{f(\alpha)}\right) = \left(\frac{f(\alpha)}{q}\right).$$

For the second part we extend this law to primary numbers $F(\alpha)$ which are the product of ideal primes $f(\alpha), f'(\alpha), f''(\alpha)\ldots$ whose norms $p, p', p''\ldots$ are primes congruent to 1 modulo λ. Let $\zeta, \zeta', \zeta''\ldots$ be primitive $p, p', p''\ldots$ roots of unity and let $g, g', g''\ldots$ be primitive roots modulo $p, p', p''\ldots$ which also satisfy

$$g^{\pi} \equiv \alpha^{-1} \mod f(\alpha),\ g'^{\pi'} \equiv \alpha^{-1} \mod f'(\alpha)\ldots$$

where $p = \lambda\pi + 1, p' = \lambda\pi' + 1\ldots$. Just as in (19) we find the relation

$$\begin{aligned}[(\alpha,\zeta)(\alpha,\zeta')\cdots]^{\lambda} &= \pm\,\alpha^{r} V \alpha^{r'} V' \cdots \\ &= \pm\,\alpha^{R} F(\alpha)^{m_1} F(\alpha^2)^{m_2} F(\alpha^3)^{m_3} \cdots F(\alpha^{\lambda-1})^{m_{\lambda-1}}\end{aligned}$$

This part of the proof is where it is important that the m_k did not depend on the primes p, and $f(\alpha)$.

This is now a relation not between ideal primes, but between actual numbers. If we raise this equation to the ϱ power where $q^{\nu} = \lambda\varrho + 1$ then the left side of the equation taken modulo an ideal prime divisor $\phi(\alpha)$ of q is congruent to the product

$$\left(\frac{q^{\nu}}{f(\alpha)}\right)\left(\frac{q^{\nu}}{f'(\alpha)}\right)\cdots = \left(\frac{q}{F(\alpha)}\right)^{\nu}.$$

Just as in the equation (20). The right side of the earlier equation after raising to the ϱ power becomes congruent modulo $\phi(\alpha)$ to

$$\left(\frac{\alpha^{R}\prod F(\alpha^k)^{m_k}}{\phi(\alpha)}\right).$$

Because $F(\alpha)$ was chosen to be primary we find that $R = 0$ and so all together this information becomes

$$\left(\frac{q^\nu}{F(\alpha)}\right) = \left(\frac{q}{F(\alpha)}\right)^\nu = \left(\frac{\prod F(\alpha^k)^{m_k}}{\phi(\alpha)}\right) = \prod \left(\frac{F(\alpha)}{\phi(\alpha^k)}\right) = \left(\frac{F(\alpha)}{q^\nu}\right) = \left(\frac{F(\alpha)}{q}\right)^\nu$$

and because p is coprime to ν we get the reciprocity

$$\left(\frac{q}{F(\alpha)}\right) = \left(\frac{F(\alpha)}{q}\right).$$

The last part of the proof is replacing the primary number $F(\alpha)$ which has only prime divisors whose norms are primes congruent to 1 mod λ, with any primary number $F(\alpha)$. The key is in finding a good replacement for the number (α, ζ). This is done by considering the sums invented by Kummer that extend the notion of a Gauss sum from a sum over a finite field with a prime number of elements to a finite field with an arbitrary number of elements. In fact it was Eisenstein that started the process of extending the base field by considering the special cases of octic reciprocity in the field $\mathbb{Q}(\zeta_8)$ and defined sums over fields $\mathbb{F}'_p$ where $p' = p^2$. We will give a definition of such a sum. Given an ideal prime $f(\alpha)$ whose norm $p' = p^f$ has degree f, we define the trace function. Notice that $\mu + \mu^p + \cdots + \mu^{p^{f-1}} \equiv \sigma \mod p$ where σ is an integer. We define $Tr(\mu)$ to be the least such positive integer and after setting ζ to be a primitive p^{th} root of unity define

$$(\alpha, \zeta) = \sum_\mu \left(\frac{\mu}{f(\alpha)}\right) \zeta^{Tr(\mu)}$$

where now the summation runs over a complete set of non-zero residues modulo $f(\alpha)$. Now using the properties of the new trace function Kummer proved a similar prime decomposition for such sums. Eisenstein then notes that if the degree, $f > 1$ then $f(\alpha) = f(\alpha^k)$ for some non-trivial value of k and so $(\frac{q}{f(\alpha)}) = (\frac{q}{f(\alpha^k)}) = (\frac{q}{f(\alpha)})^{m_k}$ from which it follows $(\frac{q}{f(\alpha)}) = 1$. We thus conclude the final reciprocity for arbitrary primary $F(\alpha)$ and extending by multiplication from q to a generic positive integer a

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

which is known as the Eisenstein reciprocity law.

The last few remarks in Eisenstein's paper are about Kummer's attacks on reciprocity, in particular that if h is the class number of the field $\mathbb{Q}(\alpha)$ then each ideal prime $f(\alpha)$ may be turned into an actual element of the field through $f(\alpha)^h$ and so if $\lambda \nmid h$ one may define the residue symbol

$$\left(\frac{f(\alpha)}{\phi(\alpha)}\right) = \left(\frac{f(\alpha)^h}{\phi(\alpha)}\right)^{h^{-1}}$$

where h^{-1} is any integer for which $hh^{-1} \equiv 1 \mod \lambda$. Kummer at this point still suspected that such primes which do not divide the corresponding class number were typical and even named them the regular primes. All of Kummer's attacks on reciprocity dealt almost

exclusively with the regular prime case and thus Eisenstein beat him to this generalization of reciprocity.

4. Stickelberger's Theorem on Ideal Class Annihilators

By the time Stickelberger's paper was published in 1890 the basic structure of algebraic number theory had been developed. Galois theory of fields was well established. Dedekind, Kummer's student had popularized the definition of algebraic integers, defined ideals in the ring of integers in a finite algebraic extension of $\mathbb{Q}$ and proven that every ideal of a number field has a unique decomposition into prime ideals. Dirichlet, who earlier had discovered the connection between L-functions and the older concept of classes of forms, had proven that the classes of fractional ideals modulo principal ideals forms a finite group and had even given an analytic formula involving an L-function that would determine the class number. Dirichlet had also proven the Dirichlet unit theorem which describes the rank of the group of units in the ring of integers in a number field. Explicit knowledge of the structure of the class group however was not (and really still is not) well understood and was the main obstacle in the way of Kummer's attacks on Fermat's Last Theorem. One of the principle difficulties in proving general statements about the class number and the class group is how little information about its structure can be obtained algebraically. One of the best, and in fact most explicit methods of dealing with the class group of a cyclotomic field is through Stickelberger's Theorem. The information comes from considering the class group as a module over the group ring $\mathbb{Z}[G]$. This essentially is completely understood by using the exponential notation.

Definition. *Let $G = \{g_1, g_2, \ldots, g_n\}$ be the Galois group of an Abelian algebraic extension K of $\mathbb{Q}$ where the group operation is denoted by multiplication. Given coefficients $a_1, a_2, \ldots, a_n \in \mathbb{Z}$ we let $\mathbb{Z}[G]$ be the additive group of formal sums*

$$\sum_{i=1}^{n} a_i g_i$$

Multiplication in $\mathbb{Z}[G]$ is defined by extending the multiplication in G linearly over $\mathbb{Z}$. The action of the additive group $\mathbb{Z}[G]$ is given by the following relation

$$k^{\sum_{i=1}^{n} a_i g_i} = \prod_{i=1}^{n} g_i(k)^{a_i}$$

where k is an arbitrary element of K.

This action naturally extends to (fractional) ideals and it is in this language best that we may state the full version of Stickelberger's theorem.

4.1. **Stickelberger's Theorem.** Let $m > 0$ be a positive integer and set $K = \mathbb{Q}(\zeta)$ where ζ is a primitive m^{th} root of unity and fix $G = Gal(\mathbb{Q}(\zeta)/\mathbb{Q})$ where we denote $\sigma_t \in G$ as the automorphism $\sigma_t \zeta = \zeta^t$. let $\mathfrak{o}$ be the ring of integers in K and let $\mathfrak{p}$ be a prime ideal in $\mathfrak{o}$ of degree f not containing m, so that the residue field has p^f elements. For $\mu \in \mathfrak{o} - \mathfrak{p}$ let $(\frac{\mu}{\mathfrak{p}})$

denote the m^{th} power residue symbol, that is the power of ζ that is congruent to $\mu^{\frac{p^f-1}{m}}$ modulo $\mathfrak{p}$, and let $\psi(\mu)$ denote the additive character $\xi^{Tr(\mu)}$ where $Tr(\mu)$ is the integer in $\{0,1,2,\ldots,p-1\}$ which is congruent to $\mu+\mu^p+\mu^{p^2}+\cdots+\mu^{p^{f-1}}$ modulo $\mathfrak{p}$ and ξ is a fixed primitive p^{th} root of unity. We may define the Gaussian sum over the residue field $\mathfrak{o}/\mathfrak{p}$ as follows:

$$G(\mathfrak{p})=\sum\left(\frac{\mu}{\mathfrak{p}}\right)^{-1}\psi(\mu)$$

where μ runs over a complete set of non-zero residues modulo $\mathfrak{p}$.

By definition $G(\mathfrak{p})$ is an element of $K(\xi)$, and we quickly see that $G(\mathfrak{p})^m$ is an element of K by noting that if an automorphism, σ leaves ζ fixed and sends ξ to ξ^c then

$$\begin{aligned}\sigma(G(\mathfrak{p}))=&\sigma\left(\sum\left(\frac{\mu}{\mathfrak{p}}\right)^{-1}\psi(\mu)\right)\\=&\sum\left(\frac{\mu}{\mathfrak{p}}\right)^{-1}\psi(c\mu)\\=&\sum\left(\frac{c^{-1}\mu}{\mathfrak{p}}\right)^{-1}\psi(\mu)\\=&\left(\frac{c}{\mathfrak{p}}\right)G(\mathfrak{p})\end{aligned}$$

From this and the fact that $(\frac{c}{\mathfrak{p}})^m=1$ we deduce that $G(\mathfrak{p})^m$ is invariant under conjugation by σ.

We expect, following the program that Gauss initiated, that it will be profitable to study the prime factorization of $G(\mathfrak{p})^m$ in the ring of integers $\mathfrak{o}$. The solution to the problem of determining this factorization is given by Stickelberger's theorem.

Theorem (Stickelberger)**.** *The prime ideal factorization of the principal ideal generated by the m^{th} power of the Gaussian sum $G(\mathfrak{p})$ is given by*

$$(G(\mathfrak{p})^m)=\mathfrak{p}^{\sum t\sigma_t^{-1}} \tag{21}$$

where the sum runs over all $0\le t<m$ such that t is coprime to m.

The importance of this theorem is that **the element $\sum t\sigma_t^{-1}$ is independent of $\mathfrak{p}$.** This truly amazing fact then gives us a universal way of connecting a prime ideal (at least an unramified one, which in this case is just one not containing m) to a principal one. The ramified primes do present some difficulties for general m, but we will postpone addressing those for now Notice that Kummer had proven this result for the case $m=\lambda$ an odd prime. Notice for this case the only ramified prime is $(1-\zeta)$, which is already a principal ideal, and thus we immediately have a result on the ideal class group. The extension of this theorem of Kummer to arbitrary m is the subject of Stickelberger's paper [26] of 1890, 40 full years after Kummer's result, and 10 years before the publication of Hilbert's Zahlbericht.

An immediate application of this theorem is to quadratic number fields. Let $\lambda > 3$ be a prime $\lambda \equiv 3 \mod 4$. If p is a prime $p \equiv 1 \mod \lambda$ then p splits completely in the two fields $\mathbb{Q}(\sqrt{-\lambda})$, $\mathbb{Q}(\zeta_\lambda)$. In the ring of integers $\mathfrak{o}$ of $\mathbb{Q}(\sqrt{-\lambda})$ write $(p) = \mathfrak{p}\overline{\mathfrak{p}}$ where the bar denotes complex conjugation. If $\widetilde{\mathfrak{o}}$ denotes the ring of integers of $\mathbb{Q}(\zeta_\lambda)$ then we have the factorization $\mathfrak{p}\widetilde{\mathfrak{o}} = \prod \mathfrak{P}^{\sigma_s}$ where $\mathfrak{P}$ is a prime ideal of $\widetilde{\mathfrak{o}}$ above $\mathfrak{p}$, and s runs over all the nonzero squares modulo λ. By Stickelberger's theorem we have the factorization $(G(\mathfrak{P})^\lambda) = \mathfrak{P}^{\sum t\sigma_t^{-1}}$, $t = 1, 2, \ldots, \lambda - 1$. Applying $\sum \sigma_s$, s a square modulo λ gives $(\alpha)\widetilde{\mathfrak{o}} = \mathfrak{p}^{\sum t\sigma_t^{-1}}\widetilde{\mathfrak{o}} = \mathfrak{p}^{\sum s}\overline{\mathfrak{p}}^{\sum n}\widetilde{\mathfrak{o}}$ where $\alpha \in \mathfrak{o}$ and n runs over the nonsquares in the interval $[1, \lambda - 1]$. Put $R = \sum s, N = \sum n$ so we have $(\alpha)\mathfrak{o} = \mathfrak{p}^R\overline{\mathfrak{p}}^N$. If $[\mathfrak{U}]$ denotes the ideal class of $\mathfrak{U}$ and 1 is the unit class then $[\mathfrak{p}]^{-1} = [\overline{\mathfrak{p}}]$. Thus $[\mathfrak{p}]^{N-R} = 1$. On the other hand if $1 \leq r \leq \lambda - 1$ we have by Stickelberger's theorem $G(\mathfrak{P})^{\sigma_r - r} = \beta$ for some $\beta \in \widetilde{\mathfrak{o}}$. Raising to the λ^{th} power we see for any square $r \neq 1$, $(\mathfrak{p}^R\overline{\mathfrak{p}}^N)^{1-r} = (\gamma)^\lambda\widetilde{\mathfrak{o}}$ for some $\gamma \in \mathfrak{o}$. It follows that $([\mathfrak{p}]^{(N-R)/\lambda})^{r-1} = 1$. Now from the above argument above $([\mathfrak{p}]^{(N-R)/\lambda})^\lambda = 1$. Since $(r - 1, \lambda) = 1$ we have proven the following

Theorem. *Let $\mathfrak{p}$ be a prime ideal of degree 1 in $\mathbb{Q}(\sqrt{-\lambda})$ for $\lambda \geq 3$ a prime such that $\lambda \equiv 3 \mod 4$. Then, $[\mathfrak{p}]^{(N-R)/\lambda} = 1$.*

By Hilbert's theorem 89 which implies that the primes of degree 1 generate the class group of a Galois number field, and which we discuss later on, we find the completely algebraically proven corollary on the class group.

Corollary. *For a prime $\lambda \equiv 3 \mod 4$ the class number of $\mathbb{Q}(\sqrt{-\lambda})$ divides $(N - R)/\lambda$.*

In fact this technique may be extended to Abelian quartic fields and so on. For a brief discussion of such generalizations consult [23]. Using analytic techniques it may be shown that in fact the class number of $\mathbb{Q}(\sqrt{-\lambda})$ is equal to $(N - R)/\lambda$, but no algebraic proof of this fact is known.

For now we consider Stickelberger's theorem to be the main ingredient in proving the following.

Corollary. *Let $\theta = \frac{1}{m}\sum t\sigma_t^{-1} \in \mathbb{Q}[G]$ be the Stickelberger element in the group ring with coefficients in $\mathbb{Q}$. Let $\beta \in \mathbb{Z}[G]$ be such that $\beta\theta \in \mathbb{Z}[G]$. Then $\beta\theta$ annihilates the ideal class group of $\mathbb{Q}(\zeta)$.*

To prove the corollary of Stickelberger's theorem we simply apply the following lemma that shows we need not worry about the ramified primes because the set of prime ideals not containing m generate the ideal class group. Thus because these kinds of ideals are annihilated by the elements $\beta\theta$ we see that the entire class group is annihilated by $\beta\theta$.

Lemma. *Let K be an algebraic number field and let M be a fixed ideal in the ring of integers of K. Then every ideal class of K contains an ideal prime to M*

This is Proposition 15.3.1 in [14] and the proof that appears there is short enough that we include it here

Proof. Let A be an ideal of K and let $\{P_1, \ldots, P_t\}$ be the set of primes dividing M which do not divide A. If P divides A then let $a(P)$ be the exponent of P occurring in the prime decomposition of A. Choose some

$$\pi(P) \in P^{a(P)} - P^{a(P)-1}$$

By the Chinese Remainder Theorem we may find an α for which

$$\alpha \equiv \pi(P) \mod P^{a(P)+1} \text{ for } P \mid A$$

$$\alpha \equiv 1 \mod P_i \text{ for } i = 1, 2, \ldots, t$$

One checks then that $(\alpha) = AC$ where $(C, M) = 1$. This shows that there is such an ideal in the inverse of the ideal class of A, and because A was arbitrary we deduce the result. □

As Washington points out in [31], Kummer may have seen a simple argument to prove the corollary during his work on reciprocity by proving that $\beta\theta$ sends any prime $\mathfrak{p}$ of degree 1 to a principal ideal and then applying the fact that primes of degree 1 generate the ideal class group, a theorem which may be proven algebraically as was done in Hilbert's Zahlbericht Theorem 89. We discuss this paper later on, but first we turn to the historically first published proof of Stickelberger's theorem.

An outline of Stickelberger's quite readable and mostly self contained paper [26] goes as follows. After this discussion a shorter proof will be given that embodies the main ideas in a somewhat more elegant presentation.

4.1.1. *(Ludwig Stickelberger).* **Ueber eine Verallgemeinerung der Kreistheilung**

Chapter 1 (Restsysteme) A few basic theorems on additive characters defined on quotients of the integers in a general finite algebraic extension of $\mathbb{Q}$ are proven. A particularly nice discussion on the discriminant of a number field is given.

Chapter 2 (Restcharaktere) For a field K which contains a primitive m^{th} root of unity, ζ and a prime ideal $\mathfrak{p}$ of degree f which does not contain m, the m^{th} power residue symbol $(\frac{\mu}{\mathfrak{p}})$ is defined to be ζ^a where a satisfies $\mu^{\frac{p^f-1}{m}} \equiv \zeta^a \mod \mathfrak{p}$. It is noted that of course, $(\frac{\mu_1\mu_2}{\mathfrak{p}}) = (\frac{\mu_1}{\mathfrak{p}})(\frac{\mu_2}{\mathfrak{p}})$.

Chapter 3 (Resolventen und Eisenstein'sche Summen) This section is devoted to defining what Stickelberger calls the 'Verallgemeinerung der Kreistheilung' which might be called the generalized Gauss sum (Kreistheilung was the word used for the various sums of roots of unity, most of which we realize now as a Gauss sum). The distinction that he is making between the regular Gauss sums is that instead of summing $\sum[\mu]\xi^\mu$ over μ in the interval $[1, p-1]$ we sum over a set of non-zero representatives of the classes modulo $\mathfrak{p}$, but we must replace the exponent μ with an appropriate value. This is accomplished by the use of the trace function, $Tr(\mu)$ mentioned earlier. Stickelberger attributes this generalization

to Eisenstein (see the remark in the Eisenstein Reciprocity section) and calls the sums of these types Eisenstein sums. He then writes

$$F_a(\xi) = \sum \left(\frac{\mu}{\mathfrak{p}}\right)^a \xi^{Tr(\mu)}$$

and proves that

$$F(\theta^a, \xi)F(\theta^b, \xi) = \left(\frac{-1}{\mathfrak{p}}\right)^a \sum_{\mu} \left(\frac{\mu}{\mathfrak{p}}\right)^{a+b} + \psi_{a,b}F_{a+b}(\xi)$$

where $\psi_{a,b}$ is the Jacobi sum

$$\psi_{a,b} = \sum_{\mu} \left(\frac{\mu}{\mathfrak{p}}\right)^a \left(\frac{1-\mu}{\mathfrak{p}}\right)^b$$

So when $m \nmid a+b$ we get the fundamental relations $F_a(\xi)F_b(\xi) = \psi_{a,b}F_{a+b}(\xi)$, and for $m \nmid a$, $F_a(\xi)F_{-a}(\xi^{-1}) = p^f$.

Chapter 4 (Arithmetische Eigenschaften der Facultäten) The main result of this section is a simple lemma:

Lemma. *If $a > 0$ then $\frac{(1-\xi)^a}{a!}$ is equal to a fraction in $\mathbb{Q}(\zeta)$ the denominator of which is coprime to p.*

Chapter 5 (Zur Theorie der Kreistheilung) This chapter gives a new proof of the the theorem in the case where $\mathfrak{p}$ is a prime of degree 1. This is done in part by considering the polynomial given by substituting $1+u$ for ξ in $F_a(\xi)$ where u is a variable in the polynomial ring $\mathbb{Z}[u]$. The congruence relation is then derived for $0 \leq a < m$

$$F_a(1+u) \equiv -\frac{u^{p-1-an}}{(p-1-an)!} \mod (\mathfrak{p}, u^{p-an})$$

where $nm = p-1$. To examine what happens to $F_a(\xi)$ we expect to substitute $\xi - 1$ for u, if we are to do this however we will need to work in the ring of integers of $K(\xi)$, denote this as $\widetilde{\mathfrak{o}}$. Then the ideal $\mathfrak{p}$ is completely ramified in $\widetilde{\mathfrak{o}}$, we denote the single prime above $\mathfrak{p}$ as $\widetilde{\mathfrak{p}}$ which contains $\xi - 1$ and $\widetilde{\mathfrak{o}}^{p-1} = \mathfrak{p}\widetilde{\mathfrak{o}}$. Thus we get the relation

$$F_a(\xi) \equiv \frac{(1-\xi)^{p-1-an}}{(p-1-an)!} \mod \widetilde{\mathfrak{p}}^{p-an}$$

The next object is to determine what happens for conjugate primes to $\widetilde{\mathfrak{p}}$. Define $a^{(k)}$ to be the least positive integer for which $a^{(k)}a \equiv ak \mod m$, and call $\mathfrak{p}_k = \sigma_{k^{-1}}(\mathfrak{p})$, with the analogous definition for $\widetilde{\mathfrak{p}}_k$. Then conjugating the above equation becomes

$$F_a(\xi) \equiv \frac{(1-\xi)^{p-1-a^{(k)}n}}{(p-1-a^{(k)}n)!} \mod \widetilde{\mathfrak{p}}_k^{p-a^{(k)}n}$$

Using this relation and the the lemma from the earlier section a counting argument derives the exact number of times that $\widetilde{\mathfrak{p}}_k$ divides $F_a(\xi)$ and thus the number of times $\mathfrak{p}_k$ divides $F_a(\xi)^m$. Finally because $F_a(\xi)F_{-a}(\xi^{-1}) = p^f$ these must be the only prime divisors of $F_a(\xi)$, and the factorization is known.

Chapter 6 (Congruenzbedingungen für die verallgemeinerten Resolventen) This chapter is devoted to applying the arguments of the last section to the generalized Gauss sums that occur when the prime $\mathfrak{p}$ does not have degree 1. Thus let f be the degree of $\mathfrak{p}$, and set $p^f - 1 = mn$

The first trick is to use properties of binomial coefficients to prove that the power residue symbol satisfies a stronger congruence

$$\left(\frac{\mu}{\mathfrak{p}}\right) \equiv \mu^n \mod \mathfrak{p}^{f+1}$$

Now similarly for some integer t

$$Tr(\mu) \equiv \mu + \mu^p + \mu^{p^2} + \cdots + \mu^{p^{f-1}} \equiv t \mod \mathfrak{p}^{f+1}$$

For any $l < s$ we have the congruence

$$\xi^s \equiv \sum_{k=0}^{l} \frac{(\xi-1)^k}{k!} s(s-1)\dots(s-k+1) \mod (1-\xi)^{l+1}$$

where each term $\frac{(\xi-1)^k}{k!}$ may be thought of as an integer in $\widetilde{\mathfrak{o}}$ by the lemma in Chapter 4. All this leads us to the fact that for $l < (f+1)(p-1)$

$$\begin{aligned}\xi^{Tr(\mu)} &\equiv \sum_{k=0}^{l} \binom{t}{k}(\xi-1)^k \mod \overline{\mathfrak{p}}^{l+1} \\ &\equiv \sum \binom{\mu}{x_0}\binom{\mu^p}{x_1}\dots\binom{\mu^{p^{f-1}}}{x_{f-1}}(\xi-1)^{x_0+x_1+\cdots+x_{f-1}} \mod \overline{\mathfrak{p}}^{l+1}\end{aligned}$$

Where the latter sum is over f-tuples of non-negative $x_0, \dots, x_{f-1}$ the sum of which is less than or equal to l. This is how the term $\xi^{Tr(\mu)}$ may be handled, the rest of the argument is somewhat straightforward, analyzing the sum

$$\sum \left(\frac{\mu}{\mathfrak{p}}\right)^a \xi^{Tr(\mu)}$$

by splitting $\xi^{Tr(\mu)}$ into the previous summation, then switching the overall order of summation and demonstrating that most of the terms cancel out. We will cover most of the rest of the details in the concise version to follow. The result appears here in a form *equivalent* to (21).

Chapter 7 (Anwendung auf die Theorie der quadratischen Formen) This chapter gives an application to the theory of quadratic forms by using essentially the remark above about

quadratic fields to prove a criterion for cases when a prime p may be written in the form

$$4p^K = C^2 + MD^2$$

Chapter 8 (Untersuchung der Eisenstein'schen Summen) This section uses the theory of Eisenstein sums to evaluate residue characters $[\mu]$ for fixed m and μ and varying $\mathfrak{p}$, a popular application of these kinds of theories as may be seen in [23].

4.1.2. *(A Proof).* We now tackle proving Stickelberger's theorem and show how it gives a result on the ideal class group of a cyclotomic field. This proof is taken from the exercises in [14] and closely resembles what we have seen thus far in this presentation

Proof. Let $K = \mathbb{Q}(\zeta_m)$ where ζ_m is a primitive m^{th} root of unity. Given a prime p which does not divide m let $q = p^f$ where f is the least positive integer form which $p^f \equiv 1 \mod m$. Let $\mathfrak{p}$ be a prime above p which will have degree f. We will also need to work in the larger field $K(\zeta_{q-1})$ where ζ_{q-1} is a primitive $(q-1)^{st}$ root of unity, so let $\mathfrak{P}$ be a prime ideal above $\mathfrak{p}$ in the larger ring of integers. To be allowed even more freedom we allow ourselves to work in the composite field $K(\zeta_{q-1}, \zeta_p)$ where ζ_p is a primitive p^{th} root of unity, the only prime ideal over $\mathfrak{P}$ in this even larger ring will be denoted by $\mathscr{P}$, which divides $\mathfrak{P}$ exactly $p-1$ times and contains the element $\lambda = 1 - \zeta_p$. Thus we have the inclusion of ideals

$$(p) \subset \mathfrak{p} \subset \mathfrak{P} \subset \mathscr{P}$$

corresponding to the inclusion of fields

$$\mathbb{Q} \subset K \subset K(\zeta_{q-1}) \subset K(\zeta_{q-1}, \zeta_p)$$

Note that the residue fields corresponding to $\mathfrak{p}$ and $\mathfrak{P}$ are isomorphic. We will prove the factorization of the Gauss sum by proving what is called the Stickelberger congruence. This congruence involves interestingly enough the p-adic expansion of a in the formula for $F_a(\xi)$ which was studied by Stickelberger. We define only a little more new notation. Let ω be the character of the residue field of $\mathfrak{P}$ which sends a nonzero element μ to the unique $(q-1)^{st}$ root of unity ζ_{q-1}^a for which $\mu \equiv \zeta_{q-1}^a \mod \mathfrak{P}$. Now define the Gauss sum

$$g_a = \sum_{\mu} \omega(\mu)^a \zeta_p^{Tr(\mu)}$$

where μ runs over a complete set of non-zero residue modulo $\mathfrak{P}$ in the ring of integers of $K(\zeta_{q-1})$. Note that the sum that appears in (21) is the sum $G(\mathfrak{p}) = g_a$ where $a = \frac{q-1}{m}$ which is so because the residue fields of $\mathfrak{p}$ and $\mathfrak{P}$ are naturally isomorphic.

Given an a with $1 < a < q-1$ write the p-adic expansion $a = a_0 + a_1 p + a_2 p^2 + \cdots + a_{f-1} p^{f-1}$, $0 \leq a_i < p$. Define

$$S(a) = a_0 + a_1 + \cdots + a_{f-1}$$

The Stickelberger congruence is written

$$g_a \equiv \frac{-(-\lambda)^{S(a)}}{a_0! a_1! \ldots a_{f-1}!} \mod \mathscr{P}^{S(a)+1} \tag{22}$$

This is essentially the congruence that appears in Stickelberger's Chapter 5, but we have unraveled the powers of λ from the product. This congruence will give us the complete prime factorization of the Gauss sum g_a. We start proving this congruence by generalizing (14) to the modern version of the Jacobi sum. Define this Jacobi sum to be

$$J(\omega^{-n}, \omega^{-m}) = \sum_{\mu} \omega(\mu)^{-n} \omega(1-\mu)^{-m}$$

where μ runs over a complete set of non-zero residues modulo $\mathfrak{P}$. Let $1 \leq n, m < q-1$ and we shall prove the congruence

$$J(\omega^{-n}, \omega^{-m}) \equiv -\frac{(m+n)!}{n!m!} \mod \mathfrak{P}$$

First notice that $J(\omega^{-n}, \omega^{-m}) = J(\omega^{q-1-n}, \omega^{q-1-m})$. Now just as in the proof of (14) we take this sum modulo $\mathfrak{P}$ and find

$$\begin{aligned} J(\omega^{q-1-n}, \omega^{q-1-m}) &\equiv \sum_{\mu} \mu^{q-1-n}(1-\mu)^{q-1-m} \mod \mathfrak{P} \\ &\equiv \sum_{\mu} \mu^{q-1-n} \Big(\sum_{k=0}^{q-1-m} (-1)^k \binom{q-1-m}{k} \mu^k \Big) \mod \mathfrak{P} \\ &\equiv \sum_{\mu} \sum_{k=0}^{q-1-m} (-1)^k \binom{q-1-m}{k} \mu^{k+q-1-n} \mod \mathfrak{P} \\ &\equiv \sum_{k=0}^{q-1-m} (-1)^k \binom{q-1-m}{k} \sum_{\mu} \mu^{k+q-1-n} \mod \mathfrak{P} \end{aligned}$$

We may evaluate the inner sum

$$\sum_{\mu} \mu^{k+q-1-n} \equiv \begin{cases} 0 \text{ if } q-1 \text{ does not divide } k-n \\ q-1 \text{ if } q-1 \text{ divides } k-n \end{cases}$$

Of course for the values of k, n, m that we are restricted to the only value of k for which $q-1$ divides $k-n$ is when $k=n$ and so we are left with

$$J(\omega^{q-1-n}, \omega^{q-1-m}) \equiv (-1)^{n+1} \binom{q-1-m}{n} \mod \mathfrak{P}$$

The last step is a straightforward calculation that

$$(-1)^{n+1} \binom{q-1-m}{n} \equiv -\frac{(m+n)!}{n!m!} \mod p$$

This is exactly Eisenstein's argument from cubic reciprocity, just applied to the more general Eisenstein sums. This result implies that for $1 < a < q-1$ where the p-adic expansion of a is given by $a_0 + a_1 p + a_2 p^2 \cdots + a_{f-1} p^{f-1}$

$$J(\omega^{-1}, \omega^{-(a-1)}) \equiv -a_0 \mod \mathfrak{P} \tag{23}$$

We now express our Gauss sums g_a in terms of the Jacobi sum via the relation that Stickelberger addressed in Chapter 3. Suppose $q-1 \nmid m+n$, then we will prove the standard relation

$$g_n g_m = J(\omega^{-n}, \omega^{-m}) g_{n+m}$$

The proof goes as follows

$$\begin{aligned} g_n g_m &= \left(\sum_{\mu} \omega^{-n}(\mu) \zeta_p^{Tr(\mu)} \right) \left(\sum_{\tau} \omega^{-m}(\tau) \zeta_p^{Tr(\tau)} \right) \\ &= \sum_{\mu,\tau} \omega^{-n}(\mu) \omega^{-m}(\tau) \zeta_p^{Tr(\mu+\tau)} \\ &= \sum_{\gamma} \left(\sum_{\mu+\tau=\gamma} \omega^{-n}(\mu) \omega^{-m}(\tau) \right) \zeta^{Tr(\gamma)} \end{aligned}$$

If $\gamma = 0$ then $\sum_{\mu+\tau=0} \omega^{-n}(\mu)\omega^{-m}(\tau) = \sum_{\mu} \omega^{-n}(\mu)\omega^{-m}(-\mu) = 0$ because $q-1 \nmid m+n$
If $\gamma \neq 0$ then setting $\mu = \gamma\mu'$ and $\tau = \gamma\tau'$ we find

$$\begin{aligned} \sum_{\mu+\tau=\gamma} \omega^{-n}(\mu)\omega^{-m}(\tau) &= \sum_{\mu'+\tau'=1} \omega^{-n}(\gamma\mu')\omega^{-m}(z\gamma\tau') \\ &= \omega^{-(n+m)}(\gamma) J(\omega^{-n}, \omega^{-m}) \end{aligned}$$

Whence the result follows.

Thus if $1 \leq a < p-1$ we may use this equality $a-1$ times and simplify (23) to the following

$$g_a \equiv \frac{(-1)^{a+1} g_1^a}{a!} \mod \mathfrak{P}$$

We have reduced our problem to that of computing g_1^a modulo $\mathscr{P}^{a+1}$. This ca be done by manipulating the sum in the definition of g_1 in the same way as was carried out in Stickelberger's Chapter 5. However, we have the fortune of a great trick. Notice that $\sum_{\mu} \omega^{-1}(\mu) = 0$. Thus,

$$\frac{g_1}{\lambda} = \sum_{\mu} \omega^{-1}(\mu) \frac{\zeta_p^{Tr(\mu)} - 1}{1 - \zeta_p}$$

We may of course evaluate

$$\frac{\zeta_p^m - 1}{1 - \zeta} = -(1 + \zeta_p + \zeta_p^2 + \cdots + \zeta_p^{m-1}) \equiv -m \mod \lambda$$

which yields the simplification

$$\begin{aligned}\frac{g_1}{\lambda} &\equiv -\sum_{\mu}\omega^{-1}(\mu)Tr(\mu) \mod \mathscr{P}\\ &\equiv -\sum_{\mu}\mu^{-1}(\mu+\mu^p+\cdots+\mu^{p^{f-1}}) \mod \mathscr{P}\\ &\equiv -\sum_{\mu}(1+\mu^{p-1}+\mu^{p^2-1}+\cdots+\mu^{p^{f-1}-1}) \mod \mathscr{P}\\ &\equiv -(q-1)\equiv 1 \mod \mathscr{P}\end{aligned}$$

From this we see quite simply that

$$g_1^a \equiv \lambda^a \mod \mathscr{P}^{a+1}$$

from which we derive the relation for $1 \le a < p-1$

$$g_a \equiv \frac{(-1)^{a+1}\lambda^a}{a!} \mod \mathscr{P}^{a+1}$$

Suppose then that the congruence (22) is true for some $1 \le a < q-1$ and that $pa < q-1$. Because $Tr(\mu^p) = Tr(\mu)$, we have $g_a = g_{pa}$. We also have $S(pa) = S(a)$ so that (22) is unchanged when a is multiplied by p. Putting together the proofs for $1 \le a < p-1$ and $pa < q-1$ we may conclude that the Stickelberger congruence is proved.

Setting $a = \frac{p^f-1}{m}$ and noting that $g_a = G(\mathfrak{p})$ we make the following conclusions

(1) $\mathfrak{p}$ divides $G(\mathfrak{p})^m$ exactly $\frac{m}{p-1}S(\frac{p^f-1}{m})$ times

(2) $\sigma_t^{-1}(\mathfrak{p})$ divides $G(\mathfrak{p})^m$ exactly $\frac{m}{p-1}S(t\frac{p^f-1}{m})$ times

The only thing that might need mentioning is that the $p-1$ term comes from the ramification of $\mathfrak{P}$ in $K(\zeta_{q-1},\zeta_p)$ and the m comes from the m^{th} power in $G(\mathfrak{p})^m$. To translate this into the more elegant form found in (21) we examine the subgroup

$$G(\mathfrak{p}) = \{\sigma \in G = Gal(K/\mathbb{Q}) \mid \sigma(\mathfrak{p}) = \mathfrak{p}\}$$

This is known as the decomposition group, and a standard result in algebraic number theory is that it is generated by the element σ_p. Now let $t_1, t_2, \ldots, t_g$ be a set of integers representing the cosets of $(\mathbb{Z}/(m))^*$ modulo the cyclic subgroup generated by p, thus if $1 \le t < m, (t,m) = 1$ then $t \equiv t_i p^j \mod m$ for a unique pair (i,j), $0 \le j < f$, $1 \le i \le g$. Thus we have the equation

$$(G(\mathfrak{p}))^m = \mathfrak{p}^{\gamma'} \text{ where } \gamma' = \frac{m}{p-1}\sum_{i=1}^{g} S(t_i\frac{p^f-1}{m})\sigma_{t_i}^{-1}$$

This is the form of the theorem that appears in Stickelberger's paper. To simplify it further takes a little more work A simple exercise shows that $S(a) = (p-1)\sum_{i=0}^{f}\left\langle \frac{p^i a}{q-1}\right\rangle$ where

$\langle x \rangle$ denotes the fractional part of x. This allows us to write

$$\gamma' = m \sum_{i=1}^{g} \left(\sum_{j=0}^{f-1} \left\langle \frac{p^j t_i}{m} \right\rangle \right) \sigma_{t_i}^{-1}$$

Because σ_p leaves $\mathfrak{p}$ unchanged γ' has the same effect on $\mathfrak{p}$ as

$$\begin{aligned} \gamma =& m \sum_{i=1}^{g} \left(\sum_{j=0}^{f-1} \left\langle \frac{p^j t_i}{m} \right\rangle \right) \sigma_{t_i}^{-1} \sigma_{p^j}^{-1} \\ =& \sum_{t \ mod \ m} \left\langle \frac{t}{m} \right\rangle \sigma_t^{-1} \\ =& \sum_{t} t \sigma_t^{-1} \text{ where } 1 \leq t < m \text{ and } (t, m) = 1 \end{aligned}$$

This proves the main theorem of this section. □

For a direct proof of the Corollary to Stickelberger's theorem, found in [31], we start with the observation that the primes of degree 1 in a number field generate the ideal class group. This is a consequence of Hilbert's theorem 89 in [13] which states

Theorem (Hilbert's Theorem 89)**.** *In each ideal class of a Galois number field there exist ideals whose prime factors are all ideals of degree 1.*

The proof of this theorem for the case of a cyclotomic field was carried out by Kummer in [20] and is a completely algebraic proof, not dissimilar to the proof of the lemma at the beginning of this Section. Thus we only need prove Stickelberger's theorem for primes of degree 1, eliminating entirely the need for the generalized Gauss sums that Stickelberger called the Verallgemeinerung der Kreistheilung.

Proof. To avoid confusion we fix $\zeta_k = e^{2\pi i/k}$ for each k. Let $p \equiv 1 \mod m$ be a prime. Then p splits completely in $\mathbb{Q}(\zeta_m)/\mathbb{Q}$ and is totally ramified in $\mathbb{Q}(\zeta_{mp})/\mathbb{Q}(\zeta_m)$. Let $\mathfrak{p}$ be a prime of $\mathbb{Q}(\zeta_m)$ above p, so that $\{\sigma_a(\mathfrak{p})\}$ is the set of all primes above p, and let $\mathfrak{Q}$ be the prime of $\mathbb{Q}(\zeta_{mp})$ above $\mathfrak{p}$. We denote the primes of $\mathbb{Q}(\zeta_{mp})$ above p as $\sigma_a\mathfrak{p}$, $(a, m) = 1$. Let $\tau : \zeta_p \to \zeta_p^g$ generate $Gal(\mathbb{Q}(\zeta_{mp})/\mathbb{Q}(\zeta_m))$, where g is a primitive root modulo p.

Let χ denote character modulo p of order exactly m and set $G(\chi) = \sum_{t=1}^{p-1} \chi(t)\zeta_p^t$. We also have the well established relations $G(\chi)\overline{G(\chi)} = p$ where the bar denotes complex conjugation and of course $\tau G(\chi) = \chi(s)^{-1} G(\chi)$. We may write

$$(G(\chi)) = \prod_{(a,m)=1} \sigma_a^{-1}(\mathfrak{Q})^{r_a},\ r_a \in \mathbb{Z}.$$

Taking the norm down to $\mathbb{Q}(\zeta_m)$, we obtain

$$(G(\chi)^{p-1}) = \prod_{(a,m)=1} \sigma_a^{-1}(\mathfrak{p})^{r_a}.$$

Note that $G(\chi)^{p-1}$ is fixed by τ and so is therefore in $\mathbb{Q}(\zeta_m)$. Since $G(\chi)\overline{G(\chi)} = p$ and $\prod \sigma_a^{-1}(\mathfrak{p}) = (p)$, it follows that $0 \leq r_a \leq p-1$ for all a.

Since $G(\chi)/(\zeta_p - 1)^{r_a}$ is relatively prime to $\sigma_a^{-1}(\mathfrak{Q})$, and since τ acts trivially mod $\sigma_a^{-1}(\mathfrak{Q})$ (because $\mathfrak{Q}$ is completely ramified in $\mathbb{Q}(\zeta_{mp})/\mathbb{Q}(\zeta_m)$), we have

$$1 \equiv \frac{G(\chi)^\tau}{(\zeta_p^g - 1)^{r_a}} \frac{(\zeta_p - 1)^{r_a}}{G(\chi)} \equiv \frac{\chi(g)^{-1}}{(1 + \zeta_p + \zeta_p^2 + \cdots + \zeta_p^{g-1})^{r_a}} \equiv \frac{\chi(g)^{-1}}{g^{r_a}} \mod \sigma_a^{-1}(\mathfrak{Q})$$

Therefore $g^{r_a} \equiv \chi(g)^{-a} \mod \mathfrak{Q}$ and hence $\mod \mathfrak{p}$. Write $\chi(g) = g^{-d} \mod \mathfrak{p}$. Then $s^{r_a} \equiv s^{da} \mod \mathfrak{p}$, hence $\mod p$. Therefore $r_a \equiv da \mod p-1$. Since $\chi(g)$ is a primitive m^{th} root of unity, we can write $d = (p-1)dc/m$ with $(c, m) = 1$, so

$$r_a \equiv \frac{p-1}{m} ca \mod p-1$$

Since $0 \leq r_a \leq p-1$, we obtain $r_a = (p-1)\left\langle \frac{ca}{m} \right\rangle$, where again $\langle \cdot \rangle$ denotes the fractional part. Therefore

$$\sum_{(a,m)=1} (p-1)\left\langle \frac{ca}{m} \right\rangle \sigma_a^{-1} = (p-1)\sigma_c\theta$$

annihilates $\mathfrak{p}$ in the class group of $\mathbb{Q}(\zeta_m) : \mathfrak{p}^{(p-1)\sigma_c\theta} = (G(\chi)^{p-1})$.

Now let $\beta \in \mathbb{Z}[G]$ and suppose that $\beta\theta \in \mathbb{Z}[G]$. Let $\gamma = G(\chi)^{\sigma_c^{-1}\beta}$. Then $\gamma^{p-1} \in \mathbb{Q}(\zeta_m)$ and

$$\mathfrak{p}^{\beta\theta(p-1)} = (\gamma^{p-1})$$

We have $\mathbb{Q}(\zeta_m) \subset \mathbb{Q}(\zeta_m, \gamma) \subset \mathbb{Q}(\zeta_m, \zeta_p)$, so the extension $\mathbb{Q}(\zeta_m, \gamma)/\mathbb{Q}(\zeta_m)$ can only be ramified at primes above p, and if this extension in non-trivial it must be ramified. But γ^{p-1} is the $(p-1)^{st}$ power of an ideal, so adjoining γ can only give ramification at primes dividing $p-1$. It follows that $\mathbb{Q}(\zeta_m) = \mathbb{Q}(\zeta_m, \gamma)$, so $\gamma \in \mathbb{Q}(\zeta_m)$. Therefore, in $\mathbb{Q}(\zeta_m)$

$$\mathfrak{p}^{\beta\theta} = (\gamma)$$

We have achieved our goal. □

5. Iwasawa's Theory and The Brumer-Stark Conjecture

This section will cover very briefly a few of the developments that build on Stickelberger's theorem from the 20th and 21st centuries. Most of these developments use the language of group ring $\mathbb{Z}[G]$ and what is now called the Stickelberger ideal of this ring, which will be defined

5.1. **The Stickelberger Ideal.** Let K be the cyclotomic field $\mathbb{Q}(\zeta_m)$. Let G be the Galois group of K and let the ring $R = \mathbb{Z}[G]$ be the group ring of G as before. If we define the Stickelberger element as before $\theta = \frac{1}{m}\sum_{(a,m)=1} a\sigma_a^{-1}$ then define the Stickelberger ideal $S \subset R$ to be the set of elements β in R for which $\beta\theta \in R$. By the corollary of the last section every element in S annihilates the class group of K. The first interesting results about the Stickelberger ideal are stated below

Theorem. *Let A be the collection of elements of R with the form $\sigma_a - a$ where a is to run over a set of representatives modulo m for which $(a, m) = 1$. Then each $\sigma_a - a \in S$ and the set A forms a basis of S. In particular if $m = p$ is an odd prime then the element $(2 - \sigma_2)\theta = \sum_{a=\frac{p+1}{2}}^{p-1} \sigma_a^{-1}$ annihilates the class group of K.*

It follows that one can gain a good bit understanding of which elements are from the Stickelberger ideal. In fact one may even answer the question of what portion of the elements of $\mathbb{Z}[G]$ are in this ideal. Let j denote the automorphism of complex conjugation and for an ideal $A \subset R$ define A^- = to be the set of elements a in A for which $ja = -a$. Then Iwasawa and Sinnott computed that the index of S^- in R^- is finite and

$$[R^- : S^-] = 2^a h^-$$

where a is computed as $2^{g-2} - 1$ where g is the number of distinct primes dividing m, and h is the minus part of the class number, defined as $h^- = \frac{h}{h^+}$ where h, h^+ are the class numbers of $\mathbb{Q}(\zeta_m)$ and $\mathbb{Q}(\zeta_m + \zeta_m^{-1})$ respectively. This result was extended to arbitrary Abelian fields by Sinnott in [25]. This is some sense says that a very large portion of elements of R annihilate the class group.

5.2. **Catalan's Conjecture.** Another spectacular use of the Stickelberger ideal was by Preda Mihăilescu around the year 2003 in his proof of Catalan's Conjecture.

Theorem (Catalan's Conjecture / Mihăilescu's Theorem)**.** *The only positive integer solution to the equation*

$$x^n - y^m = 1$$

is given by $3^2 - 2^3 = 1$.

The proof which is beautifully outlined in [2] and [3] is one of the gems of explicit algebraic number theory and has at its heart an analysis of how many elements $\sum a_i \sigma_i$ there are in S with a given bound on $\sum |a_i|$. With this argument a strictly algebraic proof of Catalan's Conjecture may be given by investigating the properties of the Mihăilescu ideal. This is defined by fixing odd primes p, q and setting $G = Gal(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ where $\zeta_p = e^{2\pi i/p}$. Then for fixed positive rational integer x, the Mihăilescu ideal is given by,

$$\mathcal{I}_M = \{\gamma \in \mathbb{Z}[G] \mid (x - \zeta_p)^\gamma = (\mathbb{Q}(\zeta_p)^*)^q\}.$$

The important fact that makes an appearance in Mihăilescu's proof is that if for some integer y we have $x^p - y^q = 1$ then $(1 - \iota)S \subset \mathcal{I}_M$, where S is the Stickelberger ideal as defined above and ι denotes the complex conjugation element in G.

The appearance of the explicit use of Stickelberger elements to finally prove such an old Diophantine problem was in large part an inspiration to the author to write this historical documentation of Stickelberger's theorem.

5.3. **Brumer-Stark Conjecture.** The results of Stickelberger's theorem may be generalized to arbitrary Abelian extensions of $\mathbb{Q}$, and a wonderful conjecture of Brumer and Stark tells us what the right generalization may be to Abelian extensions of arbitrary number fields. Quite surprisingly the conjecture makes use of special values of zeta functions associated to the field extension to generate the Stickelberger element for that field extension. We describe the generalized Stickelberger element and state the Brumer-Stark conjecture.

Let K/k be a finite Abelian extension of number fields, let $G = Gal(K/k)$ be its Galois group, and let S be a finite set of places of k containing the Archimedian places and the places that ramify in K/k, and containing at least two places. Linearly extend each character $\chi : G \to \mathbb{C}^*$ to a homomorphism of the algebra $\mathbb{C}[G]$ so that $\chi : \mathbb{C}[G] \to \mathbb{C}^*$. Following [27] we then define $\theta_{S,K/k}$ as the unique element of $\mathbb{C}[G]$ such that one has, for all characters χ of G, the equality

$$\chi(\theta_{S,K/k}) = L_S(0, \chi^{-1}, K/k)$$

where $L_S(s, \chi, K/k)$ designates the Artin L-function without its Euler factors relative to places $\mathfrak{p} \in S$ which is given by

$$L_S(s, \chi, K/k) = \prod_{\mathfrak{p} \notin S} \left(1 - \frac{\chi(\mathfrak{Frob}(\mathfrak{p}))}{N(\mathfrak{p})^{-s}}\right)^{-1}.$$

Here $\mathfrak{Frob}(\mathfrak{p})$ denotes the unique Frobenius element of the Galois group G corresponding to the prime $\mathfrak{p}$. More explicitly we may write

$$\theta_{S,K/k} = \sum_{\sigma \in G} \zeta_{S,K/k}(\sigma, 0)\sigma^{-1}$$

where $\zeta_{S,K/k}(\sigma, s)$ denotes the partial zeta function constructed from the primes ideals not in S. This zeta function is given by

$$\zeta_{S,K/k}(\sigma, s) = \sum_{\substack{\mathfrak{Frob}(\mathfrak{a})=\sigma \\ (\mathfrak{a},S)=1}} \frac{1}{N(\mathfrak{a})^s}$$

where $\mathfrak{Frob}(\mathfrak{a})$ is the product of the Frobenius elements of the primes the primes occurring in the prime factorization of $\mathfrak{a}$, i.e. the Artin symbol of the ideal $\mathfrak{a}$. Also $(\mathfrak{a}, S) = 1$ means that none of the primes dividing $\mathfrak{a}$ are in S.

As an example we fix some $m \equiv 0, 1, 3 \mod 4$ and take $k = \mathbb{Q}$, $K = \mathbb{Q}(\zeta_m)$ where $\zeta_m = e^{2\pi i/m}$. Then G is naturally isomorphic to $(\mathbb{Z}/(m))^*$ where $\sigma_a \in G$ acts via $\sigma_a(\zeta_m) = \zeta_m^a$. We let S be the infinite place in $\mathbb{Q}$ and the prime ideals that divide m. By our definition, for $a \not\equiv 0 \mod m$,

$$\zeta_{S,K/k}(\sigma_a, s) = \sum_{\substack{n=1 \\ n \equiv a \bmod m}}^{\infty} \frac{1}{n^s}$$

and it may be computed (*cf.* [14]) that $\zeta_{S,K/k}(\sigma_a, 0) = \frac{1}{2} - \frac{a}{m}$ so that

$$\theta_{S,K/k} = \sum_{\substack{a=1 \\ a \not\equiv \bmod m}}^{\infty} \left(\frac{1}{2} - \frac{a}{m}\right) \sigma_a^{-1}$$

which we can see is intimately related to the regular stickelberger element. An amazing property of this new element is that for general extensions K/k we have the two facts

(1) $\theta_{S,K/k} \in \mathbb{Q}[G]$
(2) $w\theta_{S,K/k} \in \mathbb{Z}[G]$ where w is the order of the group of roots of unity in K.

Along with these facts we have the following conjecture which by our previous observation incorporates Stickelberger's theorem.

Conjecture (Brumer-Stark)**.** *Every ideal $\mathfrak{a}$ of K has the following property:*

There exists an element $\alpha \in K$ satisfying $|\alpha|_v = 1$ for every Archimedian place v of K such that $\mathfrak{a}^{w\theta_{S,K/k}} = (\alpha)$ and such that the extension $K(\sqrt[w]{\alpha})$ is an Abelian.

6. Conclusions

Throughout this account we have found inspiringly many applications of the relationship that Gauss discovered between cyclotomy and number theory. The theory is strikingly beautiful for how much milage may be gained by direct generalizations, of the integers for one, and for the the Gauss sums in general. It has also been fruitful to see the successful use of notation to direct the generalization process, it was Eisenstein who published the most clear proofs of the smaller reciprocity laws, and it is his name now attached to the general law in Section 3. Though Stickelberger's theorem gives an amazing amount of information about the structure of a very complicated object, the proof itself is **nothing more than a generalization** of the proof that Gauss finally gave for quadratic reciprocity. It is thus with a curious and open mind that the successful number theorist must,

"Think deeply of simple things."

References

[1] É. Bezout. Sur la résolution générale des équations de tous les degrés. *Histoire de l' Academie royale des sciences, partie Memoires*, pp. 533-552.
[2] Y. Bilu. Catalan's conjecture (after Mihăilescu) *Astérisque* No. 294 (2004), vii, pp. 1–26.
[3] Y. Bilu. Catalan without logarithmic forms (after Bugeaud, Hanrot and Mihăilescu) *J. Théor. Nombres Bordeaux* 17 (2005), no. 1, pp. 69–85.
[4] G. Eisenstein. Beiträge zur Kreistheilung. (1844), In: *Mathematische Werke,* Band I, pp. 45–54. New York: Chelsea, 1975.
[5] G. Eisenstein. Beweis des Reciprocitätssatzes für die kubischen Reste in der Theorie der aus dritten Wurzeln der Einheit zusammengesetzten complexen Zahlen. (1844) In: *Mathematische Werke,* Band I, pp. 59–80. New York: Chelsea, 1975.
[6] G. Eisenstein. Beweis der allgemeinsten Reciprocittsgesetze zwischen reelen und complexen Zahlen. In: *Mathematische Werke,* Band II, pp. 189–198. New York: Chelsea, 1975.
[7] L. Euler. De extractione radicum ex quantitatibus irrationalibus, Opera Omnia I.6, pp. 31–77.
[8] L. Euler. Observationes circa residua ex divisione potestatum relicta, *Opera Omnia* 1-2, pp. 493-518.

[9] C. F. Gauss. *Disquisitiones Arithmeticae*. Transl. by A. A. Clark. New Haven, Conn.: Yale University Press, 1966.

[10] C. F. Gauss. Theory der biquadratischen Reste I, II. In *Arithmetische Untersuchen.* New York: Chelsea, 1965.

[11] C. F. Gauss. Summierung gewisser Reihen von besonderer Art. In *Arithmetische Untersuchen.* New York: Chelsea, 1965.

[12] C. F. Gauss. Neue Beweise und Erweiterungen des Fundamentalsatzes in der Lehre von den quadratischen Resten. In *Arithmetische Untersuchen.* New York: Chelsea, 1965.

[13] D. Hilbert. *The Theory of Algebraic Number Fields* , known as *die Zahlbericht*, Transl. by I. Adamson, Springer 1998.

[14] K. Ireland & M. Rosen. *A Classical Introduction to Modern Number Theory*. Springer-Verlag New York, 1990.

[15] C. Jacobi. Briefe an Gauss.*Gesammelte Werke*, vol. VII, pp 392 – 400.

[16] C. Jacobi. Über die Kreistheilung und ihre anwendung auf die zahlentheorie. *Gasammelte Werke*, vol. VI, pp. 254–264.

[17] E. Kummer. De numeris complexis qui radicibus unitatis et numeris integris realibus constant. (1847) In: *Collected Papers,* Volume I, pp. 165–192.

[18] E. Kummer. Uber die Divisoren gewisser Formen der Zahlen, welche aus der Theorie der Kreistheilung entstehen. (1846) Appears in: *Collected Papers,* Volume I, pp. 193–202.

[19] E. Kummer. Zur Theorie der complexen Zahlen. (1847) In: *Collected Papers,* Volume I, pp. 203–210.

[20] E. Kummer. Uber die Zerlegung der aus Wurzeln der Einheit gebildeten complexen Zahlen in ihre Primfactoren. (1847) In: *Collected Papers,* Volume I, pp. 211–251.

[21] J. L. Lagrange. Réflexions sur la résolution algébrique des équations. In: *Oeuvres*, Vol. 3, Gauthier-Villars, Paris, 1869, pp. 205–421.

[22] J. L. Lagrange. Traité de la résolution des équations numériques de tous les degrés, avec des notes. 4. éd. In: *Oeuvres*, Vol. 8, Gauthier-Villars, Paris, 1869.

[23] F. Lemmermeyer. *Reciprocity Laws: From Euler to Eisenstein*. Springer-Verlag Berlin Heidelberg, 2000.

[24] W. Scharlau & H. Opolka. *From Fermat to Minkowski* Springer-Verlag New York, 1985.

[25] W. Sinnott. On the Stickelberger ideal and the circular units of a cyclotomic field. *Ann. of Math.* (2) 108 (1978), no. 1, pp. 107–134.

[26] L. Stickelberger. Ueber eine Verallgemeinerung der Kreistheilung. *Math. Ann.,* **37** (1890), pp. 312–367.

[27] J. Tate. Brumer–Stark–Stickelberger. *Séminaire de Théorie des Nombres*, Univ. Bordeaux I Talence, (1980-81), exposée no. 24.

[28] J. P. Tignol. *Galois' Theory of Algebraic Equations*, Institut de Mathématique Pure et Appliquée, UCL, Louvainla-Neuve, Belgium, 1988.

[29] A. T. Vandermonde. Mémoire sur la résolution des équations, *Histoire de l'Acad. Royale des Sciences (avec les Mémoires de Math. & de Phys. pour la même année, tirés des registres de cette Acad.)* (1771), pp. 365–416.

[30] B. L. van der Wearden. *Moderne Algebra* Springer-Verlag, Berlin-Gttingen-Heidelberg, 1950.

[31] L. Washingon. Stickelbergers theorem for cyclotomic fields, in the spirit of Kummer and Thaine. *Théorie des nombres* (Quebec, PQ, 1987), pp. 990–993, de Gruyter, Berlin, 1989.