Cyclotomic Fields and Related Topics

Proceedings of the Summer School (June 7-30, 1999)

Organised by
Department of Mathematics, University of Pune and Bhaskaracharya Pratishthana, Pune

S.D. Adhikari • S.A. Katre • Dinesh Thakur
Editors

Abelian Kummer Theory
Quadratic and Cyclotomi[illegible]
Absolute Values and Com[illegible]
The Early Reciprocity Law[illegible]
Gauss-Jacobi Sums and [illegible]
An Introduction to L-function[illegible]
On the Theorem of Hasse-Ar[illegible]
The Kronecker-Weber Theor[illegible]
Fermat's Last Theorem for Re[illegible]
Main Conjecture of Iwasawa [illegible]
Elliptic Curves, Serre's Conjecture [illegible] Theorem
Notes on Ribet's Converse to Herbrand
Vandiver's Conjecture via K-theory

Bhaskaracharya Pratishthana, Pune

Cyclotomic Fields And Related Topics

Proceedings of the Summer School (June 7-30, 1999)

Organized by

Department of Mathematics, University of Pune

Bhaskaracharya Pratishthana, Pune

S. D. ADHIKARI S. A. KATRE DINESH THAKUR

Editors

BHASKARACHARYA PRATISHTHANA, PUNE

Published by:

Bhaskaracharya Pratishthana
56/14, Erandavane, Damle Path
Off Law College Road
Pune - 411 004, INDIA

Editors:

S. D. Adhikari
Mehta Research Institute
Chhatnag Road, Jhusi
Allahabad 211 019, India

S. A. Katre
Dept. of Mathematics
University of Pune
Pune-411 007, India

Dinesh Thakur
Dept. of Mathematics
University of Arizona
Tucson, AZ 85721, USA

Mathematics Subject Classification (1991) : 11Rxx, 11-01

Printed at :
Mudra, 383 Narayan Peth, Pune 411030

Price: (in India) Rs. 375/- (abroad) USD 30/-

PREFACE

Having its origins in the monumental work of Gauss (Disquisitiones Arithmeticae, 1801), the subject of Cyclotomic Fields took shape through the works of a number of mathematicians like Eisenstein, Kummer, Kronecker, Stickelberger, Hilbert, Furtwängler, Hensel and Hasse. Following the work of Iwasawa, the last fifty years have seen a substantial enrichment of the subject through the works of many researchers.

This volume represents the proceedings of the Summer School on Cyclotomic Fields held in Pune from 7th to 30th June, 1999. The Summer School, which was jointly organized by the Department of Mathematics, University of Pune and the Bhaskaracharya Pratishthana, Pune, received generous support from National Board for Higher Mathematics. In addition, support was received from the U. G. C. Special Assistance Programme in the Department of Mathematics, Pune University and also from funds of the Bhaskaracharya Pratishthana.

Starting with introductory topics on algebraic number fields (in particular quadratic and cyclotomic fields), the Summer School covered a number of important topics such as power reciprocity laws, Fermat's last theorem for regular primes, L-functions, Kronecker-Weber Theorem on abelian number fields, Stickelberger's theorem and its applications, Herbrand-Ribet theorem, Vandiver's conjecture and Iwasawa theory. In all, three different proofs of Kronecker-Weber theorem and two proofs of Stickelberger's theorem were discussed during the Summer School. A number of special topics covered include Thaine's theorem, Ramachandra units, a 'universal' torsor for a finite group and an outline of the proof of the fact that Serre's conjecture implies Fermat's Last Theorem.

We are thankful to the Bhaskaracharya Pratishthana for agreeing to publish these proceedings and deeply appreciate their co-operation. We also thank the Mehta Research Institute for extending its facilities for the preparation of this volume.

We sincerely hope that these proceedings will encourage students to make a more detailed study of this subject from text books and original research papers.

August 2000

S. D. Adhikari
S. A. Katre
Dinesh Thakur

Time Table of the Summer School

7th June - 12th June 1999

Time	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
9.30-10.45	D. Thakur	B. Sury	B. Sury	B. Sury	B. Sury	B. Sury
11.15-12.30	N.Nitsure	M.J.Narlikar	D. Thakur	S.D.Adhikari	S.A.Katre	R.Thangadurai
2.00-3.15	D. Thakur	D. Thakur	B. Sury	D. Thakur	D. Thakur	S.Venkataraman
3.35-4.50	Discussion	Discussion	Discussion	Discussion	Discussion	——

14th June - 18th June 1999

Time	Monday	Tuesday	Wednesday	Thursday	Friday
9.30-10.45	R.Raghunathan	R.Raghunathan	R.Raghunathan	R.Raghunathan	R.Raghunathan
11.15-12.30	E.Ghate	R.Sridharan	R.Sridharan	R.Balasubramanian	S.A.Katre
2.00-3.15	R.Sridharan	E.Ghate	E.Ghate	E.Ghate	E.Ghate
3.35-4.50	Discussion	R.Balasubramanian	R.Balasubramanian	Discussion	Discussion

21st June - 26th June 1999

Time	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
9.45-11.00	S.A.Katre	C.S.Rajan	C.S.Rajan	C.S.Rajan	C.S.Rajan	P.Shastri
11.30-12.45	C.S.Yogananda	C.S. Yogananda	S.D.Adhikari	S.D.Adhikari	B.Ramakrishnan	R.Sujatha
2.00-3.15	C.S.Rajan	D. Thakur	D.Thakur	D.Thakur	R.Sujatha	——
3.35-4.50	Discussion	S.V.Kanetkar	Discussion	Discussion	S. Venkataraman	——

28th June - 30th June 1999

Time	Monday	Tuesday	Wednesday
9.45-11.00	C. Khare	N. Nitsure	C. Khare
11.30-12.45	N. Nitsure	C. Khare	D. Thakur
2.00-3.15	Kirti Joshi	Kirti Joshi	——
3.35-4.50	C. Khare	Discussion	——

Organising Committee:

Patrons :

1. Prof. Arun Nigavekar, Vice Chancellor , University of Pune.
2. Prof. C. S. Inamdar, Custodian, Bhaskaracharya Pratishthana, Pune.

Members :

1. Prof. A. D. Joshi, Head, Dept. of Mathematics, Univ. of Pune, Pune.
2. Dr. H. Bhate, Dept. of Mathematics, Univ. of Pune, Pune.
3. Dr. Mrs. M. J. Narlikar, Pune.
4. Dr. Sharad V. Kanitkar, Pune.
5. Mr. D. N. Sheth, S.P.College, Pune.
6. Dr. V. V. Acharya, Fergusson College, Pune.
7. Mr. S. B. Dhavale, Trustee, Bhaskaracharya Pratishthana, Pune.
8. Mr. K. P. Jain, Trustee, Bhaskaracharya Pratishthana, Pune.

Co-ordinators :

1. Dr. Dinesh Thakur, University of Arizona, U.S.A.
2. Dr. S. A. Katre, Dept. of Mathematics, Univ. of Pune, Pune.
3. Dr. S. D. Adhikari, Mehta Research Institute, Allahabad.

Contents

Introduction

Dinesh S. Thakur

We will begin with an explanation of some of the original motivations for the study of cyclotomic fields and lay out the plan for this workshop. Many concepts just mentioned here will get explained in detail later.

In school and college, starting from (the set of) counting numbers **N**, we are led successively first to **Z**, **Q** by the need of solving equations involving simple additions and multiplications, then to **R** by looking at lengths or using limit operations, infinite sums etc.. Finally we are taught the magic of the passage from **R** to **C**, where by forcing or decreeing a solution i to one equation $x^2+1=0$, we can solve all the polynomial equations. This is the so-called Fundamental theorem of algebra.

On the other hand, if starting from **N**, we allow only algebraic operations (additions and multiplications) and hence the polynomial equations, but not the infinite or limiting processes, then we are led to (the set of) algebraic numbers $\overline{\mathbf{Q}}$. The countable (thus of measure zero) set $\overline{\mathbf{Q}}$ is much smaller than **C**.

Historically, the slow emergence of the algebraic numbers had at least three different motivating sources: (1) Basic number theory study of Diophantine equations (basically polynomial equations with integral coefficients, where we also limit ourselves to integral solutions) such as Fermat equation $x^n+y^n=z^n$, arising as a natural generalization of Pythagoras equation for the right-angled triangle sides, (2) Basic geometry study of 'cyclotomy' which means division of a circle, i.e., attempt of constructing regular polygons (by ruler and compass), (3) Study of representation of integers by quadratic forms with integral coefficients (linear forms are easy: representable numbers are the multiples of GCD of coefficients i.e, **Z** is a PID) and resulting emergence of 'reciprocity' laws.

Most of you know how to solve Pythagoras equation $x^2+y^2=z^2$ in integers, though some of you may have only seen its rational (i.e., dehomogenized) version (i.e., $a^2+b^2=1$ in rationals) as a $t=\tan(\theta/2)$ substitution in calculus which turns integrand involving rational functions of trigonometric functions of θ into one involving rational functions of t. (By school geometry, the straight line joining the point $(-1,0)$ to the point (a,b), with polar co-ordinates $(1,\theta)$, on the unit circle around origin has slope $t=\tan(\theta/2)$. Solving $b=(a+1)t$, $a^2+b^2=1$, we get $t^2/1=(1-a^2)/(a+1)^2$. Thus

$(1-t^2)/(1+t^2) = \cdots = a$ and $b = 2t/(1+t^2)$).

The usual algebra method is to reduce without loss of generality (by getting rid of common factors and by parity considerations) to $x^2 = z^2 - y^2 = (z+y)(z-y)$ where x, y, z are relatively prime, with x even and y, z odd. Then the unique factorization into primes implies that apart from the GCD, which is 2, both the factors are squares: $z+y = 2p^2$, $z-y = 2q^2$. Thus we are lead to the parametric solution $(x,y,z) = (2pq, p^2-q^2, p^2+q^2)$.

Fermat used this to show that there are no non-trivial integral solutions to $x^4+y^4 = z^4$ (I will leave out easy GCD and parity considerations): By above, a nontrivial solution $x^4+y^4 = w^2$ leads to $x^2 = 2pq$, $y^2 = p^2-q^2$ and $w = p^2+q^2$. Now $y^2+q^2 = p^2$ again gives $q = 2ab$, $y = a^2-b^2$ and $p = a^2+b^2$. Thus, $x^2 = 2pq = 4ab(a^2+b^2)$ implies by unique factorization that $a = X^2$, $b = Y^2$ and $a^2+b^2 = W^2$, thus leading to a 'smaller' solution $X^4+Y^4 = W^2$ leading to an infinite descent of positive integers, a contradiction.

Hence, the natural attempt (tried by Euler, Cauchy, Kummer etc. and successful in some special cases, but not in general) to attack the general Fermat equation $x^n+y^n = z^n$ was to try infinite descent by using factorization $x^n+y^n = z^n = \prod(x+\zeta_n^i y)$ into 'cyclotomic' integers. Here n is odd and ζ_n is a primitive n-th root of unity, e.g., $\zeta_n = e^{2\pi i/n}$. If (there is the catch) the unique factorization held for these cyclotomic integers, then apart from small GCD, each factor would be n-th power and we may try manipulations as above to get an infinite descent. Soon, once we develop these ideas a little, we will go through the proof for $n = 3$. Later, once we develop the cyclotomic machinery, we will explain Kummer's successful attempt under the condition of 'regularity', which is weaker than unique factorization condition.

As other examples, we may want to attempt solving $x^3 = y^2+1$ in integers by factoring the right side as $(y+i)(y-i)$ into 'Gaussian integers'. (Exercise: Do this, assuming unique factorization into Gaussian integers $a+ib$ ($a, b \in \mathbf{Z}$ and go through rigorous details once more after we develop Gaussian integers). Try the same for $x^2 = y^2+5 = (y+\sqrt{-5})(y-\sqrt{-5})$ and you would see it would fail: the 'integer system' $a+b\sqrt{-5}$ does not have unique factorization. Of course, this is a wrong attempt and you should have tried $5 = x^2-y^2 = (x+y)(x-y)$ instead to conclude $x = \pm 3$).

This shows that we need careful study of generalizations of concepts such as 'rational numbers', 'integers', 'primes' etc. As we see above, given a polynomial with integral coefficients, we needed to forcefully factor it by using generalized numbers: So algebraic numbers are just solutions of (non-trivial) polynomials with integral coefficients. (Easy theorem then is that $\overline{\mathbf{Q}}$: $\{$algebraic numbers$\}$ is algebraically closed, i.e., polynomial with algebraic coefficients have all roots algebraic).

What should be the algebraic integers? Those are the solutions of monic polynomials with integral coefficients: This fits with the degree one case. Another way to see why this is what we want is as follows: Want the set of algebraic integers to be closed under sums, products and conjugation (as all conjugates look the same from $\mathbf{Q}$ point of view) and hence the elementary symmetric functions of the roots, which are the coefficients of the minimal (monic) polynomial should have algebraic as well as rational and hence integral coefficients.

But, at least for the diophantine equations applications, we do not want to look at full $\overline{\mathbf{Q}}$ at once, otherwise we will have too much factorization such as $5 = (5^{1/2})^2 = (5^{1/4})^4$ and so on. So we adjoin only finitely many (which will turn out to be the same as one) algebraic numbers to $\mathbf{Q}$ at a time to get so-called number fields. Examples are $\mathbf{Q}$, $\mathbf{Q}(i)$, $\mathbf{Q}(\sqrt{-5})$, $\mathbf{Q}(\zeta_n)$ associated to equations above. If we write K for a number field, the ring of algebraic integers in it will be denoted by $\mathcal{O}_K$. We will see soon that in $\mathbf{Q}(\sqrt{-5})$, the failure of unique factorization is illustrated by $6 = 2 \times 3 = (1+\sqrt{-5})(1-\sqrt{-5})$.

Historically, there were three different (but basically the same at the end) ways to get back unique factorizations after modifying the concepts a little: 'Ideals' of Dedekind, 'ideal complex numbers' of Kummer and 'divisors' of Kronecker. We will use the ideals: The basic idea is that unique factorization fails because something is missing in the system of these generalized integers: For example, if instead of natural numbers, we just take those of the form $4n+1$, then $21 \times 21 = 9 \times 49$ is failure of unique factorization, which is restored once we add (missing) 3 and 7 in the system. When we are interested in divisibility questions, an algebraic integer enters the picture via the set of its integral multiples, so we introduce ideals which are sets of integral linear combinations of several algebraic integers (rather than just one). In other words, the set of multiples of 'something' should be closed under addition and multiplication by algebraic integers, so that ideals are such sets, i.e., $\mathcal{O}_K$ submodules of $\mathcal{O}_K$. Similarly, $\mathcal{O}_K$ submodules of K are called fractional (with the word 'fractional' being dropped sometimes, if clear from context) ideals.

We will learn basic operations on ideals and see that they have unique factorization property. But we were originally interested in algebraic numbers. So we have to study what we loose in the passage: An ideal generated by a single number is called a principal ideal and (group of non-zero fractional) ideals modulo the principal ones is called the class group. It will turn out to be finite. The numbers which are multiples of each other (i.e., they differ by 'units') give rise to the same principal ideal. The unit group will turn out to be finitely generated. So in some sense, what we loose is manageable.

So our goal in the *Dedekind domains* series in the first week is to study these notions of integers, ideals, factorizations of ideals, the structure theorems for the class group and unit group etc. This set-up is good for strong number field, function field analogies (they are thus studied together as global fields), which we will see (in the third week of the Summer School) are very useful in '*Iwasawa theory*', a theory developed by Iwasawa to study a number field (i.e., a finite extension of $\mathbf{Q}$) by studying a tower of its extensions via analytic and Galois theoretic tools.

The prototype of the *Galois theory* and cyclotomic theory was also developed by Gauss in his attempt to solve the second motivating problem: By ruler and compass we get lines (linear equations) and circles (quadratic equations) and their successive intersection points. Successive degree 2 extensions lead to degree 2^k extensions of $\mathbf{Q}$. Construction of a regular n-gon corresponds basically to $e^{2\pi i/n}$ whose minimal polynomial has degree $\phi(n)$. Also, $\phi(n) = 2^k$ easily implies $n = 2^r$ times product of distinct Fermat primes, i.e., the primes of the form $2^{2^s}+1$. So the construction is impossible for n not of this form. On the other hand to show the construction is possible for such n, we have to have good control on subfield structure of $\mathbf{Q}(\zeta_n)$ to realize it by successive ruler and compass construction. This is where Gauss developed Galois and cyclotomic theory he needed in this context. We will give some details later.

To understand how the usual integers factor into ideals in a number field, we need to understand how the usual primes factor into prime ideals in number fields. We will see that such prime decomposition laws are quite simple in cyclotomic fields, but not in general. The deeper reason behind this will turn out to be that the Galois groups of the cyclotomic fields are abelian (i.e., commutative). We will prove famous *Kronecker-Weber theorem* which says that any finite abelian extension of $\mathbf{Q}$ is contained in some $\mathbf{Q}(\zeta_n)$. So the study of cyclotomic extensions (which by definition are subfields of the basic cyclotomic extensions $\mathbf{Q}(\zeta_n)$ is the study of finite abelian extensions of $\mathbf{Q}$. Replacing the base $\mathbf{Q}$ by a number field leads to a neat generalization called *Class field theory.* We will only state and illustrate the main theorems.

We will then prove Kummer's results on *Fermat's last theorem* by using cyclotomic and *Kummer theory.* Here we already go a little beyond abelian theory. After extensive work by several mathematicians, the Fermat's last theorem was finally proved by Wiles (and Taylor) by going still further in handling number fields with non-abelian Galois groups. We will be content with showing historical continuity for motivation, central ideas, themes and some techniques.

Two important, but hard to compute, structures are class group and unit group. We will see that special values of simply defined analytic functions, called *zeta and L-functions*, reveal a lot of information on them. Concurrently with the Dedekind domains series, we will also have example oriented series on *quadratic and cyclotomic fields* to illustrate the theory and to build up an example base. We will see the importance in cyclotomic theory of index 2 and of degree 2 (quadratic) sub-extensions of the basic cyclotomic fields.

We said that we would deal with $\mathbf{Q}(\zeta_n)$ rather than $\overline{\mathbf{Q}}$ to deal with the n-th Fermat equation, but it turns out that we need to take further extensions, such as *Kummer extensions*, which are obtained by adjoining a n-th root of some number in a number field already containing ζ_n. We also motivated $\overline{\mathbf{Q}}$ by saying that in number theory we do not want analytic operations. But the truth is that we use all the tools we can get even to study rational or algebraic numbers and so we use not only $\mathbf{R}$, which is the completion of $\mathbf{Q}$ for the usual absolute values, but we study all possible *absolute values and completions*. These concepts of p-adic sizes and completions form a 'local' approach and then we also briefly study *local-global principles* such as *Hasse theorem* for quadratic forms.

Even at the school level, we study some analogies between integers and rational numbers on one hand and polynomials and rational functions on the other. These analogies are even better, more useful and deep when we only allow coefficients from a finite field to our rational functions. This, for example, forces only finitely many remainders (residue classes) when we divide, just as in the integer case. Again the basic Dedekind domains theory and zeta and L-functions can be developed. One of the main unsolved problems in number theory is Riemann hypothesis, whose function field analog, due to Artin, was proved by Weil. As we will see, Iwasawa theory got started when Iwasawa attempted to carry over the successful tools in function fields to number field case.

Apart from the size and group structure of the class group, the relations (called *Stickelberger relations*) with respect to Galois action on it also encode a lot of useful arithmetic information. They will be proved by studying the ideal factorization of the *Gauss and Jacobi sums*.

As for the third motivating question, Fermat, Euler and many other mathematicians played with the questions such as which natural numbers are of the form $x^2 + y^2$ or $x^2 + 5y^2$ etc. Factoring, as before, in $\mathbf{Q}(i)$ or $\mathbf{Q}(\sqrt{-5})$, this relates to the question of which numbers are norms (basically products of conjugates) from those number fields. Because of the multiplicativity of norms, it quickly reduces to the question of which primes are thus representable. If a prime p is $x^2 + y^2$, then $x^2 \equiv -1$ modulo p. This leads to the question when -1 is a square modulo p (for a given p, this is easy to check, but we want to characterize all such p's: they are primes of the form $4n + 1$ and 2) or more generally for which primes a given number is a quadratic residue. Playing with many examples, Euler quickly found, for example, that if p and q are primes of the form $4n+1$, then p is a square modulo q if and only if q is a square modulo p. This is an instance of quadratic reciprocity, which was generalized to any contexts such as *Power reciprocity* laws. We will see many proofs, most natural ones occuring in factorization laws. We will also give proofs of these using Gauss-Jacobi sums.

We have tried here to give a quick plan. It will make more sense as we go along and master the terminology through the lectures and more importantly, through the problem sessions and discussions.

Dinesh Thakur
Department of Mathematics
University of Arizona
Tucson, AZ 85721
USA
e-mail: thakur@math.arizona.edu

Abelian Kummer Theory

M. J. Narlikar

In this article, we study abelian extensions of exponent m when the underlying field k has characteristic coprime to m and it contains all the m-th roots of unity. When the field k has characteristic $p > 0$, we shall also discuss the abelian extensions of k of exponent p.

§1. Independence of Group Characters

Let G be a group. A character ψ of G in a field K is a homomorphism $\psi : G \to K^*$, where K^* is the set of all nonzero elements of the field K. Let μ_m be the group of m^{th} roots of unity in K.

Examples (1) $f : \mathbf{R}/\mathbf{Z} \to \mathbf{C}^*$ given by $f(x) = e^{2\pi imx}$ for some fixed integer m where $\mathbf{R}$ (resp. $\mathbf{C}$) is the field of real (resp. complex) numbers and $\mathbf{Z}$ is the ring of rational integers.

(2) For a cyclic group $G = < a >$ of order n, the character $f : G \to \mu_n$ given by $f(a) = \alpha$, for some α in μ_n.

Theorem 1.1 (Artin) Let $\psi_1, \psi_2, \cdots, \psi_n$ be distinct group characters of G in K. Then they are linearly independent over K.

Proof : Suppose $\psi_1, \psi_2, \cdots, \psi_n$ are linearly dependent over K. Then there are $a_1, \cdots, a_n$ in K, not all 0, such that $a_1\psi_1 + a_2\psi_2 + \cdots + a_n\psi_n = 0$. Without loss of generality, assume that n is the smallest natural number with all a_i's non-zero.

Since ψ_1, ψ_2 are distinct characters, there is a z in G such that $\psi_1(z) \neq \psi_2(z)$. Also, $a_1\psi_1(zx) + a_2\psi_2(zx) + \cdots + a_n\psi_n(zx) = 0$, for all $x \in G$. Hence, $a_1\psi_1(x) + a_2\frac{\psi_2(z)}{\psi_1(z)}\psi_2(x) + \cdots + a_n\frac{\psi_n(z)}{\psi_1(z)}\psi_n(x) = 0$, for any x in G.

Thus, together with $a_1\psi_1(x) + a_2\psi_2(x) + \cdots + a_n\psi_n(x) = 0$, we arrive at

$$[a_2 - a_2\frac{\psi_2(z)}{\psi_1(z)}]\psi_2(x) + [a_3 - a_3\frac{\psi_3(z)}{\psi_1(z)}]\psi_3(x) + \cdots [a_n - a_n\frac{\psi_n(z)}{\psi_1(z)}]\psi_n(x) = 0,$$

for any x in G, which contradicts the minimality of n.

Theorem 1.2 (Hilbert's theorem 90) : If K/k is a cyclic extension of degree n with its Galois group $G = < \sigma >$, then for $\beta \in K,\ N(\beta) = 1 \Leftrightarrow \exists\ \alpha$ in K such that $\beta = \alpha/\sigma\alpha$. In other words, the kernel of the norm map from K^* to K^* consists of elements of the form $\alpha/\sigma\alpha, \alpha \in K$.

Proof : The implication ($\Leftarrow$) is obvious.

We shall prove the other implication. Let $N(\beta) = 1$, for $\beta \in K$. Consider the homomorphisms $\sigma, \sigma^2, \cdots, \sigma^n =$ id of K^* into itself and apply the theorem of independence of characters (Theorem 1.1) to deduce that

$$\text{id } + \beta\sigma + \beta^{1+\sigma}\sigma^2 + \cdots + \beta^{1+\sigma+\cdots+\sigma^{n-2}}\sigma^{n-1}$$

is not identically zero. Here $\beta^\sigma = \sigma(\beta), \beta^{\sigma^2} = \sigma^2(\beta)$, etc.. Hence, there is θ in K such that

$$\alpha = \theta + \beta\theta^\sigma + \beta^{1+\sigma}\theta^{\sigma^2} + \cdots + \beta^{1+\sigma+\cdots\sigma^{n-2}}\theta^{\sigma^{n-1}} \neq 0.$$

Therefore, $\beta\alpha^\sigma = \alpha$, and hence, $\beta = \dfrac{\alpha}{\sigma(\alpha)}$.

Theorem 1.3 (Hilbert's theorem 90 in additive form). If K/k is a cyclic extension of degree n with its Galois group $G = \ < \sigma >$, then for $\beta \in K$, $\text{Tr}(\beta) = 0 \Leftrightarrow \exists\ \alpha$ in K such that $\beta = \alpha - \sigma\alpha$. In other words, the kernel of the trace map from K to K consists of elements of the form $\alpha - \sigma\alpha,\ \alpha \in K$.

Proof : The implication ($\Leftarrow$) is obvious.

For the other implication, we take $\beta \in K$ with $\text{Tr}(\beta) = 0$. By applying the theorem of independence of characters, we have θ in K such that $\text{Tr}(\theta) \neq 0$. Let

$$\alpha = \frac{1}{\text{Tr}(\theta)}\left\{\beta\theta^\sigma + (\beta + \beta^\sigma)\theta^{\sigma^2} + \cdots + (\beta + \beta^\sigma + \cdots + \beta^{\sigma^{n-2}})\theta^{\sigma^{n-1}}\right\}.$$

Then $\beta = \alpha - \sigma\alpha$.

§2. Cyclic Extensions

Theorem 2.1 Let k be a field and n an integer > 0, prime to the characteristic of k and assume that the n^{th} roots of unity are in k.
(i) Let K/k be a cyclic extension of degree n. Then there is α in K such that $K = k(\alpha)$ and α is a root of $X^n - a = 0$, for some a in k.
(ii) Conversely if $a \in k$ and α is a root of $X^n - a$, then $k(\alpha)$ is a cyclic extension of k of degree d where $d \mid n$ and α^d is in k.

Proof : (i) Let ζ be a primitive n^{th} root of unity in k and K/k be cyclic of order n. We know that $N(\zeta^{-1}) = \zeta^{-n} = 1$. Hence by Hilbert's theorem 90, there is α in K such that $\sigma\alpha = \zeta\alpha$. Then $\sigma^2\alpha = \zeta^2\alpha$ and so on. Also $\alpha, \sigma\alpha, \cdots, \sigma^{n-1}\alpha$ are all distinct. Hence, $[k(\alpha) : k] \geq n$. Since $\alpha \in K$, we get $k(\alpha) = K$. Also, $\sigma(\alpha^n) = \zeta^n\alpha^n = \alpha^n$. Thus, $\alpha^n \in k$ and we take $a = \alpha^n$.

(ii) If $a \in k$ and α is a root of $X^n - a$, then $\alpha\zeta^j$ is a root for each j and all the roots of $X^n - a$ are in $k(\alpha)$. Hence, $k(\alpha) = K$ is a cyclic extension of degree d which divides n.

If σ is an automorphism of K, then $\sigma\alpha = \omega_\sigma\alpha$, where ω_σ is an n^{th} root of unity. If ω_σ is not a primitive n^{th} root of unity, then $\omega : G \to \mu_n$ is injective, not surjective and G is isomorphic to a subgroup (cyclic) of μ_n.

Theorem 2.2 Let k be a field of characteristic p.
(i) Let K/k be a cyclic extension of degree p. Then there is α in K such that $K = k(\alpha)$ and α is a root of $X^p - X - a = 0$, for some a in k.
(ii) Conversely, for $a \in k$, if the polynomial $f(X) = X^p - X - a$ has a root in k then all the roots are in k; or else it is irreducible. In the latter case, $k(\alpha)$ is cyclic of degree p over k.

Proof : (i) Let K/k be cyclic of degree p. We know that $\mathrm{Tr}_k^K(-1) = 0$. Hence, by Hilbert's theorem 90, there is α in K such that $-1 = \alpha - \sigma\alpha$, that is, $\sigma\alpha = \alpha + 1$. Hence, $\sigma^j\alpha = \alpha + j$ and all $\sigma\alpha, \sigma^2\alpha, \cdots, \sigma^n\alpha$ are distinct. Thus $K = k(\alpha)$. Now, $\sigma(\alpha^p - \alpha) = (\sigma\alpha)^p - (\sigma\alpha) = (\alpha+1)^p - (\alpha+1) = \alpha^p - \alpha$. Hence, $\alpha^p - \alpha \in k$.

(ii) If α is a root of f, then $\alpha + j$ is a root for $1 \leq j \leq p$. Hence, the first part of (ii). Assume that there is no root in k. Then f is irreducible. Otherwise, $f(X) = g(X)h(X)$ where g and h have degrees strictly less than p. Since $f(X) = \prod_{j=1}^{p}(X - \alpha - j)$ for any root α, $g(X)$ is a product of some $(X - \alpha - j)$. Let $d = \deg g$. The coefficient of X^{d-1} is the sum of terms $-(\alpha + j)$ over d of the integers. Hence it is $-d\alpha + m$ for some integer m. Hence $d\alpha \in k$, and as $d \neq 0, \alpha \in k$. Contradiction.

Since $f(X)$ has no multiple roots, $k(\alpha)$ is Galois and moreover cyclic with $\sigma : \alpha \to \alpha + 1$ as a generator of the Galois group.

§3. The Duality Theorem

Let A be an abelian group of exponent m (i.e. $a^m = 1$ for any $a \in A$). For the cyclic group $\mathbf{Z}_m$ of order m, let $\hat{A} = \mathrm{Hom}(A, \mathbf{Z}_m)$. Then $\hat{A}$ is called the dual of A. If $f : A \to B$ is a homomorphism between two groups of exponent m, then we have a natural homomorphism $\hat{f} : \hat{B} \to \hat{A}$ such that $(\widehat{f \circ g}) = \hat{g} \circ \hat{f}$.

Theorem 3.1 If A is a finite abelian group and it is a direct product $A = B \times C$, then $\hat{A} \cong \hat{B} \times \hat{C}$. Any finite abelian group is isomorphic to its own dual.

Proof : We have

$$B \xleftarrow{f} B \times C \xrightarrow{g} C$$

where f (leftarrow) and g (rightarrow) are the projections onto the first and the second components. Then $\psi_1 \in \hat{B}$ and $\psi_2 \in \hat{C}$ generate (ψ_1, ψ_2) in $(\widehat{B \times C})$ by

$$(\psi_1, \psi_2)(x, y) = \psi_1(x) + \psi_2(y).$$

Thus

$$(\widehat{f, g}) : \hat{B} \times \hat{C} \to (\widehat{B \times C})$$

is a map which has inverse $\phi \to (\phi_1, \phi_2)$ as

$$\phi(x, y) = \phi_1(x, e) + \phi_2(e, y),$$

where e is the identity element.

We know that any finite abelian group is a direct product of cyclic groups. Hence it is enough to show that a cyclic group is isomorphic to its own dual.

Let $A = \ < a > \ $ be of order n. Then any $f \in \hat{A}$ is determined by $f(a)$ and $f(a)$ can take precisely n different values. Let $t \in \mathbf{Z}_n$ be a primitive element in $\mathbf{Z}_n$. Then ψ defined by $\psi(a) = t$ is seen to generate $\hat{A} =< \psi >$. Thus, the result follows.

If C, A and A' are abelian groups, then a map $F : A \times A' \to C$ is called bilinear if $F(a, a')$ is linear in each component, i.e., $F(a_1 a_2, a') = F(a_1, a') + F(a_2, a')$ (the group operation on C is written additively) and $F(a, a'_1 a'_2) = F(a, a'_1) + F(a, a'_2)$. The kernel of F on the right is $\{x' \in A' \mid F(a, x') = 0 \ \forall \ a \in A\}$ and the kernel of F on the left is $\{x \in A \mid F(x, a') = 0, \ \forall \ a' \in A\}$.

There are natural injections

$$0 \to A'/B' \to \mathrm{Hom}(A/B, C) \qquad \cdots \quad I$$

and

$$0 \to A/B \to \mathrm{Hom}(A'/B', C) \qquad \cdots \quad II,$$

where B' is the kernel on the right and B is the kernel on the left. If C is cyclic of order m, then A'/B' and A/B have exponent m.

Theorem 3.2 Let $F : A \times A' \to C$, be a bilinear map of two abelian groups A and A' into a cyclic group C. With the same notation as above, A'/B' is finite if and only if A/B is finite and in that case, $A'/B' \cong (\widehat{A/B})$.

Proof : The sequences I and II give the result immediately. We have to note that A/B is finite implies that Hom $(A/B, C)$ and hence A'/B' are finite. The last part follows from the Theorem 3.1 above.

§4. Abelian Extensions

Theorem 4.1 Let k be a field and $m > 0$ an integer coprime to the characteristic of k and assume that all the m^{th} roots of unity are in k. Let B be a subgroup of k^* such that $k^{*m} \subset B$ and let $K_B = k(B^{1/m})$. Then K_B is Galois and abelian of exponent m. Let $G = \mathrm{Gal}(K_B/k)$.

Then we have a bilinear map $< \ , \ >: G \times B \to \mu_m$ as described below: $\sigma \in G,\ \ a \in B$ and $\alpha^m = a \Rightarrow \ \ < \sigma, a > \ \ = \sigma\alpha/\alpha$. The kernel on the left is 1 and the kernel on the right is k^{*m}.

The extension K_B/k is finite if and only if $[B : k^{*m}]$ is finite. In that case,

$$B/k^{*m} \cong \hat{G} = \mathrm{Hom}(G, \mu_m).$$

In particular, $[K_B : k] = [B : k^{*m}]$.

Proof : We have K_B Galois as $X^m - a$ splits completely in K_B for each $a \in B$. It can be checked that $< \sigma, a >= \sigma\alpha/\alpha$ is independent of the m-th root α of a, and the properties mentioned above are easily verified. Now, in the case of the kernel on the right, if $< \sigma, a > \ \ = 1$ for all σ in G, consider the field $k(a^{1/m})$. If $a^{1/m}$ is not in k, then there is an automorphism τ of $k(a^{1/m})$ over k which is not identity and it has an extension to K_B. Call the extension $\bar{\tau}$. Check that $< \bar{\tau}, a >\neq 1$.

In the duality theorem, let $A = G,\ \ A' = B$. Then we get, the injections

$$0 \to G \to \mathrm{Hom}(B/k^{*m}, \mu_m) \text{ and } \ 0 \to B/k^{*m} \to \mathrm{Hom}(G, \mu_m).$$

Thus the result follows.

Theorem 4.2 In the notation of Theorem 4.1, the map $B \to K_B$ gives a bijection of the subgroups of k^* containing k^{*m} and the abelian extensions of k of exponent m.

Proof : Let B_1, B_2 be subgroups of k^* as above and $B_1 \subset B_2$. Then $B_1^{1/m} \subset B_2^{1/m}$ so that $K_{B_1} \subset K_{B_2}$. Conversely, if $K_{B_1} \subset K_{B_2}$, then let $b \in B_1$, we shall prove that $b \in B_2$. Since $k(b^{1/m}) \subset k(B_2^{1/m})$ and $b^{1/m}$ is in some finitely generated subextension of K_{B_2}, we may assume that B_2/k^{*m} is finitely generated. Let $B_3 = \ \ < B_2, b >$. Then $k(B_2^{1/m}) = k(B_3^{1/m})$ and from Theorem 4.1 above, $[K_{B_2} : k] = [B_2 : k^{*m}] = [K_{B_3} : k] = [B_3 : k^{*m}]$. Hence, $B_2 = B_3$.

Now, if K is an abelian extension of k of exponent m, we have $\sigma^m = 1$ for any σ in G. Any finite subextenstion is a compositum of cyclic extensions. By Theorem 2.1, each cyclic extension of exponent m is obtained by adjoining an m-th root. Hence, K is obtained by adjoining a family of m^{th} roots of

$\{b_j\}_{j\in J}, b_j \in k^*$. If B is the subgroup of k^* generated by $\{b_j\}$ and k^{*m}, then $k(B^{1/m}) = K_B = K$. If $b' = ba^m$, for $a, b \in k$, then $k(b'^{1/m}) = k(b^{1/m})$.

If k has characteristic p and the operator P on k is defined as $P(x) = x^p - x$, then P is an additive homomorphism of k into itself. $P(k)$ now plays the role of k^{*m} in the last theorem. A root of the polynomial $x^p - x - a$ for $a \in k$ will be denoted by $P^{-1}(a)$. If B is an additive subgroup of k which contains $P(k)$, let $K_B = k(P^{-1}(B))$ be the field obtained by adjoining $P^{-1}(a)$ to k for all a in B.

Theorem 4.3 Let k be a field of characteristic p. The map $B \to k(P^{-1}(B))$ is a bijection between subgroups of k containing $P(k)$ and abelian extensions of k of exponent p. Let $K = K_B = k(P^{-1}(B))$ and G be its Galois group. For $\sigma \in G$ and $b \in B$, let $< \sigma, a > \ = \sigma\alpha - \alpha$ with $P(\alpha) = a$. Then, there is a bilinear map $C \times B \to \mathbf{Z}/p\mathbf{Z}$ given by $< \sigma, a > \ = \sigma\alpha - \alpha$, and its kernel on the left is 1 and the kernel on the right is $P(k)$. The extension K_B/k is finite if and only if $[B : P(k)]$ is finite and in that case $[K_B : k] = [B : P(k)]$.

Proof : Very similar to that of Theorem 4.2 above. We need to use Theorem 2.2 above and note $< \sigma, a >$ is a rational integer. Also, $\sigma\alpha - \alpha = 0$, for all α with $\alpha^p - \alpha = a$, implies $\alpha \in k$ and $a \in P(k)$.

REFERENCE

1. Serge Lang, *Algebra*, (Ch. VI), 3rd Ed. (Addison-Wesley), 1994.

M. J. Narlikar
1, Akashganga, IUCAA Housing
Ganeshkhind, Pune - 411007, India

Introduction to Number Fields

B.Sury

1. Integral extensions

Definition. An element x of a ring B is *integral* over a subring A if it satisfies a monic polynomial with coefficients from A. One says *B is integral over A* if all elements of B are so.

Examples. For any n, the n-th roots of unity are integral elements of $\mathbf{C}$ over $\mathbf{Z}$. The two square roots of $1/2$ are *not* integral over $\mathbf{Z}$.

Proposition. *For rings $A \subset B$, the following are equivalent for an element x of B:*
(a) x is integral over A.
(b) The subring $A[x]$ of B generated by A and x is finitely generated as an A-module.
(c) There exists a subring C of B such that $A[x] \subset C$ and C is finitely generated as an A-module.

Proof. The assertions $(a) \Rightarrow (b)$ and $(b) \Rightarrow (c)$ are obvious. To prove the assertion $(c) \Rightarrow (a)$, start with A-module generators $y_1, \cdots, y_n$ for C. As $x \in C$, we can write $xy_i = \sum_j a_{ij} y_j$ for certain $a_{ij} \in A$.
This can be rewritten as a matrix equation $My = 0$ where y is the column made up of the y_i's and $m_{ij} = \delta_{ij}x - a_{ij}$. Multiplying on the left by the adjoint of M, we get $dy_i = 0$ where $d = \det(M)$. As y_i's generate C, we have $dC = 0$. In particular, as C is a ring, $1 \in C$, we have $d.1 = d = 0$.

But, $d = \det(\delta_{ij}x - a_{ij})$ is a monic polynomial in x over A. This proves the proposition.

Corollary. *$A \subset B$ rings. Let $x_1, \cdots, x_n \in B$. Suppose x_1 is integral over A and x_i is integral over $A[x_1, \cdots, x_{i-1}]$ for $2 \leq i \leq n$. Then, $A[x_1, \cdots, x_n]$ is finitely generated as an A-module.*

Subcorollary. *For rings $A \subset B$, the set C of elements of B integral over A is a subring of B which is integral over A.*

Definition. In the notation above, C is referred to as the *integral closure* of A in B. If A is an integral domain, it is said to be integrally closed if it equals its integral closure in its quotient field.

Examples/Exercises. *(a) Any UFD is integrally closed.*
(b) For what d is the ring $\mathbf{Z}[\sqrt{d}]$ integrally closed?

(c) If C is integral over B and B is integral over A, then C is integral over A.
(d) If C is the integral closure of A in B, then C is integrally closed in B.
(e) If $A \subset B$ and $B \setminus A$ is closed under multiplication then A is integrally closed in B.
(f) If B is integral over A and I is a non-zero ideal, then $I \cap A$ is a non-zero ideal of A and B/I is integral over $A/(I \cap A)$.

Proposition. *(a) If B is integral over A and $S \subset A$ is a multiplicative subset, then $S^{-1}B$ is integral over $S^{-1}A$.*
(b) If C is the integral closure of A in B, and $S \subset A$ is a multiplicative subset, then $S^{-1}C$ is the integral closure of $S^{-1}A$ in $S^{-1}B$.

Proof. (a) is clear. For (b), start with any $b/s \in S^{-1}B$ which is integral over $S^{-1}A$. Write

$$\frac{b^n}{s^n} + \frac{a_1}{s_1}\frac{b^{n-1}}{s^{n-1}} + \cdots + \frac{a_n}{s_n} = 0 \text{ in } S^{-1}B.$$

This means that there exists $t \in S$ such that $t((s_1 \cdots s_n b)^n + a_1(s_1 \cdots s_n b)^{n-1} s s_2 \cdots s_n + \cdots + a_n s^n s_1^n \cdots s_{n-1}^n s_n^{n-1}) = 0$ in B. Multiply by t^{n-1} to conclude that $ts_1 \cdots s_n b \in C$. This proves the proposition.

Proposition. *Let A be an integral domain. Then, the following are equivalent:*
(a) A is integrally closed.
(b) For each prime ideal P, the local ring A_P is integrally closed.
(c) For each maximal ideal M, A_M is integrally closed.

Proof. $(b) \Rightarrow (c)$ is evident. The implication $(a) \Rightarrow (b)$ is immediate from the above proposition. Finally, we prove $(c) \Rightarrow (a)$. Since all the A_M's are contained in the quotient field of A, it suffices to show that $A = \cap_M A_M$. To prove this latter statement, let us call $B = \cap_M A_M$. As $A \subset B$ is a subring with the property that $A_M \subset S^{-1}B \subset A_M$ where $S = A \setminus M$, we get $A_M = S^{-1}B$. Therefore, viewing B/A as an A-module, we have $S^{-1}(B/A) = 0$. Now, if $0 \neq b \in B/A$, then look at the ideal $I = Ann(b) := \{a \in A : ab = 0 \in B/A\}$. As $b \neq 0$, $1 \notin I$. Let $M \supset I$ be a maximal ideal of A. As $S_M^{-1}(B/A) = 0$, the image of b is zero; in other words, there exists $s \in S_M$ with $sb = 0$ in B/A. But, then $s \in I$ by the very definition of I. This is a contradiction to the assumption that $M \supset I$. The proof is complete.

Lemma. *Let $A \subset B$ be integral domains such that B is integral over A. Then, B is a field if, and only if, A is a field.*

Proof. If A is a field, consider any $0 \neq b \in B$. Writing $b^n + a_1 b^{n-1} + \cdots + a_n = 0$ with $a_n \neq 0$, we have $-a_n^{-1}(b^{n-1} + a_1 b^{n-2} + \cdots + a_{n-1}) = b^{-1}$. Conversely, let B be a field. Let $0 \neq a \in A$. As $a^{-1} \in B$, we may write $a^{-n} + a_1 a^{-(n-1)} + \cdots + a_n = 0$ for some $a_i \in A$. Multiplying by a^{n-1}, we get $a^{-1} \in A$.

Corollary. *Let B be integral over A. Suppose that $Q \subset B$ is a prime ideal. Then, $Q \cap A$ is a prime ideal of A which is maximal if, and only if, Q is maximal. Moreover, if $Q_0 \supset Q$ is a prime ideal of B such that $Q_0 \cap A = Q \cap A$, then $Q = Q_0$.*

Proof. The first statement is a restatement of the lemma modulo the exercise (f) above. To prove the final assertion, write $P = Q \cap A$ and $S = A \setminus P$. Then, $S^{-1}B$ is integral over $S^{-1}A$. Now, $N := S^{-1}Q \subset N_0 := S^{-1}Q_0$ are prime ideals such that $M \subset N \cap S^{-1}A \subset N_0 \cap S^{-1}A \subset S^{-1}A$. As $S \cap Q_0$ is empty, $N_0 = S^{-1}Q_0 \neq S^{-1}B$ and so $N_0 \cap S^{-1}A = S^{-1}A$. As the ring $S^{-1}A$ is a local ring with the unique maximal ideal $S^{-1}P$, we must have $M = N \cap S^{-1}A = N_0 \cap S^{-1}A$. As $S^{-1}B$ is integral over $S^{-1}A$, the prime ideals N and N_0 must be maximal as M is. But, $N \subset N_0$ so that $N = N_0$ and we get $Q = Q_0$. The proof is complete.

Going-up theorem. *Let B be integral over A. Then,*
(a) for each prime ideal P of A, there exists a prime ideal Q of B lying over P (i.e. such that $Q \cap A = P$),
(b) If $P_1 \subset P_2$ are prime ideals of A and Q_1 is a prime ideal of B lying over P_1, then there exists a prime ideal $Q_2 \supset Q_1$ of B lying over P_2.

Proof. (a) Let us localize at P i.e. let $S = A \setminus P$. Then $S^{-1}B$ is integral over $S^{-1}A$. Start with *any* maximal ideal $N \subset S^{-1}B$. Then, $N \cap S^{-1}A$ is a maximal ideal of $S^{-1}A$. Therefore, it is the unique maximal ideal $S^{-1}P$ of the local ring $S^{-1}A$. If Q is the inverse image of N in B, it is a prime ideal and must lie over P (as the composites $A \to S^{-1}A \subset S^{-1}B$ and $A \subset B \to S^{-1}B$ are equal).
(b) Write $\bar{A} = A/P_1, \bar{B} = B/Q_1$. Then, $\bar{B}$ is integral over $\bar{A}$. If $\bar{P}_2$ denotes the image of P_2 in $\bar{A}$, there is (by (a)) a prime ideal $\bar{Q}_2$ of $\bar{B}$ lying over $\bar{P}_2$. If Q_2 is the inverse image of $\bar{Q}_2$ in B, it is a prime ideal of B lying over P_2 (as the composites $A \to \bar{A} \subset \bar{B}$ and $A \subset B \to \bar{B}$ are equal).

Definition. The *dimension* of a ring A is the largest integer d for which there is a strictly increasing chain of prime ideals $P_0 \subset P_1 \subset \cdots \subset P_d$.

Examples/exercises.
(a) Any field is of dimension 0.
(b) **Z** *has dimension* 1. *In fact, the integral closure of* **Z** *in any finite field*

extension of $\mathbf{Q}$ *is of dimension* 1*; this follows from the next corollary.*
(c) The polynomial ring $K[X_1, \cdots, X_n]$ *over a field* K *has dimension* n*.*

Corollary. *If* B *is integral over* A*, then their dimensions are equal.*

2. Dedekind domains

Definition. A *Dedekind domain* is a Noetherian, integrally closed domain of dimension 1.

Remark. Sometimes one regards fields also as Dedekind domains; in that case the above definition must be refined to include dimension zero also. Note that a ring A has dimension 1 if, and only if, it is not a field and all non-zero prime ideals are maximal.

Examples/exercises.
(a) Any PID is a Dedekind domain (we shall write DD for short).
(b) $\mathbf{Z}[X]$ *is not a DD (Why?).*

Scholium. *If* L *is a finite separable extension of fields, then the 'trace form'* $Tr : L \times L \to K; (x, y) \mapsto Tr_K^L(xy)$ *is non-degenerate.*

Proposition. *Let* A *be a DD and let* L *be a finite, separable extension of the quotient field* K *of* A*. Then,* B *is a DD.*

Proof. We already know that B must have dimension 1 and must be integrally closed. To show that B is Noetherian, we prove the stronger statement that B is an A-submodule of a free A-module of rank $n = [L : K]$. If this is proved, it would follow that B is a Noetherian A-module. Any ideal of B is, in particular, an A-submodule of B and, therefore, finitely generated as an A-module (and therefore as a B-module). Thus, it suffices to show that B is an A-submodule of a free A- module of rank n. To see this, let $e_1, \cdots, e_n$ be any K-basis of L lying in B (Why is it possible to choose such a basis?). Then, if $e_1^*, \cdots, e_n^*$ is its dual basis with respect to the trace form i.e., if $Tr_K^L(e_i e_j^*) = \delta_{ij}$, then any $x \in L$ is of the form $\sum_i Tr(xe_i)e_i^*$. If $x \in B$, then all the coefficients $Tr(xe_i) \in A$ as they are integral over A. Therefore, $B \subset \sum_i Ae_i$ which is a free A-module of rank n (as e_i's are linearly independent over K). Thus, the proof is complete.

Remarks. The hypothesis of separability is not needed for the conclusion above and can be proved in this generality using the so-called Krull-Akizuki theorem. However, in the proof above, we had the intermediary assertion that B is a finitely generated A-module and this may not be true in general.

Definition. If A is an integral domain and if K denotes its quotient field, one defines a *fractional ideal* to be a non-zero A-submodule I of K such that $I \subset d^{-1}A$ for some $d \neq 0$ in A.

Examples/exercises.
(a) Each finitely generated A-submodule of K is a fractional ideal.
(b) If A is Noetherian, each fractional ideal is finitely generated as an A-module.
(c) If I, J are fractional ideals, then so are $I \cap J, I+J, IJ$. Moreover, $IJ = JI$ and $I(JK) = (IJ)K$.

Lemma. *Let A be a Noetherian, integrally closed domain, $I \neq 0$ an ideal. If $x \in K \setminus A$, then $xI \not\subset I$.*

Proof. If $x \in K$ is so that $xI \subset I$, then $x^n I \subset I$ for all n. So, $A[x]$ is an A-submodule of K which satisfies $A[x] \subset d^{-1}A$ for some $d \neq 0$ in A (in fact, any $d \neq 0$ in I). As A is Noetherian, so is $d^{-1}A$ and thus $A[x]$ is a finitely generated A-module. This means that x is integral over A i.e. $x \in A$.

Proposition. *Let A be a DD and let P be a non-zero prime (= maximal) ideal. If K denotes the quotient field of A, then the set*

$$P' := \{x \in K : xP \subset A\}$$

is a fractional ideal of A and properly contains A. Further, P' is the unique fractional ideal such that $PP' = P'P = A$.

Proof. It is trivial to see that P' is an A-module. Moreover, evidently $P' \subset d^{-1}A$ for any $d \neq 0$ in P. Thus, P' is a fractional ideal and clearly contains A. We shall show now that $A \neq P'$. For this, we make use of the following:
Claim: Every non-zero ideal of A contains a finite product of non-zero prime ideals.
The claim is proved as follows. If there are exceptions to the claim made above, consider the family of ideals which fail to contain a product as claimed. As A is Noetherian, there exists a maximal such ideal M. So, M itself cannot be prime. If $ab \in M$ with neither a nor b in M, then the ideals $M + (a)$ and $M + (b)$ contain products of prime ideals. As M is contained in their product, M contains a product of prime ideals, which contradicts our assumption. Therefore, the claim is indeed true. Now, let $a \neq 0$ be in P. Then, the ideal $(a) \supset P_1 P_2 \cdots P_n$ with n minimal possible and P_i's non-zero primes. So, $P \supset P_1 \cdots P_n$. As P is prime, we have $P \supset P_i$ for some i, say $P \supset P_1$. As P_i are maximal, we obtain $P = P_1$. Writing $I = P_2 \cdots P_n$ or A according as $n > 1$ or $n = 1$, we get $I \not\subset (a)$ by the minimality of n. Choose any $b \in I \setminus (a)$. Then, $ba^{-1} \notin A$. Now, $PI \subset (a) \Rightarrow Pb \subset (a)$ i.e., $ba^{-1} \in P'$.

Hence, we have shown that $A \neq P'$. Further, we have $P = PA \subset PP' \subset A$ so that PP' is an (actual) ideal of A containing P. It must, therefore, be either equal to P or to the unit ideal A. If $x \in P' \setminus A$, we must have (by the above lemma) $xP \not\subset P$. This means that $xP \subset P'P \setminus P$. Thus, $PP' = A$. Finally, if P_0 is any fractional ideal such that $PP_0 = P_0P = A$, then $P' = AP' = (P_0P)P' = P_0(PP') = P_0A = P_0$ which proves uniqueness also.

Notation. One uses the notation P^{-n} instead of P'^n for any n. Then, (like ideals) one has $AP^{-n} = P^{-n}$.

Theorem. *Let A be a DD. Then, any fractional ideal $I \neq A$ can be uniquely written as $I = P_1^{n_1} \cdots P_k^{n_k}$ where n_i are non-zero integers and P_i are distinct prime ideals.*

Proof. The uniqueness is easy to prove as follows.

If $P_1^{n_1} \cdots P_k^{n_k} = Q_1^{m_1} \cdots Q_r^{m_r}$, then one can shift all the negative powers on each side to the other side to obtain an equality where all powers are positive. Then, a simple induction on the sum of the exponents yields uniqueness.

We prove the existence of the prime ideal decomposition by contradiction. First, we assume that there is an (actual) ideal I which is not expressible as a product of prime ideals. By using the fact that A is Noetherien, we obtain an ideal I which is maximal with respect to this property. Of course, I is not a prime ideal. If $I \subset P$ with P maximal, then $I = AI \subset P^{-1}I \subset P^{-1}P = A$. Now, if $x \in P^{-1} \setminus A$, then $xI \not\subset I$ and so $xI \subset P^{-1}I \setminus I$. Hence $P^{-1}I$ is an (actual) ideal which contains I properly. By the choice of I, we obtain that $P^{-1}I$ must be a product of prime ideals. Therefore, clearly I itself is such a product, which manifestly contradicts the choice of I. Therefore, every ideal in A is, indeed, a product of prime ideals.

Finally, if J is any fractional ideal, there is some $d \neq 0$ in A such that dJ is an ideal of A. So, if $(d) = P_1^{a_1} \cdots P_r^{a_r}$ and $dJ = Q_1^{b_1} \cdots Q_s^{b_s}$, then $J = P_1^{-a_1} \cdots P_r^{-a_r} Q_1^{b_1} \cdots Q_s^{b_s}$. This proves the theorem.

Examples/Exercises. (a) In any DD, $P \supset P^2 \supset P^3 \cdots$ is a strictly decreasing chain.

(b) Every fractional ideal in a DD can be generated by two elements one of which can be taken to be any arbitrary element.

Hint: Enough to prove this for ideals I; in this case if $a \in I$ and if $(a) = P_1^{a_1} \cdots P_r^{a_r}$ and $I = P_1^{b_1} \cdots P_r^{b_r}$, then $a_i \geq b_i \geq 0$. Use the Chinese remainder theorem to choose an appropriate element b in I so that $I = (a, b)$.

(c) A DD which has only finitely many prime ideals is a PID.

Hint : If $P_1, \cdots P_n$ are all the prime ideals, use the Chinese remainder theorem to choose $x_i \in P_i$, $x_i \notin P_i^2$ and $x_i \equiv 1 \bmod P_j$ for $i \neq j$. Then, $P_i = (x_i)$.
(d) Use the fact that $\mathbf{Z}[\sqrt{-5}]$ is not a PID and (c) above to prove that there are infinitely many prime numbers (!)

3. Prime decomposition in extension fields

Let A be a DD with quotient field K and let L be a finite, separable extension of K. Then, we have seen that the integral closure B of A in L is again a DD. If $A = \mathbf{Z}$, then L is called an *algebraic number field* and B is called the *ring of integers of L.*

Exercises. *(a) Show that if $K \subset L$ are algebraic number fields, then the ring of integers of L is the integral closure of the ring of integers of K in L.*
(b) Find the ring of integers of the field $\mathbf{Q}(\sqrt{d})$ for any square-free integer d.

Definition. For a field extension L/K of degree n, and an n-tuple of elements $v_1, \cdots, v_n$ of L, one defines *the discriminant of the n-tuple $v_1, \cdots, v_n$* to be the element $D^L_K(v_1, \cdots, v_n) = \det(M)$ of K where $M_{ij} = Tr^L_K(v_i v_j)$. This is an important concept, and let us start with a few easy exercises to see its use.

Exercises. *Let L, K, v_i be as above.*
(a) Show that $D^L_K(v_1, \cdots, v_n) \neq 0$ if, and only if, $v_1, \cdots, v_n$ form a K-basis of L.
(b) If $K = \mathbf{Q}$ and v_i form a $\mathbf{Z}$-basis of the ring of integers (this always exists as we observed), then $D^L_K(v_1, \cdots, v_n)$ is an integer which is independent of the choice of the $\mathbf{Z}$-basis.
(c) If $\sigma_1, \cdots, \sigma_n$ are the K-embeddings of L in $\mathbf{C}$, then $D^L_K(v_1, \cdots, v_n) = \det(N)^2$ where $N_{ij} = \sigma_i(v_j)$.

Definition. The *discriminant D_K of an algebraic number field K* is the discriminant of any $\mathbf{Z}$-basis of its ring of integers. By the exercise (b) above, it is well-defined. Moreover, it is clear that if $\{v_1, \cdots, v_n\}$ are in $\mathcal{O}_K$ and satisfy $D_K = D^K_{\mathbf{Q}}(v_1, \cdots, v_n)$, then $\{v_i\}$ form an integral basis (Why?).

Exercise. *(a) For a square-free integer d, show that the discriminant of $\mathbf{Q}(\sqrt{d})$ is d or $4d$ according as whether $d \equiv 2, 3$ mod 4 or $d \equiv 1$ mod 4.*
(b) Let $K = \mathbf{Q}(\alpha)$ be an algebraic number field. Suppose the minimal

(monic) polynomial of α is $f(X) = \prod_{i=1}^{n}(X - \alpha_i)$. Then, prove that

$$D_{\mathbf{Q}}^{K}(1, \alpha, \cdots, \alpha^{n-1}) = \prod_{i<j}(\alpha_i - \alpha_j)^2 = (-1)^{n(n-1)/2} N_{\mathbf{Q}}^{K} f'(\alpha)$$

where N denotes the norm map.
(c) Use (b) to show that for any n, and $K = \mathbf{Q}(\zeta_n)$ with ζ_n a primitive n-th root of unity, one has $D_{\mathbf{Q}}^{K}(1, \zeta, \cdots, \zeta^{\phi(n)-1})$ divides $n^{\phi(n)}$.
(d) Let K be an algebraic number field and let $\alpha_1, \cdots, \alpha_n$ be a $\mathbf{Q}$-basis of K contained in $\mathcal{O}_K$, the ring of integers of K. Then,

$$\mathcal{O}_K \subset \{\sum m_i\alpha_i/d : m_i \in \mathbf{Z}, d|m_i^2\}$$

Here d stands for $D_{\mathbf{Q}}^{K}(\alpha_1, \cdots, \alpha_n)$.
Hint: Write any $\alpha \in \mathcal{O}_K$ as $\sum_i t_i\alpha_i$ with $t_i \in \mathbf{Q}$. Apply the various embeddings of K to this equation and solve the system of linear equations by Cramer's rule.
(e) If K, L have degrees m, n over $\mathbf{Q}$ and if KL has degree mn, then $\mathcal{O}_{KL} \subset \frac{1}{d}\mathcal{O}_K\mathcal{O}_L$ where d is the GCD of D_K and D_L.
Hint: Use the fact (implied by the hypothesis $[KL : \mathbf{Q}] = mn$) that each embedding of K in $\mathbf{C}$ has a unique extension as an embedding of KL which restricts to the identity on L. Then, use the same idea as for (d).

Lemma. *For any positive integer n, consider the field $K = \mathbf{Q}(\zeta)$ where $\zeta = e^{2i\pi/n}$. Then, $\mathcal{O}_K = \mathbf{Z}[\zeta]$.*

Proof. By the exercises (c) and (e) above, and the fact that Euler's phi-function is multiplicative, it suffices to prove the lemma when n is a power of a prime.

Let us use the notation disc(α) when we talk about $D_{\mathbf{Q}}^{K}(1, \alpha, \cdots, \alpha^{m-1})$ for some number field $K = \mathbf{Q}(\alpha)$ of degree m. Let $n = p^r$ and ζ be a primitive n-th root of unity. From an earlier exercise, we have disc(ζ) = disc($1 - \zeta$). Moreover, $p = \prod_{(k,p)=1}(1 - \zeta^k)$ as seen by evaluating the corresponding cyclotomic polynomial at 1. Evidently, $1 - \zeta^k$ is an associate of $1 - \zeta$ for any k coprime to p. Therefore, p divides $(1 - \zeta)^{\phi(p^r)}$ in $\mathbf{Z}[\zeta]$. Now, by an exercise above, every element of $\mathcal{O}_K$ is of the form

$$\sum_{i<\phi(p^r)} m_i(1 - \zeta)^{i-1}/d,$$

where $d = \text{disc}(\zeta)$. Note that d is a power of p. If $\mathcal{O}_K \neq \mathbf{Z}[1 - \zeta]$, then there exists an element $x \in \mathcal{O}_K$ for which not all m_i are divisible by d. If all the

m_i's are divisible by p, we can divide them all by p and proceeding this way we finally arrive at an element in $\mathcal{O}_K$ of the form $x = \sum_{i \geq j} m_i(1-\zeta)^{i-1}/p$ with $j \geq 1$ and m_j not a multiple of p. Now, we noted in the beginning of the proof that p is an associate of $(1-\zeta)^{\phi(p^r)}$ in $\mathbf{Z}[\zeta]$. This means, in particular, that $p/(1-\zeta)^j \in \mathbf{Z}[\zeta] \subset \mathcal{O}_K$. Hence, we have $xp/(1-\zeta)^j \in \mathcal{O}_K$. Hence, we get from the expression for x that $m_j/(1-\zeta) \in \mathcal{O}_K$. So, $N^K_{\mathbf{Q}}(1-\zeta)$ divides $N^K_{\mathbf{Q}}(m_j) = m_j^{\phi(p^r)}$ i.e., $p|m_j$, which is a contradiction. This proves the lemma.

Definition. Let A be a DD, K its quotient field and L a finite, separable extension. Let B denote the integral closure of A in L. For any non-zero prime ideal P of A, as B is a DD, one can write $PB = P_1^{e_1} \cdots P_g^{e_g}$ where all $e_i > 0$. The integer e_i is called the *ramification index of* P_i and sometimes denoted by $e(P_i/P)$ to make its dependence clear. P is said to be *unramified in* B if each $e_i = 1$; otherwise it is said to be ramified. P is said to be *totally ramified* if $g = 1$ and $e_i > 1$. The primes P_i lie over P and these are all the primes lying over P (*Why*?). The degree f_i (denoted by $f(P_i/P)$) of the field extension $B/P_i \supset A/P$ is evidently (*why?*) at the most equal to the degree of L over K. The finite field A/P (*why is it finite?*) is called the residue field of K at P. The field extension $B/P_i \supset A/P$ is called the residue field extension at P_i and f_i is called the *residue field degree of* P_i.

Exercises. *Answer the three why's in the above definition.*

Proposition. *Let A be a DD, K its quotient field and L a finite separable extension. Let B denote the integral closure of A in L. For a non-zero prime ideal P of A, writing $PB = P_1^{e_1} \cdots P_g^{e_g}$ we have $\sum_{i=1}^{g} e_i f_i$ where $f_i = [B/P_i : A/P]$.*

Proof. The trick is to localize at P i.e. consider $S^{-1}A$ and $S^{-1}B$ where $S = A \setminus P$. Now $S^{-1}B$ is the integral closure of $S^{-1}A$ in L, and $S^{-1}A/S^{-1}P \cong A/P$. Note also that $PS^{-1}B = Q_1^{e_1} \cdots Q_g^{e_g}$ where $Q_i = P_i S^{-1}B$ and that $S^{-1}B/Q_i \cong B/P_i$ Thus, to prove the proposition we may replace A, B by $S^{-1}A, S^{-1}B$. In this case, A, B are PIDs as they are DDs with only finitely many primes! Therefore, B which is a submodule of a free A-module is, itself, free of rank n (the rank is n as B contains a K-basis of L). Let $v_1, \cdots, v_n$ be an A-basis of B. If $\bar{v_i}$ denotes the image of v_i modulo PB, we have $B/PB = \sum_{i=1}^{n}(A/P)\bar{v_i}$. Moreover, if $\sum_{i=1}^{n} \bar{a_i}\bar{v_i} = 0$ in B/PB, then $\sum_{i=1}^{n} a_i v_i \in PB$. This forces each a_i to be in P since v_i's form a basis of B. Thus, $\bar{v_1}, \cdots, \bar{v_n}$ is a basis of the A/P-vector space B/PB. Thus, $\dim_{A/P} B/PB = n$. Let us count this same dimension in another way. By the Chinese remainder theorem, one has $B/PB = B/\prod P_i^{e_i} \cong \oplus B/P_i^{e_i}$ as rings as well as as A/P-vector spaces. We need to count the dimension of each $B/P_i^{e_i}$. Now, since $P \subset P_i$, we have $PP_i^r \subset P_i^{r+1}$ for all $r \geq 1$. Hence, P_i^r/P_i^{r+1} is an A/P-vector space. Thus, as A/P-vector spaces, we have

$$B/P_i^{e_i} \cong B/P_i \oplus P_i/P_i^2 \oplus \cdots \oplus P_i^{e_i-1}/P_i^{e_i}$$

Further, as B is a PID, one can write $P_i = (\pi_i)$. Then, for each r, the multiplication by π_i^r gives an A/P-isomorphism from B/P_i onto P_i^r/P_i^{r+1}. Hence, we have $\dim_{A/P} B/P_i^{e_i} = e_i f_i$ which gives that $n = \sum e_i f_i$.

Definition. With A, B as before, a maximal ideal P of A is said to *split completely in B* if $e_i = 1 = f_i$; so PB is a product of n distinct primes.

Examples/Exercises. *(a) Show that the e's and the f's multiply in towers.*
(b) Let p be a prime, $\zeta = e^{2i\pi/p}$ and $K = \mathbf{Q}(\zeta)$. Then, p is totally ramified in K.
Hint: Show that $p = \prod_{i=1}^{p-1}(1-\zeta^i)$ and that each $1-\zeta^i$ is a unit times $1-\zeta$.

Corollary. *Let the notations be as in the above proposition. Assume, in addition, that L/K is a Galois extension. Then, all the e_i's are equal and all the f_i's are equal. Hence $n = efg$ for some positive integers e, f, g.*

Proof. We shall show that the Galois group Gal(L/K) acts transitively on the set $\{P_1, \cdots, P_g\}$. If it does not, then there exist $i \neq j$ such that $gP_i \neq P_j$ for all $g \in$ Gal(L/K). Then, choosing by the Chinese remainder theorem, an element $b \in P_j, b \equiv 1 \bmod gP_i$ for each $g \in G$. But then the element $a = N_K^L(b) = \prod_g g(b)$ is in A on the one hand, and is in P_j on the other. As $A \cap P_j = P$, this means that $\prod_g g(b) \in P \subset P_i$ i.e. some $g(b) \in P_i$, which contradicts the choice of b. Hence, it follows that the Galois group acts transitively. Then, if $gP_i = P_j$, the observation $PB = g(PB)$ along with the uniqueness of decomposition into prime ideals in B yields $e_i = e_j$. Therefore, all the e_i's are equal. Finally, if $g(P_i) = P_j$, then g induces an A/P-isomorphism from B/P_i to B/P_j and so $f_i = f_j$. The corollary is proved.

Definitions. With notations as above, the *decomposition group* of P_i is the subgroup $D_{P_i} := \{g \in \text{Gal}(L/K) : g(P_i) = P_i\}$. The Galois group induces a natural homomorphism θ_{P_i} from D_{P_i} to Gal$((B/P_i)/(A/P))$. The kernel T_{P_i} is called the *inertia group of P_i*. If the inertia group T_{P_i} is trivial, one defines the *Frobenius element* Fr_{P_i} at P_i as the inverse image under the isomorphism θ_{P_i} of the Frobenius automorphism $t \mapsto t^{\#(A/P)}$ which generates Gal$((B/P_i)/(A/P))$.

Exercises.
(a) Show that the above homomorphism from D_{P_i} to $Gal((B/P_i)/(A/P))$ is surjective.
Hint: Use the Chinese remainder theorem.
(b) Show that the D_{P_i}'s are mutually conjugate and that $\#D_{P_i} = ef$, $\#T_{P_i} = e$ for all i.
Hint: D_{P_i} is the stabiliser at P_i for the action of $Gal(L/K)$ on the set $\{P_1, \cdots, P_g\}$.

Definition. For any algebraic number field K and a non-zero ideal I, the *norm $N(I)$ of I* is defined to be the cardinality of the finite ring $\mathcal{O}_K/I$.

Corollary. *Let K be an algebraic number field. Then,*
(a) if I, J are non-zero ideals, $N(IJ) = N(I)N(J)$.
(b) if P is a maximal ideal, $N(P) = p^f$ where p is the prime number lying below P and $f = f(P/p)$.
(c) if L/K is an extension of degree n, then for any non-zero ideal I of $\mathcal{O}_K$, $N(I\mathcal{O}_K) = N(I)^n$.
(d) if $a \neq 0$ is in $\mathcal{O}_K$, $N((a)) = \mid N_{\mathbf{Q}}^K(a) \mid$.

Examples/exercises. *Let $K = \mathbf{Q}(\sqrt{d})$ where d is a square-free integer. For any odd prime p, denote by (a/p) the Legendre symbol. Then,*
(a) if $p|d$, p is (totally) ramified i.e. $p\mathcal{O}_K = P^2$ where the prime ideal $P = (p, \sqrt{d})$,
(b) if p is odd and coprime to d, it is unramified and splits completely or remains a prime according as whether $(d/p) = 1$ or not,
(b)′ if $d = q$ is a prime $\equiv 1$ mod 4, and p is an odd prime, prove that $(q/p) = 1 \Leftrightarrow$ the polynomial $X^2 - X + \frac{1-q}{4}$ has a solution mod $p \Leftrightarrow \mathbf{Q}(\sqrt{q})$ is fixed by the Frobenius $Fr_p \Leftrightarrow (p/q) = 1$.
(c) if d is odd, 2 is ramified if $d \equiv 3$ mod 4, splits completely if $d \equiv 1$ mod 8 and remains a prime if $d \equiv 5$ mod 8.
(d) Prove the whole of quadratic reciprocity law by proving a corresponding version of (b)′ for primes $\equiv 3$ mod 4.

Remarks. The exercise above provides a nice criterion to decide when a prime splits completely in a quadratic extension. The criterion is in terms of some congruences. One of the principal aims of ramification theory (in fact, of algebraic number theory itself!) is to give a 'nice' criterion for a prime to split completely in a given extension; one often calls such a criterion to be a reciprocity law. The reason that one is interested in a criterion to decide which primes split completely is that given K, the set of primes of K which split in L determine L uniquely. The last fact mentioned is deep and the proof requires the so-called class field theory.

An interesting exercise - Why is Fermat's last theorem not trivial to prove?

(a) Let p be an odd prime and $\zeta = e^{2i\pi/p}$. Show that the element $S = \sum_{i=1}^{p-1}(i/p)\zeta^i$ of $K = \mathbf{Q}(\zeta)$ satisfies $S^2 = (-1/p)p$. Hence conclude that every quadratic extension of $\mathbf{Q}$ is contained in a cyclotomic extension.

(b) Let $K = \mathbf{Q}(\sqrt{-23}), L = \mathbf{Q}(\zeta)$ where $\zeta = e^{2i\pi/23}$. Show that $\mathcal{O}_L$ is not a PID.

Hint : $K \subset L$ by (a). Also, $2\mathcal{O}_K = P\bar{P}$ where $P = (2, \frac{1+\sqrt{-23}}{2})$ and $\bar{P} = (2, \frac{1-\sqrt{-23}}{2})$. If a prime Q in L lying over P is principal, then P^f is principal where $f = f(Q/P)$. As P is not principal and $P^3 = (\frac{-3+\sqrt{-23}}{2})$, P^f cannot be principal as f divides $[L : K]$.

Theorem (A Cyclotomic reciprocity law). *Let n be a positive integer and p be a prime not dividing n. Denote by ζ a primitive n-th root of unity. Then, p is unramified in $K = \mathbf{Q}(\zeta)$ and splits into $\phi(n)/f$ primes where f is the order of p in the unit group of $\mathbf{Z}/n$ and ϕ is Euler's phi function. In particular, p splits completely in K if, and only if, $p \equiv 1$ mod n.*

Proof. We already know that p is unramified as the minimal polynomial of ζ (indeed, its multiple $X^n - 1$ itself) has distinct roots mod p. Let P be a prime in K which lies over p. First, we observe that the powers $\zeta^i, 0 \leq i \leq n-1$ are distinct modulo P. This is a consequence of the identity $n = \prod_{i=1}^{n-1}(1-\zeta^i)$ and the observation that $n \notin P$; these imply that $1 - \zeta^i \notin P$. Now, the Frobenius element Fr_p of $\mathrm{Gal}(K/\mathbf{Q})$ satisfies $\mathrm{Fr}_p(x) \equiv x^p \bmod P$ for all $x \in \mathcal{O}_K$. But, $\mathrm{Fr}_p(\zeta)$ is obviously again an n-th root of unity. In view of the observation made above, it follows that $\mathrm{Fr}_p(\zeta) = \zeta^p$. ¿From this, it follows easily that the order $f(P/p)$ of Fr_p is just the order f of p in $(\mathbf{Z}/n)^*$.

Remarks. When K is the quotient field of a DD A, and L is a finite, separable extension of K and B the integral closure of A in L, the following theorem of Kummer provides a way to read off the decomposition of a prime ideal in terms of the decomposition of the minimal polynomial of α modulo P. Here $L = K(\alpha)$ and $\alpha \in B$ and the theorem is valid under a mild assumption.

Theorem (Kummer). *Let $A, K, L = K(\alpha), B, P, f$ be as above. Assume, in addition, that $B = A[\alpha]$. Write $\bar{f} = \bar{p}_1^{e1} \cdots \bar{p}_g^{e_g}$ where $\bar{p}_i$ are irreducible polynomials in $(A/P)[X]$ and $\bar{f}$ denotes the image of f mod P. Then,*

$$PB = P_1^{e_1} \cdots P_g^{e_g}$$

where P_i's are prime ideals and $f(P_i/P) = deg(\bar{p}_i)$. Indeed, $P_i = PB + (p_i(\alpha))$ where p_i's are arbitrary lifts of $\bar{p}_i$'s.

Before proving the theorem, let us look at its applications to see really how powerful it is.

Applications of Kummer's theorem

I. Prime decomposition in quadratic fields

As we saw earlier, if $K = \mathbf{Q}(\sqrt{d})$ with d square-free, then $\mathcal{O}_K = \mathbf{Z}[\alpha]$ where $\alpha = \sqrt{d}$ or $\frac{1+\sqrt{d}}{2}$ according as $d \equiv 2, 3 \bmod 4$ or $d \equiv 1 \bmod 4$. The minimal polynomial f is $X^2 - d$ in the first case and $X^2 - X + \frac{1-d}{4}$ in the second. If $d \equiv 2$ or $3 \bmod 4$, $f(X) = X^2 - d$ is a square modulo any prime p dividing d and also modulo 2. Thus, 2 and primes dividing d are (totally) ramified. If an odd prime p does not divide d, then f modulo p is reducible or irreducible according as whether d is a square modulo p or not. Thus, these primes, respectively, split completely and remain inert. Similarly, one can argue for the case $X^2 - X + \frac{1-d}{4}$ corresponding to $d \equiv 1 \bmod 4$.

II. Discriminant criterion for ramification

Theorem. *Suppose $K = \mathbf{Q}(\alpha)$ is an algebraic number field and assume that $\mathcal{O}_K = \mathbf{Z}[\alpha]$ for some α. Then, a prime p ramifies in K if, and only if, p divides Disc(K).*

Proof. Let $f(X) = \prod_i (X - \alpha_i)$ be the minimal polynomial of α. We have seen that disc(K) = disc$(f) = \pm\prod_{i \neq j}(\alpha_i - \alpha_j)$. By Kummer's theorem, a prime ramifies in K if, and only if, f has a multiple root modulo p. This is so if, and only if, disc$(\bar{f}) \equiv 0 \bmod p$ i.e. if, and only if, p divides disc(f). Here $\bar{f}$ denotes the reduction of f modulo p.

Proof of Kummer's theorem. Consider the ring homomorphisms

$$A[X] \to (A/P)[X] \to (A/P)[X]/(\bar{p}_i(X))$$

Call the composite map ϕ_i. Note that $(A/P)[X]/(\bar{p}_i(X)) \cong (A/P)[\alpha_i]$ for any root α_i of $\bar{p}_i$. Therefore, Ker(ϕ_i) is a maximal ideal as ϕ_i is evidently surjective. Moreover, it is clear that $P \subset$ Ker(ϕ_i) and $p_i(X) \in$ Ker(ϕ_i) for any arbitrary $p_i \in A[X]$ which maps to $\bar{p}_i$. Further, it is clear from the definition of ϕ_i that Ker(ϕ_i) is the ideal generated by P and p_i in $A[X]$. Now, by the hypothesis, $\bar{f} = \bar{p}_1^{e_1} \cdots \bar{p}_g^{e_g}$ which implies that $f \in (P, p_i) =$ Ker(ϕ_i). Therefore, ϕ_i factors through (f) to give a surjective homomorphism $\theta_i : A[X]/(f) \to (A/P)[X]/(\bar{p}_i(X))$. Note that we have assumed that $B = A[\alpha]$ which gives that $A[X]/(f) \cong B$ where X maps to α. So, we have obtained $\theta_i : B \to (A/P)[X]/(\bar{p}_i(X))$ which is surjective and has kernel Ker$(\theta_i) = PB + p_i(\alpha)B$. Thus, $P_i :== PB + p_i(\alpha)B =$ Ker(θ_i) are maximal ideals in

B. As they contain P, they lie over P. Note that $f(P_i/P) = [B/P_i : A/P] = \dim_{A/P}(A/P)[X]/(\bar{p}_i(X)) = \deg\bar{p}_i$. We shall prove now that P_i exhaust all the maximal ideals of B lying over P and have ramification indices equal to e_i.
Note first that the assumption $\bar{f} = \bar{p}_1^{e_1} \cdots \bar{p}_g^{e_g}$ gives, on comparing degrees that $\sum_i e_i f_i = \deg(f) = [L : K]$. The same thing also gives for arbitrary lifts p_i that $f - p_1^{e_1} \cdots p_g^{e_g} \in P[X]$ which, in turn gives, on evaluation at α, that $p_1(\alpha)^{e_1} \cdots p_g(\alpha)^{e_g} \in PA[\alpha] = PB$. So, if Q is any prime ideal of B lying over P, we have $p_1(\alpha)^{e_1} \cdots p_g(\alpha)^{e_g} \in PB \subset Q$. Then, $p_i(\alpha) \in Q$ for some i. But then, $P_i = PB + p_i(\alpha) \subset Q$ and, as both are maximal ideals, they must be equal.
Finally, let $PB = P_1^{d_1} \cdots P_g^{d_g}$. Then,

$$\begin{aligned} P_1^{e_1} \cdots P_g^{e_g} &= (P, p_1(\alpha))^{e_1} \cdots (P, p_g(\alpha))^{e_g} \\ &\subset PB + (p_1(\alpha)^{e_1} \cdots p_g(\alpha)^{e_g}) = PB = P_1^{d_1} \cdots P_g^{d_g}. \end{aligned}$$

Thus, $e_i \geq d_i$. As $\sum e_i f_i = [L : K] = \sum d_i f_i$, this forces $d_i = e_i$. The proof is complete.

The last application was generalised by Dedekind to the situation when the base field is the quotient field K of any DD A and when the integral closure B of A in a finite, separable extension L may not satisfy the condition $B = A[\alpha]$ for any α.

The following example shows that the condition $B = A[\alpha]$ may not hold for any α.

Example. *Let K denote the unique subfield K of $L = \mathbf{Q}(\zeta_{31})$ of degree 6 over $\mathbf{Q}$. Then, $\mathcal{O}_K \neq \mathbf{Z}[\alpha]$ for any α.*

Reason: In general, if E/F is a finite Galois extension, and D is the decomposition group at some prime Q of E, then, $P = Q \cap \mathcal{O}_F$ splits completely in E^D (Why?).
Returning to our situation, look at the prime 2 which is unramified. As the order of 2 modulo 31 is 5, 2 splits in $\mathcal{O}_L$ into $\phi(31)/5 = 6$ primes. Therefore, the decomposition group D at any prime of L lying over 2 has order 5. As $\mathrm{Gal}(L/\mathbf{Q})$ is cyclic, it has a unique subgroup of order 5 (indeed, of order any divisor of 30). Thus the fixed field L^D is of degree 6 over $\mathbf{Q}$ and must be K. By the observation made in the beginning, it follows that 2 splits completely (into 6 primes) in K. Hence, if $\mathcal{O}_K$ were of the form $\mathbf{Z}[\alpha]$, it would follow by Kummer's theorem that the minimal polynomial of α would split modulo $Q \cap \mathbf{Z}$ into six distinct linear factors. However, over $\mathbf{Z}/2$, there

are only two linear polynomials! This contradiction establishes the validity of the example.

Before stating and proving Dedekind's theorem, we introduce the concept of the discriminant ideal.

Definition. Let A be a DD with quotient field K, L be a finite, separable extension and B be the integral closure of A in L. The *discriminant ideal* $D_{B/A}$ is defined to be the ideal of A generated by the elements $\mathrm{disc}(v_1, \cdots, v_n)$ as $\{v_1, \cdots, v_n\}$ runs through K-bases of L which are contained in L.

Exercise. *If B is free over A, then $D_{B/A}$ is the principal ideal generated by the discriminant of any A-basis of B.*
Hint: For any A-basis $\{e_1, \cdots, e_n\}$ of B and any K-basis $\{v_1, \cdots, v_n\}$ of L which is contained in A, write $v_j = \sum_i a_{ij} e_i$ with $a_{ij} \in A$. Then, $\mathrm{disc}(v_1, \cdots, v_n) = \det(a_{ij})^2 \mathrm{disc}(e_1, \cdots, e_n)$.

Exercise. *For any n, let Φ_n denote the nth cyclotomic polynomial (i.e. minimal polynomial of $e^{2i\pi/n}$ over $\mathbf{Q}$). Note that $X^n - 1 = \prod_{d|n} \Phi_d(X)$. Let p be a prime not dividing n and $a \in \mathbf{Z}$. Show that p divides $\Phi_n(a)$ if, and only if, a has order n in $(\mathbf{Z}/p)^*$. Moreover, this happens for some p, a if, and only if, $p \equiv 1 \bmod n$. Hence, show that there are infinitely many primes $p \equiv 1 \bmod n$.*

Exercise. *For any n, and any prime $p \equiv 1 \bmod n$, show that p splits completely in the cyclotomic field $\mathbf{Q}(\zeta_n)$ into the prime ideals $P_i = (p, \zeta_n - i)$, where i has order n in $(\mathbf{Z}/p)^*$.*

Exercise. *Let K be the quotient field of a DD A, and suppose that L is a finite, Galois extension of K. Let B denote the integral closure of A in L and let $P \subset A$ be a maximal ideal. If $PB = (P_1 \cdots P_g)^e$ in B, then show that there are fields E, F such that $K \subset F \subset E \subset L$ with $[L : E] = e$, $[E : F] = f$, $[F : K] = g$. Further, prove that such E, F exist with the properties: (i) P splits completely in F into the product of the primes of F lying below $P_1, \cdots, P_g$,*
(ii) each prime of F lying above P remains a prime in E,
and (iii) each prime of F lying above P totally ramifies in L.

Hint: Look at the fixed fields under the decomposition group and the inertia group of any P_i.

Lemma. *Let $S \subset A$ be a multiplicative subset. Then, $D_{S^{-1}B/S^{-1}A} = S^{-1}(D_{B/A})$. In particular, for a prime P of A and $S = A \setminus P$, one has*

$P \supset D_{B/A} \Leftrightarrow S^{-1}(P) \supset D_{S^{-1}(B)/S^{-1}(A)}$.

Proof. If $\{v_i\}$ is a K-basis contained in B, then v_i's are also in $S^{-1}(B)$. So, $D_{B/A} \subset D_{S^{-1}B/S^{-1}A}$. Therefore, $S^{-1}(D_{B/A}) \subset D_{S^{-1}B/S^{-1}A}$. Conversely, if $\{w_i\}$ is a K-basis contained in $S^{-1}B$, then there exists $s \in S$ such that $sw_i \in B$ for all i. Therefore, $\text{disc}(sw_1, \cdots, sw_n) = s^{2n}\text{disc}(w_1, \cdots, w_n)$. As the left hand side is in $D_{B/A}$, it follows that $\text{disc}(w_1, \cdots, w_n) \in S^{-1}(D_{B/A})$ which proves the other part of the equality asserted.

Theorem (Dedekind). *Let A, K, L, B be as before. Assume that every finite extension of A/P (for any maximal ideal P) is separable - this is true when K is an algebraic number field, for then, A/P is a finite field. Then P ramifies in L if, and only if, $P \supset D_{B/A}$.*

Proof. By the lemma, one can, without loss of generality, localise at P. Then, A, B etc. get replaced by $S^{-1}A, S^{-1}B$ which are PID's (Why?). Then, B is A-free with a basis $\{v_1, \cdots, v_n\}$ say. As observed earlier, this means that the images $\bar{v_i}$ of v_i give a basis of the A/P-vector space B/PB.
Claim: If $b \in B$, then $\overline{Tr_{L/K}(b)} = tr(\bar{b})$ where $\bar{b}$ is regarded as an A/P-endomorphism of B/PB.
To see why this is so, let us look at the endomorphism $\rho_b : B \to B; x \mapsto xb$. Write M for the matrix of ρ_b with respect to the basis $\{v_i\}$. Then, $v_i b = \sum_j m_{ij} v_j$. Reading this modulo PB, we get the fact that $\bar{M}$ is the matrix of $\bar{b}$. This gives $tr(\bar{b}) = \sum_i \bar{m_{ii}} = \overline{tr(\rho_b)} = \overline{Tr_{L/K}(b)}$ which was claimed above. Hence, $D_{B/A} = (\text{disc}(v_1, \cdots, v_n)) \subset P$ if, and only if, $\text{disc}(\bar{v_1}, \cdots, \bar{v_n}) = 0$. Let us write $PB = P_1^{e_1} \cdots P_g^{e_g}$; then $B/PB \cong B/P_1^{e_1} \oplus \cdots \oplus B/P_g^{e_g}$. To prove the theorem, let us first assume that P is unramified in B; then all the e_i are 1. Thus, B/PB is a direct sum of fields B/P_i which are separable by our hypothesis. Choose a new A/P-basis $\{\bar{b_1}, \cdots, \bar{b_n}\}$ of B/PB which is compatible with the direct sum decomposition (What does that mean?). Then, for each $\bar{b} = b^{(1)} + \cdots + b^{(g)} \in B/PB$, the matrix of $\rho_{\bar{b}}$ consists of diagonal blocks $M_1, \cdots, M_g$ where $M_i = \rho_{b^{(i)}}$. Therefore, $tr(\bar{b}) = \sum_i tr^{(i)}(b^{(i)}$ where $tr^{(i)}$ denotes the trace from B/P_i to A/P. Consequently, $\text{disc}^{B/PB}_{A/P}(\bar{b_1}, \cdots, \bar{b_n}) = \prod_i \text{disc}^{B/P_i}_{A/P}(b_1^{(i)}, \cdots, b_n^{(i)}) \neq 0$. Hence, for the original A/P-basis $\{\bar{v_i}\}$, one has $\text{disc}(\bar{v_1}, \cdots, \bar{v_n}) = d^2\text{disc}(\bar{b_1}, \cdots, \bar{b_n}) \neq 0$ in A/P, where d is the determinant of the change of basis. This proves that $P \not\supset D_{B/A}$ as observed earlier.
Conversely, suppose that some $e_i > 1$. Then, $B/P_i^{e_i}$ (and so B/PB itself) has a nilpotent element, say u_1. Extend it to a basis $\{u_1, \cdots, u_n\}$ of B/PB. As $u_1 u_i$ is nilpotent, one has $tr(u_1 u_i) = 0$ for all i. Therefore, $\text{disc}(u_1, \cdots, u_n) = 0$ and so for the other basis too, one has $disc(\bar{v_1}, \cdots, \bar{v_n})$

$= 0$. In other words, $P \supset D_{B/A}$. This completes the proof.

4. Finiteness of class number and Minkowski's bound

In this section, we shall show that the class group of an algebraic number field is finite. Its order, called the *class number*, gives a measure of the deviation from the unique factorisation property. Although the finiteness is easy to establish, the easy proof gives a somewhat large bound. A much better bound was obtained by Minkowski using a geometric method. We shall discuss Minkowski's method and in the next section, we shall apply it to prove a theorem of Dirichlet on the structure of units of a number field.

Theorem. *For an algebraic number field K, the class group is finite.*

Proof. Fix an integral basis $\{v_1, \cdots, v_n\}$ of $\mathcal{O}_K$. Let $I \neq 0$ be any ideal and consider the subset S of $\mathcal{O}_K$ consisting of all $\sum_{i=1}^{n} m_i v_i$ with $0 \leq m_i \leq N(I)^{1/n}$. Evidently, $\# \ S > N(I) = \# \ (\mathcal{O}_K/I)$. Therefore, there exist $a \neq b \in S$ such that $a - b \in I$. Notice that $a - b = \sum_i m_i v_i$ for some integers m_i which satisfy $\mid m_i \mid \leq N(I)^{1/n}$. Let us compute its norm over $\mathbf{Q}$. We have $N_{K/\mathbf{Q}}(a-b) = \prod_i \sigma_i(\sum_j m_j v_j)$ where σ_i's are the embeddings of K in $\mathbf{C}$. Therefore,

$$\mid N_{K/\mathbf{Q}}(a-b) \mid \ = \prod_i \mid \sum_j m_j \sigma_i(v_j) \mid \ \leq \prod_i \sum_j \mid m_j \mid \ \mid \sigma_i(v_j) \mid \ \leq N(I)C,$$

where $C = \prod_i \sum_j \mid \sigma_i(v_j) \mid$ is a constant independent of the ideal I; it depends only on K. Now $a - b \in I \Rightarrow (a-b) = IJ$ for some non-zero ideal J. Thus $N_{K/\mathbf{Q}}(a-b) = N(I)N(J) \leq N(I)C$ and we get $N(J) \leq C$. As J is just the inverse of I in the class group, it runs through the class group when I does. Therefore, we have shown that any element of the class group has a representative ideal whose norm is at the most the constant C. As there are only finitely many ideals with the norm bounded by an absolute constant, the theorem follows.

Example. Let $K = \mathbf{Q}(\sqrt{2})$. Then, $\mathcal{O}_K = \mathbf{Z}[\sqrt{2}]$ has $\{1, \sqrt{2}\}$ as a $\mathbf{Z}$-basis. The constant C above is $C = (1+\sqrt{2})^2 = 5.8....$ So, every ideal has a representative I with norm at the most 5. Thus, the prime ideals dividing I must have norm ≤ 5 which means that they are among those lying over $2, 3$ and 5. Now, $3, 5$ are unramified and must, therefore, be either inert or split. As 2 is not a square mod 3, 3 remains prime. So is the case with 5 also. Finally, 2 is the square of the prime ideal $(\sqrt{2})$. Thus, we have shown that every ideal class contains a representative ideal which is principal. Thus, the class group is trivial, i.e. $\mathcal{O}_K$ is a PID.

The bound given above is somewhat large. One can do somewhat better; proceeding as in the proof of the theorem, one can write out the matrix M of $a-b$ with respect to the basis $\{v_1, \cdots, v_n\}$. $M = \sum_i m_i M_i$ where M_i is the matrix of v_i with respect to the same ordered basis. Note that all the entries of M_i are integers whose absolute values are bounded by a constant depending only on the basis $\{v_i\}$ and not on the ideal I. Then, by definition, $\mid N_{K/\mathbf{Q}}(a-b) \mid \;=\; \mid \det(M) \mid \;\leq C_0 N(I)$. This constant C_0 is better than the constant C in the proof of the theorem. For example, when $K = \mathbf{Q}(\sqrt{-5})$, we have $C = 10,\ C_0 = 6$. But, in fact, the method we shall discuss below, due to Minkowski, gives a much better bound. In this example, it will give a constant less than 3 which will enable us to conclude quite easily that the class number is 2.

Definitions. A *lattice* Λ in the Euclidean space $\mathbf{R}^n$ is the $\mathbf{Z}$-span of an $\mathbf{R}$-basis of $\mathbf{R}^n$. Clearly, the group $GL_n(\mathbf{R})$ of invertible $n \times n$ matrices acts transitively on the set of all lattices. Thus, any lattice can be identified with $g\mathbf{Z}^n$ for some $g \in GL_n(\mathbf{R})$. Given a lattice Λ, *a fundamental parallelotope* for it is the set of vectors $\{\sum_i t_i e_i : 0 \leq t_i < 1\}$ for any basis $\{e_i\}$ of Λ. As any two $\mathbf{Z}$-bases are transforms of each other under a matrix in $GL_n(\mathbf{Z}) = \{\gamma \in M_n(\mathbf{Z}) : det(\gamma) = \pm 1\}$, the *volume of the lattice* $\Lambda = g\mathbf{Z}^n$ is the well-defined non-zero real number $\mid det(g) \mid$. We write $\mathrm{Vol}(\mathbf{R}^n/\Lambda)$ for the volume of Λ.

Lemma. *Let K be an algebraic number field. Let $\sigma_1, \cdots, \sigma_r, \tau_1, \cdots, \tau_s$, $\bar{\tau}_1, \cdots, \bar{\tau}_s$ be the embeddings of K in* $\mathbf{C}$. *Here, the σ_i's take real values and the τ_j's take nonreal values. Then, the map $\theta : t \mapsto$*

$$(\sigma_1(t), \cdots, \sigma_r(t), Re(\tau_1(t)), \cdots, Re(\tau_s(t)), Im(\tau_1(t)), \cdots, Im(\tau_s(t)))$$

from K to $\mathbf{R}^n$ embeds $\mathcal{O}_K$ as a lattice. Its volume is $\sqrt{\mid \mathrm{disc}(K) \mid}/2^s$. In particular, K embeds densely in $\mathbf{R}^n$.

Proof. Let $v_1, \cdots, v_n$ be a $\mathbf{Z}$-basis of $\mathcal{O}_K$. We show that $\theta(v_1), \cdots, \theta(v_n)$ are linearly independent. If we write $\theta = (\theta_1, \cdots, \theta_n)$ to mean the obvious, look at the matrix M with $m_{ij} = \theta_i(v_j)$. Elementary column operations transform M to the matrix whose i-th row is

$$(1/2i)^s(\sigma_1(v_i), \cdots, \sigma_r(v_i), \tau_1(v_1), \bar{\tau}_1(v_i), \cdots, \tau_s(v_i), \bar{\tau}_s(v_i))$$

This gives the result that the determinant of M is $(1/2i)^s\sqrt{\mathrm{disc}(K)}$; so $\mathrm{Vol}(\mathbf{R}^n/\theta(\mathcal{O}_K)) = \sqrt{\mid \mathrm{disc}(K) \mid}/2^s$.

Definition and Remarks. Given a positive integer n and non-negative integers r, s such that $r + 2s = n$, define a *norm on* $\mathbf{R}^n$ by $N_{r,s}(x) =$

$x_1 \cdots x_r(x_{r+1}^2 + x_{r+2}^2) \cdots (x_{n-1}^2 + x_n^2)$. Thus, in the situation of a number field K of degree n over $\mathbf{Q}$ and r, s, θ as above, we have $N_{r,s}(\theta(t)) = N_{K/\mathbf{Q}}(t)$ for all $t \in \mathcal{O}_K$.

Theorem (Minkowski). *Every lattice Λ in $\mathbf{R}^n$ contains $x \neq 0$ with $N_{r,s}(x) \leq \frac{n!}{n^n}(\frac{8}{\pi})^s \mathrm{Vol}(\mathbf{R}^n/\Lambda)$.*

We shall give the proof of this important theorem after pointing out some very useful consequences of it.

Corollary. *Let $[K : \mathbf{Q}] = n$ and r, s have the usual meaning. Then,*
(a) Every non-zero ideal I contains $x \neq 0$ with

$$| N(x) | \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s \sqrt{| \operatorname{disc}(K) |}\, N(I).$$

(b) Every ideal class contains an ideal I with

$$| N(I) | \leq \frac{n!}{n^n}\left(\frac{4}{\pi}\right)^s \sqrt{| \operatorname{disc}(K) |}.$$

(c) $\operatorname{disc}(K) > 1$ *if $K \neq \mathbf{Q}$.*
(d) If $K \neq \mathbf{Q}$, then some prime number p ramifies in K.

Proof. Using the lemma above, $\mathcal{O}_K$ can be viewed as a lattice in $\mathbf{R}^n$ whose volume has also been computed. Therefore, both (a) and (b) are direct consequences of Minkowski's theorem. To prove (c), just observe that the number $\frac{n^n}{n!}(\frac{\pi}{4})^s > \frac{1}{n!}(\frac{n\pi}{4})^n > 1$ for $n > 1$. Finally, (d) follows from Dedekind's theorem which showed that prime numbers which divide the discriminant of K must ramify in K.

Example/Exercise. Let $K = \mathbf{Q}(\sqrt{-5})$. Then, the above constant (called Minkowski's constant) on the right hand side of (b) shows that each ideal class contains a representative ideal I of norm $N(I) \leq \frac{4\sqrt{5}}{\pi} < 3$. So, one need only consider the ideals lying above 2 viz., $(2, 1 \pm \sqrt{-5})$. It is easy to see that these are not principal and thus it follows that K has class number 2.
Using this fact, show that the equation $x^2 + 5 = y^3$ has no integral solutions.

For the proof of Minkowski's theorem, one needs the following beautiful lemma on convex bodies which is of independent interest:

Minkowski's lemma. *Let Λ be a lattice in $\mathbf{R}^n$, E a convex, measurable, centrally symmetric subset of $\mathbf{R}^n$ such that* $\mathrm{Vol}(E) > 2^n \mathrm{Vol}(\mathbf{R}^n/\Lambda)$. *Then, E contains some non-zero point of Λ. Further, if E is also compact, then the strict inequality in the hypothesis can be weakened to $\geq$.*

Proof. Let F be a fundamental parallelotope for Λ. Then, we have $\mathbf{R}^n = \bigsqcup_{x\in\Lambda}(x+F)$. Now, $\frac{1}{2}E = \bigsqcup_{x\in\Lambda}(\frac{1}{2}E \cap (x+F))$. By the hypothesis,

$$\begin{aligned}\mathrm{Vol}(F) < \mathrm{Vol}(E)/2^n = \mathrm{Vol}(E/2) &= \sum_{x\in\Lambda}\mathrm{Vol}(\frac{1}{2}E \cap (x+F)) \\ &= \sum_{x\in\Lambda}\mathrm{Vol}((\frac{1}{2}E - x)\cap F)\end{aligned}$$

Therefore, as x runs over Λ, the sets $(\frac{1}{2}E - x)\cap F$ are not all disjoint. Thus, we get $x \neq y$ in Λ so that $\frac{1}{2}a - x = f = \frac{1}{2}b - y$ for some $a, b \in E, f \in F$. Clearly, then we get $0 \neq x - y = \frac{1}{2}a + \frac{1}{2}(-b) \in E \cap \Lambda$. This proves the main assertion. For the case when E is also compact, one may consider the sets $(1+\frac{1}{n})E$ and obtain lattice points $x_n \neq 0$ as above. Evidently, then all the $x_n \in 2E \cap \Lambda$ which is a finite set. Thus, for some n_0, $x_{n_0} \in (1+\frac{1}{n})E$ for infinitely many n i.e. $x_{n_0} \in \bar{E} = E$. The proof is complete.

Corollary. *Suppose that Ω is a compact, convex, centrally symmetric subset of $\mathbf{R}^n$ such that $Vol(\Omega) > 0$ and such that $\mid N_{r,s}(a) \mid \leq 1 \;\; \forall\, a \in \Omega$. Then, every lattice Λ contains a non-zero vector x with*

$$\mid N_{r,s}(x) \mid \leq 2^n \frac{Vol(\mathbf{R}^n/\Lambda)}{Vol(\Omega)}.$$

The proof is immediate from Minkowski's lemma applied to the set $E = t\Omega$ where $t^n = 2^n \frac{Vol(\mathbf{R}^n/\Lambda)}{Vol(\Omega)}$.

Proof of Minkowski's theorem. Let Ω be the subset of $\mathbf{R}^n$ defined by the inequality $\sum_{i=1}^r \mid x_i \mid +2\sqrt{(x_{r+1}^2 + x_{r+2}^2)} + \cdots + 2\sqrt{(x_{n-1}^2 + x_n^2)} \leq n$. We shall prove that Ω is convex, and that $\mid N_{r,s}(a) \mid \leq 1 \;\forall\, a \in \Omega$. Then, we shall compute its volume and apply the above corollary.

Step I: Ω is convex

¿From the definition of Ω, it is easy to see that if midpoints of any two points of Ω are in Ω, then Ω is convex. Let $(x_1, \cdots, x_n), (y_1, \cdots, y_n) \in \Omega$. Then, we have

$$\sum_{i=1}^r \mid x_i \mid +2\sqrt{(x_{r+1}^2 + x_{r+2}^2)} + \cdots + 2\sqrt{(x_{n-1}^2 + x_n^2)} \leq n,$$

$$\sum_{i=1}^r \mid y_i \mid +2\sqrt{(y_{r+1}^2 + y_{r+2}^2)} + \cdots + 2\sqrt{(y_{n-1}^2 + y_n^2)} \leq n.$$

Adding and using the triangle inequality

$$\sqrt{(a^2+b^2)} + \sqrt{(c^2+d^2)} \geq \sqrt{((a+c)^2 + (b+d)^2)}$$

one concludes that $(\frac{x_1+y_1}{2}, \cdots, \frac{x_n+y_n}{2}) \in \Omega$.
Step II: $\mid N_{r,s}(a) \mid \leq 1 \; \forall \; a$.
This is clear from the usual inequality $A.M \geq G.M$.
Step III: $Vol(\Omega) = \frac{(2n)^n}{n!}(\frac{\pi}{8})^s$.
Let $V_{r,s}(t)$ denote the volume of the set Ω_t defined in a similar fashion to Ω but with n replaced by the real number $t > 0$. It is easy to see from the definition that $V_{r,s}(t) = V_{r,s}(1)t^{r+2s}$. Now, if $r > 0$, then

$$\begin{aligned} V_{r,s}(1) &= 2\int_0^1 V_{r-1,s}(1-x)dx \\ &= 2V_{r-1,s}(1)\int_0^1 (1-x)^{r-1+2s}dx = \frac{2}{r+2s}V_{r-1,s}(1). \end{aligned}$$

Proceeding inductively, one obtains finally that $V_{r,s}(1) = \frac{2^r}{(r+2s)\cdots(2s+1)}$.
Similarly, if $s > 0$, then

$$\begin{aligned} V_{0,s}(1) &= \int\int_{x^2+y^2 \leq 1/4} V_{0,s-1}\left(1 - 2\sqrt{(x^2+y^2)}\right) dxdy \\ &= \int_0^{2\pi}\int_0^{1/2} V_{0,s-1}(1-2\rho)\rho d\rho d\theta. \end{aligned}$$

Once again, iterating inductively, one finally obtains $V_{0,s}(1) = (\frac{\pi}{2})^s\frac{1}{(2s)!}$. Then, $Vol(\Omega_t) = t^n V_{r,s}(1) = t^n 2^{r-s}\pi^s\frac{1}{n!}$ which gives that $Vol(\Omega = \Omega_n) = n^n\frac{2^n}{2^{3s}}\pi^s\frac{1}{n!} = \frac{(2n)^n}{n!}(\frac{\pi}{8})^s$. The proof of Step III and, along with it, that of Minkowski's theorem, is complete.

5. Dirichlet's unit theorem

In this section, we use Minkowski's method to find the structure of the units in any algebraic number field K.

Recall that we embedded $\mathcal{O}_K$ as a lattice Λ_0 in $\mathbf{R}^n$ by means of $\theta : a \mapsto (\sigma_1(a), \cdots, \sigma_r(a), Re\tau_1(a), Im\tau_1(a), \cdots, Re\tau_s(a), Im\tau_s(a))$. Here $n = [K : \mathbf{Q}]$ and $\sigma_1, \cdots, \sigma_r, \tau_1, \bar{\tau}_1, \cdots, \tau_s, \bar{\tau}_s$ are the distinct embeddings of K in $\mathbf{C}$. Clearly, if a is a unit in $\mathcal{O}_K$, then both u and u^{-1} map to vectors which are linearly dependent. Thus, one needs to go to a subspace of $\mathbf{R}^n$ to be sensitive to the units.

Lemma. *Consider the composite map L in*

$$\mathcal{O}_K^* \subset \mathcal{O}_K \setminus 0 \xrightarrow{\theta} \Lambda_0 \setminus 0 \to \mathbf{R}^{r+s}$$

where the last map is $(x_1, \cdots, x_n) \mapsto$
$(log(\mid x_1 \mid), \cdots, log(\mid x_r \mid), log(x_{r+1}^2 + x_{r+2}^2), \cdots, log(x_{n-1}^2 + x_n^2)\ )$. *Then,*

(i) the image of $L : \mathcal{O}_K^* \to \mathbf{R}^{r+s}$ is contained in the hyperplane H of vectors $(x_1, \cdots, x_{r+s})$ such that $\sum_{i=1}^{r+s} x_i = 0$.
(ii) L is a homomorphism.
(iii) $Im(L) \cong \mathbf{Z}^d$ for some $d \leq r + s - 1$.
(iv) $\mathrm{Ker}(L) \cong \mu(K)$, the group of roots of unity in K and $\mathcal{O}_K^* \cong \mu(K) \times \mathbf{Z}^d$ for some $d \leq r + s - 1$.

Proof. (i) follows since units must have norm ±1. (ii) is obvious. To see that (iii) holds, let R be any bounded region in $H \subset \mathbf{R}^{r+s}$ and let $L(u) \in R$. Then, all the conjugates of u have absolute values bounded by a constant depending on R. As the coefficients of the minimal polynomial of u are symmetric functions of the various conjugates of u, this means that there are only finitely many polynomials satisfied by units whose images under L lie in the bounded region R. In other words, $R \cap \mathrm{Im}(L)$ is finite i.e. $\mathrm{Im}(L)$ is discrete in H. Now, (iii) follows by the easy exercise below. The first assertion of (iv) is trivial and the second one follows because one can check easily that units $u_1, \cdots, u_d$ mapping under L to a basis of $\mathrm{Im}(L)$ have to generate a free abelian group.

Exercise. *Show by induction on n that a discrete subgroup of $\mathbf{R}^m$ is isomorphic to $\mathbf{Z}^d$ for some $d \leq m$.*

Dirichlet's unit theorem. *$\mathcal{O}_K^* = \mu(K) \times V$ where $V \cong \mathbf{Z}^{r+s-1}$.*

In other words, the image of $\mathcal{O}_K^*$ under L is actually a lattice in H. This will be seen by actually showing the existence of $r+s-1$ units whose images under L are linearly independent.

Lemma. *Fix any $k \leq r + s$. Then, $\forall$ $\alpha \neq 0$ in $\mathcal{O}_K$, there exists $\beta \in \mathcal{O}_K$ with $| N(\beta) | \leq (\frac{2}{\pi})^s \sqrt{| \operatorname{disc}(K) |}$ and satisfies $\beta_i < \alpha_i$ $\forall$ $i \neq k$. Here α_i, β_i denote the co-ordinates of their images under L.*

Proof. Let c_i be constants such that $0 < c_i < e^{\alpha_i}$ $\forall$ $i \neq k$ and $c_k = (\frac{2}{\pi})^s \sqrt{| \operatorname{disc}(K) |} / \prod_{i \neq k} c_i$. Then, consider the set $\Omega \subset \mathbf{R}^n$ defined by $| x_i | \leq c_i$, $\forall$ $i \leq r$ and $x_{r+1}^2 + x_{r+2}^2 \leq c_{r+1}, \cdots, x_{n-1}^2 + x_n^2 \leq c_{r+s}$. $\mathrm{Vol}(\Omega) = (2c_1) \cdots (2c_r)(\pi c_{r+1}) \cdots (\pi c_{r+s}) = 2^n \mathrm{Vol}(\mathbf{R}^n/\Lambda_0)$. Applying Minkowski's lemma, one gets some $t \neq 0$ in $\Omega \cap \Lambda_0$. Then, choose $\beta \in \mathcal{O}_K$ corresponding to t.

Lemma. *Fix any $k \leq r+s$. Then, $\exists$ $u \in \mathcal{O}_K^*$ such that $L(u) = (u_1, \cdots, u_{r+s})$ satisfies $u_i < 0$ $\forall$ $i \neq k$.*

Proof. Start with any $\alpha_1 \neq 0$ in $\mathcal{O}_K$ and apply the previous lemma to get some β as above; call that α_2. Repetitively, one gets a sequence $\{\alpha_n\}$ in $\mathcal{O}_K$ such that for all $i \neq k$, the i-th co-ordinate of $L(\alpha_{n+1})$ is less than that of

$L(\alpha_n)$. By the lemma, $\mid N(\alpha_n) \mid$ are bounded above as $n \to \infty$. Therefore, the principal ideals (α_n) are only finitely many. Taking any $n < m$ so that $(\alpha_n) = (\alpha_m)$, we have $\alpha_m = \alpha_n u$ for some unit u. Evidently, u does the job.

The proof of Dirichlet's unit theorem is completed as follows. Observe that the units $u_k, k \leq r + s$, obtained by the previous lemma have the property that the $(r+s) \times (r+s)$ matrix $A = (a_{ij})$ whose k-th row is $L(u_k)$ satisfies $a_{ij} < 0$ for all $i \neq j$ and each row sums to 0. It is an easy elementary exercise to see that the rank of A must be $r + s - 1$.

REFERENCES

1. Gerald J. Janusz, *Algebraic Number Fields*, Graduate Studies in Mathematics, Vol. 7, Second Edition, American Mathematical Society, (1996).

2. Daniel A. Marcus, *Number Fields*, Springer-Verlag (1977).

3. Raghavan Narasimhan, S. Raghavan, S.S. Rangachari and Sundar Lal, *Algebraic Number Theory*, TIFR pamphlet (1966).

B. Sury
Indian Statistical Institute
Bangalore
e-mail: bsury@isibang.ac.in

Quadratic and Cyclotomic fields

Dinesh S. Thakur

As a complement to Sury's lectures on Dedekind domains, we will now give an example oriented introduction to quadratic and cyclotomic fields. In the workshop, the two series of talks went in parallel. So we might use terminology defined carefully in Sury's talks. We have omitted simple details which were usually worked out in problem sessions and are also given in many texts. Since the repetition usually helps, we have not tried for an efficient or a general treatment.

Apart from $\mathbf{Q}$, the simplest class of number fields are the *quadratic* fields i.e., the degree 2 extensions (so that there are no non-trivial subfields) obtained by solving a quadratic and hence (by completing the square) of the form $K = \mathbf{Q}(\sqrt{a/b})$. Multiplying by b and getting rid of the squares under the square-root, we can write it as $K = \mathbf{Q}(\sqrt{m})$, where m is square-free. (So these are special Kummer extensions). These fields are then distinct for distinct m. These are Galois extensions, with Galois conjugate of a general element $r + s\sqrt{m}$ $(r, s \in \mathbf{Q})$ being $r - s\sqrt{m}$. The norm and the trace are essentially just the coefficients of the minimal polynomial in this case.

What are the *algebraic integers*? By making a common denominator, we can write a general element of the field as $(a + b\sqrt{m})/c$, with integral a, b, c with the GCD $(a, b, c) = 1$. Since the trace and norms are usual integers, we have $c|2a$ and $c^2|a^2 - mb^2$. So if we let $d = (a, c)$, then d^2 divides a^2, b^2 and $a^2 - mb^2$ and hence mb^2. Since m is square-free, d divides b, so that $d = 1$. Therefore, c divides $2a$ now implies c divides 2 and hence c is 1 or 2 without loss of generality. (Another way to see this is that if x is an algebraic integer in the field, $x^2 + Bx + C = 0$, hence $x = (-B \pm \sqrt{B^2 - 4C})/2)$. If $c = 2$, then a is odd and $mb^2 \equiv a^2 \equiv 1$ modulo 4 and hence $m \equiv 1$ modulo 4 and b is odd.

So (exercise: finish the details) the $\mathbf{Z}$-*basis* of $\mathcal{O}_K$ is $1, (1 + \sqrt{m})/2$ or $1, \sqrt{m}$ depending on whether $m \equiv 1$ modulo 4 or not. If we write $K = \mathbf{Q}(\sqrt{d})$, where d is the discriminant, which is always $\equiv 0, 1$ modulo 4 (and is either m or $4m$), then we can say that the basis is always $1, (d + \sqrt{d})/2$.

What are the *units*? If α is a unit, it divides 1 and hence its norm being a rational integer dividing 1 is ± 1. Further, if $d < 0$ (the imaginary quadratic fields), the norm is positive, so is 1.

So (exercise) the only units in the imaginary quadratic fields are ± 1, except for $\mathbf{Q}(\sqrt{-1}) = \mathbf{Q}(\zeta_4)$, $\mathbf{Q}(\sqrt{-3}) = \mathbf{Q}(\zeta_6)$, where we have 4 and 6 units (which are the obvious roots of unity) respectively.

In the real quadratic case, we are led to equations such as $x^2 - my^2 = \pm 1$, called Pell equation (or rather Brahmagupta-Bhaskara-Fermat-Pell equation). For example, for $\mathbf{Q}(\sqrt{2})$, we see a solution $x = y = 1$ and get corresponding unit $w = 1 + \sqrt{2}$. Since this is not a root of unity, we get infinitely many units $\pm(1+\sqrt{2})^n$, $n \in \mathbf{Z}$. In fact, these are all the units in this case: In general, if a unit other than ± 1 exists, then a smallest unit $\epsilon > 1$ exists (otherwise both the conjugates $x \pm y\sqrt{d}$ get close to 1 which forces x close to 1 and y close to 0, so equal to it) and is called the *fundamental unit.* It is easy to see that all the units are then given by $\pm\epsilon^n$. In our case, if $1 < x + y\sqrt{2} = \epsilon < w$, then $x^2 - 2y^2 = \pm 1$ implies $-1 < x - y\sqrt{2} < 1$. Adding the two, we get $0 < 2x < 1 + w$, so that $x = 1$ giving a contradiction $1 < 1 + y\sqrt{2} < 1 + \sqrt{2}$.

Dirichlet's theorem guarantees existence of a fundamental unit in the real quadratic case. We will just state a recipe: The continued fraction of $\sqrt{d} + \lfloor\sqrt{d}\rfloor$ is purely periodic with period vector $(a_0, \cdots a_{r-1})$ and p_{nr-1} and q_{nr-1} (the numerators and denominators of the convergents), are all the solutions of $x^2 - dy^2 = 1$ for even r or for odd r with even n and are solutions of $x^2 - dy^2 = -1$ (which has no solutions, if r is even) when both r and n are odd. The fundamental solution corresponds to $n = 1$.

(Exercise): For the imaginary quadratic fields, the integers sit discretely in $\mathbf{C}$, where as they are dense in $\mathbf{R}$ in the real quadratic fields.

Usual proof of the fact that $\mathbf{Z}$ is a principal ideal domain and unique factorization domain uses the *division algorithm*: The smallest positive element in the ideal is its generator by the division algorithm. Let us see how often the division algorithm works in the imaginary quadratic case, where the size comparisons are now done using the norm: Let $0 > m = -\mu$. If there is a division algorithm, given a, b we get a quotient q such that $a = qb + r$, with a 'smaller' remainder. This translates to norm of $a/q - b$ being smaller than 1. Translating to usual integers, given rationals r and s, we can find x and y such that $|(r-x)^2 - m(s-y)^2| < 1$, with x, y integers, if $m \not\equiv 1$ modulo 4 and half-integers otherwise. Choosing $r = s = 1/2$, in the first case, we see that $1/4 + \mu/4 < 1$, so that $\mu = 1$ or 2. In the second case, similarly, we get $1/16 + \mu/16 < 1$ giving $\mu = 3$, 7 or 11. So there are exactly 5 imaginary (and in fact, 16 real) quadratic fields which are *Euclidean* (for the size given by the norm function). There are four more unique factorizations domains, with $\mu = 163$ being the largest.

As an exercise in manipulations with ideals, let us see how ideals restore unique factorization in our example:

$$\begin{aligned} 6 &= 2\times 3 \\ &= (1+\sqrt{-5})(1-\sqrt{-5}) \\ &= (2,1+\sqrt{-5})(2,1-\sqrt{-5})(3,1+\sqrt{-5})(3,1-\sqrt{-5}). \end{aligned}$$

Note $I|J$ if and only if $J\subset I$, i.e., the multiples of J are contained in the multiples of I. Verify that $(2,1+\sqrt{-5})(3,1+\sqrt{-5})=(1+\sqrt{-5})$. Also note that $(2,1+\sqrt{-5})=(2,1-\sqrt{-5}+2\sqrt{-5})=(2,1-\sqrt{-5})$, so that 2 is a square (and a norm) of an ideal (it ramifies). (The discriminant of $\mathbf{Q}(\sqrt{-5})$ is -20, so that 2 and $5=-(\sqrt{-5})^2$ ramify.) There is no element of norm 2, otherwise we would have integral solution to $2=x^2+5y^2$. Hence, our ideal is non-principal. In fact, the class group is of order 2 in this case (exercise).

Let us use this fact to show that $x^3=y^2+5$ has no integral solutions: Looking at modulo 4 possibilities for an assumed solution, we see that y (which is seen to be prime to 5 also) is even, and so the GCD of the two factors $y\pm\sqrt{-5}$, which has to divide 2, can in fact be assumed to be 1. As the class number is prime to 3, each factor is a cube of an ideal. Since the units here are ± 1, which are also cubes, we get $y+\sqrt{-5}=(a+b\sqrt{-5})^3=a^3-5b^2+\sqrt{-5}b(3a^2-5b^2)$, where a and b are (rational) integers. Comparing the imaginary parts, we see that $b=\pm 1$, so that $1=\pm(3a^2-5)$, which is a contradiction.

Now let us look at the basic properties about the *cyclotomic* fields: We denote a primitive n-th root of unity by ζ_n. As a complex number, it is $e^{2\pi ik/n}$, with $(k,n)=1$. Since $1=kr+cn$, each of this is a power of any other, so that $K=\mathbf{Q}(\zeta_n)$ is a Galois extension. The minimal polynomial of ζ_n is called the n-th cyclotomic polynomial and is given by (verify irreducibility) $\Phi_n(x)=\prod_{1\le k\le n,(k,n)=1}(x-\zeta_n^k)$. Its degree is $\phi(n)$. Recall that for $n=\prod p_i^{n_i}$, we have $\phi(n)=\prod(p_i-1)p_i^{n_i-1}$.

The *Galois group* can be identified with $(\mathbf{Z}/n\mathbf{Z})^*$, with the action the automorphism σ_k corresponding to k being defined by $\sigma_k(\zeta_n)=\zeta_n^k$. In Galois theory, we have learned the correspondence between subgroup structure of the Galois group and the *subfield structure* of the field. In particular, since the group is abelian, all the subfields are Galois (with abelian Galois group). Let us find them explicitly for $\mathbf{Q}(\zeta_p)$, for p a prime: We know (exercise) that the Galois group $(\mathbf{Z}/p\mathbf{Z})^*$ is cyclic, say with generator g. So for $e|(p-1)$, we will have a unique sub-extension of degree e. For example, if $e=(p-1)/2$, then it is $\mathbf{Q}(\zeta_p)^+ := \mathbf{Q}(\zeta_p+\zeta_p^{-1})=\mathbf{Q}(\cos(2\pi/p))$, which is the maximal real subfield (and is obtained by averaging with respect to the complex conjugate). In general, we write $ef = p-1$ and define the so-called *periods* $\eta_i := \sum_{j=0}^{f-1} \zeta_p^{g^{ej+i}}$. We see that $\sigma_g(\eta_i) = \eta_{i+1}$, where i runs modulo e. So that there are e periods, all conjugate and distinct (as the minimal equation of ζ_p has all the $p-1$ powers occurring in it, whereas $\eta_i - \eta_j$ has fewer). Since each is left invariant with respect to the subgroup $H = \langle \sigma_g^e \rangle$, $\mathbf{Q}(\eta_i) = \mathbf{Q}(\eta_0)$ is the degree e sub-extension we want. The periods are useful in the construction of regular polygons, because they give explicit subfield structure needed in the problem.

What is the *quadratic subfield* of $\mathbf{Q}(\zeta_p)$? It is $\mathbf{Q}(\sqrt{(-1)^{(p-1)/2}p})$. One way to see this is to evaluate the corresponding period: $\sum \zeta_p^{g^{2j}} = \sum \zeta_p^k$, where k runs through quadratic residues and this relates to quadratic Gauss sum evaluated in Adhikari's lectures. Another way is to note that, in general, the discriminant which is square of the product of the differences of the conjugates, and it belongs to the base, its square root belongs to the field (Galois), giving a quadratic extension, if it is not a square. In our case, the discriminant is $\prod_{i<j}(\zeta_p^i - \zeta_p^j)^2$. Taking out the roots of unity, we see that the power of $(1-\zeta_p)$ is $2(1+2+\cdots+(p-2)) = (p-1)(p-2)$, so that the discriminant is (using $(p) = (1-\zeta_p)^{p-1}$ proved below) $\pm p^{p-2}$. Easy way to fix (and remember) the sign is to note that the maximal real subfield $\mathbf{Q}(\zeta_p)^+$ has has degree $(p-1)/2$, which is odd if $p \equiv 3$ modulo 4 and hence can not have quadratic subfield. So in this case, the subfield is imaginary quadratic.

What are the *algebraic integers* in $K = \mathbf{Q}(\zeta_n)$? In fact, $\mathcal{O}_K = \mathbf{Z}[\zeta_n]$. (See the proof in Sury's notes in this volume or in Washington[1]). One way inclusion is clear and since K is a quadratic field for $n \leq 4$ and $n = 6$, we can already verify the claim in those cases.

Let us write $\zeta = \zeta_p$ for this paragraph. We have $\Phi_p(x) = \prod(x-\zeta^i) = (x^p-1)/(x-1) = x^{p-1} + \cdots + 1$, so that putting $x = 1$, we see that $p = \prod(1-\zeta^i) = Norm(1-\zeta)$. Now $(1-\zeta^i)/(1-\zeta)$ clearly belongs to $\mathbf{Z}[\zeta]$ and so does its inverse (which is obtained by just relabelling the primitive root!), so these are units. Hence we get $p\mathcal{O}_K = (1-\zeta)^{p-1}$ as ideals, so that p is totally ramified in $\mathbf{Q}(\zeta)$ with $(1-\zeta)$ being the prime above p. The story for $n = p^m$ is similar (exercise). On the other hand if two distinct primes p and q divide n, then since $1-\zeta_n$ divides $1-\zeta_p$ and $1-\zeta_q$, it divides p and q and hence is a unit. Together with the roots of unity and monomials in these, we get a readily available supply of units called cyclotomic units.

More precisely, the group of *cyclotomic units* for $K = \mathbf{Q}(\zeta_n)$ or $K = \mathbf{Q}(\zeta_n)^+$ is defined to be $C_n = \langle \pm\zeta_n, 1-\zeta_n^k \rangle \cap \mathcal{O}_K^*$. (For the intermediate

[1] For such references the reader may look into the general bibliography at the end of this volume.

extensions it is better to modify the definition by taking norms from the full cyclotomic extension, see Washington). For $n = p^m$, Washington Lemma 8.1 shows that the cyclotomic units are generated by

ζ_{p^m}, -1 and $\zeta_{p^m}^{(1-a)/2}(1-\zeta_{p^m}^a)/(1-\zeta_{p^m})$, for $1 < a < p^m/2$, $(a,p) = 1$. The latter are $\phi(p^m)/2-1 = r_2-1$ of them and in fact they are independent, giving the full rank of the unit group given by the Dirichlet theorem. In fact, the index of the cyclotomic units subgroup in the full unit group is the size of the class group of $\mathbf{Q}(\zeta_{p^m})^+$. So in some sense, the amount of failure of unique factorization is linked to amount of failure of capturing all units from these readily available cyclotomic ones! For n not a prime power, the story is more complicated. We will study the Ramachandra units for the general case in R. Balasubramanian's lectures[2].

Now we look at *how the usual primes factor*, when we go up in quadratic or cyclotomic extensions:

Claim: For the *quadratic field* K of discriminant d, and for an odd prime p, we have (i) $p\mathcal{O}_K = \wp^2$, $\wp$ prime if and only if $p|d$ i.e., $(d/p) = 0$, (ii) $p\mathcal{O}_K = \wp_1\wp_2$, $\wp_i$ distinct primes if and only if $(d/p) = 1$ and (iii) $p\mathcal{O}_K = \wp$ prime if and only if $(d/p) = -1$, where (d/p) is the Legendre symbol.

For the proof as well as $p = 2$ case, see TIFR pamphlet, pp. 63-64. Note that $\wp = (p, \sqrt{d})$ in case (i) and $\wp_i = (p, a \pm \sqrt{d})$ in case (ii), where $a^2 \equiv d$ modulo p.

Claim: For the *cyclotomic field* $K = \mathbf{Q}(\zeta_n)$, and for a prime p, we have (i) p is ramified if and only if $p|n$ and (ii) If p does not divide n, then $p\mathcal{O}_K = \wp_1 \cdots \wp_g$, with $\wp_i$ distinct primes of residue degree f and $g = \phi(n)/f$. Here f is the order of p modulo n i.e., f is the smallest positive integer such that $p^f \equiv 1$ modulo n.

Proof: Since p divides n if and only if it divides the discriminant, we know from the general theory covered in Sury's lectures that (i) holds (we have also seen that if $p|n$, then p is ramified, at least for n a prime power) and (ii) holds except possibly for the last statement. Suppose the residue degree is f_1, so that $Norm(\wp_i) = p^{f_1} = |\mathcal{O}_K/\wp_i|$. Then by Fermat's little theorem we have $\alpha^{p^{f_1}} \equiv \alpha$ modulo $\wp_i$, and f_1 is smallest with this property, as the multiplicative group of the finite field $\mathcal{O}_K/\wp_i$ is cyclic. On the other hand, by the definition of f, we have $\zeta_n^{p^f} = \zeta_n$ and since any $\alpha \in \mathcal{O}_K$ can be written as $\sum a_i\zeta_n^i$, we have $\alpha^{p^f} \equiv \alpha$ modulo $\wp_i$. So $f_1 \leq f$. But if $f_1 < f$, then $\zeta_n^{p^{f_1}}$ is distinct from ζ_n and $\wp_i$ divides $\zeta_n^{p^{f_1}} - \zeta_n$ which occurs as a factor of the discriminant, and hence p divides a discriminant, which is a contradiction proving $f_1 = f$ as claimed.

[2] The text of these lectures was not available for these proceedings. – Editors

Let us see how the *quadratic reciprocity law*: $(p/q)(q/p) = (-1)^{(p-1)(q-1)/4}$ for odd distinct primes p and q gets a natural proof when you compare the factorization of a prime in a quadratic field above with the recipe of the cyclotomic field factorization applied to its quadratic sub-extension:

We have, $1 = ((-1)^{(p-1)/2}p/q) = (-1)^{(p-1)(q-1)/4}(p/q)$ if and only if q splits in the quadratic field K of discriminant $d := (-1)^{(p-1)/2}p$ if and only if K is the unique quadratic subfield of the decomposition field, which is of degree g if and only if g is even if and only if $q^{(p-1)/2} \equiv (q^f)^{g/2} \equiv 1^{g/2} \equiv 1$ modulo p if and only if $(q/p) = 1$.

We can also see this without using the concepts of the decomposition field, its degree and so on: If q splits in quadratic extension, then since each of those two conjugates primes decompose the same way in the cyclotomic field containing this quadratic field, the corresponding g is even (in other words, g multiplies in tower) and hence q is a quadratic residue modulo p as above. By symmetry, this takes care of all cases except possibly $p \equiv q \equiv 3$ modulo 4 and (p/q) and (q/p) are 1. We leave this as an exercise (use that f also multiplies in tower together with $q^{(p-1)/2} \equiv 1$ and p does not split in $\mathbf{Q}(\sqrt{-q})$ to get a contradiction).

Another 'quick' and natural way: Let $q = Q_1Q_2$. Let h be the class number of K and write $Q_1^h = ((x + \sqrt{d}y)/2)$. Multiplying by the conjugate (i.e., taking norm), we get $\pm 4q^h = x^2 - dy^2 \equiv x^2$ modulo p. Now 4 is a square, -1 is a square modulo p, if $p \equiv 1$ modulo 4 and otherwise the right hand side is positive, so that the sign on left is also positive. Hence q^h is a square modulo p. Now it is a (hard) fact of the genus theory of quadratic fields (see TIFR pamphlet for the details) that h is odd (when the discriminant of the quadratic field consists of a single prime). So this implies that q is a square modulo p as required. This proof is harder, but shows the reciprocity connection quickly and uses only quadratic theory.

Another way uses index 2 (rather than degree 2) subfields of the cyclotomic fields and norms: Note that $(p-1)/2$ occurring in the quadratic reciprocity law is the degree of the index 2 subfield $\mathbf{Q}(\zeta_p)^+$ and consider the compositum L of $\mathbf{Q}(\zeta_p)^+$ and $\mathbf{Q}(\zeta_q)^+$. Then $\pi_p := (1-\zeta_p)(1-\zeta_p^{-1})$ is the norm of $1-\zeta_p$ from L to $\mathbf{Q}(\zeta_p)^+$. Define π_q similarly. Let $N(x)$ denote the norm of x from L to $\mathbf{Q}$.

Claim: $N(\pi_p - \pi_q) = (p/q)$. Assuming this, the quadratic reciprocity follows by interchanging p and q, as $N(-1) = (-1)^{[L:\mathbf{Q}]} = (-1)^{(p-1)(q-1)/4}$.

Proof: We have $\eta := \pi_p - \pi_q \equiv \pi_p$ modulo (π_q). For any $\sigma \in Gal(L/\mathbf{Q})$, we have $(\pi_q)^\sigma = (\pi_q)$ as ideal, as $\zeta_q^\sigma = \zeta_q^a$ for some a prime to q. Hence $N(\eta) \equiv N(\pi_p)$ modulo (π_q) and hence modulo q, as both the sides are in $\mathbf{Q}$.

But now $N(\pi_p)$ is the norm from $\mathbf{Q}(\zeta_p)^+$ to $\mathbf{Q}$ of $\pi_p^{(q-1)/2}$ so equals $p^{(q-1)/2}$ which is congruent to (p/q) modulo q. This congruence implies equality, because $N(\eta) = \pm 1$, as $\eta = \zeta_q^{-1}(1 - \zeta_p\zeta_q)(1 - \zeta_p^{-1}\zeta_q)$ is a unit.

We will see one more proof using quadratic Gauss sums in Adhikari's lectures.

Finally, we *compare unit groups and class groups* of $K := \mathbf{Q}(\zeta_n)$ to those of $K^+ := \mathbf{Q}(\zeta_n)^*$: For a number field K, let U_K denote its unit group, μ_K its subgroup consisting of all roots of unity in K, C_K its class group and $C_K^{(p)}$ its p-primary part.

Claim: The index $[U_K : \mu_K U_{K^+}]$ is 1 or 2 according as whether n is a prime power or not.

Proof: Consider the homomorphism $\psi : U_K \to \mu_K$ defined by $\psi(u) = \overline{u}/u$. (Recall that algebraic integer with all its absolute values being one is a root of unity). This induces injective homomorphism from U_K/U_{K^+} to μ_K. It follows from the definition that $\psi(U_{K^+}\mu_K) = \mu_K^2$. Since μ_K/μ_K^2 is of order 2, it follows that the index is 2 or 1 according as whether ψ is surjective or not. If n is not a prime power, then $1 - \zeta_n$ is a unit which maps to $-\zeta_n^{-1}$, and hence ψ is surjective. On the other hand, suppose n is a prime power, and ψ is surjective, with $\psi(u) = -\zeta_n^{-1}$. Put $\alpha := (1 - \zeta_n)/u$. Then $\overline{\alpha} = \alpha$, so that α is real. But α being a prime element of K can not lie in K^+. This contradiction finishes the proof of the claim.

The unit groups in K and K^+ are thus not much different and thus the regulators are essentially the same and thus taking ratio of the corresponding zeta functions we get a formula for the relative class number, by getting rid of the usually hard to handle regulators.

To start comparing the class groups, we start with a weaker result in more general situation:

Claim: If L is a Galois extension of degree d of a number field K, and if a prime p does not divide d, then the natural map $C_K^{(p)} \to C_L^{(p)}$ is injective and the map $C_L^{(p)} \to C_K^{(p)}$ induced by the norm is surjective.

Proof: If I is an ideal of $\mathcal{O}_K$ representing a class in $C_K^{(p)}$ such that $I\mathcal{O}_L = (\alpha)$, then $I^d = (Norm_K^L(\alpha))$ is principal, which implies I is principal, as p does not divide d. This proves the first part. On the other hand, if p does not divide d, every element in $C_K^{(p)}$ is a d-th power, this proves surjectivity, since $I^d = Norm(I\mathcal{O}_L)$.

Claim: Natural map $C_{K^+} \to C_K$ is injective (so h^+ divides h).

Proof: If I is an ideal of $\mathcal{O}_{K^+}$, such that $I\mathcal{O}_K = (\alpha)$, then $\overline{I\mathcal{O}_K} = (\overline{\alpha}) = I\mathcal{O}_K$. Hence $\alpha/\overline{\alpha}$ is a unit with all its conjugates having absolute value 1 and hence it is a root of unity. If n is not a prime power, since ψ

is surjective, we can write it as $\overline{u}/u$. This implies that αu is real, but then $I = (\alpha u)$ proves what we want. If n is a prime power, put $\lambda := 1 - \zeta_n$. Then $\lambda/\overline{\lambda} = -\zeta_n$, which is a generator of μ_K. Hence $\overline{\alpha}/\alpha = (\lambda/\overline{\lambda})^d$ for some d. Now λ-adic valuation takes even values on K^+, and $\alpha\lambda^d$ and I are real. Hence, $d = v_\lambda(\alpha\lambda^d) - v_\lambda(\alpha) = v_\lambda(\alpha\lambda^d) - v_\lambda(I)$ is even. This implies $\alpha/\overline{\alpha} = (-\zeta_n)^d \in \mu_K^2$ and hence equals $\zeta/\overline{\zeta}$ by above. This means $\alpha\zeta$ is real and $I = (\alpha\zeta)$ finishes the proof.

(Exercise) Compare the questions and arguments above with those encountered in Hilbert 90, in Narlikar and Nitsure lectures.

Dinesh Thakur
Department of Mathematics
University of Arizona
Tucson, AZ 85721
USA
e-mail: thakur@math.arizona.edu

Absolute Values and Completions

B.Sury

This article is in the nature of a survey of the theory of complete fields. It is not exhaustive but serves the purpose of familiarising the readers with the basic notions involved. Hence, complete (!) proofs will not be given here. It is no surprise that algebraic number theory benefits a lot from introducing analysis therein. The familiar notion of construction of real numbers is just one aspect of this facility.

§ 1. **Discrete valuations**

Definition 1.1. Let K be any field. A surjective map $v : K^* \to \mathbf{Z}$ is called a *discrete valuation* if:

$$v(xy) = v(x) + v(y),$$

$$v(x+y) \geq \mathrm{Inf}(v(x), v(y))$$

Here, for notational purposes, one also defines $v(0) = \infty$. Note also that one must have $v(1) = 0 = v(-1)$.

Premier example 1.2. For each prime number or, more generally, for any non-zero prime ideal P in a Dedekind domain A, one has the P*-adic valuation* v_P given by the prescription $v_P(x) = a$ where the fractional principal ideal $(x) = P^a I$ with I coprime to P. This is a discrete valuation on the quotient field K of A.

Lemma 1.3. *(a) If v is a discrete valuation on a field K, then $A_v := \{x \in K : v(x) \geq 0\}$ is a local PID. Its maximal ideal is $P_v = \{x \in K : v(x) > 0\}$.* (A_v is called the *valuation ring* of v).
(b) For a discrete valuation v on a field K, if k_v denotes the residue field A_v/P_v and $U_i = 1+P_v^i$ for $i > 0$, then $A_v^*/U_1 \cong k_v^*$ and $U_i/U_{i+1} \cong P_v^i/P_v^{i+1} \cong k_v^+$.
(c) If A is a Dedekind domain, v a discrete valuation on its quotient field K and, $A \subset A_v$, then $P := A \cap P_v$ is a non-zero prime ideal of A. Moreover, $v = v_P, PA_v = P_v, A/P \cong A_v/P_v$.

Proof. Quite easy.

Exercise 1.4. *Let v be a discrete valuation on a number field K. Then $\mathcal{O}_K \subseteq A_v$.*

Proposition 1.5. *On a number field K, the map $P \mapsto v_P$ sets up a bijection between non-zero prime ideals of $\mathcal{O}_K$ and discrete valuations.*

Indication of proof. The proof follows from the easily proved step: If $A \subseteq B$ are discrete valuation rings with the same quotient field K. Then $A = B$.

Proposition 1.6. *Let F be any field and $K = F(X)$, the function field in one variable over F. Define $v_\infty(f/g) = deg(g) - deg(f)$. Then,*
(a) v_∞ defines a discrete valuation on K which is zero on F^*,
(b) v_P, as P runs through the prime ideals of $F[X]$ along with v_∞ exhaust all the possible discrete valuations on K that are trivial on F^*,
(c) (Product formula) For each $f \in K^*$, one has

$$v_\infty(f) + \sum_P f_P v_P(f) = 0$$

where P runs through the non-zero prime ideals of $F[X]$ and $f_P = [F[X] : P]$ is the degree of any polynomial generating P.

Proof. (a) is obvious.
(b) Let v be any discrete valuation on K which is trivial on F^*. First, suppose that $v(X) \geq 0$. Then, $v(f) \geq 0 \; \forall \; f \neq 0 \in F[X]$. As v surjects onto integers, there is some monic irreducible polynomial f such that $v(f) > 0$. If $v(g) > 0$ for another monic, irreducible polynomial g, then $v(1) = v(sf + tg) > 0$, which is a contradiction. Thus, $v(g) = 0$ for all monic irreducible polynomials $g \neq f$. Thus, writing any $h \in F[X]$ as a product of irreducibles, one gets $v(h) \in v(f)\mathbf{Z}$. As v is surjective, $v(f) = 1$ i.e., $v = v_f$. Therefore, we have shown that if $v(X) \geq 0$, then $v = v_P$ for some non-zero prime ideal P.

If $v(X) < 0$, it is easy to see by induction on the degree that $v(h) = v(X)\deg(h)$ for any $h \in F[X]$. By surjectivity again, one gets $v(X) = -1$ and so $v = v_\infty$.
(c) Finally, writing any $f \in K^*$ as $f = u \prod_i p_i^{v_{p_i}(f)}$ and comparing degrees, one gets the product formula.

§ 2. **Absolute values**

Definition 2.1. On a field K, an *absolute value* is a function $| \; |: K \to \mathbf{R}^{\geq 0}$ such that
(a) $| \, x \, | \; = 0 \Leftrightarrow x = 0$,
(b) $| \, xy \, | \; = \; | \, x \, || \, y \, |$, and
(c) $| \, x + y \, | \; \leq \; | \, x \, | + | \, y \, |$.

Remarks and examples 2.2. (a) Clearly, an absolute value on a field defines a metric on it.

We shall always omit from consideration the *trivial absolute value* which is $\equiv 1$ on K^*.

Easy exercise: On a finite field, show that the only absolute value is the trivial one. What does this give in relation to proposition 1.6?

(b) $| \quad |$ is called *a non-archimedean absolute value* if

$$| x + y | \quad \leq \text{Max}(| x |, | y |).$$

This is stronger than the property 2.1(c).

Trivial exercise: Why is the word non-archimedean used here?
An absolute value which is not non-archimedean is called archimedean!

(c) If v is a discrete valuation on K, then for any fixed positive $\lambda < 1$, the prescription $| x |= \lambda^{v(x)}$ gives a non-archimedean absolute value. Note that the value group $| K^* |$ is discrete in $\mathbf{R}^{\geq 0}$.
Exercise: An absolute value on a field K has a value group $| K^* |$ which is discrete if, and only if, it arises from a discrete valuation on K. (*Hint:* If $| K^* |$ is discrete, choose the maximal element $\lambda \in | K^* | \cap (0, 1)$.)

If $| \quad |$ is a discrete absolute value on K, one notes that the corresponding valuation ring and its maximal ideal are, respectively, $\{x \in K :| x |\leq 1\}$ and $\{x \in K :| x |< 1\}$. A generator of P is often called a *uniformising parameter.*

(d) If K is any field and $\sigma : K \to \mathbf{C}$ any embedding, then $| x |_\sigma := \quad | \sigma(x) |$ defines a nontrivial absolute value on K. Here the right side has the usual absolute value on $\mathbf{C}$. This is archimedean.

(e) The square of the usual absolute value on $\mathbf{C}$ is *not* an absolute value. However, if $| \quad |$ is a non-archimedean absolute value on a field K, so is $| \quad |^t$ for any positive real t.

Definition 2.3. Two absolute values $| \quad |_1$ and $| \quad |_2$ on K are said to be *equivalent* if $\exists\ t > 0$ such that $| x |_1 =| x |_2^t$ for all $x \in K$.

Exercise: Two absolute values are equivalent if, and only if, they define equivalent topologies.

2.4. Product formula over Q. Let us apply the above generalities to $\mathbf{Q}$. We have the archimedean absolute value $| \quad |_\infty$ coming from the inclusion of $\mathbf{Q}$ in $\mathbf{R}$. For each prime number p, we have the p-adic absolute value which we normalise as follows. Define $| p |_p = 1/p$ i.e., we have taken $\lambda = 1/p$ in

2.2(c). Then, we have, for each $x \in \mathbf{Q}^*$,

$$| x |_\infty \prod_p | x |_p = 1.$$

That this is a product formula analogous to 1.6(c) for function fields is justified by the following easy result:

Theorem (Ostrowski) 2.5. *Any non-trivial absolute value on* $\mathbf{Q}$ *is equivalent exactly to one of* $| \ \ |_\infty$ *or* $| \ \ |_p$ *for some prime* p.

Sketch of proof. Suppose $| \ \ |$ is any absolute value. If $| n | \leq 1$ for all integers n, it is easy to prove that $| \ \ | = | \ \ |_p$ for some prime p. This is just as in the proof of 1.6. Now, suppose that there is a positive integer n with $| n | > 1$. Write $| n | = | n |_\infty^t = n^t$ for some $t > 0$. Use the n-adic expansion to show this holds (with the same t) for any integer in place of n.

Exercise 2.6. *(a) An absolute value on a field* K *is non-archimedean if, and only if,* $| \mathbf{Z}.1_K |$ *is bounded.*
(b) If $Char(K) > 0$*, then any absolute value on* K *is non-archimedean.*
(c) Any discrete absolute value is non-archimedean.
(d) The restriction of a nontrivial absolute value on a number field to $\mathbf{Q}$ *is again nontrivial.*
(e) An absolute value $| \ \ |$ *is non-archimedean if, and only if,* $| z | \ < 1$ *implies that* $| 1 + z | \ < 1$.

Corollary 2.7. *Any nontrivial absolute value on an algebraic number field* K *is equivalent to exactly one of the archimedean ones coming from the various embeddings of* K *in* $\mathbf{C}$ *or to a discrete one coming from a prime ideal of* $\mathcal{O}_K$.

Proof. This follows from 1.5,2.5 and,2.6(d).

Remarks 2.8. The non-archimedean absolute values have proprties which look strange in the first instance as we are used to the usual notion of absolute value coming from the reals which is archimedean. For instance, a series converges if, and only if, its n-th term tends to 0 (!) Any triangle is isosceles (!) Every point inside a circle is its centre (!) etc.

§ 3. **Completions**

Definition 3.1. Let $(K, | \ \ |)$ be a field with an absolute value. A *completion* of $(K, | \ \ |)$ is an absolute-valued field $(L, | \ \ |_L)$ which is complete as a metric space and has the property that there is some embedding $i : K \to L$ with the image of K dense and $| x | = | i(x) |_L$ for $x \in K$.

Proposition 3.2. *Each $(K, |\quad|)$ has a completion. Further, if $(L, |\quad|)$ and $(L', |\quad|')$ are two completions where $i : K \to L$ and $i' : K \to L'$ are corresponding embeddings, then there is an isomorphism $\sigma : (L, |\quad|) \to (L', |\quad|')$ of absolutely-valued fields such that $i' = \sigma \circ i$.*

The proof will not be given here but the argument is entirely analogous to the construction of the reals from the rationals in terms of Cauchy sequences.

Corollary 3.3. *Let $(K, |\quad|)$ be an absolutely-valued field and $(\hat{K}, |\quad|_0)$ its completion. Then, $|\quad|$ is non-archimedean if, and only if, $|\quad|_0$ is so. Moreover, in this case, the value groups of K and $\hat{K}$ are the same.*

Proof. The proof is a direct consequence of the construction of $\hat{K}$.

Exercise: Prove this without using the construction.

Theorem 3.4. (Gelfand-Tornheim-Ostrowski) *Any field k which is complete with respect to an archimedean absolute value is isomorphic to* $\mathbf{R}$ *or* $\mathbf{C}$ *as absolutely-valued fields.*

Proof. For a proof, see Cassels' *Local fields.*

Proposition 3.5. (Series expansion) *Suppose $(k, |\quad|)$ is complete with respect to a discrete absolute value. Denote by $A$and P the corresponding valuation ring and its maximal ideal. Fix a set of representatives Σ in A for the residue field A/P. Then, for any uniformising parameter π, elements α of k admit Laurent series expansions of the form $\sum_{i=-n}^{\infty} a_i \pi^i$ where the 'digits' $a_i \in \Sigma$ of α are uniquely determined.*

Proof

For any $\alpha \in k^*$, one has $\pi^n \alpha \in A$ for some n. So, it suffices to show that each $\alpha \in A$ has an expansion as claimed. By the very definition of Σ, there is $a_0 \in \Sigma$ such that $\alpha - a_0 \in P$. So, $\alpha = a_0 + \pi\alpha_1$. Continuing with α_1 and so on, one gets a series expansion. It makes sense as the n-th term tends to 0. Uniqueness is easy to prove.

Example 3.6. Look at the completion $\mathbf{Q}_p$ of $\mathbf{Q}$ with the p-adic absolute value. Its valuation ring is usually denoted by $\mathbf{Z}_p$. One calls $\mathbf{Q}_p$ and $\mathbf{Z}_p$ the p-adic numbers and the p-adic integers respectively. Note that p is a uniformising parameter and Σ can be taken to be the finite set $\{0, 1, \cdots, p-1\}$. Thus, every p-adic number has a unique expansion as $\sum_{-n}^{\infty} a_i p^i$ where its 'digits' a_i are between 0 and $p-1$. Note the analogy with the decimal expansions of real numbers. The only difference here is that there are infinitely many positive powers of p and only finitely many negative powers. So, it is worthwhile to think of p as $1/10$.

Lemma 3.7. *Suppose $(k, |\ \ |)$ is complete with respect to a discrete absolute value. Denote by A and P the corresponding valuation ring and its maximal ideal. Then, k is locally compact if, and only if, A is compact which is again if, and only if, A/P is finite.*

Proof. If A is compact, then evidently k is locally compact since $k = \cup_n \pi^{-n} A$. Assume k is locally compact. Let C be a compact neighbourhood of 0. Then, for large enough n, $\pi^n A \subseteq C$. As $\pi^n A$ is closed, it is compact also. Thus, we have shown the equivalence of compactness of A and local compactness of k.
If A is compact, then from the openness of P in k, we get that A/P is compact as well as discrete and therefore, finite. To prove finally that the finiteness of A/P implies the compactness of A, it suffices to prove sequential compactness as A is a metric space. Let $\{a^{(n)}\}$ be any infinite sequence in A. Write the series expansion $a^{(n)} = \sum_{r \geq 0} a_{n,r} \pi^r$. As n varies, the elements $a_{n,0}$ run over a finite set (viz., a set of representatives of A/P). Thus, they are all equal for infinitely many n. Replace the original sequence with a subsequence for which the terms $a_{n,0}$ are all the same, say a_0. Proceeding this way, one finally concludes that there is a subsequence of the original sequence which converges to an element of A.

Hensel's lemma 3.8. *Suppose $(k, ||)$ is complete with respect to a discrete absolute value. Denote by A and P the corresponding valuation ring and its maximal ideal. If $f(X) \in A[X]$ is a polynomial which factors modulo P into two coprime polynomials $\bar{g}, \bar{h}$, then there exist $g, h \in A[X]$ such that $f = gh$ and $deg(g) = deg(\bar{g})$.*

Exercises 3.9. *(a) Prove Hensel's lemma.*
(b) Find the order and structure of $\mathbf{Q}_p^*/(\mathbf{Q}_p^*)^2$.
(c) Prove that the only automorphism of $\mathbf{Q}_p$ is the identity.

Let K be an algebraic number field. Start with a discrete absolute value on it (this will come from a prime ideal). Let A be the corresponding valuation ring and P its maximal ideal. If L is a finite extension of K and B the integral closure of A in L, one can write $PB = P_1^{e_1} \cdots P_g^{e_g}$. Let Q denote one of the P_i's. Let K_P and L_Q denote the completions of K and L with respect to the P-adic and the Q-adic absolute values. If $\hat{A}, \hat{B}$ denote their valuation rings and $\hat{P}, \hat{Q}$ their maximal ideals, it is routine to prove:
Exercise 3.10. *(a) $\hat{P} = P\hat{A}, \hat{Q} = Q\hat{B}$,*
(b) $\hat{P}\hat{B} = \hat{Q}^{e_1}$,

(c) $[L_Q : K_P] = e_1 f_1$.

The next proposition is crucial to many of the results to follow.

Proposition 3.11. (Extensions of valuations over complete fields) *Let $(K, |\ |)$ be a complete field. If L is a finite extension of K, then there is exactly one absolute value on L which extends $|\ |$. Moreover, L is complete with respect to it.*

Proof. The archimedean case is taken care of by Theorem 3.4. So, we assume that the absolute value is non-archimedean. Let us first prove the existence of an extension. Define $| x |_L = | N_{L/K}(x) |$ for any $x \in L$. The first two properties are clear and we only need to prove that if $x \in L$ satisfies $| x |_L \leq 1$, then $| 1 + x |_L \leq 1$. In other words, if $| N_{L/K}(x) | \leq 1$, then $| N_{L/K}(1 + x) | \leq 1$. Let $f(T) = a_0 + a_1 T + \cdots + T^n$ be the minimal polynomial of x over K. Now, $| a_0 | = | N_{L/K}(x) | \leq 1$ i.e., $a_0 \in A$, the valuation ring of K. Now, $g(T) = f(T-1)$ is clearly the minimal polynomial of $1 + x$ over K. Therefore,

$$| N_{L/K}(1 + x) | = | g(0) | = | f(-1) | = | a_0 - a_1 + a_2 - \cdots | .$$

So, if we show that $| a_i | \leq 1$, it would follow that $| N_{L/K}(1 + x) | \leq 1$. Suppose the contrary. Let a_r be such that $| a_r | > 1$, that $| a_r | = M :=$ Max $(| a_i |)$ and that r is the maximal index i so that $| a_i | = M$. With this notation, we have $a_r^{-1} \in A$ and $a_i a_r^{-1} \in A$ for all i and $a_i a_r^{-1} \in P$ for all $i > r$, where P is the maximal ideal of A. Thus, the polynomial $a_r^{-1} f(T) \in A[X]$ reduces modulo P to the polynomial $\bar{h}(T) = X^r +$ smaller degree terms. Applying Hensel's lemma to the factorisation $\bar{h}(T) U(T)$ where U is the constant polynomial 1, we have $f = h(T) u(T)$ for some lifts such that $h(T)$ mod P is $\bar{h}(T)$. But, as $r < n$, this means that f is reducible, a contradiction, which implies that all $a_i \in A$. This proves the existence.
We prove the uniqueness when K is locally compact, which is the main case of interest to us. The general case is not too difficult and one can look at Cassels's book (loc. cit.). Let $\{v_1, \cdots, v_n\}$ be a K-basis of L. We claim that any extension $|\ |_L$ is equivalent to *the sup-norm* $|\ |_0$ with respect to this basis.

Firstly, $| x |_L = | a_1 v_1 + \cdots + a_n v_n |_L \leq n \ \sup_i (| a_i |_0) \ | x |_0$. Here, we haven't used the local compactness but we shall use it for the opposite implication. By the local compactness of K, there is some $y \in L$ such that $| y |_L = \text{Min}(| x |_L : | x |_0 = 1)$. Now, let $0 \neq x \in L$. Write $x = a_1 v_1 + \cdots + a_n v_n$. If $| a_r |_L = | x |_0 = Max(| a_i |_L)$, then $x = a_r z$ with $| z |_0 = 1$.

So, $| y |_L \leq | z |_L = | x / a_r |_L = | x |_L \ / \ | a_r |_L = \ | x |_L \ / \ | x |_0$. In other words, $\mid x\mid_0 \leq \mid x \mid_L (1/\mid y\mid_L)$. This proves that $\mid \quad \mid_L$ and the sup-norm $\mid \quad \mid_0$ are equivalent and proves the proposition.

Corollary 3.12. (Unramified extensions)
Suppose $(k, \mid \quad \mid)$ is complete with respect to a discrete absolute value. Denote by A and P the corresponding valuation ring and its maximal ideal. Let l be a finite extension of degree n over k. Let f and F denote the residue fields of k, l respectively. Then, the association $e \mapsto (e \cap A)$ mod P is a bijection from $\{e : k \subset e \subset l$ and e unramified over $k\}$ to $\{E : f \subset E \subset F\}$.

In particular, there is a unique (upto isomorphism) unramified extension of any degree d viz., the splitting field over k of $X^{q^d} - X$ where $q = \#f$.
The proof is a consequence of Hensel's lemma (*Exercise:* What is the polynomial factorisation to which Hensel is applied?) and the fact that over finite fields there is a unique extension, upto isomorphism, of a given degree.

Definition 3.13. If $\mid \quad \mid$ is a discrete absolute value on k, an *Eisenstein polynomial* is a polynomial $f \in k[X]$ of the form $\sum_{i=0}^n a_i X^i$ with $a_i \in P$ for $i < n$, a_n a unit and $a_0 \in P \setminus P^2$. It is an easy exercise to show that such a polynomial is irreducible.

Proposition 3.14. (Totally ramified extensions) *Suppose $(k, \mid \quad \mid)$ is complete with respect to a discrete absolute value. Let A, P, π have the usual meaning. Then, an extension of k is totally ramified if, and only if, it is obtained by attaching a root of an Eisenstein polynomial.*

Proof. Suppose that α is a root of an Eisenstein polynomial $f(X) = \sum_{i=0}^n a_i X^i$. Then $\sum_{i=0}^n a_i \alpha^i = 0$ and so

$$\mid \alpha^n \mid \quad = \quad \mid a_n \alpha^n \mid \quad = \quad \mid \textstyle\sum_{i=0}^{n-1} a_i \alpha^i \mid \quad = \quad \mid a_0 \mid \quad = \quad \mid \pi \mid.$$

Thus, $k(\alpha)$ is totally ramified extension of k.
Conversely, suppose K is totally ramified over k of degree n. If Π_K is a uniformising parameter for K, then the powers $\Pi_K^i, i < n$, must be linearly independent over k as total ramification forces their absolute values to be in distinct cosets of the value groups of k in K. Thus, they form a k-basis of K. Write $\Pi_K^n + a_{n-1}\Pi_K^{n-1} + \cdots + a_0 = 0$ with $a_i \in k$. But, the various roots of this polynomial give extensions of $\mid \quad \mid$ to K and must coincide by the uniqueness of such an extension. In other words, the roots of this polynomial have absolute value $\mid \Pi_K \mid$. As each a_i is a sum of roots, we have $\mid a_i \mid < 1$ and $\mid a_0 \mid = \mid$ product of the roots $\mid \quad = \quad \mid \Pi_K^n \mid \quad = \quad \mid \pi \mid$. In other words, the polynomial $\sum a_i X^i$ is an Eisenstein polynomial. The proposition is proved.

Krasner's lemma 3.15. *Suppose $(k, \mid \quad \mid)$ is complete with respect to a discrete absolute value. Let α, β be algebraic over k and suppose that α is* *separable over $k(\beta)$. Assume that β is 'very close' to α in the sense that $\mid \beta - \alpha \mid < \mid \sigma(\alpha) - \alpha \mid$ for all k-isomorphisms of $k(\alpha)$. Then, $k(\alpha) \subseteq k(\beta)$.*

Proof. By the separability assumption, it suffices to show the conclusion that each $k(\beta)$-isomorphism τ of $k(\alpha, \beta)$ fixes α. Note that any such τ gives a new absolute value on $k(\alpha, \beta)$ by $\mid x \mid_\tau = \quad \mid \tau(x) \mid$. By the uniqueness, this gives that the hypothesis implies $\mid \tau(\beta - \alpha) \mid < \mid \sigma(\alpha) - \alpha \mid$. That is, $\mid \beta - \tau(\alpha) \mid < \mid \sigma(\alpha) - \alpha \mid$. So, $\mid \tau(\alpha) - \alpha \mid < \mid \sigma(\alpha) - \alpha \mid$. In other words, $\tau(\alpha) = \alpha$. The lemma follows.

Definitions and remarks 3.16 (continuity of roots) With k as before, let $f(X) \in k[X]$ be a monic polynomial of degree n which factorises as $\prod_{i=1}^{t}(X - a_i)^{r_i}$ in the algebraic closure of k. Let us define $\mid f \mid$ to be the maximum of the absolute values of the coefficients of f. Clearly, if $g \in k[X]$ is close to f i.e., if $\mid f - g \mid$ is small, then for any root b of g, the value $\mid f(b) \mid = \mid f(b) - g(b) \mid$ is small. In other words, as g comes close to f, any root of g comes close to some root of f. It is an easy exercise to see that if g is sufficiently close to f and if $b_1, \cdots, b_r$ are the roots (with multiplicity) of g which come close to a root a_i of f, then $r = r_i$.

Corollary 3.17. *With k, f as above, if f is irreducible and separable, then any monic g which is sufficiently close to f is irreducible too. Moreover, if b is a root of g coming close to a root a of f, then $k(a) = k(b)$ if f, g are sufficiently close.*

Proof. The proof is immediate from 3.15 and 3.16.

Corollary 3.18. *Any finite extension $k = \mathbf{Q}_p(\alpha)$ arises as the closure of a finite extension K of $\mathbf{Q}$ where $[k : \mathbf{Q}_p] = [K : \mathbf{Q}]$.*

Proof. The proof is immediate from choosing a polynomial $g \in \mathbf{Q}[X]$ which is close in the p-adic topology to the minimal polynomial of α over $\mathbf{Q}_p$ and applying 3.17.

Now, we can prove a remarkable theorem (contrast it with the situation of number fields !)

Theorem 3.19. *Any finite extension k of $\mathbf{Q}_p$ has only finitely many extension fields (upto isomorphism) of a given degree.*

Proof. As there is a unique unramified extension of a given degree over any finite extension of $\mathbf{Q}_p$ by Corollary 3.12, it suffices to prove the finiteness of the number of totally ramified extensions of a given degree n. In this case, Proposition 3.14 tells us that any such extension arises from an Eisenstein polynomial of degree n. As such a polynomial has a unit as the top coefficient and other coefficients coming from the maximal ideal P, we have a mapping from the product $U \times P \times \cdots \times P$ to the set of totally ramified extensions of degree n. Here, the factor P is repeated $n-1$ times. The crucial observation is that by 3.17, a neighbourhood of a point in this product determines fields which are all isomorphic. By the compactness of U and P, the theorem follows.

REFERENCES

1. J. W. S. Cassels, *Global Fields, Chapter II in Algebraic Number Theory*, Cassels and Fröhlich Eds., Academic Press (1967).

2. J. W. S. Cassels, *Local Fields*, LMS Student Series 3, Cambridge University Press (1986).

3. Serge Lang, *Algebraic Number Theory*, Springer, Second Edition (1994).

B. Sury
Indian Statistical Institute,
Bangalore 560 059
e-mail: bsury@isibang.ac.in

The Early Reciprocity Laws: From Gauss to Eisenstein

Sukumar Das Adhikari

1. Introduction. We shall start with the law of quadratic reciprocity which was guessed by Euler and Legendre and whose first complete proof was supplied by Gauss. A result central to number theory, the law of quadratic reciprocity, apart from being fascinating on its own, has led to very important generalizations.

The main aim of this article is to sketch a proof of the Eisenstein reciprocity law. Having many applications and being beautiful on its own, the Eisenstein reciprocity law related to the cyclotomic fields, is a precursor of the more general reciprocity laws. Before we move on to discuss about the Eisenstein reciprocity law, we shall have a brief discussion on cubic reciprocity as well. There we shall sketch Williams' proof [10] of Eisenstein's supplement to the law of cubic reciprocity. While for the Stickelberger relation we shall refer to the article of S. A. Katre [6] in this volume, for a deduction of some of the early reciprocity laws from Artin's we refer to that of Parvati Shastri [9]. For the proofs of some results on Gauss and Jacobi sums and, in fact, for many details about the early reciprocity laws including the biquadratic case, we refer to the beautiful book [5] of Ireland and Rosen. We also refer to the interesting expository article of Wyman: "What is a reciprocity law?"[11].

In what follows, for any prime power q, $\mathbf{F}_q$ will denote the finite field with q elements. The symbols $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{C}$ will denote respectively the set of integers, the set of rationals and the complex numbers. For a finite set S, $|S|$ will denote the number of elements of S. For a field K, K^* will denote the multiplicative group of its non-zero elements.

2. Quadratic reciprocity. The problem of solving a general polynomial congruence reduces to that of solving congruences with prime power moduli plus a set of linear congruences. The problem of solving a quadratic congruence $f(x) \equiv 0 \pmod{p}$, where $f(x)$ is a quadratic polynomial with integer coefficients and p is an odd prime, by 'completing the square' reduces to the problem of solving the congruence

$$x^2 \equiv d \pmod{p}, \;\; d \in \mathbf{Z}, \quad p \text{ a rational prime.} \tag{1}$$

The laws of quadratic reciprocity, one of the most celebrated results in all of number theory, give an algorithm for knowing the existence of solutions to the congruence (1). What we shall see is that the laws of quadratic reciprocity describe the set of primes modulo which a quadratic polynomial in $\mathbf{Z}[x]$ splits. In general, results giving similar informations (See [11], for instance) are known as reciprocity laws. However, the term reciprocity attached to the early reciprocity laws had its obvious meaning.

For a rational prime p, other than 2, and for $x \in \mathbf{F}_p^*$, the *Legendre symbol* $\left(\frac{x}{p}\right)$ is defined to be $x^{(p-1)/2}$. It is easy to see that (see Serre [8] or Adhikari [1] for instance) $\left(\frac{x}{p}\right) = 1$ or -1 according as x is a square mod p or not, i.e., $y^2 \equiv x \pmod{p}$ has a solution or not. One says that x is *quadratic residue* or *quadratic non-residue mod* p respectively.

For a rational prime p, other than 2, observing that the index of $\mathbf{F}_p^{*2}$ in $\mathbf{F}_p^*$ is 2, there are as many residues as non-residues mod p. Also, $\left(\frac{x}{p}\right) \cdot \left(\frac{y}{p}\right) = \left(\frac{xy}{p}\right)$, i.e., the Legendre symbol is a character of the mutiplicative group $\mathbf{F}_p^*$.

The definition of $\left(\frac{x}{p}\right)$ is extended to all of $\mathbf{F}_p$ by putting $\left(\frac{0}{p}\right) = 0$ and we can view $\left(\frac{x}{p}\right)$ as a function on $\mathbf{Z}$ in the obvious way.

We now state the laws of quadratic reciprocity where part (iii) is the proper reciprocity law and the first two parts are known as supplementary laws.

Theorem 2.1. (Laws of Quadratic Reciprocity). If p and l are two distinct odd primes,

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

ii) $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.

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

The result is amazing because there is no obvious reason to expect any connection between the symbols $\left(\frac{p}{l}\right)$ and $\left(\frac{l}{p}\right)$, or, in other words, between the congruences $x^2 \equiv p \pmod{l}$ and $x^2 \equiv l \pmod{p}$. We shall now indicate a proof of part (iii); for complete proofs one may refer to [1] or [8] mentioned above. In fact, almost any number theory text will contain one or more proofs of it. The book [5] of Ireland and Rosen mentioned above contains at least three different proofs of the theorem. We also refer to the interesting book [2] of Cox, where for a given positive integer n, the various reciprocity laws are seen to bear upon answering the question of finding the primes p which can be expressed in the form $p = x^2 + ny^2$. In fact, Euler's discovery of quadratic reciprocity was prompted by such questions.

Let ω denote a primitive l-th root of unity in an algebraic closure Ω of $\mathbf{F}_p$. We consider the sum $S = \sum_{x \in \mathbf{F}_l^*} \left(\frac{x}{l}\right) \omega^x$.

$$\begin{aligned} \text{We have, } S^2 &= \sum_{x,y \in \mathbf{F}_l^*} \left(\frac{xy}{l}\right) \omega^{x+y} \\ &= \sum_{y,z \in \mathbf{F}_l^*} \left(\frac{y^2 z}{l}\right) \omega^{y(z+1)} \quad \text{(Putting x = yz)} \\ &= \sum_{y,z \in \mathbf{F}_l^*} \left(\frac{z}{l}\right) \omega^{y(z+1)} \\ &= \sum_{y \in \mathbf{F}_l^*} \left(\frac{-1}{l}\right) \omega^0 + \sum_{z \neq -1} \left(\frac{z}{l}\right) \sum_{y \in \mathbf{F}_l^*} \omega^{y(z+1)} \\ &= \left(\frac{-1}{l}\right)(l-1) + (-1) \sum_{z \neq -1} \left(\frac{z}{l}\right) \end{aligned}$$

(Since, $\sum_{y \in \mathbf{F}_l^*} \omega^{y(z+1)} + 1 = 1 + \omega + \cdots + \omega^{l-1} = 0$),

$$\text{so,} \quad S^2 = l \left(\frac{-1}{l}\right) - \sum_{z \in \mathbf{F}_l^*} \left(\frac{z}{l}\right).$$

Now, there are as many squares as non-squares in $\mathbf{F}_l^*$, so $\sum_{z \in \mathbf{F}_l^*} \left(\frac{z}{l}\right) = 0$ and hence

$$S^2 = l \left(\frac{-1}{l}\right). \tag{2}$$

and

$$S^{p-1} = \left(\frac{p}{l}\right). \tag{3}$$

From (2) and (3),

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

This proves the equality in (iii) modulo p. Since p is odd, (iii) follows.

The following remark is not out of place.

Remark 2.1. In the proof above, if we replace ω by a primitive l-th root of unity in an algebraic closure of the rationals $\mathbf{Q}$, then defining S in the same way, S will again satisfy equation (2), that is, $S^2 = \pm l$. Thus, observing that $\mathbf{Q}(\sqrt{-1}) = \mathbf{Q}(\zeta_4)$ (where ζ_n is a primitive n-th root of unity in $\bar{\mathbf{Q}}$), square root of any odd prime is contained in $\mathbf{Q}(\zeta)$ for some root of unity ζ. Further, observing that $\sqrt{2} \in \mathbf{Q}(\zeta_8)$ (for $2 = -i(1+i)^2$), it follows that any quadratic extension K of $\mathbf{Q}$ is contained in $\mathbf{Q}(\zeta)$ for a root of unity ζ, thus giving an easy special case of the Kronecker-Weber theorem (see Ghate [3], in this volume).

3. Cubic reciprocity. Questions regarding solutions of the congruence $x^n \equiv a \pmod{p}$ for rational primes p for larger n's led Gauss to formulate the cubic and biquadratic reciprocities corresponding to $n = 3$ and 4 respectively. In 1844, Eisenstein was first to publish complete proofs of these theorems. In this section, we give a quick sketch of the cubic reciprocity law. On our way, we shall come across Gauss and Jacobi sums. For the details not supplied here, one may look into [5].

Writing $\omega = (-1 + \sqrt{-3})/2$, we consider the ring

$$\mathbf{Z}[\omega] = \{a + b\omega \mid a, b \in \mathbf{Z}\}.$$

For $a+b\omega \in \mathbf{Z}[\omega]$, with the norm function defined by $N(a+b\omega) = a^2-ab+b^2$, $\mathbf{Z}[\omega]$ is a Euclidean domain. The units in $\mathbf{Z}[\omega]$ are elements α with $N(\alpha) = 1$ and they are $\pm 1, \pm\omega, \pm\omega^2$. If p is a rational prime such that $p \equiv 2 \pmod{3}$, then p remains prime in $\mathbf{Z}[\omega]$. The rational primes $p \equiv 1 \pmod{3}$, split into a product of a pair of primes complex conjugate to each other. The rational prime 3 has the factorization $3 = -\omega^2(1-\omega)^2$ where $1-\omega$ is a prime in $\mathbf{Z}[\omega]$.

If $\pi \in D = \mathbf{Z}[\omega]$ is a prime, then $D/\pi D$ is a finite field with $N(\pi)$ elements and for an element $\alpha \in D$ coprime to π,

$$\alpha^{N(\pi)-1} \equiv 1 \pmod{\pi}.$$

If the norm of π is different from 3, it is not difficult to see that the residue classes of the elements $1, \omega$ and ω^2 are distinct mod πD and therefore, $\{1, \omega, \omega^2\}$ being a subgroup of order 3 of the multiplicative group $(D/\pi D)^*$, $(N(\pi)-1)/3$ is an integer. Now, with π and α as above, $1, \omega$ and ω^2 being distinct mod πD, from the identity

$$\alpha^{N(\pi)-1} - 1 = (\alpha^{(N(\pi)-1)/3} - 1)(\alpha^{(N(\pi)-1)/3} - \omega)(\alpha^{(N(\pi)-1)/3} - \omega^2),$$

it follows that $\alpha^{(N(\pi)-1)/3}$ is congruent to exactly one of the elements $1, \omega$ or ω^2 modulo π in D.

If π is a prime in D with $N(\pi) \neq 3$, then, if π is coprime to α, the unique element to which $\alpha^{(N(\pi)-1)/3}$ is congruent modulo π, is defined to be the *cubic residue character* of α modulo π and we use the notation $(\alpha/\pi)_3$ or $\chi_\pi(\alpha)$ for it. If π divides α, we define $(\alpha/\pi)_3 = 0$.

One observes that $(\cdot/\pi)_3$ is a character of the multiplicative group $(D/\pi D)^*$, and for $\alpha \in (D/\pi D)^*$, $(\alpha/\pi)_3 = 1$ if and only if the congruence $x^3 \equiv \alpha \pmod{\pi}$ is solvable.

Because there are six units in the ring D, a non-zero element in D has six associates. For a given prime π of norm not equal to 3, we single out one among its six associates. This is done in the following way. A prime π in D, is said to be *primary* if $\pi \equiv 2 \pmod{3}$ in D. If $\pi = a + b\omega$, it amounts to say that $a \equiv 2 \pmod{3}$ and $b \equiv 0 \pmod{3}$ in $\mathbf{Z}$.

It is clear that neither the prime $1-\omega$ nor any of its associates is primary. Rational primes $p \equiv 2 \pmod{3}$, which remain prime in D are primary and their other associates are not. For a prime π of norm $p \equiv 1 \pmod{3}$, again it is not difficult to see that among the associates of π, there is exactly one which is primary. With this, we are ready to state the law of cubic reciprocity.

Theorem 3.1. (The Law of Cubic Reciprocity). Consider two primes π_1 and π_2 in D such that neither of them is of norm 3 and both are primary. We also assume that $N(\pi_1) \neq N(\pi_2)$. Then

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

Remark 3.1. As in the case of quadratic reciprocity, there are supplementary laws for the cubic residue character of the units and the prime $1 - \omega$. It is easy to see that $\chi_\pi(-1) = 1$ and if π is not of norm 3, by definition, $\chi_\pi(\omega) = 1, \omega$, or ω^2 respectively for the cases $N(\pi) \equiv 1, 4$, or $7 \pmod{9}$.

Regarding $1 - \omega$, if $N(\pi) \neq 3$, then

$$\chi_\pi(1-\omega) = \omega^{2m}, \tag{4}$$

where the integer m is defined as follows. If $\pi = q$ is a rational prime, then m is defined by $q = 3m - 1$. If $\pi = a + b\omega$ is a primary complex prime, then m is defined by $a = 3m - 1$.

At the end of this section, we shall sketch Williams' proof [10] of Eisenstein's supplement to the law of cubic reciprocity (equation (4) above).

Before proceeding to prove Theorem 3.1, we obtain certain results on Gauss and Jacobi sums.

Let p be a rational prime. For a finite field $\mathbf{F}_p$, characters of the multiplicative group $\mathbf{F}_p^*$, that is, homomorphisms $\mathbf{F}_p^* \to \mathbf{C}^*$ will be referred to as multiplicative characters on $\mathbf{F}_p$. We shall denote the trivial character by ϵ, that is, $\epsilon(a) = 1$ for all $a \in \mathbf{F}_p^*$. If $\chi \neq \epsilon$, we define $\chi(0) = 0$. The trivial character is extended by defining $\epsilon(0) = 1$.

If χ is a multiplicative character on $\mathbf{F}_p$, and $\zeta = e^{2\pi i/p}$, then for an element a of $\mathbf{F}_p$, the sum $\sum_{t\in\mathbf{F}_p} \chi(t)\zeta^{at}$ is called a *Gauss sum on* $\mathbf{F}_p$ and is denoted by $g_a(\chi)$. For $g_1(\chi)$, we shall simply write $g(\chi)$.

If χ_1 and χ_2 are two multiplicative characters of $\mathbf{F}_p$, then

$$\sum_{a+b=1} \chi_1(a)\chi_2(b)$$

is called a *Jacobi sum* and is denoted by $J(\chi_1, \chi_2)$.

If χ is a multiplicative character on $\mathbf{F}_p$, then we know that $\sum_{t\in\mathbf{F}_p} \chi(t) = p$ or 0, according as χ is the trivial character ϵ or not. Also, if $a \in \mathbf{F}_p^*$, then $\sum_\chi \chi(a) = p - 1$ or 0, according as a is the identity element 1 or not. We proceed to prove some results on Gauss and Jacobi sums.

Proposition 3.1.

i) If a is a non-zero element of $\mathbf{F}_p$ and χ a non-trivial multiplicative character on $\mathbf{F}_p$, then $g_a(\chi) = \chi(a^{-1})g_1(\chi)$.

ii) If a is a non-zero element of $\mathbf{F}_p$ and χ is the trivial multiplicative character ϵ, then $g_a(\epsilon) = 0$.

iii) $g_0(\chi) = p$ or 0 according as χ is the trivial character ϵ or not.

Proof: All these statements follow directly from the definition of a Gauss sum. We prove only part (i) here. To prove (i), we just observe that $\chi(a)g_a(\chi) = \chi(a)\sum_{t\in\mathbf{F}_p} \chi(t)\zeta^{at} = \sum_{t\in\mathbf{F}_p} \chi(at)\zeta^{at} = g_1(\chi)$.

Remark 3.2. In Remark 2.1, the sum S was a particular Gauss sum. This was the particular case corresponding to $a = 1$ of the *quadratic Gauss sum* $\sum_{x\in\mathbf{F}_p^*} \left(\frac{x}{p}\right)\zeta^{ax}$. We remark that for a general multiplicative character $\chi \neq \epsilon$ on $\mathbf{F}_p$, one has $g(\chi)g(\overline{\chi}) = \chi(-1)p$.

Proposition 3.2.

i) $J(\epsilon, \epsilon) = p$.

ii) If χ is a non-trivial multiplicative character on $\mathbf{F}_p$, then

$$J(\epsilon, \chi) = 0 \text{ and } J(\chi, \chi^{-1}) = -\chi(-1).$$

iii) If χ_1 and χ_2 are non-trivial multiplicative characters on $\mathbf{F}_p$ such that $\chi_2 \neq \chi_1^{-1}$, then

$$J(\chi_1, \chi_2) = \frac{g(\chi_1)g(\chi_2)}{g(\chi_1\chi_2)}.$$

Proof: We prove only part (iii) here. We have

$$g(\chi_1)g(\chi_2) = \sum_{a,b \in \mathbf{F}_p} \chi_1(a)\chi_2(b)\zeta^{a+b} = \sum_{t \in \mathbf{F}_p} \left(\sum_{a+b=t} \chi_1(a)\chi_2(b) \right) \zeta^t.$$

We observe that the inner sum survives only when $t \neq 0$ and in that case, substituting $a = ta'$ and $b = tb'$, a straightforward calculation shows that the inner sum is $(\chi_1\chi_2)(t)J(\chi_1, \chi_2)$. Therefore, from the above equation we have

$$g(\chi_1)g(\chi_2) = \sum_{t \in \mathbf{F}_p} (\chi_1\chi_2)(t)J(\chi_1, \chi_2)\zeta^t = J(\chi_1, \chi_2)g(\chi_1\chi_2), \quad \text{as desired.}$$

Proposition 3.3. Let $n > 2$ be an integer. Let p be a rational prime such that $p \equiv 1 \pmod{n}$ and χ a multiplicative character of $\mathbf{F}_p$ of order n. Then

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

Proof: From part (iii) of Proposition 3.2, we have $g(\chi)^2 = J(\chi, \chi)g(\chi^2)$. Multiplying by $g(\chi)$ and applying part (iii) of Proposition 3.2 again we have $g(\chi)^3 = J(\chi, \chi)J(\chi, \chi^2)g(\chi^3)$. By repeating this process, we get

$$g(\chi)^{n-1} = J(\chi, \chi)J(\chi, \chi^2) \cdots J(\chi, \chi^{n-2})g(\chi^{n-1}). \tag{5}$$

But, $g(\chi)g(\chi^{n-1}) = g(\chi)g(\overline{\chi}) = \chi(-1)p$, by Remark 3.2. Hence by (5), we obtain our result.

Now we shall proceed to prove the law of cubic reciprocity. We shall need the following lemmas.

Lemma 3.1. Let π be a prime in D with $N(\pi) \neq 3$. Then

i) $\overline{\chi_\pi(\alpha)} = (\chi_\pi(\alpha))^2 = \chi_\pi(\alpha^2)$.

ii) $\chi_{\overline{\pi}}(\overline{\alpha}) = \overline{\chi_\pi(\alpha)}$.

iii) If $\pi = q$ is a rational prime of D, $\chi_q(\overline{\alpha}) = \chi_q(\alpha^2)$.

Proof: For part (i), since the squares of the numbers $1, \omega$, and ω^2 are equal to their corresponding conjugates, $\overline{\chi_\pi(\alpha)} = (\chi_\pi(\alpha))^2 = \chi_\pi(\alpha^2)$.

Next, observing that $N(\pi) = N(\overline{\pi})$, from the definition it follows that

$$\chi_{\overline{\pi}}(\overline{\alpha}) \equiv \overline{\chi_\pi(\alpha)} \pmod{\overline{\pi}}.$$

Since $1, \omega$, and ω^2 are distinct modulo $\overline{\pi}$, we get part (ii).

Finally, for a rational prime q in D, $\overline{q} = q$ and hence from part (ii) and part (i), $\chi_q(\overline{\alpha}) = \overline{\chi_q(\alpha)} = \chi_q(\alpha^2)$.

Lemma 3.2. Let π be a complex prime in D such that $N(\pi) = p \equiv 1 \pmod 3$ in D. We also assume π to be primary. Identifying $D/\pi D$ with $\mathbf{F}_p$, and therefore considering χ_π as a multiplicative character on $\mathbf{F}_p$, if we consider the Jacobi sum $J(\chi_\pi, \chi_\pi)$, then we have

$$J(\chi_\pi, \chi_\pi) = \pi.$$

Proof: First we note that χ_π being a cubic character, $\chi_\pi(-1) = 1$. Therefore, by Proposition 3.3, we have

$$g(\chi_\pi)^3 = pJ(\chi_\pi, \chi_\pi). \tag{6}$$

We now claim another result on $J(\chi_\pi, \chi_\pi)$, which again is true if we replace χ_π by a general cubic character. We claim that if we write $J(\chi_\pi, \chi_\pi) = a + b\omega$ with $a, b \in \mathbf{Z}$, then we have the following congruences in $\mathbf{Z}$:

$$a \equiv 2 \pmod 3 \quad \text{and} \quad b \equiv 0 \pmod 3. \tag{7}$$

We now establish the claim (7) above.

Observing that $\chi_\pi(0) = 0$ and $\chi_\pi(t)^3 = 1$ for $t \in \mathbf{F}_p^*$, we have the following congruence in the ring of algebraic integers $\mathbf{O}$:

$$g(\chi_\pi)^3 \equiv \sum_{t \in \mathbf{F}_p^*} \zeta^{3t} \pmod 3.$$

Since the last sum is -1, by (6) above, we have

$$pJ(\chi_\pi, \chi_\pi) = g(\chi_\pi)^3 \equiv -1 \pmod 3.$$

Therefore, since $p \equiv 1 \pmod 3$,

$$a + b\omega = J(\chi_\pi, \chi_\pi) \equiv -1 \pmod 3. \tag{8}$$

Working with $\overline{\chi}_\pi$ instead of χ_π, and observing that $\overline{g(\chi_\pi)} = g(\overline{\chi}_\pi)$ one obtains

$$a + b\overline{\omega} \equiv -1 \pmod 3. \tag{9}$$

Subtracting (9) from (8),

$$\begin{array}{rrcl} & b(\omega - \overline{\omega}) & \equiv & 0 \pmod 3 \\ \Rightarrow & b\sqrt{-3} & \equiv & 0 \pmod 3 \\ \Rightarrow & b\sqrt{-3} & = & 3\eta, \text{ for some } \ \eta \in \mathbf{O} \\ \Rightarrow & -3b^2 & = & 9\eta^2 \\ \Rightarrow & 3 & | & b \end{array}$$

Therefore, (8) implies that

$$a + 1 = 3\theta \text{ for some } \ \theta \in \mathbf{O}.$$

Since $\theta \in \mathbf{Q}$ as well, θ must be a rational integer. Hence, $a \equiv -1 \pmod 3$. Therefore, claim (7) is established.

Now, from Proposition 3.2 (iii) and Remark 3.2, we have

$$J(\chi_\pi, \chi_\pi)\overline{J(\chi_\pi, \chi_\pi)} = p$$

and therefore by (7), $J(\chi_\pi, \chi_\pi)$ is a primary prime.

Writing $J(\chi_\pi, \chi_\pi) = \gamma$, we observe that

$$\pi\overline{\pi} = p = \gamma\overline{\gamma}.$$

This gives $\pi|\gamma$ or $\pi|\overline{\gamma}$. Since all the primes involved in the above equation are primary,

$$\pi = \gamma \text{ or } \ \pi = \overline{\gamma}.$$

We rule out the second possibility.

We have

$$J(\chi_\pi, \chi_\pi) = \sum_{x \in \mathbf{F}_p} \chi_\pi(x)\chi_\pi(1-x) \equiv \sum_{x \in \mathbf{F}_p} x^{(p-1)/3}(1-x)^{(p-1)/3} \pmod \pi. \tag{10}$$

Now, observing that $x^{(p-1)/3}(1-x)^{(p-1)/3}$ is a polynomial in x of degree $2(p-1)/3 < (p-1)$, and from elementary number theory, recalling the congruence

$$1^k + 2^k + \cdots + (p-1)^k \equiv 0 \pmod p, \text{ when } (p-1) \not| k,$$

we have
$$\sum_{x\in\mathbf{F}_p} x^{(p-1)/3}(1-x)^{(p-1)/3} \equiv 0 \pmod{p}.$$

Therefore, from (10), we have $J(\chi_\pi, \chi_\pi) \equiv 0 \pmod{\pi}$, showing thereby that the second possibility does not occur. Lemma 3.2. is thus established.

Proof of Theorem 3.1: We consider the following three possible cases:

A) Both π_1 and π_2 are rational.

B) One of them, say π_1, is rational and π_2 is complex.

C) Both π_1 and π_2 are complex.

Case A. In this case, since $\pi_1 = q_1$ and $\pi_2 = q_2$ are rational primes coprime to each other, observing that $\overline{q}_1 = q_1$ and $\overline{q}_2 = q_2$, from Lemma 3.1 above,
$$\chi_{\pi_1}(\pi_2) = \overline{\chi_{\pi_1}(\pi_2)} = \chi_{\pi_1}(\pi_2^2).$$

Since $\chi_{\pi_1}(\pi_2) \neq 0$, this implies that $\chi_{\pi_1}(\pi_2) = 1$. For the same reason, $\chi_{\pi_2}(\pi_1)$ is also 1 and therefore, Theorem 3.1 is established in this case.

Case B.

Here in order to simplify notations, we write $\pi_1 = q$ and $\pi_2 = \pi$. Now, $q \equiv 2 \pmod{3}$ and $N(\pi) = p \equiv 1 \pmod{3}$.

From Remark 3.1, Proposition 3.3 and Lemma 3.2, we have
$$g(\chi_\pi)^3 = p\pi. \tag{11}$$

Now, (11) implies
$$\begin{aligned} g(\chi_\pi)^{q^2-1} &= (p\pi)^{(q^2-1)/3}, \\ \text{and hence,}\quad g(\chi_\pi)^{q^2-1} &\equiv \chi_q(p\pi) \pmod{q}, \quad \text{by definition.} \end{aligned}$$

Since by Lemma 3.1, $\chi_q(p) = \overline{\chi_q(p)}$ and the only real value of the character χ_q is 1, we have $\chi_q(p) = 1$ and therefore, from above,
$$g(\chi_\pi)^{q^2} = \chi_q(\pi) g(\chi_\pi) \pmod{q}. \tag{12}$$

Now, by definition
$$g(\chi_\pi)^{q^2} = \left(\sum_{t\in\mathbf{F}_p} \chi_\pi(t)\zeta^t\right)^{q^2} \equiv \sum_{t\in\mathbf{F}_p} \chi_\pi(t)^{q^2}\zeta^{q^2 t} \pmod{q}.$$

Since $q^2 \equiv 1 \pmod{3}$, we have $\chi_\pi(t)^{q^2-1} = 1$.

Therefore,

$$\begin{array}{rrcl} & g(\chi_\pi)^{q^2} & \equiv & \displaystyle\sum_{t \in \mathbf{F}_p} \chi_\pi(t)\zeta^{q^2 t} \pmod{q} \\ \Rightarrow & g(\chi_\pi)^{q^2} & \equiv & g_{q^2}(\chi_\pi) \pmod{q}, \quad \text{(by definition)} \\ \Rightarrow & g(\chi_\pi)^{q^2} & \equiv & \chi_\pi(q^{-2})g(\chi_\pi) \pmod{q} \quad \text{(by Proposition 3.1 (i))} \\ \Rightarrow & g(\chi_\pi)^{q^2} & \equiv & \chi_\pi(q)g(\chi_\pi) \pmod{q}. \end{array}$$

Hence, by (12),

$$\begin{array}{rrcl} & \chi_\pi(q)g(\chi_\pi) & \equiv & \chi_q(\pi)g(\chi_\pi) \pmod{q} \\ \Rightarrow & \chi_\pi(q)p & \equiv & \chi_q(\pi)p \pmod{q} \quad \text{(multiplying both sides by} \\ & & & \qquad\qquad\qquad\qquad g(\overline{\chi}_\pi) \ \text{ and by Remark 3.2)} \\ \Rightarrow & \chi_\pi(q) & \equiv & \chi_q(\pi) \pmod{q} \\ \Rightarrow & \chi_\pi(q) & = & \chi_q(\pi). \end{array}$$

Case C. Write $N(\pi_1) = p_1$ and $N(\pi_2) = p_2$. Then, $p_i \equiv 1 \pmod{3}$ for $i = 1, 2$.

Let $\gamma_1 = \overline{\pi}_1$ and $\gamma_2 = \overline{\pi}_2$.

Then γ_i's are primary and $p_i = \pi_i\gamma_i$ for $i = 1, 2$.

As in Case B above, we start with the relation

$$g(\chi_{\gamma_1})^3 = p_1\gamma_1, \tag{13}$$

which implies

$$\begin{array}{rrcl} & g(\chi_{\gamma_1})^{p_2-1} & = & (p_1\gamma_1)^{(p_2-1)/3}, \\ \text{and hence,} & g(\chi_{\gamma_1})^{p_2-1} & \equiv & \chi_{\pi_2}(p_1\gamma_1) \pmod{\pi_2}, \ \text{by definition.} \end{array}$$

Now, going back to the definition of $g(\chi_{\gamma_1})$ and proceeding as in case B, from (13) we obtain

$$\chi_{\gamma_1}(p_2^2) = \chi_{\pi_2}(p_1\gamma_1). \tag{14}$$

Similarly,

$$\chi_{\pi_2}(p_1^2) = \chi_{\pi_1}(p_2\pi_2). \tag{15}$$

Again, by Lemma 3.1 (i),

$$\chi_{\gamma_1}(p_2^2) = \overline{\chi_{\gamma_1}(p_2)}.$$

Therefore, by Lemma 3.1 (ii)

$$\chi_{\gamma_1}(p_2^2) = \chi_{\pi_1}(p_2). \tag{16}$$

$$\begin{aligned}
\text{Now,}\quad \chi_{\pi_1}(\pi_2)\chi_{\pi_2}(p_1\gamma_1) &= \chi_{\pi_1}(\pi_2)\chi_{\gamma_1}(p_2^2) \qquad \text{(by (14))}\\
&= \chi_{\pi_1}(\pi_2)\chi_{\pi_1}(p_2) \qquad \text{(by (16))}\\
&= \chi_{\pi_1}(p_2\pi_2)\\
&= \chi_{\pi_2}(p_1^2) \qquad \text{(by (15))}\\
&= \chi_{\pi_2}(p_1\pi_1\gamma_1)\\
&= \chi_{\pi_2}(\pi_1)\chi_{\pi_2}(p_1\gamma_1)
\end{aligned}$$

and hence $\chi_{\pi_1}(\pi_2) = \chi_{\pi_2}(\pi_1)$, as desired.

Williams' proof of equation (4): We extend the definition of the cubic residue character so that we are able to work with non-prime integers of D in the denominator of the symbol.

Let $\alpha, \tau \in D$. Also assume that $\tau \not\equiv 0 \pmod{(1-\omega)}$, in D.

We define $\chi_\tau(\alpha)\,(= (\alpha/\tau)_3)$

$$= \begin{cases} 1 & \text{if } \tau \text{ is a unit of } D,\\ \chi_{\tau_1}(\alpha)\chi_{\tau_2}(\alpha)\cdots\chi_{\tau_r}(\alpha) & \text{when } \tau \text{ is a non-unit and}\\ & \ \tau = \tau_1\tau_2\cdots\tau_r \text{ is the factorization}\\ & \ \text{into products of primes of } D. \end{cases}$$

Since it is easy to check that

$$\frac{N(\pi_1)-1}{3} + \frac{N(\pi_2)-1}{3} \equiv \frac{N(\pi_1\pi_2)-1}{3} \pmod 3$$

holds for any two primes π_1 and π_2 of D not of norm 3, by the above definition, for any $\tau \in D$ with $\tau \not\equiv 0 \pmod{(1-\omega)}$, we have

$$(\omega/\tau)_3 = \omega^{(N(\tau)-1)/3}. \tag{17}$$

If π is a rational prime of D, let $\pi = 3m-1$. Now,

$$\begin{aligned}
\chi_\pi(1-\omega) &= \chi_\pi\left((1-\omega)^4\right) = \left(\chi_\pi\left((1-\omega)^2\right)\right)^2 \\
&= (\chi_\pi(-3\omega))^2 = (\chi_\pi(-3))^2\,(\chi_\pi(\omega))^2 \\
&= (\chi_\pi(\omega))^2 \text{ (see Case A in the proof of Theorem 3)} \\
&= \omega^{2(N(\pi)-1)/3} = \omega^{6m^2-4m} = \omega^{2m},
\end{aligned}$$

and therefore, (4) is established when π is a rational prime of D.

Now, let $\pi = a + b\omega$ be a complex prime where $a = 3m-1$ and $b = 3n$. In this case,

$$\begin{aligned}
\chi_\pi(1-\omega) &= \chi_a(b)\chi_\pi(1-\omega) \qquad (\text{since } \chi_a(b)=1) \\
&= (\chi_a(\omega))^2\,\chi_a(b\omega)\chi_\pi(1-\omega) \\
&= (\chi_a(\omega))^2\,\chi_a(\pi)\chi_\pi(1-\omega) \qquad (\text{since } b\omega \equiv \pi \pmod{a}) \\
&= \omega^{2(a^2-1)/3}\chi_a(\pi)\chi_\pi(1-\omega) \qquad (\text{by (17)}) \\
&= \omega^{6m^2-4m}\chi_a(\pi)\chi_\pi(1-\omega) \\
&= \omega^{6m^2-4m}\chi_\pi(a)\chi_\pi(1-\omega) \qquad (\text{by Theorem 3.1 and our extended definition of the cubic residue character}) \\
&= \omega^{2m}\chi_\pi(a-a\omega) \\
&= \omega^{2m}\chi_\pi(-(a+b)\omega) \qquad (\text{since } b\omega \equiv -a \pmod{\pi}) \\
&= \omega^{2m}\chi_\pi(\omega)\chi_\pi(a+b) \\
&= \omega^{2m+(p-1)/3}\chi_{a+b}(\pi) \\
&= \omega^{n}\chi_{a+b}(\pi) \qquad (\text{observing that } p = a^2-ab+b^2) \\
&= \omega^{n}\chi_{a+b}(b(1-\omega)) \qquad (\text{since } b(1-\omega) \equiv -\pi \pmod{(a+b)}) \\
&= \omega^{n}\chi_{a+b}(b)\chi_{a+b}(1-\omega) \\
&= \omega^{n+2(m+n)} \\
&= \omega^{2m}.
\end{aligned}$$

4. Eisenstein reciprocity law. We shall here state and give a sketch of the proof of the Eisenstein reciprocity law. In many places, for details we shall refer to Ireland and Rosen [5]. For applications of Eisenstein reciprocity law, we again refer to [5] as well as Ribenboim's book [7] "13 Lectures on Fermat's Last Theorem". One may also look into the recent book of Esmonde and Murty [4] for that purpose.

Let m be a positive integer and $\zeta_m = e^{2\pi i/m}$. Let D_m denote the ring of integers of $\mathbf{Q}(\zeta_m)$.

We recall some facts about the ring D_m. Ideals appearing in our discussion will be non-zero. First, $D_m = \mathbf{Z}[\zeta_m]$. Now, let P be a prime ideal in D_m not containing m and write $q = N(P) \overset{\text{def}}{=} |D_m/P|$. Then, $q \equiv 1 \pmod{m}$ and the elements $1, \zeta_m, \cdots, \zeta_m^{m-1}$ are distinct mod P.

We also recall that the only roots of unity in $\mathbf{Q}(\zeta_m)$ are $\pm\zeta_m^i$, $i = 1, 2, \cdots, m$.

Now, if $\alpha \in D_m$ such that $\alpha \notin P$, then

$$\alpha^{q-1} \equiv 1 \pmod{P}.$$

$$\text{Therefore,} \quad \prod_{i=0}^{m-1} \left(\alpha^{(q-1)/m} - \zeta_m^i\right) \equiv 0 \pmod{P}.$$

Since P is a prime ideal, $\exists\ i, 0 \leq i < m$ such that

$$\alpha^{(q-1)/m} \equiv \zeta_m^i \pmod{P}. \tag{18}$$

Since $\zeta_m^i \not\equiv \zeta_m^j \pmod{P}$ for $i \not\equiv j \pmod{m}$, the integer i in (18) is unique mod m.

If $\alpha \in D_m$ is such that $\alpha \notin P$, the unique element ζ_m^i to which $\alpha^{(N(P)-1)/m}$ is congruent modulo P, is defined to be the *m-th power residue symbol* and is denoted by $(\alpha/P)_m$. This gives us a multiplicative character of the finite field D_m/P of q elements. If $\alpha \in P$, we define $(\alpha/P)_m = 0$.

Once again, it is not difficult to check that $(\alpha/P)_m = 1$ if and only if $x^m \equiv \alpha \pmod{P}$ is solvable in D_m.

Now, let $A \subset D_m$ be any proper ideal prime to m. Let $A = P_1 \cdots P_n$ be a decomposition into product of prime ideals of D_m; P_i's are not necessarily distinct. For $\alpha \in D_m$, we have the following definition.

$$(\alpha/A)_m \overset{\text{def}}{=} \prod_i (\alpha/P_i)_m.$$

If $\beta \in D_m$ is such that β is prime to m, then we define

$$(\alpha/\beta)_m \overset{\text{def}}{=} (\alpha/(\beta))_m\,,$$

where (β) denotes the principal ideal βD_m.

Let $l > 0$ be an odd prime in $\mathbf{Z}$. Then a non-zero element $\alpha \in D_l$ is called *primary* if it is prime to l and congruent to a rational integer modulo $(1-\zeta_l)^2$. We claim that for $\alpha \in D_l$ such that α prime to l there is an integer $c \in \mathbf{Z}$, unique mod l, such that $\zeta_l^c \alpha$ is primary.

We know that in D_l, the principal ideal $(l) = lD_l = (1-\zeta_l)^{l-1}$ and the principal ideal $(1-\zeta_l)$ is prime of degree 1.

Hence there exists $a \in \mathbf{Z}$ such that

$$\alpha \equiv a \ (\mathrm{mod}\ (1-\zeta_l)). \tag{19}$$

From (19),

$$\frac{\alpha - a}{1-\zeta_l} \in D_l.$$

Therefore, applying the same argument once again, there exists $b \in \mathbf{Z}$ such that

$$\frac{\alpha - a}{1-\zeta_l} \equiv b \ (\mathrm{mod}\ (1-\zeta_l)).$$

The above implies,

$$\alpha \equiv a + b(1-\zeta_l) \ (\mathrm{mod}\ (1-\zeta_l)^2). \tag{20}$$

Again, writing $\zeta_l = 1-(1-\zeta_l)$, we observe that for $d \in \mathbf{Z}$,

$$\zeta_l^d \equiv 1 - d(1-\zeta_l) \ (\mathrm{mod}\ (1-\zeta_l)^2). \tag{21}$$

From (20) and (21),

$$\zeta_l^d \alpha \equiv a + (b-ad)(1-\zeta_l) \ (\mathrm{mod}\ (1-\zeta_l)^2). \tag{22}$$

From (19), we observe that if l divides a in $\mathbf{Z}$, then $(1-\zeta_l)$ would divide α in D_l contradicting the assumption that α is prime to l. Therefore, l does not divide a in $\mathbf{Z}$. Therefore, there exists $d \in \mathbf{Z}$ such that $ad \equiv b \ (\mathrm{mod}\ l)$.

With this d, from (22) we have

$$\zeta_l^d \alpha \equiv a \ (\mathrm{mod}\ (1-\zeta_l)^2).$$

From the proof, a given α determines uniquely a and b, and hence determines a unique d mod l and our claim is established.

Now, we state the Eisenstein reciprocity law.

Theorem 4.1. (The Eisenstein Reciprocity Law). Let l be an odd prime and $a\ (\neq \pm 1) \in \mathbf{Z}$ is such that $(l,a) = 1$. Let $\alpha \in D_l$ be such that α is a primary non-unit element of D_l and α and a are prime to each other. Then

$$(\alpha/a)_l = (a/\alpha)_l.$$

We go through a sequence of definitions and propositions before we take up the proof of Theorem 4.1.

First we set some notations. If σ is an element of the group $G = \mathrm{Gal}(\mathbf{Q}(\zeta_m)/\mathbf{Q})$, for an element $\alpha \in \mathbf{Q}(\zeta_m)$, we shall write α^σ instead of $\sigma(\alpha)$. Similarly, for an ideal A of D_m, we shall use the notation A^σ to denote $\sigma(A)$. It is known that $\prod_{\sigma \in G} \sigma(A) = (N(A))$. The proof of the following proposition is straightforward.

Proposition 4.1. *Let A be a proper ideal of D_m, prime to m. Let σ be an element of the group G defined above. Then*

$$(\alpha/A)_m^\sigma = (\alpha^\sigma/A^\sigma)_m.$$

We start with generalizing the notion of Gauss sums. Also, we work on arbitrary finite fields. Let $\mathbf{F}$ be a finite field such that $|\mathbf{F}| = p^f$, where p is a rational prime. Let $\chi : \mathbf{F}^* \to \mathbf{C}^*$ be a character. Let the order of χ be m. Let $\psi : \mathbf{F} \to \mathbf{C}^*$ be a non-trivial additive character. Then values of χ are m-th roots of unity while those of ψ are p-th roots of unity.

As before, the trivial multiplicative character ϵ is extended by defining $\epsilon(0) = 1$ and if $\chi \neq \epsilon$, we define $\chi(0) = 0$.

We consider the *Gauss sum*

$$g(\chi, \psi) \overset{\text{def}}{=} \sum_{t \in \mathbf{F}} \chi(t)\psi(t).$$

We observe that $g(\chi, \psi) \in \mathbf{Q}(\zeta_m, \zeta_p)$.

We now specify the characters χ and ψ we shall be working with.

Let P be a prime ideal in D_m not containing m. Let P lie over the rational prime p, that is, $p\mathbf{Z} = P \cap \mathbf{Z}$. Let $N(P) = q = p^f$ and $\mathbf{F} = D_m/P$.

We know that $p^f \equiv 1 \pmod{m}$. Let t be a non-zero element of $\mathbf{F}$. Let $t = \gamma + P$ for $\gamma \in D_m$. We define

$$\chi_P(t) = (\gamma/P)_m^{-1}.$$

Now, we describe the additive character ψ. Consider the trace function $\mathrm{tr} \ : \mathbf{F} \to \mathbf{Z}/p\mathbf{Z}$ defined by $\mathrm{tr}(t) = t + t^p + t^{p^2} \cdots + t^{p^{f-1}}$. We define our additive character ψ by $\psi(t) = \zeta_p^{\mathrm{tr}(t)}$.

Now we consider the corresponding Gauss sum

$$g(P) \overset{\text{def}}{=} g(\chi_P, \psi) \overset{\text{def}}{=} \sum_{t \in \mathbf{F}} \chi_P(t)\psi(t).$$

We define

$$\Phi(P) \overset{\text{def}}{=} g(P)^m.$$

As before (see Remark 3.2),

$$|g(P)|^2 = q. \tag{23}$$

Now both $\mathbf{Q}(\zeta_m)$ and $\mathbf{Q}(\zeta_p)$ are subfields of $\mathbf{Q}(\zeta_m, \zeta_p) = \mathbf{Q}(\zeta_{mp})$ and

$$\mathrm{Gal}(\mathbf{Q}(\zeta_{mp})/\mathbf{Q}) = \left\{\sigma_r : \zeta_{mp} \mapsto \zeta_{mp}^r,\ (r, mp) = 1\right\}.$$

We observe that σ_r leaves $\mathbf{Q}(\zeta_m)$ elementwise fixed if and only if $r \equiv 1 \pmod{m}$. Similarly, σ_r leaves $\mathbf{Q}(\zeta_p)$ elementwise fixed if and only if $r \equiv 1 \pmod{p}$.

Let $c \in \mathbf{Z}$ be such that $c \equiv 1 \pmod{m}$.

$$\begin{aligned}
\sigma_c(g(P)) &= \sigma_c\left(\sum_{t\in\mathbf{F}} \chi_P(t)\psi(t)\right) \\
&= \sum_{t\in\mathbf{F}} \chi_P(t)\psi(t)^c \qquad (\text{since, } \chi_P(t) \in \mathbf{Q}(\zeta_m),\ \sigma_c(\chi_P(t)) = \chi_P(t)) \\
&= \sum_{t\in\mathbf{F}} \chi_P(t)\psi(ct) \\
&= \chi_P(c)^{-1} g(P)
\end{aligned}$$

Therefore,

$$\sigma_c\left(\Phi(P)\right) = \sigma_c\left((g(P)^m\right) = g(P)^m = \Phi(P).$$

Thus, $\Phi(P)$ is invariant under σ_c and hence the following proposition.

Proposition 4.2.

$$\Phi(P) \in \mathbf{Q}(\zeta_m).$$

Now, we state the following result, and, as had been mentioned already, we refer to [5] or [6] for its proof.

Proposition 4.3. (Stickelberger) Let P be a prime ideal in D_m such that $m \notin P$. Then the principal ideal

$$(\Phi(P)) = (g(P)^m) = P^{\sum t\sigma_t^{-1}},$$

where the sum is over all $1 \leq t < m$ which are relatively prime to m.

Let $A \subset D_m$ be a proper ideal prime to m. Let $A = P_1P_2 \cdots P_n$ where P_i's are prime ideals. Then we define

$$\Phi(A) \overset{\text{def}}{=} \Phi(P_1)\Phi(P_2)\cdots\Phi(P_n).$$

If $A = (\alpha)$ is a principal ideal, we write $\Phi(\alpha)$ for $\Phi((\alpha))$.

Proposition 4.4. *Let A, B be proper ideals of D_m, both prime to m. Let $\alpha \in D_m$ be prime to m. Let $\gamma = \sum t\sigma_t^{-1}$, where the sum is over all $1 \le t < m$ which are relatively prime to m. Then*

i) $\Phi(A)\Phi(B) = \Phi(AB)$.

ii) $|\Phi(A)|^2 = N(A)^m$.

iii) *The principal ideal $(\Phi(A))$ is equal to A^γ.*

iv) $\Phi(\alpha) = \epsilon(\alpha)\alpha^\gamma$ *for some unit $\epsilon(\alpha)$ of D_m.*

Proof: Whereas (i) follows directly from definition, (ii) follows from (23). Part (iii) follows from Proposition 4.3 (Stickelberger).

Now, by Part (iii), we get that the principal ideal $(\Phi(\alpha)) = (\alpha)^\gamma = (\alpha^\gamma)$. That is, $\Phi(\alpha)$ and α^γ generate the same principal ideal and that implies (iv).

If $\alpha \in D_m$ is as in Proposition 4.4, we now proceed to have more precise informations about $\epsilon(\alpha)$ appearing in part (iv). By part (ii) of the above proposition, $|\Phi(\alpha)|^2 = |N(\alpha)|^m$. On the other hand, it is not very difficult to see (see [5]) that $|(\alpha)^\gamma|^2 = |N(\alpha)|^m$. It then follows that $|\epsilon(\alpha)| = 1$. Similarly, by using Proposition 4.1, one derives that $|\epsilon(\alpha)^\sigma| = 1$ for all $\sigma \in G$. Therefore, $\epsilon(\alpha)$ is a root of unity. Since $\epsilon(\alpha) \in \mathbf{Q}(\zeta_m)$, we have the following proposition.

Proposition 4.5. $\epsilon(\alpha) = \pm\zeta_m^i$ *for some i.*

Let $m = l$, an odd prime and $\alpha \in D_l$ be primary. Then one can obtain (see [5]) the following more precise information.

Proposition 4.6. $\epsilon(\alpha) = \pm 1$.

Now, let $\alpha \in D_l$ be primary and a non-unit and B a proper ideal of D_l such that B is prime to l and $N(B)$ is prime to α. Since l is odd, -1 is not an l-th root of unity and hence by the above proposition

$$(\epsilon(\alpha)/B)_l = 1. \tag{24}$$

For prime ideals P, P' of D_l, both prime to l, such that $N(P)$ and $N(P')$ are relatively prime, it is not very difficult (see [5]) to observe that

$$(\Phi(P)/P')_l = (N(P')/P)_l.$$

From this, using part (iv) of Proposition 4.4 and Proposition 4.1 one derives

$$(\epsilon(\alpha)/B)_l(\alpha/N(B))_l = (N(B)/\alpha)_l.$$

Therefore, by (24),

$$(\alpha/N(B))_l = (N(B)/\alpha)_l \tag{25}$$

Proof of Theorem 4.1: Let p be a rational prime other than l such that p is prime to α. Let P be a prime ideal of D_l containing p. Then $N(P) = p^f$ and therefore, by (25),

$$(\alpha/p)_l^f = (p/\alpha)_l^f.$$

Since f divides the degree of the extension $\mathbf{Q}(\zeta_l)$ over $\mathbf{Q}$, which is $l-1$, l does not divide f and hence from the above,

$$(\alpha/p)_l = (p/\alpha)_l$$

and the theorem follows by multiplicativity.

Acknowledgements. I would like to thank Professors M. Ram Murty and D. S. Nagaraj for carefully going through the manuscript.

REFERENCES

1. S. D. Adhikari, *An Introduction to Commutative Algebra and Number Theory*, Narosa Publishing House, (1999).

2. David A. Cox, *Primes of the form* x^2+ny^2, John Wiley & Sons, (1989).

3. Eknath Ghate, *The Kronecker-Weber Theorem*, This volume.

4. Jody Esmonde and M. Ram Murty, *Problems in Algebraic Number Theory*, Springer-Verlag, (1999).

5. Kenneth Ireland and Michael Rosen, *A Classical Introduction to Modern Number Theory*, 2nd edition, Springer-Verlag, (1990).

6. S. A. Katre, *Gauss-Jacobi sums and Stickelberger's Theorem*, This volume.

7. P. Ribenboim, *13 Lectures on Fermat's Last Theorem*, Springer-Verlag, (1979).

8. J. P. Serre, *A Course in Arithmetic*, Springer-Verlag, (1973).

9. Parvati Shastri, *Reciprocity laws: Artin - Hilbert*, This volume.

10. Kenneth S. Williams, *On Eisenstein's supplement to the law of cubic reciprocity*, Bull. Cal. Math. Soc., **69**, 311–314 (1977).

11. B. F. Wyman, *What is a reciprocity law?*, Am. Math. Monthly, **79**, 571–586 (1972).

Sukumar Das Adhikari
Mehta Research Institute
Chhatnag Road, Jhusi
Allahabad 211 019, India
e-mail: adhikari@mri.ernet.in

Gauss-Jacobi Sums and Stickelberger's Theorem

S. A. Katre

In this article we shall prove Stickelberger's theorem using factorisation of Gauss sums. This theorem tells us about certain elements of the integral group ring of the Galois group of an abelian number field which annihilate the ideal class group of the number field. We shall then apply Stickelberger's theorem to prove Herbrand's theorem. Herbrand's theorem is a stronger version of Kummer's theorem : " $p \mid h(\mathbb{Q}(\zeta_p)) \Rightarrow p \mid$ some Bernoulli number". Our main reference is [8].

§ 1. Gauss and Jacobi Sums

Let p be an odd prime and q be a power of p. Let $\mathbb{F}_q$ be the finite field of q elements. Let ζ_p be a fixed primitive p^{th} root of 1. The Galois group $\text{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ is cyclic generated by the Frobenius automorphism σ_p of $\mathbb{F}_q$ given by $x \mapsto x^p$.

For $q = p^f$, let $\text{Tr} : \mathbb{F}_q \to \mathbb{F}_p$ be the trace map, $\text{Tr}(\alpha) = \alpha + \alpha^p + \cdots + \alpha^{p^{f-1}}$. Then for $g(x) = x + x^p + \cdots + x^{p^{f-1}} \in \mathbb{F}_p[x]$,

$$\prod_{i=0}^{p-1}(g(x) - i) = g(x)^p - g(x) = x^q - x.$$

For every $a \in \mathbb{F}_q$, $a^q - a = 0$, so for every $i \in \mathbb{F}_p$, $g(x) - i$ has p^{f-1} zeros in $\mathbb{F}_q$. This shows that Tr is onto. Hence $\psi : \mathbb{F}_q \to \mathbb{C}^\times, \psi(x) = \zeta_p^{\text{Tr}(x)}$ is a well-defined nontrivial additive character of $\mathbb{F}_q$. Let $\chi : \mathbb{F}_q^\times \to \mathbb{C}^\times$ be a multiplicative character of $\mathbb{F}_q^\times$. Extend χ to all of $\mathbb{F}_q$ by setting $\chi(0) = 0$ (even if χ is the trivial character[1] $\mathbf{1}$). As $\chi^{q-1} = \mathbf{1}$, the order of χ is coprime to p.

Definition 1 *Let χ, χ_1, χ_2 be multiplicative characters on $\mathbb{F}_q$. The Gauss sum corresponding to χ is defined as*

$$g(\chi) = -\sum_{a \in \mathbb{F}_q} \chi(a)\psi(a).$$

[1] This convention is different from the one in the article of S. D. Adhikari [1] in these proceedings, where $\mathbf{1}(0)=1$. Also the definition of Gauss and Jacobi sums in [1] differs from ours in sign.

The Jacobi sum corresponding to χ_1 *and* χ_2 *is defined as*

$$J(\chi_1,\chi_2) = -\sum_{a\in\mathbb{F}_q} \chi_1(a)\chi_2(1-a).$$

Proposition 1 (a) $g(\mathbf{1}) = 1, J(\mathbf{1},\mathbf{1}) = 2-q$.

(b) *If* χ,χ_1,χ_2 *have orders dividing* m *then* $g(\chi)\in\mathbb{Q}(\zeta_{mp})$ *and* $J(\chi_1,\chi_2)\in\mathbb{Q}(\zeta_m)$. $g(\chi)$ *and* $J(\chi_1,\chi_2)$ *are algebraic integers.*

(c) $J(\mathbf{1},\chi) = J(\chi,\ \mathbf{1}) = 1$ *if* $\chi \neq \mathbf{1}$.

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

(e) $J(\chi,\overline{\chi}) = \chi(-1)$ *if* $\chi\neq\mathbf{1}$.

(f) $\dfrac{g(\chi_1)g(\chi_2)}{g(\chi_1\chi_2)} = J(\chi_1,\chi_2)$ *if* $\chi_1\chi_2\neq\mathbf{1}$.
$g(\chi)g(\overline{\chi}) = \chi(-1)q$ *if* $\chi\neq\mathbf{1}$.
Thus if χ_1,χ_2 *are characters of order dividing* m*, then* $g(\chi_1)g(\chi_2)/g(\chi_1\chi_2)$ *is an algebraic integer in* $\mathbb{Q}(\zeta_m)$.

(g) *If* $\chi\neq\mathbf{1}, g(\chi)\overline{g(\chi)} = q$.
If $\chi_1,\chi_2,\chi_1\chi_2\neq\mathbf{1}, J(\chi_1,\chi_2)\overline{J(\chi_1,\chi_2)} = q$.

(h) *If* χ^m *is trivial and* $(b,m)=1$,

$$g(\chi)^{b-\sigma_b} := \frac{g(\chi)^b}{g(\chi)^{\sigma_b}} \in \mathbb{Q}(\zeta_m)$$

where $\sigma_b \in Gal(\mathbb{Q}(\zeta_m,\zeta_p))/\mathbb{Q}$ *is defined by* $\zeta_p\mapsto\zeta_p$ *and* $\zeta_m\mapsto\zeta_m^b$. *In particular taking* $b=1+m$, $g(\chi)^m\in\mathbb{Q}(\zeta_m)$.

(i) $g(\chi^p) = g(\chi)$.

Proof.

(e) For $\chi\neq\mathbf{1}$, $J(\chi,\overline{\chi}) = -\sum\chi(c)\overline{\chi}(1-c) = -\sum_{c\neq1}\chi\left(\dfrac{c}{1-c}\right) = \chi(-1)$.

(f)
$$\begin{aligned} g(\chi_1)g(\chi_2) &= \sum_{a,b}\chi_1(a)\chi_2(b)\psi(a+b)\\ &= \sum_{a,b}\chi_1(a)\chi_2(b-a)\psi(b)\\ &= \sum_{\substack{a,b\\ b\neq0}}\chi_1(a)\chi_2(b-a)\psi(b) + \sum_a\chi_1(a)\chi_2(-a)\\ &= S_1+S_2 \text{ say.}\end{aligned}$$

If $\chi_1\chi_2 \neq \mathbf{1}$, $S_2 = 0$. If $\chi_1\chi_2 = \mathbf{1}$, $S_2 = \chi_1(-1)(q-1)$.
In S_1, put $a = bc$. Then

$$S_1 = \sum_{\substack{b,c \\ b \neq 0}} \chi_1(b)\chi_2(b)\chi_1(c)\chi_2(1-c)\psi(b) = g(\chi_1\chi_2)J(\chi_1, \chi_2).$$

If $\chi_1\chi_2 = \mathbf{1}$, $S_1 = g(\mathbf{1})J(\chi_1, \overline{\chi_1}) = \chi_1(-1)$ and so $g(\chi)\overline{g(\chi)} = \chi(-1)q$.

(h) First, $g(\chi)^{\sigma_b} = g(\chi^b)$. Also, if for $(c,p) = 1, \tau_c \in \mathrm{Gal}(\mathbb{Q}(\zeta_m, \zeta_p)/\mathbb{Q})$ is defined by $\zeta_m \mapsto \zeta_m,\ \zeta_p \mapsto \zeta_p^c$, then $g(\chi)^{\tau_c} = \chi(c)^{-1}g(\chi)$ and similarly for $g(\chi^b)$. Hence τ_c fixes $g(\chi)^{b-\sigma_b}$.
(i) $g(\chi^p) = -\sum \chi^p(a)\zeta_p^{\mathrm{Tr}(a)} = -\sum \chi(a^p)\zeta_p^{\mathrm{Tr}(a^p)} = g(\chi)$. □

§ 2. Stickelberger's Theorem.

Let ζ_m denote a primitive m^{th} root of unity. Let $M/\mathbb{Q}$ be a finite abelian extension, so by Kronecker-Weber theorem, $M \subset \mathbb{Q}(\zeta_m)$ for some m. Assume m minimal. $G = \mathrm{Gal}(M/\mathbb{Q})$ may be regarded as a quotient of $(\mathbb{Z}/m\mathbb{Z})^\times = \mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$.

For $(a,m) = 1, \sigma_a$ denotes the element $\zeta_m \mapsto \zeta_m^a$ of $\mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$ as well as its restriction to M. For $x \in \mathbb{R}$, let $\{x\}$ = the fractional part of x, so $x - \{x\} \in \mathbb{Z}$ and $0 \leq \{x\} < 1$. Define the Stickelberger element θ in the group-ring $\mathbb{Q}[G]$ by

$$\theta = \theta(M) = \sum_{\substack{a \ (\mathrm{mod}\ m) \\ (a,m)=1}} \left\{\frac{a}{m}\right\} \sigma_a^{-1} = \sum_{\substack{1 \leq a \leq m \\ (a,m)=1}} \frac{a}{m}\sigma_a^{-1}.$$

Let $I(M) = \mathbb{Z}[G] \cap \theta\mathbb{Z}[G]$ consist of $\mathbb{Z}[G]$-multiples of θ which have integer coefficients. Then $I(M)$ is an ideal of the group-ring $\mathbb{Z}[G]$ and is called the Stickelberger ideal. Clearly, $m\theta \in I(M)$.

Lemma 1 *Let I' be the ideal of $\mathbb{Z}[G]$ generated by all the elements of the form $c - \sigma_c$ with $(c,m) = 1$. Equivalently I' is (finitely) generated by m and $c - \sigma_c$ with $(c,m) = 1,\ 1 \leq c < m$. Let $\beta \in \mathbb{Z}[G]$. If $\beta \in I'$ then $\beta\theta \in \mathbb{Z}[G]$, so that $I'\theta \subset I$. If, moreover, $M = \mathbb{Q}(\zeta_m)$, then $I = I'\theta$.*

Proof. We have

$$(c - \sigma_c)\theta = \sum_a \left(c\left\{\frac{a}{m}\right\} - \left\{\frac{ac}{m}\right\}\right)\sigma_a^{-1}.$$

$$\text{This is in } \mathbb{Z}[G], \quad \text{as} \quad c\left\{\frac{a}{m}\right\} - \left\{\frac{ac}{m}\right\} \equiv c \cdot \frac{a}{m} - \frac{ac}{m} \pmod 1$$
$$\equiv 0 \pmod 1.$$

Hence $I'\theta \subset I$.

Note also that $m = (1+m) - \sigma_{1+m} \in I'$, so that I' is generated by m and $c - \sigma_c$ with $(c, m) = 1,\ 1 \le c < m$.

Next consider the case $M = \mathbb{Q}(\zeta_m)$. Suppose

$$(\sum_a x_a \sigma_a)\theta \in \mathbb{Z}[G],$$

where $x_a \in \mathbb{Z}$. The coefficient of identity in $(\sum_a x_a \sigma_a)\left(\sum_c \left\{\frac{c}{m}\right\} \sigma_c^{-1}\right)$ is $\sum_a x_a \left\{\frac{a}{m}\right\}$ and this $\in \mathbb{Z}$. Hence $\sum x_a \cdot \frac{a}{m}$ is in $\mathbb{Z}$, i.e. $m \mid \sum x_a \cdot a$, so that $\sum x_a \cdot a \in I'$. Hence $\sum x_a \sigma_a = \sum x_a(\sigma_a - a) + \sum x_a a \in I'$.
Thus $I \subset I'\theta$. □

Example. The above result $I = I'\theta$ is not necessarily true for a proper subfield of $\mathbb{Q}(\zeta_m)$. Let $M = \mathbb{Q}(\sqrt{3}) = \mathbb{Q}(\zeta_{12} + \zeta_{12}^{11})$ (the maximal real subfield) $\subset Q(\zeta_{12})$. Then on M, $\sigma_1 = \sigma_{11} = 1$ and $\sigma_5 = \sigma_7 = \sigma$, say. Then $\theta(M) = 1 + \sigma$ itself is in $\mathbb{Z}[G]$, i.e. $1 \cdot \theta \in \mathbb{Z}[G]$, so $\theta \in I$. But I' is generated by $5 - \sigma,\ 7 - \sigma$ and $11 - 1$ i.e. by 2 and $1 + \sigma$, hence $I'\theta$ is generated by 2θ and $\theta^2 = 2 + 2\sigma = 2\theta$ i.e. by 2θ. Hence $\theta = 1 + \sigma \notin I'\theta$, i.e. $I'\theta \subsetneqq I$.

This example also tells us that although θ corresponding to a cyclotomic field does not belong to the corresponding Stickelberger ideal, θ corresponding to a proper subfield M of $\mathbb{Q}(\zeta_m)$ may belong to $I(M)$.

Action of $\mathbb{Z}[G]$ on ideals and ideal classes : If $x = \sum x_\sigma \sigma \in \mathbb{Z}[G]$, then x acts on ideals of M by

$$A^x := \Pi(A^\sigma)^{x_\sigma}$$

This also gives rise to an action on ideal classes.

Stickelberger's Theorem. *The Stickelberger ideal of an abelian number field M annihilates the ideal class group of M. In other words, if A is an ideal of M and $x \in I(M)$, the Stickelberger ideal of M, then A^x is a principal ideal of M.*

§ 3. Prime Factorisation of Gauss Sums.

Let p be an odd prime and let $q = p^f$ be a power of p. Thus $p^f \equiv 1 \pmod{q-1}$ with f least. Let $\mathcal{P}$ be one of the $\phi(q-1)/f$ prime ideals of $\mathbb{Q}(\zeta_{q-1})$ lying above p. Now $\mathbb{Z}[\zeta_{q-1}] \pmod{\mathcal{P}}$ is a finite field of p^f elements. Also the $(q-1)^{\text{st}}$ roots of unity are distinct $\pmod{\mathcal{P}}$. There is a group isomorphism

$$\omega = \omega_{\mathcal{P}} : \mathbb{F}_q^\times \to (q-1)^{\text{st}} \text{ roots of } 1$$

satisfying for $a \in \mathbb{F}_q^\times$, $\mathbb{F}_q = Z[\zeta_{q-1}]/\mathcal{P}$,

$$\omega(a) \pmod{\mathcal{P}} = a.$$

i.e. $\omega(a)$ is that $(q-1)^{\text{st}}$ root of unity which lies in the coset $\pmod{\mathcal{P}}$ corresponding to a. Then ω is a character on $\mathbb{F}_q^\times$ called the Teichmüller character (corresponding to $\mathcal{P}$) and it depends upon the model of $\mathbb{F}_q$ given by $\mathbb{Z}[\zeta_{q-1}]/\mathcal{P}$. The character $\omega : (\mathbb{Z}[\zeta_{q-1}]/\mathcal{P})^\times \to \mathbb{Q}(\zeta_{q-1})$ can thus be described by saying that for $\beta \in \mathbb{Z}[\zeta_{q-1}], \omega(\beta + \mathcal{P})$ is that unique root of unity ζ_{q-1}^k for which $\zeta_{q-1}^k \equiv \beta \pmod{\mathcal{P}}$.

As $\mathbb{F}_q^\times$ is cyclic, the characters on $\mathbb{F}_q^\times$ form a cyclic group. Now, $\omega(\zeta_{q-1} + \mathcal{P}) = \zeta_{q-1}$, so ω has order $q-1$, i.e. ω generates the character group on $\mathbb{F}_q^\times$. Any character on $\mathbb{F}_q^\times$ may be written as $\omega^{-\alpha}$ for an integer $\alpha \pmod{q-1}$. We now obtain a factorisation of the Gauss sum $g(\omega^{-\alpha})$ in $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$.

Since $g(\chi)g(\overline{\chi}) = q$ for any character $\chi \neq \mathbf{1}$ on $\mathbb{F}_q^\times$, any prime divisor of $g(\omega^{-\alpha})$ for $\alpha \not\equiv 0 \pmod{q-1}$ comes from a prime divisor of p in $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$. p has $\phi(q-1)/f$ distinct prime divisors in $\mathbb{Q}(\zeta_{q-1})$ ($\phi(.)$ being the Euler ϕ function) and p totally ramifies in $\mathbb{Q}(\zeta_p)$, $(p) = (1-\zeta_p)^{p-1}$. As $(p, q-1) = 1$, the cyclotomic fields $\mathbb{Q}(\zeta_{q-1})$ and $\mathbb{Q}(\zeta_p)$ are linearly disjoint, and so a prime of $\mathbb{Q}(\zeta_{q-1})$ lying above p totally ramifies in $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$.

Let $\tilde{\mathcal{P}}$ be the (unique) prime of $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$ lying above $\mathcal{P}$. For $\alpha \in \mathbb{Z}$, we want to know the exponent, say $s(\alpha)$, of the power of this prime ideal $\tilde{\mathcal{P}}$ occurring in $g(\omega^{-\alpha})$. We have that

$$s(\alpha) = v_{\tilde{\mathcal{P}}}(g(\omega^{-\alpha})) \text{ depends only on } \alpha \pmod{q-1}.$$

We shall show that $s(\alpha) =$ the sum of the digits of α when α is expressed in base p. In terms of the function s we can also obtain the exponents corresponding to other prime ideals in $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$ occurring in $g(\omega^{-\alpha})$ and this will give us the prime ideal decomposition of $g(\omega^{-\alpha})$ in $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$.

Lemma 2 *Let* $\alpha,\ \beta \in \mathbb{Z}$.
(a) $s(0) = 0$.

(b) $0 \leq s(\alpha+\beta) \leq s(\alpha)+s(\beta)$.
(c) $s(\alpha+\beta) \equiv s(\alpha)+s(\beta) \pmod{p-1}$.
(d) $s(p\alpha) = s(\alpha)$.
(e) $\sum_{\alpha=1}^{q-2} s(\alpha) = (q-2)(f)(p-1)/2$.

Proof. (a) : $g(\omega^{-0}) = g(\mathbf{1}) = 1$.
(b) and (c) : By Prop. 1(f), $g(\omega^{-\alpha})g(\omega^{-\beta})/g(\omega^{-\alpha-\beta})$ is an algebraic integer in $\mathbb{Q}(\zeta_{q-1})$. Also $\mathcal{P} = \tilde{\mathcal{P}}^{p-1}$, so the values of $v_{\tilde{\mathcal{P}}}$ on $\mathbb{Q}(\zeta_{q-1})$ are divisible by $p-1$.
(d) : $g(\chi^p) = g(\chi)$.
(e) : As $g(\omega^{-\alpha})g(\omega^{\alpha}) = \pm q = \pm p^f$, we get

$$s(\alpha) + s(q-1-\alpha) = v_{\tilde{\mathcal{P}}}(p^f) = (p-1)f.$$

Adding for $\alpha = 1, 2, \cdots, q-2$ gives the result. □

Lemma 3 $s(\alpha) > 0$ *if* $\alpha \not\equiv 0 \pmod{q-1}$. $s(1) = 1$.

Proof. As $\pi = \zeta_p - 1 \in \tilde{\mathcal{P}}$, $\zeta_p \equiv 1 \pmod{\tilde{\mathcal{P}}}$, so

$$g(\omega^{-\alpha}) = -\sum \omega^{-\alpha}(a)\zeta_p^{\mathrm{Tr}(a)} \equiv -\sum \omega^{-\alpha}(a) \pmod{\tilde{\mathcal{P}}}.$$

But $\sum_{a \in \mathbb{F}_q} \omega^{-\alpha}(a) = 0$, as $\alpha \not\equiv 0 \pmod{q-1}$. Thus $g(\omega^{-\alpha}) \equiv 0 \pmod{\tilde{\mathcal{P}}}$. Hence $s(\alpha) > 0$. Next,

$$\begin{aligned} g(\omega^{-1}) &= -\sum \omega^{-1}(a)\zeta_p^{\mathrm{Tr}(a)} \\ &= -\sum \omega^{-1}(a)(1+\pi)^{\mathrm{Tr}(a)} \equiv -\sum \omega^{-1}(a)(1+\pi\mathrm{Tr}(a)) \pmod{\tilde{\mathcal{P}}^2} \\ &\equiv -\pi \sum \omega^{-1}(a)\mathrm{Tr}(a) \pmod{\tilde{\mathcal{P}}^2}. \end{aligned}$$

Regarding $\mathbb{F}_q$ as $\mathbb{Z}[\zeta_{q-1}] \pmod{\mathcal{P}}$ and noting that $a \mapsto a^p$ generates the Galois group of $\mathbb{Z}[\zeta_{q-1}] \pmod{\mathcal{P}}$ over $\mathbb{Z}/p\mathbb{Z}$, we have $\mathrm{Tr}(a) = a + a^p + \cdots + a^{p^{f-1}} \pmod{\mathcal{P}}$. Hence

$$\sum_{a \in \mathbb{F}_q} \omega^{-1}(a)\mathrm{Tr}(a) \equiv \sum_{\substack{a \not\equiv 0 \\ a \,(\mathrm{mod}\ \mathcal{P})}} a^{-1}(a + a^p + \cdots + a^{p^{f-1}}) \pmod{\mathcal{P}}.$$

For $0 < b < f$, $\sum_{a \not\equiv 0} a^{p^b-1} \equiv 0 \pmod{\mathcal{P}}$,
so this sum reduces to $\sum_{a \not\equiv 0} 1 = q-1 \equiv -1 \pmod{\mathcal{P}}$.

$$\text{This gives } g(\omega^{-1}) \equiv \pi \pmod{\tilde{\mathcal{P}}^2}.$$

But $\mathbb{Q}(\zeta_{q-1}, \zeta_p)/\mathbb{Q}(\zeta_p)$ is unramified at π. So $s(1) = v_{\tilde{\mathcal{P}}}(\pi) = 1$. □

Proposition 2 *Let $0 \leq \alpha < q-1$ and let the base-p expansion of α be $\alpha = a_0 + a_1 p + \cdots + a_{f-1}p^{f-1}$, $0 \leq a_i \leq p-1$. Then*

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

Proof. We have, $s(0) = 0$. As $s(1) = 1$ and $0 \leq s(\alpha+\beta) \leq s(\alpha) + s(\beta)$, and $s(\alpha+\beta) \equiv s(\alpha) + s(\beta) \pmod{p-1}$ we get $s(\alpha) = \alpha$ for $0 \leq \alpha \leq p-2$. This gives the result if $q = p$. If $q > p$, consider $s(p-1) \leq p-1$ and as $s(p-1) > 0$ and $s(p-1) \equiv p-1 \pmod{p-1}$, we get $s(p-1) = p-1$.

$$\begin{aligned} \text{Also } s(\alpha) &\leq s(a_0) + s(a_1 p) + \cdots + s(a_{f-1}p^{f-1}) \\ &= s(a_0) + s(a_1) + \cdots + s(a_{f-1}) \\ &= a_0 + a_1 + \cdots + a_{f-1}. \end{aligned}$$

Now as α runs over the integers from 0 to $q-1$, inclusive, each coefficient in the base-p expansion takes each of the values $0, 1, \cdots, p-1$ exactly p^{f-1} times, so

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

As $q-1 = (p-1) + (p-1)p + \cdots + (p-1)p^{f-1}$, omitting $\alpha = q-1$ we get

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

Hence the result follows. □

In summary, given a prime divisor $\mathcal{P}$ of p in $\mathbb{Q}(\zeta_{q-1})$ and given the corresponding Teichmüller character $\omega = \omega_{\mathcal{P}}$,

$$\begin{aligned} v_{\tilde{\mathcal{P}}}(g(\omega^{-\alpha})) &= s(\alpha) = a_0 + a_1 + \cdots + a_{f-1} \\ &= \text{ the sum of digits of } \alpha \text{ in its base-}p\text{ expansion.} \end{aligned}$$

We now get the prime ideal decomposition of a Gauss sum $g(\chi)$.

Let m be a fixed positive integer, p a prime, $(p, m) = 1$. Let f be the order of $p \pmod{m}$, so m divides $p^f - 1 = q - 1$. Let $\wp_0$ be any fixed prime of $\mathbb{Q}(\zeta_m)$ lying above p. Let $\tilde{\wp}_0$ be the unique prime of $\mathbb{Q}(\zeta_m, \zeta_p)$ lying above $\wp_0$, so $\tilde{\wp}_0{}^{p-1} = \wp_0$. Let $\mathcal{P}_0$ be a prime of $\mathbb{Q}(\zeta_{q-1})$ lying above $\wp_0$, and let $\tilde{\mathcal{P}}_0$ be the prime of $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$ lying above $\mathcal{P}_0$ (and $\tilde{\wp}_0$).

Let $\omega = \omega_{\mathcal{P}_0}$ be the Teichmüller character on $\mathbb{F}_q$ corresponding to $\mathcal{P}_0$. Let $\chi = \omega^{-d}$, where $d = (q-1)/m$. Then χ is a character on $\mathbb{F}_q$ of order m associated with the Teichmüller character ω obtained from $\mathcal{P}_0$. It may be observed that χ depends only on $\wp_0$ and it is in fact the reciprocal of the character $\omega^d = \omega^{(q-1)/m}$ which can be identified with the m^{th} power residue character associated with $\wp_0$ under the natural isomorphism $\mathbb{F}_q \cong \mathbb{Z}[\zeta_{q-1}]/\mathcal{P} \cong \mathbb{Z}[\zeta_m]/\wp_0$. (See [1].)

Now $\chi^m = 1$, so $g(\chi) \in \mathbb{Q}(\zeta_m, \zeta_p)$. Since $g(\chi)\overline{g(\chi)} = q = p^f$, the factorisation of $g(\chi)$ involves only primes of $\mathbb{Q}(\zeta_m, \zeta_p)$ above p, i.e. the conjugates over $\mathbb{Q}$ of $\tilde{\wp}_0$. For $(a, m) = 1$, let $\sigma_a : \zeta_m \mapsto \zeta_m^a$ be the element of $\mathrm{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$. For each such a, fix an extension of σ_a to $\mathbb{Q}(\zeta_{q-1}, \zeta_p)$ such that $\zeta_p^{\sigma_a} = \zeta_p$.

$$\begin{tikzcd}
 & \mathbb{Q}(\zeta_{q-1},\zeta_p) \quad \tilde{\mathcal{P}}_0 \arrow[dl, dash] \arrow[dr, dash] & \\
\mathbb{Q}(\zeta_{q-1}) \quad \mathcal{P}_0 \arrow[dr, dash] & & \mathbb{Q}(\zeta_m,\zeta_p) \quad \tilde{\wp}_0 \arrow[dl, dash] \\
 & \mathbb{Q}(\zeta_m) \quad \wp_0 \arrow[d, dash] & \\
 & \mathbb{Q} \quad p &
\end{tikzcd}$$

As $(p, m) = 1$, p is unramified and the decomposition group for p in $(\mathbb{Z}/m\mathbb{Z})^\times$ is the cyclic group generated by p (mod m), thus it is represented by $\{1, p, \cdots, p^{f-1}\}$.

Let R denote a set of representatives for $(\mathbb{Z}/m\mathbb{Z})^\times$ modulo this decomposition group. Then $\{\wp_0^{\sigma_a^{-1}} \mid a \in R\}$ is the set of conjugates of $\wp_0$. Now $\tilde{\wp}_0$ is the unique prime of $\mathbb{Q}(\zeta_m, \zeta_p)$ lying above $\wp_0$. Hence, the conjugates of $\tilde{\wp}_0$ are $\tilde{\wp_0}^{\sigma_a^{-1}}, a \in R$. Let $\tilde{\wp} = \tilde{\wp_0}^{\sigma_a^{-1}}$ be any of these. Then for $\chi = \omega^{-d}, d = \frac{q-1}{m}$,

$$v_{\tilde{\wp}}(g(\chi)) = v_{\tilde{\wp}_0}(g(\chi)^{\sigma_a}) = v_{\tilde{\wp}_0}(g(\chi^a)) = v_{\tilde{\mathcal{P}}_0}(g(\chi^a)) = v_{\tilde{\mathcal{P}}_0}(g(\omega^{ad})) = s(ad),$$

noting that $v_{\tilde{\wp}_0} = v_{\tilde{\mathcal{P}}_0}$ as $\tilde{\mathcal{P}}_0/\tilde{\wp}_0$ is unramified. We have thus proved

Proposition 3 *For* $\chi = \omega^{-d}, d = \frac{q-1}{m}$, $(g(\chi)) = \tilde{\wp}_0^{\sum_R s(ad)\sigma_a^{-1}}$.

Arguing as above we also have,

Proposition 3′.(Prime factorisation of Gauss sums) *For* $k \not\equiv 0 \,(\mathrm{mod}\ m)$,

$$(g(\chi^k)) = \tilde{\wp}_0^{\sum_R s(kad)\sigma_a^{-1}}.$$

§ 4. Proof of Stickelberger's Theorem

We first obtain a factorisation of $(g(\chi)^m)$ in terms of the Stickelberger element θ. Recall that $g(\chi)^m \in \mathbb{Q}(\zeta_m)$ as $\chi^m = 1$.

In the following lemma we have an alternative expression for the sum $s(h)$ of the base-p digits of an integer h.

Lemma 4 *Let* $0 \leq h < q-1$. *Then*

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

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

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

As $h < q-1$, some digit $a_i < p-1$, so RHS $< q-1$. Hence

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

Summing over i, we get

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

This gives the result. □

Proposition 4 *If* $\chi = \chi_{\wp_0}$ *is the reciprocal of the* m^{th} *power residue character corresponding to a prime* $\wp_0$ *of* $\mathbb{Q}(\zeta_m)$, *lying above* p, *then*

(i) $$(g(\chi)^m) = \wp_0^{m\theta} = \wp_0^{\sum_{1 \leq a \leq m, (a,m)=1} a\sigma_a^{-1}},$$

as ideals in $\mathbb{Q}(\zeta_m)$.

(ii) (Prime factorisation of Jacobi sums) *Let* j, k *be integers such that* $jk(j+k) \not\equiv 0 \,(\text{mod } m)$. *Let*

$$\theta_{j,k} = \sum_{\substack{a \,(\text{mod } m) \\ (a,m)=1}} \left(\left[\frac{(j+k)a}{m}\right] - \left[\frac{ja}{m}\right] - \left[\frac{ka}{m}\right] \right) \sigma_a^{-1}$$

Then $\left(J(\chi^j, \chi^k)\right) = \wp_0^{\theta_{j,k}}$ *as ideals in* $\mathbb{Q}(\zeta_m)$.

Proof. We have, by Lemma 4,

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

Hence by Proposition 3,

$$(g(\chi)^m) = \tilde{\wp}_0^{m\sum_R s(ad)\sigma_a^{-1}} = \tilde{\wp}_0^{m(p-1)\sum_{i=0}^{f-1}\sum_R \left\{\frac{p^i a}{m}\right\}\sigma_a^{-1}}.$$

As $\tilde{\wp}_0{}^{p-1} = \wp_0$ and $\sigma_{p^i}(\wp_0) = \wp_0$ (since $p^i \in$ the decomposition group of $\wp_0$), we get $(g(\chi)^m) = \wp_0^{m\sum_{b\,(\mathrm{mod}\ m),(b,m)=1}\left\{\frac{b}{m}\right\}\sigma_b^{-1}} = \wp_0^{m\theta}$ as ideals in $\mathbb{Q}(\zeta_m, \zeta_p)$.

Due to unique factorisation of ideals in $\mathbb{Q}(\zeta_m)$, $(g(\chi)^m) = \wp_0^{m\theta}$ holds also in $\mathbb{Q}(\zeta_m)$.

Next write $\theta_j = \sum\limits_{\substack{a\,(\mathrm{mod}\ m)\\(a,m)=1}} \left\{\dfrac{ja}{m}\right\} \sigma_a^{-1}$. Then as before, $(g(\chi^j)^m) = \wp_0^{m\theta_j}$.

Let j, k be such that $j, k, j+k \not\equiv 0\,(\mathrm{mod}\ m)$. Then

$$m\theta_j + m\theta_k - m\theta_{j+k} = m \sum_{(a,m)=1} \left(\left[\frac{(j+k)a}{m}\right] - \left[\frac{ja}{m}\right] - \left[\frac{(ka}{m}\right] \right) \sigma_a^{-1}$$

so that, using Prop. 1(f),

$$\Big(J(\chi^j, \chi^k)^m\Big) = \wp_0^{m\theta_{j,k}}, \text{ where } \theta_{j,k} \in \mathbb{Z}[G].$$

Hence $(J(\chi^j, \chi^k)) = \wp_0^{\theta_{j,k}}$, by unique factorisation in $\mathbb{Q}(\zeta_m)$. □

Remark : The above proposition shows that if $\wp_0$ is a prime of $\mathbb{Q}(\zeta_m)$ such that $\wp_0 \not| \, m$, taking p to be the prime of $\mathbb{Z}$ lying below $\wp_0$, $\wp_0^{m\theta}$ and $\wp_0^{\theta_{j,k}}$ are principal in $\mathbb{Q}(\zeta_m)$.

We next prove such a result for an abelian number field. We first have

Proposition 5 *Let M be a subfield of $\mathbb{Q}(\zeta_m)$. Let A be an ideal of M such that $(A, m) = 1$. Let $\theta = \theta(M)$. Then $A^{m\theta}$ is principal in $\mathbb{Z}[\zeta_m]$. Further, for $\beta \in \mathbb{Z}[G], A^{\beta\theta}$ is principal in $\mathbb{Z}[\zeta_m]$.*

Proof. Let A be an ideal of $M \subset \mathbb{Q}(\zeta_m)$ such that $(A, m) = 1$. Write $A = \prod \wp_i$, $\wp_i$ being prime ideals of $\mathbb{Q}(\zeta_m)$ not necessarily distinct. Let p_i be primes of $\mathbb{Z}$ lying below $\wp_i$. Let $P =$ the square-free part of $\prod p_i$.

Then, by Proposition 4, $A^{m\theta} = \prod \wp_i^{m\theta}$ is principal in $\mathbb{Q}(\zeta_{mP})$.

Extending elements of $\mathrm{Gal}(M/\mathbb{Q})$ to $\mathrm{Gal}(\mathbb{Q}(\zeta_{mP}/\mathbb{Q})$ by fixing ζ_P, we assume that $m\theta \in \mathbb{Z}[\mathrm{Gal}(\mathbb{Q}(\zeta_{mP})/\mathbb{Q})]$.

Let $\chi_{\wp_i}$ be the character of order m on the finite field $\mathbb{Z}[\zeta_m]/\wp_i$ defined using $\wp_i$. Then $g(\chi_{\wp_i}) \in \mathbb{Z}[\zeta_{mp_i}]$ and as ideals of $\mathbb{Z}[\zeta_{mP}]$,

$A^{m\theta} = \prod \wp_i^{m\theta} = \prod (g(\chi_{\wp_i})^m) = (\gamma^m)$ where $\gamma = \prod g(\chi_{\wp_i}) \in \mathbb{Z}[\zeta_{mP}]$.

As each $g(\chi_{\wp_i})^m \in \mathbb{Z}[\zeta_m]$, $A^{m\theta}$ is principal in $\mathbb{Z}[\zeta_m]$ itself.

Next suppose $\beta \in \mathbb{Z}[G], G = \mathrm{Gal}(M/\mathbb{Q})$ such that $\beta\theta \in \mathbb{Z}[G]$, θ being the Stickelberger element for M. Then, as above, extending the elements of $\mathrm{Gal}(M/\mathbb{Q})$ to $\mathrm{Gal}(\mathbb{Q}(\zeta_{mP})/\mathbb{Q})$ assume that $\beta\theta \in \mathrm{Gal}(\mathbb{Q}(\zeta_{mP})/\mathbb{Q})$.

Then

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

as ideals of $\mathbb{Z}(\zeta_{mP})$.

We next prove that γ^β in fact $\in \mathbb{Q}(\zeta_m)$, so that $A^{\beta\theta}$ is principal in $\mathbb{Z}[\zeta_m]$.

Firstly, $\gamma^{m\beta} = (\gamma^m)^\beta \in \mathbb{Q}(\zeta_m)$ and $(\gamma^{m\beta}) = (\gamma^\beta)^m = (A^{\beta\theta})^m$ is the m^{th} power of an ideal of $\mathbb{Z}[\zeta_m]$.

If $\wp$ is a prime ideal of $L = \mathbb{Q}(\zeta_m)$ such that $(\wp, m) = 1$, locally, $(\gamma^{m\beta}) = (\pi_\wp^{mt}), t \geq 0, \pi_\wp$ being a local uniformizer. So $\gamma^{m\beta} = \epsilon\pi_\wp^{mt}, \epsilon$ being a local unit. So $\gamma^\beta = \epsilon^{1/m}\pi_\wp^t$. This gives $L_\wp(\gamma^\beta) = L_\wp(\epsilon^{1/m})$ as a Kummer extension of $L_\wp$ where $\pi_\wp \not| \epsilon$. Hence the extension is unramified. Thus $\mathbb{Q}(\zeta_m, \gamma^\beta)/\mathbb{Q}(\zeta_m)$ is unramified at $\wp$.

Also $\mathbb{Q}(\zeta_m) \subset \mathbb{Q}(\zeta_m, \gamma^\beta) \subset \mathbb{Q}(\zeta_m, \zeta_P)$, so ramification can occur only at primes dividing P. But $(P, m) = 1$, so the extension is unramified.

In view of the following lemma , we see that $\gamma^\beta \in \mathbb{Q}(\zeta_m)$:

Lemma 5 *Let $m, n \geq 1$ and $m|n$. If $K/\mathbb{Q}(\zeta_m)$ is unramified at all primes and $K \subset \mathbb{Q}(\zeta_n)$, then $K = \mathbb{Q}(\zeta_m)$.*

Proof. Let p be a prime dividing n/m. Then $\mathbb{Q}(\zeta_{mp})/\mathbb{Q}(\zeta_m)$ is totally ramified at the primes above p. Hence $K \cap \mathbb{Q}(\zeta_{mp})/\mathbb{Q}(\zeta_m)$ is totally ramified and unramified at a prime above p. Thus, $K \cap \mathbb{Q}(\zeta_{mp}) = \mathbb{Q}(\zeta_m)$. So $[K(\zeta_{mp}) : \mathbb{Q}(\zeta_{mp})] = [K : \mathbb{Q}(\zeta_m)]$. Now a lift of an unramified extension is unramified at all primes. Hence we can argue similarly with m replaced by mp. Continuing like this, finally,

$$[K : \mathbb{Q}(\zeta_m)] = [K(\zeta_n) : \mathbb{Q}(\zeta_n)] = [\mathbb{Q}(\zeta_n) : \mathbb{Q}(\zeta_n)] = 1.$$

Hence $K = \mathbb{Q}(\zeta_m)$. □

Next to prove that $\gamma^\beta \in M$. Let $\wp_i$ be a prime divisor of A in $\mathbb{Q}(\zeta_m)$. Let p be the rational prime lying below $\wp_i$. Let $q = p^f$ where f is the residue class degree of $\wp_i$. Let $\mathcal{P}$ be a prime of $\mathbb{Q}(\zeta_{q-1})$ lying over $\wp_i$. Now $\chi_{\wp_i}$ can be defined in terms of $\mathcal{P}$ and hence write $\chi_{\wp_i} = \chi_\mathcal{P}$.

Let $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_{q-1})/M)$. Then σ gives rise to a map

$$\sigma : \mathbb{Z}[\zeta_{q-1}] \pmod{\mathcal{P}} \to \mathbb{Z}[\zeta_{q-1}] \pmod{\mathcal{P}^\sigma},$$

and we have, if $\chi_\mathcal{P}(a) = \zeta$, then $\chi_{\mathcal{P}^\sigma}(a) = \zeta^\sigma$. Thus $\chi_{\mathcal{P}^\sigma} = \chi_\mathcal{P}^\sigma$. But $\chi_\mathcal{P}^m = 1$, so $\chi_{\mathcal{P}^\sigma} = \chi_\mathcal{P}$ for $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_{q-1})/\mathbb{Q}(\zeta_m))$. Thus $\chi_\mathcal{P}$ depends only on $\wp_i$, so we can in fact use the notation $\chi_{\wp_i}$.

As above, for $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_m)/M)$, $\chi_{\wp_i}^\sigma = \chi_{\wp_i^\sigma}$. Extending σ to $\mathbb{Q}(\zeta_{mp})$ by letting $\zeta_p \mapsto \zeta_p$, $g(\chi_{\wp_i})^\sigma = g(\chi_{\wp_i}^\sigma) = g(\chi_{\wp_i^\sigma})$. Now for $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_m)/M)$, $A^\sigma = A$, so σ permutes all the prime divisors $\wp_i$ of A in $\mathbb{Q}(\zeta_m)$.

Hence

$$\gamma^{\beta\sigma} = \prod g(\chi_{\wp_i})^{\beta\sigma} = \prod g(\chi_{\wp_i^\sigma})^\beta = \gamma^\beta.$$

But already $\gamma^\beta \in \mathbb{Q}(\zeta_m)$. Hence $\gamma^\beta \in M$. So $A^{\beta\theta} = (\gamma^\beta)$ is principal in M.

Finally, let A be any ideal of M. Write $A = BC$ where $(C, m) = 1$ and a prime $\wp$ divides B if and only if $\wp$ divides (A, m). By approximation theorem, first get an integer $b \in M$ such that for every prime divisor $\wp$ of m in M, $v_\wp(b) = v_\wp(A)$, so that $(b) = BD$ with $(D, m) = 1$. Now there is an ideal E of M such that $(E, m) = 1$ and DE is principal $= (c)$, say, where c is an integer in M. (Use the factorisation in M of primes below D.) Then, $(b)CE = BDCE = ADE = A(c)$. Take $A_1 = CE$ and $a = b/c$. Thus we have $A = (a)A_1$, with $a \in M$, and $(A_1, m) = 1$. Then $A^{\beta\theta} = (a^{\beta\theta})A_1^{\beta\theta}$, which is principal. This completes the proof of Stickelberger's theorem .

(For a simpler proof of Stickelberger's theorem in the case of the full cyclotomic field $\mathbb{Q}(\zeta_m)$, see the article of C. S. Yogananda [9].)

§ 5. Herbrand's Theorem

Herbrand's theorem is an interesting application of Stickelberger's theorem for the cyclotomic field $\mathbb{Q}(\zeta_p)$. Herbrand's theorem and its converse characterise the Bernoulli numbers $B_2, B_4, \cdots, B_{p-3}$ divisible by p in terms of the structure of the $p-$Sylow subgroup of the ideal class group of $\mathbb{Q}(\zeta_p)$.

Bernoulli Numbers. The Bernoulli numbers $B_n,\ n \geq 0,$ were first defined by Jakob (James) Bernoulli (1654-1705) and were so designated by L. Euler. They are defined by the exponential generating function (the series being convergent for $|\,t\,| < 2\pi$)

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

$$\text{We have,} \quad \frac{t}{e^t - 1} + \frac{t}{2} \quad = \quad \frac{t(e^t + 1)}{2(e^t - 1)} = \frac{t}{2} \cdot \frac{e^{t/2} + e^{-t/2}}{e^{t/2} - e^{-t/2}},$$

which is an even function of t. Hence $B_{2k-1} = 0$ for $k \geq 2$. First few values of B_n are : $B_0 = 1,\ B_1 = -\dfrac{1}{2},\ B_2 = \dfrac{1}{6},\ B_4 = -\dfrac{1}{30},\ B_6 = \dfrac{1}{42},\ B_8 = -\dfrac{1}{30},$ $B_{10} = \dfrac{5}{66},\ B_{12} = -\dfrac{691}{2730},\ B_{14} = \dfrac{7}{6}$. These can be successively obtained from the recurrence relation for Bernoulli numbers viz.

$$(m+1)B_m = -\sum_{k=0}^{m-1} \binom{m+1}{k} B_k.$$

Bernoulli proved that the sum $S_r(n)$ of r^{th} powers of first n natural numbers is given in terms of the B_m as

$$(r+1)S_r(n) = n^{r+1} - \binom{r+1}{1} B_1 n^r + \binom{r+1}{2} B_2 n^{r-1} + \cdots + \binom{r+1}{r} B_r n.$$

Euler obtained the values of the Riemann zeta function $\zeta(s)$ at positive even integral values of s in terms of Bernoulli numbers as

$$\zeta(2n) = (-1)^{n+1} \frac{2^{2n-1} B_{2n}}{(2n)!} \pi^{2n}.$$

Thus the even-indexed Bernoulli numbers are nonzero and they alternate in sign. As $\zeta(2n) > 1$ and $e^n > \frac{n^n}{n!}$ we get $|\frac{B_{2n}}{2n}| > \frac{1}{e\pi}\left(\frac{n}{e\pi}\right)^{2n-1}$, so that $|\frac{B_n}{n}| \to \infty$ as $n \to \infty$, (n even).

It was proved by Von Staudt Clausen that the denominator of a Bernoulli number B_n, with n positive and even, is square-free. More precisely, he proved that for n even $> 0, B_n + \sum\limits_{(p-1)|n} \frac{1}{p}$ is an integer, (thus 2, 3 always appear in the denominator of each such B_n). (See [2], [4], [6].)

Exercise. If $n = 2q$ where q is a prime of the form $3k+1$, then $B_n \equiv \frac{1}{6}$ (mod 1).

One also has the following congruence for Bernoulli numbers :

Kummer Congruence. Suppose m, n are positive even integers such that $m \equiv n \not\equiv 0 \pmod{(p-1)}$. Then

$$\frac{B_m}{m} \equiv \frac{B_n}{n} \pmod{p}.$$

More generally, if m, n are positive even integers and

$$m \equiv n \pmod{(p-1)p^\alpha} \text{ and } n \not\equiv 0 \pmod{(p-1)},$$

then

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

as $n \to \infty$, (n even.)

A prime p is called irregular if $p|B_j$ for some $j = 2, 4, \cdots, p-3$. Otherwise p is called a regular prime.

As $|\frac{B_n}{n}| \to \infty$ with n (n even). Suppose $p_1, \cdots, p_r$ are irregular primes and N is large so that $m = N(p;_1 -1) \cdots (p_r - 1)$ satisfies $|\frac{B_m}{m}| > 1$. Let p be a prime in the numerator of $\frac{B_m}{m}$. Now $(p_i - 1)|m$, so all the p_i appear in the denominator of b_m; so $p \neq p_i$ for all i. Also $(p-1) \nmid m$, for otherwise p would be in the denominator of B_m. Let $m' \equiv m \pmod{p-1}$, $0 < m' < p-1$. Then by the Kummer congruence, $p|B_{m'}$. Therefore p is irregular. This shows that the number of irregular primes is infinite. At present it is not known whether there are infinitely many regular primes or not. However, numerical evidence (W. Johnson) and probabilistic arguments (C. L. Siegel) indicate that about 61% of all primes are regular.

Kummer proved that if $p \not| $ the class number h of $\mathbb{Q}(\zeta_p)$, then the equation $x^p + y^p = z^p$ has no positive integral solution. (See [6], [7].)

Using his results on the class number formulas for cyclotomic fields, Kummer also proved that the condition $p \not| \, h(\mathbb{Q}(\zeta_p))$ is actually equivalent to the regularity of p, i.e. $p \not| \, B_2, B_4, \cdots, B_{p-3}$. (See [6].)

Generalised Bernoulli Numbers. A Dirichlet character $\pmod n$ is a multiplicative homomorphism $\chi : (\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^\times$. If $n|m$, it induces a homomorphism $(\mathbb{Z}/n\mathbb{Z})^\times \to \mathbb{C}^\times$ by composition with the natural map from $(\mathbb{Z}/m\mathbb{Z})^\times \to (\mathbb{Z}/n\mathbb{Z})^\times$. We choose n minimal which induces such a Dirichlet character $\pmod m$ and call n to be the conductor of the character denoted by f or f_χ.

The character $\chi \pmod 8$ defined by $\chi(1) = \chi(5) = 1$ and $\chi(3) = \chi(7) = -1$ has conductor 4 as it can be induced from the character $\pmod 4$ defined by $\chi(1) = 1, \chi(3) = -1$, and 4 is minimal. The character $\pmod 6$ defined by $\chi(1) = 1, \chi(5) = -1$ has conductor 3 as it can be induced from the character $\pmod 3$ defined by $\chi(1) = 1, \chi(3) = -1$, and 3 is minimal.

Given a Dirichlet character χ of modulus and conductor f, the generalised Bernoulli numbers $B_{n,\chi}$ are defined by

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

For $\chi = \mathbf{1}$, the character of conductor 1, we have $\displaystyle\sum_{n=0}^{\infty} B_{n,\mathbf{1}} \, \frac{t^n}{n!} = \frac{te^t}{e^t-1} = \frac{t}{e^t-1} + t$. Thus $B_{n,\mathbf{1}} = B_n$ except when $n = 1$, when we have $B_{1,\mathbf{1}} = \frac{1}{2}, B_1 = -\frac{1}{2}$. Also note that if $\chi \neq \mathbf{1}$, $\sum_{a=1}^{f} \chi(a) = 0$, and hence,

$$\begin{aligned} \text{i)} \quad B_{0,\chi} &= 0, \\ \text{ii)} \quad B_{1,\chi} &= \frac{1}{f}\sum_{a=1}^{f} \chi(a)a. \end{aligned}$$

Let G be a finite abelian group and $\widehat{G}$ its character group. Let R be a commutative ring with unity which contains the values of all $\chi \in \widehat{G}$ and in which $|G|$ is invertible (e.g. $R = \overline{\mathbb{Q}}$). For $\chi \in \widehat{G}$, define

$$\varepsilon_\chi = \frac{1}{|G|}\sum_{\sigma\in G}\chi(\sigma)\sigma^{-1} \in R[G].$$

Then we have

(a) $\varepsilon_\chi^2 = \varepsilon_\chi$,

(b) $\varepsilon_\chi\varepsilon_\psi = 0$ if $\chi \neq \psi$,

(c) $\mathbf{1}= \sum\limits_{\chi\in\widehat{G}}\varepsilon_\chi$,

(d) $\varepsilon_\chi\sigma = \chi(\sigma)\varepsilon_\chi$.

ε_χ are called as the orthogonal idempotents of the group ring $R[G]$. Let M be a module over $R[G]$. Let $M_\chi = \varepsilon_\chi M$. Then

$$\varepsilon_\chi M_\chi = \varepsilon_\chi^2 M = \varepsilon_\chi M = M_\chi.$$

Using $\mathbf{1}= \sum_{\chi\in\widehat{G}}\varepsilon_\chi$, we see that $M = \sum\limits_\chi M_\chi$. Next suppose $0 = \sum m_\chi$ with $m_\chi \in M_\chi$. Then as $m_\chi = \varepsilon_\chi m'_\chi$ with $m'_\chi \in M$, multiplying by ε_ψ we get $0 = \varepsilon_\psi m'_\psi = m_\psi$. Hence $M = \oplus M_\chi$. Also for $\sigma \in G$, and $m_\chi \in M_\chi$, writing $m_\chi = \varepsilon_\chi\ m'_\chi$ with $m'_\chi \in M$,

$$\sigma m_\chi = \sigma\varepsilon_\chi m'_\chi = \chi(\sigma)\varepsilon_\chi m'_\chi = \chi(\sigma)m_\chi.$$

Thus for the action of $\sigma \in G$ on the $R[G]$-module M, the elements of M_χ are eigenvectors with eigenvalue $\chi(\sigma)$.

Now let $R = \mathbb{Z}_p$, the ring of p-adic integers. Let $G = \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}) \approx (\mathbb{Z}/p\mathbb{Z})^\times$. As $\mathbb{Z}_p \supset \mathbb{Z}_{(p)}$ (the localization of $\mathbb{Z}$ at p), $|G| = p-1$ is invertible in $\mathbb{Z}_p$. Also $\mathbb{Z}_p$ contains all the $(p-1)^{\text{st}}$ roots of unity and these are congruent (mod p) to the elements $1, 2, \cdots, p-1$. $\widehat{G}$ denotes the group of characters of G with values in $\mathbb{Z}[\zeta_p]$.

Write elements of G as $\sigma_a : \zeta_p \mapsto \zeta_p^a, \quad (a,p) = 1$. Define $\omega \in \widehat{G}$ as follows. Write for convenience $\omega(\sigma_a)$ as $\omega(a)$. For $a \in (\mathbb{Z}/p\mathbb{Z})^\times, a$ is a $(p-1)^{\text{st}}$ root of unity. Define $\omega(a)$ to be the $(p-1)^{\text{st}}$ root of unity in $\mathbb{Z}_p$, which comes from a; i.e. $\omega(a)$ is such that $\omega(a) \equiv a \pmod{p}$. Then $\omega \in \widehat{G}$ and $\widehat{G} = \{\omega^i | 0 \leq i \leq p-2\}$. Then the orthogonal idempotents of $\mathbb{Z}_p[G]$ are

$$\varepsilon_i = \varepsilon_{\omega^i} = \frac{1}{p-1}\sum_{a=1}^{p-1}\omega^i(a)\sigma_a^{-1}, 0 \leq i \leq p-2.$$

Define

$$\begin{aligned}\varepsilon_- &= \sum_{i \text{ odd}} \varepsilon_i = \frac{1}{p-1}\sum_{a=1}^{p-1}\left(\sum_{i \text{ odd}} \omega^i(a)\right)\sigma_a^{-1} = \frac{1-\sigma_{-1}}{2},\\ \text{and}\\ \varepsilon_+ &= \sum_{i \text{ even}} \varepsilon_i = \frac{1+\sigma_{-1}}{2}.\end{aligned}$$

Then for any $\mathbb{Z}_p[G]$-module A, $A = A^- \oplus A^+$, where $A^- = \varepsilon_- A$ and $A^+ = \varepsilon_+ A$.

Let $\theta = \dfrac{1}{p}\displaystyle\sum_{a=1}^{p-1} a\sigma_a^{-1}$ be the Stickelberger element for $\mathbb{Q}(\zeta_p)$. Using (d) above, we get,

$$\begin{aligned}\varepsilon_i\theta &= \frac{1}{p}\sum_{a=1}^{p-1} a\omega^i(a)\varepsilon_i = B_{1,\omega^{-i}}\varepsilon_i\\ \text{and}\quad \varepsilon_i(c-\sigma_c)\theta &= (c-\omega^i(c))B_{1,\omega^{-i}}\varepsilon_i\end{aligned}$$

Any abelian p-group A is a $\mathbb{Z}_p$-module by the action $(\displaystyle\sum_{j=0}^{\infty} b_j p^j)a = \sum_{j=0}^{\infty} b_j(p^j a)$, since the latter sum is finite.

From now on, let A be the p-Sylow subgroup of the ideal class group of $\mathbb{Q}(\zeta_p)$. Then the Galois group G also acts on A, so A is a $\mathbb{Z}_p[G]$-module. Write $A = \displaystyle\bigoplus_{i=0}^{p-2} A_i$, where $A_i = \varepsilon_i A$. By Stickelberger's theorem, for $(c,p) = 1, (c-\sigma_c)\theta$, which is in the Stickelberger ideal of $\mathbb{Q}(\zeta_p)$, annihilates in particular A, and hence each $A_i = \varepsilon_i A$. Thus, $(c-\omega^i(c))B_{1,\omega^{-i}}$ annihilates A_i.

Note. $p\theta = \sum_{(a,p)=1} a\sigma_a^{-1}$ and $(p-1)\varepsilon_1 = \sum_{(a,p)=1}\omega(a)\sigma_a^{-1}$. As $w(a) \equiv a \pmod{p}, p\theta \equiv (p-1)\epsilon_1 \pmod{p}$)). As A_i are p-groups, it may be possible to accept that $p\theta$ annihilates A_i for $i \neq 1$. We require Stickelberger's theorem to conclude that $p\theta$ annihilates A_1.

Consider $0 \leq i \leq p-2$.

Case 1. If $i \neq 0$ is even,

$$B_{1,\omega^{-i}} = \frac{1}{p}\sum_{(a,p)=1}\omega^{-i}(a)a = \frac{1}{p}\cdot\frac{1}{2}\sum_{(a,p)=1} w^{-i}(a)\{a+(p-a)\} = 0,$$

so we do not get any information in this case.

Case 2. If $i = 0, B_{1,\mathbf{1}} = \dfrac{1}{2}$, so $\frac{c-1}{2}$ annihilates A_0.

Taking some $1 < c \leq p-1$, $\frac{c-1}{2}$ is invertible in $\mathbb{Z}_p$, so $A_0 = 0$. This is otherwise obvious, because $A_0 = \epsilon_0 A$ and $\epsilon_0 = \text{Norm}/(p-1)$.
Case 3. Let i be odd. Suppose $i = 1$. Let $c = 1+p$. Then

$$(c - \omega(c))B_{1,\omega^{-1}} = pB_{1,\omega^{-1}} = \sum_{a=1}^{p-1} a\omega^{-1}(a) \equiv p - 1 \not\equiv 0 \pmod{p},$$

so $(c-\omega(c))B_{1,\omega^{-1}}$ is unit of $\mathbb{Z}_p$, so that $A_1 = 0$. Next suppose $i \neq 1$. Choose an integer c (e.g. a primitive root mod p) such that $c \not\equiv c^i \equiv \omega^i(c) \pmod{p}$. Then $c - \omega^i(c)$ is a unit of $\mathbb{Z}_p$, so $B_{1,\omega^{-i}}$ is in $\mathbb{Z}_p$ and it annihilates A_i. This gives

Proposition 6 $A_0 = A_1 = 0$. *For* $i = 3, 5, \cdots, p-2, B_{1,\omega^{-i}} \in \mathbb{Z}_p$ *and it annihilates* A_i.

Herbrand's Theorem. *Let i be odd,* $3 \leq i \leq p-2$. *If* $A_i \neq 0$, *then* $p \mid$ *the Bernoulli number* B_{p-i}.

Proof. Suppose $A_i \neq 0$. Then $B_{1,\omega^{-i}}$ must be a non-unit in $\mathbb{Z}_p$, i.e. $B_{1,\omega^{-i}} \equiv 0 \pmod{p}$. Now it can be proved that (see Cor. 5.15, [6]) if n is odd and $n \not\equiv -1 \pmod{p-1}$, then $B_{1,\omega^n} \equiv \dfrac{B_{n+1}}{n+1} \pmod{p}$, and both the sides are p-integral. Hence, $B_{1,w^{-i}} \equiv \dfrac{B_{p-i}}{p-i} \pmod{p}$, Hence $p|B_{p-i}$. This proves Herbrand's theorem. □

The converse of Herbrand's theorem is

Ribet's Theorem. Let i be odd, $3 \leq i \leq p-2$. If $p|B_{p-i}$, then $A_i \neq 0$.

For an elementary proof of Ribet's Theorem see Chapter 15 of [6]. See also [5]. For irregular primes, Herbrand-Ribet give a piece-by-piece information about which Bernoulli numbers are divisible by p in terms of the $\mathbb{Z}_p[G]$-module structure of the p-Sylow subgroup of the ideal class group of $\mathbb{Q}(\zeta_p)$.

The number $i(p)$ of $B_j, j = 2, 4, \cdots, p-3$, which are divisible by p is called the index of irregularity of p. As a consequence of Ribet's theorem we get that the p-rank of the ideal class group of $\mathbb{Q}(\zeta_p)$ (i.e.the number of summands when A is written as a direct sum of cyclic groups of prime power order) is at least $i(p)$, i.e. the number of Bernoulli numbers divisible by p.

Vandiver's Conjecture. This conjecture says that $p \nmid h^+(\mathbb{Q}(\zeta_p))$, the class number of the maximal real subfield $\mathbb{Q}(\zeta_p)^+$ of $\mathbb{Q}(\zeta_p)$. The conjecture has already been checked by computer for all primes up to 4,000,000 and even if it is false it is expected to hold for most primes.

It is known that if Vandiver's conjecture holds, then the p-rank of the ideal class group of $\mathbb{Q}(\zeta_p)$ equals the number of Bernoulli numbers divisible by p. (See also the article of E. Ghate [3].)

Acknowledgements. I thank S. V. Kanetkar and Vijay Patankar for carefully going through the article and Dinesh Thakur for suggesting this topic for my lectures.

REFERENCES

1. S. D. Adhikari, The Early Reciprocity Laws: From Gauss to Eisenstein, These proceedings.
2. G. H. Hardy and E. M. Wright, An Introduction to the Theory of Numbers, Fourth Ed., Cambridge Uni. Press, 1975.
3. Eknath Ghate, Vandiver's Conjecture via K-theory, These proceedings.
4. K. Ireland and M. Rosen, A Classical Introduction to Modern Number Theory, GTM 84, Springer-Verlag, New York Inc., 1982.
5. C. Khare, Notes on Ribet's Converse to Herbrand, These proceedings.
6. Paulo Ribenboim, 13 Lectures on Fermat's Last Theorem, Springer-Verlag New York Inc., 1979.
7. Dinesh S. Thakur, Fermat's Last Theorem for Regular Primes, These proceedings.
8. L. C. Washington, Introduction to Cyclotomic Fields, Second Ed., GTM 83, Springer-Verlag New York Inc.,1997.
9. C. S. Yogananda, Stickelberger Revisited, These proceedings.

S. A. Katre
Department of Mathematics,
University of Pune,
Pune-411 007.
e-mail: sakatre@@math.unipune.ernet.in

An Introduction to L-functions

Ravi Raghunathan

Introduction.

These lecture notes on L-functions consist of two parts. Part I deals with complex L-functions and is supposed to motivate Part II which deals with p-adic L-functions. I have attempted to make the notes as self-contained as possible and most of the more difficult exercises have hints. Solutions to the exercises can be found in one or the other of the references in the bibliography. I am grateful to Professor Dipendra Prasad for carefully reading an initial version of these notes, pointing out several errors and making numerous suggestions for improving these notes. I would also like to thank Professor S. D. Adhikari for pointing out several errors in the original version.

Part I: Complex valued L-functions

1. The Riemann zeta-function. Let $\{a_n\}_{n=1}^{\infty}$ be a sequence of complex numbers whose properties we wish to understand. One very useful way in which to study such a sequence is to associate to it a function of a *complex variable* $D(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$, or even more generally, $D(s) = \sum_{n=1}^{\infty} \frac{a_n}{\lambda_n^s}$, where $\{\lambda_n\}_{n=1}^{\infty}$ is a sequence of positive real numbers such that $\lambda_n \to \infty$. Such a series is called a Dirichlet series. The basic philosophy is that interesting properties of the numbers a_n are mirrored in the properties of the function $D(s)$, so studying the latter gives us insights into the former. Because $D(s)$ is a function of a complex variable we can use tools from complex analysis to study this function and historically this approach has proved very successful (This approach is not dissimilar to forming generating functions from sequences to obtain *real analytic functions* or power series in a real variable. One then studies the properties of this auxiliary series - which often satisfies a differential equation - to obtain information about the original sequence.). In number theory questions about prime numbers and their distributions can often be answered by such techniques.

The simplest possible example arises when we take $\lambda_n = n$ and $a_n \equiv 1$. This gives us the famous Riemann ζ-function, $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$, which serves as the proto-typical example for Dirichlet series arising in number theory. We will discuss some of its properties below. We will see a number of other examples later which will also have similar properties.

(a) Half-plane of convergence. For $s = \sigma + it$ we have

$$|\zeta(s)| \leq \sum_{n=1}^{\infty} \frac{1}{|n^s|} = \sum_{n=1}^{\infty} \frac{1}{|n^{\sigma} n^{it}|} = \sum_{n=1}^{\infty} \frac{1}{n^{\sigma}}.$$

We know that the last series converges for $Re(s) > 1$, so one checks easily that the series defining $\zeta(s)$ converges uniformly and absolutely on compact sets in this half-plane. Hence, (see Lemmas 2.1 to 2.3 below) $\zeta(s)$ is *holomorphic* (equivalently *analytic*) in this half-plane. In fact, we shall see that $(s-1)\zeta(s)$ can be analytically continued to the whole complex plane $\mathbf{C}$.

(b) Euler product. For $Re(s) > 1$, consider the product

$$\prod_{p\ prime} \frac{1}{\left(1 - \frac{1}{p^s}\right)}. \tag{1}$$

Expanding each individual factor as a geometric series (in p^{-s}) we see that the above expression is nothing but

$$\prod_{p\ prime} \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \dots\right) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

where the last equality follows by opening the parentheses and using the unique factorisation theorem for primes! The expression (1) is called an *Euler product* and makes explicit the connection between the ζ-function and prime numbers. Using (1), we can give Euler's proof of the infinitude of primes as follows: If the number of primes is finite, (1) defines an analytic function at $s = 1$. On the other hand, the series expression for $\zeta(s)$ shows that $\lim_{s \to 1} \zeta(s) = \infty$, which gives us a contradiction.

Exercise 1. Make the above "proof" rigorous.

The above argument shows that knowledge about the ζ-function just at $s = 1$ already gives some information about the distribution of prime numbers. Perhaps understanding the behaviour of the ζ-function at other values of s leads to even more information about primes. A famous theorem (and its even more famous corollary) of Hadamard and de la Vallée Poussin demonstrates this.

Theorem 1.1. *(Hadamard, de la Vallée Poussin, 1898)* $\zeta(1 + it) \neq 0$ *for all* t *in* $\mathbf{R}$.

Corollary 1.2. *(The Prime Number Theorem) Let*

$$\pi(x) = card\{p\ prime \mid p \leq x\}.$$

Then, as $x \to \infty$, we have the asymptotic formula

$$\pi(x) \sim \frac{x}{logx}. \tag{2}$$

Corollary (1.2) was first conjectured by Gauss. Finding a precise error term in the formula for $\pi(x)$ inspired Riemann to his famous Riemann Hypothesis (see below). For a proof of Theorem 1.1 we refer the reader to [L1].

(c) Analytic continuation and functional equation. Recall that for $Re(s) > 0$, we define

$$\Gamma\left(\frac{s}{2}\right) = \int_{-\infty}^{\infty} e^{-x^2} x^s d^{\times} x.$$

(Here $d^{\times} x$ denotes the measure $\frac{dx}{|x|}$. The usual definition of the Γ-function in most textbooks is given by the formula

$$\Gamma(s) = \int_0^{\infty} e^{-x} x^{s-1} dx,$$

from which the previous one can be easily deduced. It is well known that the Γ-function can be continued to a meromorphic function on the whole complex plane with (simple) poles at the negative integers and zero, and analytic elsewhere.)

Theorem 1.3. *The function*

$$\mathcal{Z}(s) = \pi^{-\frac{s}{2}} \Gamma(\frac{s}{2}) \zeta(s) \tag{3}$$

can be extended to a meromorphic function on all of the complex plane $\mathbf{C}$*. $\mathcal{Z}(s)$ is analytic at all points of* $\mathbf{C}$ *except at $s = 0$ and $s = 1$ where it has a simple pole. Further, we have the functional equation*

$$\mathcal{Z}(s) = \mathcal{Z}(1-s). \tag{4}$$

Now that we have seen that $\zeta(s)$ can be meromorphically continued to all of $\mathbf{C}$ we can state the

Riemann Hypothesis. *For $0 \leq Re(s) \leq 1$, $\zeta(s) = 0$ only if $Re(s) = \frac{1}{2}$.*

Exercise 2. Show that $\zeta(s) \neq 0$ for $Re(s) > 1$ (Hint: Show that $\log \zeta(s)$ is analytic for $Re(s) > 1$.).

Theorem 1.1 guarantees that $\zeta(s)$ is not zero on the line $Re(s) = 1$. Using the functional equation, one sees immediately that $\zeta(s) \neq 0$ for s such that $Re(s) \leq 0$ and $s \notin -2\mathbb{N} \bigcup \{0\}$.

Remark 1. $\zeta(s)$ does have zeros in the half-plane $Re(s) < 0$. From the fact that the Γ-function has simple poles at the negative integers but that (3) defines a function analytic at those points, we deduce that $\zeta(s)$ has zeros at all the even negative integers. These are called the *trivial* zeros.

Remark 2. The functional equation (4) shows that $Z(s)$ has a line of symmetry at $s = \frac{1}{2}$.

Remark 3. The function $\mathcal{Z}(s)$ is often called an L-function. The terminology L-function usually denotes the product of a Dirichlet series and suitable gamma functions so that the product satisfies a functional equation. Frequently, though, the two terms are used interchangeably.

(d) The value or residue at $s = 1$. $\zeta(s)$ has a simple pole at $s = 1$ with residue 1. We shall see in later examples of L-functions that the value or residue at $s = 1$ carries a lot of information especially about the class numbers of number fields.

(e) The value at other integers. The value of $\zeta(s)$ at other integers also carries a lot of number theoretic information and involves the *Bernoulli numbers.* Using properties of these numbers, Kummer was able to prove Fermat's Last Theorem in a very large number of cases. In lecture 2, we shall see that the values of $\zeta(s)$ at integral points are central to even defining p-adic L-functions - that is, functions of a p-adic variable with p-adic numbers as the range as well. Just to give an idea of what may be involved we give a heuristic calculation ignoring all issues of convergence. We first define the *Bernoulli numbers* B_n (see also [Ka], this volume) by the equation

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

Note that the Bernoulli numbers are rational. For an integer k we can write

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

Of course, the above expression is convergent only for $k < 0$, but we ignore this fact and proceed formally as follows:

$$\begin{aligned}\zeta(1-k) &= \sum_{n=1}^{\infty}\left(\frac{d}{dt}\right)^{k-1} e^{-nt}\Bigg|_{t=0}\\ &= \left(\frac{d}{dt}\right)^{k-1}\sum_{n=1}^{\infty} e^{nt}\Bigg|_{t=0}\\ &= \left(\frac{d}{dt}\right)^{k-1}\left(\frac{1}{1-e^{t}}-1\right)\Bigg|_{t=0}\\ &= \left(\frac{d}{dt}\right)^{k-1}\left(\frac{-1}{t}\sum_{n=1}^{\infty} B_n\frac{t^n}{n!}\right)\Bigg|_{t=0}\\ &= -\frac{B_k}{k}.\end{aligned}$$

In fact, it turns out there is rigorous way in which to establish the above equality which shows that the values of the ζ-function at negative integers are given by the Bernoulli numbers. It is easy to check that $B_{2k+1} = 0$, for any $k > 0$ (since $\zeta(1-k)$ vanishes for odd values of k). By the functional equation we find that the Bernoulli numbers give the values of the ζ-function at the positive even integers but yield no information at the positive odd integers. For example, it is known that $\zeta(3)$ is irrational but not even whether it is algebraic or transcendental.

2. Generalities on Dirichlet series. We state four lemmas which will be of great use in determining the domain of convergence of Dirichlet series and their analyticity. For proofs we refer the reader to Chapter 1 of Part II of [S]. Let $D(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$ be a Dirichlet series. Let $A_N = \sum_{n=1}^{N} a_n$, C_1, C_2 be positive constants and k any real number.

Lemma 2.1. *If $D(s)$ converges for $s = s_0$, then $D(s)$ converges for all s in* $\mathbf{C}$ *such that $Re(s) > Re(s_0)$.*

Lemma 2.1 says that if a Dirichlet series converges at one point in $\mathbf{C}$ it automatically converges in the half-plane to the right of the point. The union of all the half-planes where $D(s)$ converges is itself a half-plane and is called *the half-plane of convergence.*

Lemma 2.2. *If $|a_n| \le C_1 n^k$, then $D(s)$ converges for $Re(s) > k+1$. If $|A_N| \le C_2 N^k$ then $D(s)$ converges for $Re(s) > k$.*

Lemma 2.3. *If $f_N(s)$, $N = 1, 2, \ldots$ is a sequence of holomorphic functions converging uniformly on compact subsets of a domain D in* $\mathbf{C}$, *then $f(s) = \lim_{N\to\infty} f_N(s)$ is holomorphic on D.*

Applying the above lemma to $f_N(s) = \sum_{n=1}^{N} \frac{a_n}{n^s}$, we see that $D(s)$ is holomorphic in the domain of convergence of its series.

Lemma 2.4. *Suppose that $a_n \geq 0$ for all n and that s_0 is in* $\mathbf{R}$. *If $D(s)$ converges for $Re(s) > s_0$ and can be analytically continued to a function on a neighbourhood of s_0, then there is an $\epsilon > 0$ such that $D(s)$ converges for s such that $Re(s) > s_0 - \epsilon$.*

Exercise 3. Prove Lemma 2.4. Notice the connection with Exercise 1. (Hint: Express the derivatives of $D(s)$ as Dirichlet series. Then consider the Taylor series expansion for $D(s)$ about s_0 and deduce the convergence of the Dirichlet series from the convergence of the Taylor expansion.)

Our goal is to now introduce a number of Dirichlet series that arise in number theory and study these with respect to the properties (a) through (e) listed in Section 1.

3. More examples. We introduce some more Dirichlet series arising in number theory and discuss their domains of convergence and the existence of an Euler product.

Example 1. The Hurwitz zeta-function. Let $0 < b \leq 1$. We define

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

Note that if $b = 1$ we recover the Riemann ζ-function. Lemmas 1, 2 and 3 imply that $\zeta(s)$ converges to a holomorhphic function in the right half-plane $Re(s) > 1$. By comparison with $\zeta(s)$ we can see that $H(s,b)$ also converges in this half-plane, and applying Lemma 3 we may conclude that $H(s,b)$ is analytic in this domain.

If a and f are positive integers and $1 \leq a \leq f$, set $b = \frac{a}{f}$.

$$H(s,b) = \sum_{n=0}^{\infty} \frac{1}{(n+\frac{a}{f})^s} = \sum_{n=0}^{\infty} \frac{f^s}{(nf+a)^s} = f^s \sum_{m \equiv a \pmod{f}} \frac{1}{m^s}.$$

In the next section we shall obtain a meromorphic continuation of $H(s,b)$ (and hence of $\zeta(s)$!) to all of $\mathbf{C}$. In general, $H(s,b)$ does not possess an Euler product.

Example 2. The Dirichlet L-series. Let χ be a Dirichlet character of conductor f (by conductor, we mean the minimal integer needed to define χ. See p. 20 of [W], or [Ka] for a more precise explanation of our convention) We can extend χ to a function on $\mathbb{N}$ (also called χ) by setting $\chi(n) = 0$ if $(n,f) > 1$. The *parity* $\delta \in \{0,1\}$ of χ is determined by the equation $\chi(-1) = (-1)^\delta$. We say that χ is even if $\delta = 0$ and odd otherwise. We define

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

Note that if $f = 1$, we recover $\zeta(s)$. We will assume from now on that χ is not trivial.

Exercise 4. Show that $|\sum_{n=1}^{N} \chi(n)| < C$ for some $C > 0$ and any $N \in \mathbb{N}$.

Using Exercise 4 and Lemmas 2.1, 2.2 and 2.3, one sees that $L(\chi, s)$ converges to an analytic function in the half-plane $Re(s) > 0$. We also have an Euler product

$$L(\chi, s) = \prod_{p\ prime} \frac{1}{\left(1 - \frac{\chi(p)}{p^s}\right)},$$

which follows from the multiplicativity of χ. The above formula is valid for $Re(s) > 1$. Note that we can write

$$L(\chi, s) = f^{-s} \sum_{a=1}^{f} H(s, \frac{a}{f}). \tag{5}$$

A meromorphic continuation for $H(s, b)$ will thus yield one for $L(\chi, s)$. In fact, we will see that $L(\chi, s)$ is an entire function.

Example 3. The Dedekind zeta-function. Let K be a number field and $N_{\mathbf{Q}}^{K}$ be the norm from K to $\mathbf{Q}$. We define

$$\zeta_K(s) = \sum_{\mathfrak{a} \neq 0} \frac{1}{(N_{\mathbf{Q}}^{K} \mathfrak{a})^s},$$

where $\mathfrak{a}$ runs over all non-zero ideals of $\mathcal{O}_K$, the ring of integers of K. It is not so easy to show for arbitrary K that $\zeta_K(s)$ is convergent and holomorphic for $Re(s) > 1$, has an analytic continuation to all $s \neq 1$ and has a simple pole at $s = 1$. However, in the specific case when $K = \mathbf{Q}(\zeta_f)$, a cyclotomic field, we will be able to prove this by using the formula

$$\zeta_K(s) = \prod_{\chi} L(\chi, s), \tag{6}$$

where the product runs over all characters of $(\mathbb{Z}/f\mathbb{Z})^{\times} \simeq G(K/\mathbf{Q})$, and $G(K/\mathbf{Q})$ denotes the Galois group of K over $\mathbf{Q}$. Note that for arbitrary fields K we have the Euler product

$$\zeta_K(s) = \prod_{\mathfrak{P}} \frac{1}{\left(1 - \frac{1}{(N_{\mathbf{Q}}^{K} \mathfrak{P})^s}\right)}, \tag{7}$$

where the product runs over all prime ideals in $\mathcal{O}_K$. The expression (7) is valid for $Re(s) > 1$. The last equality follows from the fact that in Dedekind domains we have unique factorisation of ideals into prime ideals.

Exercise 5. Prove formula (6) above.

3. Analytic continuation and the functional equation.

We first obtain a meromorphic continuation for $H(s,b)$. By the remarks above this will automatically yield a meromorphic continuation for $\zeta(s)$, $L(\chi,s)$ and $\zeta_K(s)$, where K is a cyclotomic field. Our method will also be very useful in studying the values of these functions at integer points. We first set

$$F(z)=\frac{ze^{z(1-b)}}{e^z-1}. \tag{8}$$

Note that for $z\neq 2\pi im$, m in $\mathbb{Z}$, (8) defines a holomorphic function. We now integrate $F(z)z^{s-2}$ on the contour $\mathcal{C}$ described by $\mathcal{C}=(\infty,\epsilon]\cup S_\epsilon\cup[\epsilon,\infty)$, where S_ϵ denotes the circle of radius ϵ centred at the origin. We follow the convention that $z^s=e^{s\log z}$ is real and positive on the positive real axis. We treat $(\infty,\epsilon]$ as the "top" of the real axis where $z^s=e^{s\log|z|}$ and $[\epsilon,\infty)$ as the "bottom" of the real axis where $z^s=e^{s(\log|z|+2\pi i)}$. Hence, we have

$$\begin{aligned}G(s)&=\int_{\mathcal{C}}F(z)z^{s-2}dz\\&=\int_{(\infty,\epsilon]}F(z)z^{s-2}dz+\int_{S_\epsilon}F(z)z^{s-2}dz+\int_{[\epsilon,\infty)}F(z)z^{s-2}dz\\&=(e^{2\pi is}-1)\int_{(\epsilon,\infty]}F(t)t^{s-2}dt+\int_{S_\epsilon}F(z)z^{s-2}dz.\end{aligned} \tag{9}$$

We remark that when $z=t$, $t\in\mathbf{R}$, $F(t)$ clearly decays exponentially as $t\to\infty$ (recall that $b\neq 0$!). Hence, the both terms in (9) define analytic functions for all s (exercise: differentiate under the integral sign and check this). On the other hand, when $Re(s)>1$, if $\epsilon\to 0$, one sees easily that the second term of (9) goes to zero. Thus, we see that

$$\int_{\mathcal{C}}F(z)z^{s-2}dz=(e^{2\pi is}-1)\int_0^\infty F(t)t^{s-2}dt, \tag{10}$$

for $Re(s)>1$. We now expand $F(t)$ as a geometric series to obtain

$$F(t)=t\sum_{n=0}^{\infty}e^{-(b+n)t}.$$

Substituting this in (10) we get

$$(e^{2\pi is}-1)\int_0^\infty\sum_{n=0}^{\infty}e^{-(b+n)t}t^{s-1}dt.$$

We may switch the sum and the integral above to get

$$(e^{2\pi is}-1)\sum_{n=0}^{\infty}\int_0^{\infty}e^{-(b+n)t}t^{s-1}dt,$$

which after a change in variables yields

$$(e^{2\pi is}-1)\sum_{n=0}^{\infty}\frac{1}{(n+b)^s}\int_0^{\infty}e^{-t}t^{s-1}dt=(e^{2\pi is}-1)\Gamma(s)H(s,b). \qquad (11)$$

We caution once again that (11) is valid only for $Re(s)>1$. We can, however, use (11) to get a meromorphic continuation for $H(s,b)$ for all s, since (9) defines an analytic function $G(s)$ for all values of s. Indeed, we can now set

$$H(s,b)=((e^{2\pi is}-1)\Gamma(s))^{-1}G(s),$$

where $G(s)$ is the analytic function defined by the integral in (9). The expression above gives the Hurwitz ζ-function in terms of the Gamma function and a *Mellin transform* $G(s)$ (see Exercise 13 below). We already know that $H(s,b)$ is analytic for $Re(s)>1$. On the other hand, the function $(e^{2\pi is}-1)\Gamma(s)$ is *non-vanishing* for $Re(s)<1$ (note that $\Gamma(s)$ has a simple pole at zero and all the negative integers, and $(e^{2\pi is}-1)$ has a simple zero at those points). Hence, $H(s,b)$ is actually analytic for all $s\neq 1$. At $s=1$ it has a simple pole. We have thus proved

Theorem 3.1. *$H(s,b)$ can be continued to a meromorphic function on all of $\mathbf{C}$ which is analytic if $s\neq 1$ and has a simple pole at $s=1$.*

Corollary 3.2. *$\zeta(s)$ can be continued to a function on all of $\mathbf{C}$ which is analytic if $s\neq 1$ and has a simple pole at $s=1$. $\mathcal{Z}(s)$ has simple poles at $s=0$ and $s=1$, and is analytic for all other values of s.*

Proof. Choose $b=1$.

Corollary 3.3. *If $\chi\neq 1$, $L(\chi,s)$ is entire.*

Proof. This is immediate if we use the expression (5).

Corollary 3.4. *If $K=\mathbf{Q}(\zeta_f)$, then $\zeta_K(s)$ can be continued to an analytic function for $s\neq 1$. If $s=1$, $\zeta_K(s)$ has a simple pole.*

Proof. Again, this is immediate from Corollaries 3.3 and 3.4 and Exercise 5.

Let K be a number field and set $N=\deg(K/\mathbf{Q})$. Let r_1 be the number of real embeddings and r_2 be half the number of complex embeddings of K. We let d_K be the discriminant of K. We also introduce the functions

$$\Gamma_{\mathbf{R}}(s)=\pi^{-\frac{s}{2}}\Gamma(\frac{s}{2}) \quad\text{and}\quad \Gamma_{\mathbf{C}}(s)=(2\pi)^{-s}\Gamma(s).$$

Theorem 3.5. *Let*

$$\mathcal{Z}_K(s) = |d_K|^{\frac{s}{2}}(\Gamma_{\mathbf{R}}(s))^{r_1}(\Gamma_{\mathbf{C}}(s))^{r_2}\zeta_K(s).$$

Then $\mathcal{Z}_K(s)$ can be continued to a meromorphic function on $\mathbf{C}$, with simple poles at $s = 0$ and $s = 1$, and analytic elsewhere. Further, $\mathcal{Z}_K(s)$ satisfies the functional equation

$$\mathcal{Z}_K(s) = \mathcal{Z}_K(1 - s). \tag{12}$$

We will not prove this in general. In the specific case when $K = \mathbf{Q}(\zeta_f)$, we will be able to deduce this from the functional equations for Dirichlet L-functions stated below.

Let χ be a Dirichlet character of conductor f and parity δ (as in Example 2 of this section). Let $\tau(\chi) = \sum_{a=1}^{f} \chi(a)e^{\frac{2\pi i a}{f}}$ denote as usual the *Gauss sum* associated to χ. Let $W_\chi = \frac{\tau(\chi)}{i^\delta\sqrt{f}}$.

Exercise 6. Show that $|W_\chi| = 1$.

Exercise 7. (Hard!) Prove that

$$\prod_{\chi} W_\chi = 1, \tag{13}$$

where the product runs over all Dirichlet characters of conductors dividing f.

Theorem 3.6. *Let χ be a primitive Dirichlet character of conductor f.*

$$\Lambda(\chi, s) = \Gamma_{\mathbf{R}}(s + \delta)L(\chi, s)$$

satisfies the functional equation

$$\Lambda(\chi, s) = W_\chi \Lambda(\bar{\chi}, 1 - s). \tag{14}$$

Exercise 8. Use Exercises 5,6 and 7 together with Theorem 3.6 to prove Theorem 3.5 when $K = \mathbf{Q}(\zeta_f)$ (Caution: Theorem 3.6 is valid only for *primitive* Dirichlet characters.)

It remains to prove Theorem 3.6. This we do through the series of exercises given below. A good reference for this material is [L1], where a much more general result is proved.

Exercise 9. Let $\phi(x)$ be a *Schwartz function* on $\mathbf{R}$, that is, a smooth (i.e., C^∞), real valued function such that for any n in $\mathbb{N}$ and any polynomial $P(x)$, there is a constant $C_P > 0$ such that

$$|\phi^{(n)}(x)| < C_P \frac{1}{|P(x)|},$$

($\phi^{(n)}$ denotes the n^{th} derivative of ϕ) provided $|x|$ is sufficiently large. For this class of functions, called the *Schwartz space* and denoted $\mathcal{S}(\mathbf{R})$, prove the *Poisson summation formula*

$$\sum_{n=-\infty}^{\infty} \phi(n) = \sum_{n=-\infty}^{\infty} \hat{\phi}(n), \tag{15}$$

where $\hat{\phi}$ denotes the Fourier transform of ϕ and is given by

$$\hat{\phi}(y) = \int_{-\infty}^{\infty} \phi(x) e^{-2\pi i x y} dx.$$

(Hint: Notice that $\sum_{n=-\infty}^{\infty} \phi(x+n)$ converges uniformly to a periodic function $F(x)$. Compute its Fourier series and set $x = 0$.)

Exercise 10. Recall the formula for character inversion for *primitive* Dirichlet characters of conductor f.

$$\chi(n) = \frac{\chi(-1)\tau(\chi)}{f} \sum_{a \pmod f} \bar{\chi}(a) e^{\frac{2\pi n a}{f}}, \tag{16}$$

where $\bar{\chi}$ denotes the complex conjugate of χ. Notice that this gives us a method to extrapolate χ to a function on all of $\mathbf{R}$ (replace n by x in the above formula). Apply the Poisson summation formula (see Exercise 9) to the function $\phi(x) = \chi(x)\psi(x)$ to prove the *twisted* Poisson summation formula

$$\sum_{n=-\infty}^{\infty} \chi(n)\psi(n) = \frac{\chi(-1)\tau(\chi)}{f} \sum_{n=-\infty}^{\infty} \bar{\chi}(n) \hat{\psi}\left(\frac{n}{f}\right). \tag{17}$$

Exercise 11. We define the *theta function* associated to χ and a Schwartz function ϕ by

$$\theta_{\chi,\phi}(t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} n^{\delta} \chi(n) \phi(n^2 t) = \frac{1}{2}\chi(0) + \frac{1}{2} \sum_{n=1}^{\infty} n^{\delta} \chi(n) \phi(n^2 t). \tag{18}$$

Note that this is nothing but the left-hand side of (17) with $\psi(x) = g_t(x) = x^{\delta}\phi(x^2 t)$. Hence, by (17) we see that

$$\theta_{\chi,\phi}(t) = \frac{i^{\delta}}{(N t^{\frac{1}{2}})^{1+\delta}} \theta_{\bar{\chi},\hat{\phi}}\left(\frac{1}{f^2 t}\right). \tag{19}$$

When $f = 1$ and $\phi(x) = g_t(x) = e^{-\pi x^2 t}$, $\theta_{\chi,\phi}(t)$ is the familiar *Heat kernel.*

Exercise 12. When χ is not trivial, use (19) to show that $\theta_{\chi,\phi}(t)$ decays rapidly as $t \to \infty$ (i.e, faster than any negative power of t) and also as $t \to 0$ (i.e., faster than any positive power of t). If χ is trivial, show the same for $\theta_{\chi,\phi}(t) - \frac{1}{2}$.
Exercise 13. The *Mellin transform* of a function $h(t)$ is given by the formula

$$M(h, s) = \int_0^\infty h(t) t^{\frac{s}{2}} \frac{dt}{t}. \tag{20}$$

(Actually, this is the Mellin transform at $s/2$ but for the sake of less complicated formulae we will stick to the notation above.) We see that $\pi^{-s/2}\Gamma(s/2)$ is $M(e^{-\pi t^2}, s)$ in the above notation. Show that if χ is not trivial and $Re(s) > 1$

$$M(\theta_{\chi,\phi}, s) = \Lambda(\chi, s), \tag{21}$$

while if χ is trivial

$$M(\theta_{\chi,\phi} - 1/2, s) = \Lambda(1, s) = \mathcal{Z}(s). \tag{22}$$

On the other hand, check that by Exercise 12, the left-hand sides of (21) and (22) define analytic functions for all values of s. This gives another proof of the analytic continuation of $L(\chi, s)$.
Exercise 14. Use (19) in equations (21) and (22) to prove that

$$M(\phi, s) L(\chi, s) = M(\hat{\phi}, 1 - s) L(\bar{\chi}, 1 - s). \tag{23}$$

Now, if $\phi(x) = e^{-\pi x^2}$ we recover the functional equation (14).

Exercise 14 shows that one obtains a functional equation for any choice of Schwartz function $\phi(x)$ and that (14) is merely a special case of (23) for a specific choice of $\phi(x)$! Using two different choices of $\phi(x)$, say $\phi_1(x)$ and $\phi_2(x)$ in (23) (assume $\delta = 0$) and taking a quotient gives us Tate's *local functional equation at infinity*

$$\frac{M(\phi_1, s)}{M(\phi_2, s)} = \frac{M(\hat{\phi}_1, 1 - s)}{M(\hat{\phi}_2, 1 - s)}. \tag{24}$$

This can be very easily proved directly by Fubini's theorem.

4. The value or residue at $s = 1$. It is easy to check

that $\zeta(s)$ has a simple pole at $s = 1$ with residue 1. More generally, we have

Theorem 4.1. *$\zeta_K(s)$ has a simple pole at $s = 1$ with residue given by*

$$\frac{2^{r_1} (2\pi)^{r_2} h_K R}{w |d_K|^{\frac{1}{2}}},$$

where r_1, r_2 and d_K are as before, w is the number of roots of unity in K, h_K is the class number of K and R is the regulator of K.

For a definition of the regulator of a number field see p. 41 of [W]. We will not be able to prove this in the general case. In the cyclotomic case $K = \mathbf{Q}(\zeta_f)$ we note that

$$\zeta_K(s) = \prod_p \prod_{\mathfrak{P}|p} \frac{1}{\left(1 - \frac{1}{(N_{\mathbf{Q}}^K \mathfrak{P})^s}\right)} = \prod_p \frac{1}{\left(1 - \frac{1}{p^{\alpha s}}\right)^{\gamma}}, \tag{25}$$

where $N_{\mathbf{Q}}^K \mathfrak{P} = p^{\alpha}$ and γ is the number of primes $\mathfrak{P}$ lying above p.
Exercise 15. For $K = \mathbf{Q}(\zeta_f)$ show that $\zeta_K(s)$ has a simple pole at $s = 1$ (Hint: Use Exercise 3 together with (25)).

Exercise 15 immediately yields Theorem 4.2 below as a relatively simple corollary. This special case of Theorem 4.1 and its celebrated corollary were proved by Lejeune Dirichlet. They mark the entry of complex analytic techniques into number theory.

Theorem 4.2. *If χ is a non-trivial primitive character, $L(\chi, 1) \neq 0$.*

Proof. Because of Exercise 15, we see that both sides of (5) have a simple pole at $s = 1$. It follows that all the other factors on the right-hand side of (5) must be non-vanishing at $s = 1$.

In fact, using Theorem 4.1, we have the formula

$$\frac{2^{r_1}(2\pi)^{r_2} h_K R}{w|d_K|^{\frac{1}{2}}} = \prod_{\chi} L(\chi, 1). \tag{26}$$

Since the residue of $\zeta(s)$ at 1 is 1, (26) is immediate.

Corollary 4.3. *Let $(a, b) = 1$. There are infinitely many primes of the form $an + b$.*

Exercise 16. Prove Corollary 4.3 as follows:
Step 1. Show that for $Re(s) > 1$ we have

$$\sum_{\chi \,(\mathrm{mod}\ n)} \chi(a^{-1}) \log L(\chi, s) = \sum_{p \equiv a \,(\mathrm{mod}\ n)} \frac{\phi(n)}{p^s} + g(s), \tag{27}$$

where $\phi(n)$ denotes the Euler ϕ-function and $g(s)$ is an analytic function for $Re(s) > \frac{1}{2}$.
Step 2. Use Theorem 4.2 and take limits as $s \to 1$ to get the desired result.

In fact, we have much stronger results than Theorem 4.2. One can prove

Theorem 4.3. *Let $D(s)$ denote either a Dirichlet L-series $L(\chi, s)$ or $\zeta_K(s)$ for some number field K. Then*

$$D(1+it) \neq 0,$$

for all t in $\mathbf{R}$.

As was discussed earlier, the case $K = \mathbf{Q}$ already implies the Prime Number Theorem. When one considers other number fields one obtains finer information about the distribution of primes.

It is not hard to determine the values of $L(\chi, 1)$ more explicitly. Specifically, we can prove

Theorem 4.4.

$$L(\chi,1) = \begin{cases} -\frac{\tau(\chi)}{f} \sum_{a=1}^{f} \bar{\chi}(a) log|1 - e^{\frac{2\pi a}{f}}| & \textit{if } \chi(-1) = 1 \qquad (28a) \\ \pi i \frac{\tau(\chi)}{f} \frac{1}{f} \sum_{a=1}^{f} \bar{\chi}(a) a & \textit{if } \chi(-1) = 1 \qquad (28b) \end{cases}$$

The proof is left as an exercise. Alternatively, we refer to pp. 37-39 of [W]. We will return to this subject in the next section after introducing the Bernoulli numbers $B_{n,\chi}$.

5. The values of L-functions at integer points. Let χ be a Dirichlet character of conductor f. We define the *Bernoulli numbers* B_n and the *twisted Bernoulli numbers* $B_{n,\chi}$ by the formulae

$$\frac{t}{e^t - 1} = \sum_{n=0}^{\infty} B_n \frac{t^n}{n!} \quad \text{and} \quad \sum_{a=1}^{f} \frac{\chi(a) t e^{at}}{e^t - 1} = \sum_{n=0}^{\infty} B_{n,\chi} \frac{t^n}{n!}. \tag{29}$$

The *Bernoulli polynomials* $B_n(X)$ are given by

$$\frac{te^{tX}}{e^t - 1} = \sum_{n=0}^{\infty} B_n(X) \frac{t^n}{n!}.$$

Clearly $B_{n,1} = B_n$ for all $n > 1$ while $B_{1,1} = \frac{1}{2}$ and $B_1 = -\frac{1}{2}$. We also note that $B_n(0) = B_n$ for all n. Let F denote any multiple of f. We record the following the identities which are easy consequences of the definitions:

$$B_n(1-X) = (-1)^n B_n(X). \tag{30}$$

$$B_n(X) = \sum_{i=0}^{n} \binom{n}{i} B_i X^{n-i}. \tag{31}$$

$$B_{n,\chi} = F^{n-1} \sum_{a=1}^{F} \chi(a) B_n\left(\frac{a}{F}\right). \tag{32}$$

In particular, we see that $B_{1,\chi} = \frac{1}{f}\sum_{a=1}^{f}\chi(a)a$. Using (28b), we see that for odd characters χ the value of $L(\chi, 1)$ is nothing but $\frac{\pi i \tau(\chi)}{f} B_{1,\bar{\chi}}$.

Theorem 4.2 shows that $B_{1,\chi}$ is *non-zero* and at least for odd characters there is no other elementary proof of this fact. We have thus expressed the value of $L(\chi, s)$ at $s = 1$ in terms of a Bernoulli number. We proceed to generalise this.

Theorem 5.1. *For $n \geq 1$ and $0 < b \leq 1$ we have*

$$H(1-n, b) = -\frac{B_n(b)}{n}. \tag{33}$$

Corollary 5.2.

$$L(\chi, 1-n) = -\frac{B_{n,\chi}}{n}. \tag{34}$$

The corollary follows immediately from the Theorem and equation (5).

Proof of Theorem 5.1. We recall that equation (9) yields

$$G(s) = (e^{2\pi i s} - 1)\int_{(\epsilon,\infty]} F(t) t^{s-2} dt + \int_{S_\epsilon} F(z) z^{s-2} dz.$$

Substituting $s = 1 - n$ we see that the first term in the expression above vanishes so we get

$$G(1-n) = \int_{S_\epsilon} F(z) z^{-n-1} dz = (2\pi i)\frac{B_n(1-b)}{n!}. \tag{35}$$

Notice that

$$\lim_{s \to 1-n} (e^{2\pi i s} - 1)\Gamma(s) = \frac{(2\pi i)(-1)^{n-1}}{(n-1)!},$$

whence follows

$$H(1-n, b) = (-1)^{n-1}\frac{B_n(1-b)}{n} = -\frac{B_n(b)}{n}.$$

The fact that $\Lambda(\chi, s)$ is entire for $s = 1 - n$ for all non-trivial χ and for all $n \geq 1$ tells us that

$$L(\chi, 1-n)\begin{cases} \neq 0 & \text{if} \quad n \equiv \delta \pmod 2 & (36a) \\ = 0 & \text{if} \quad n \not\equiv \delta \pmod 2. & (36b) \end{cases}$$

Hence, we find by (34) that for $n \geq 1$

$$B_{n,\chi} = 0 \quad \text{if} \quad n \not\equiv \delta \pmod 2. \tag{37}$$

It remains only to dispose off the case $\chi = 1$ and $n = 1$ and in this case it is trivially verified that $\zeta(0) = -\frac{1}{2}$.

Part II: p-adic L-functions

6. The p-adic ζ-function. In Part I we studied L-functions whose domains and ranges are subsets of the complex numbers. In this part we will study analogues with domains and ranges which are subsets of the p-adic numbers. We motivate the study of such L-functions using the simplest case - namely that of the Riemann ζ-function. In this section we will assume that the prime p is odd. This is just to simplify the arguments we give below.

Recall that in Section 5 we defined the Bernoulli numbers B_n and showed that they were related to the values of the Riemann ζ-function at the negative integers. We showed that for $n \geq 1$,

$$\zeta(1-n) = -\frac{B_{n,1}}{n}.$$

For $n > 1$, we have $-\frac{B_{n,1}}{n} = -\frac{B_n}{n}$. Note that $-\mathbb{N}$ is dense in $\mathbf{Z}_p$. Hence, a continuous function on $\mathbf{Z}_p$ would be completely determined by its values on $-\mathbb{N}$. We now state two remarkable properties of Bernoulli numbers that we will prove later.

Theorem 6.1. *(von Staudt-Claussen) For n even and positive,*

$$B_n + \sum_{(p-1)|n} \frac{1}{p} \in \mathbb{Z}, \tag{38}$$

where the sum runs over primes p such that $(p-1)$ divides n.

We already know that B_n is rational but Theorem 6.1 asserts that even more is true. In fact, pB_n is in $\mathbf{Z}_p$ for all n and all p. We next state

Theorem 6.2. *(Kummer's Congruences) Suppose $m \equiv n \pmod{(p-1)p^a}$ and $n \not\equiv 0 \pmod{p-1}$, then*

$$(1-p^{m-1})\frac{B_m}{m} \equiv (1-p^{n-1})\frac{B_n}{n} \pmod{p^{a+1}}. \tag{39}$$

By restricting our attention to $-\mathbb{N}$, we may view $\zeta(s)$ as a $\mathbf{Q}$ -valued function, and since $\mathbf{Q}$ is embedded in $\mathbf{Q}_p$ we may view it as a $\mathbf{Q}_p$ -valued function. Theorem 6.2 says, roughly, that if m and n are close to each other in the p-adic metric, then $(1-p^{m-1})\zeta(1-m)$ and $(1-p^{n-1})\zeta(1-n)$ are also close to each other in $\mathbf{Q}_p$! This statement is interesting only when m and n are even integers; otherwise $\zeta(1-m)=0$ and $\zeta(1-n)=0$, so we get nothing particularly new. The function $(1-p^{m-1})\zeta(m)$ is thus continuous on arithmetic progressions of the form $a+(p-1)\mathbb{Z}$ $(a \neq 0)$ endowed with the p-adic topology and we can extend it to all of $\mathbf{Z}_p$, since the union of these sets is *dense* in $\mathbf{Z}_p$. Hence, we have obtained a continuous $\mathbf{Q}_p$-valued function on $\mathbf{Z}_p$ by *p-adic interpolation* which we shall denote $\zeta_p(s)$, where s ranges over $\mathbf{Z}_p$. $\zeta_p(s)$ is called the p-adic ζ-function.

We will study the function $\zeta_p(s)$ and try and produce interesting number-theoretic results from its properties, just as we did in the complex-valued case in Part I. *A priori*, $\zeta_p(s)$ is merely a continuous function from $\mathbf{Z}_p$ to $\mathbf{Q}_p$. However, we shall show that $\zeta_p(s)$ is actually p-adic analytic for all $s \neq 1$ in $\mathbf{Z}_p$, i.e., it can be expressed as a power series in $(s-s_0)$ in a neighbourhood of any point $s_0 \neq 1$. (At $s_0 = 1$ it will have a simple pole.) Indeed, we shall not define $\zeta_p(s)$ as above at all. Instead, we will prescrible it using an analogue of the Mellin transform in the p-adic setting and prove that the resultant function is analytic and that its values at the integers are precisely those given by equation (33). We will also be able to prove the Kummer Congruences as byproducts of our definitions.

Naturally, there are p-adic analogues of the other L-functions that we considered in Part 1. In each case we could define them by interpolation as above, or by means of the Mellin transform, and once again they will converge in an open set in $\mathbf{Q}_p$. These L-functions do not possess Euler products or satisfy functional equations. Their values at integer points are naturally of great interest. In particular, their values at 1 give p-adic analogues of the class number formula (formula (25)) and Dirichlet's theorem on the rank of units in a number field (Leopoldt's conjecture - see Chapter 5 of [W]).

We also remark in passing that it is better to study p-adic L-functions as functions from an open set in $\mathbf{C}_p$ to $\mathbf{C}_p$, where $\mathbf{C}_p$ denotes the field of "p-adic complex numbers" (see Section 12). The field $\mathbf{C}_p$ contains $\mathbf{Q}_p$, and has the advantage of being both complete and algebraically closed, making it a natural analogue of the usual complex numbers. The main reference for our exposition below is [Ko]. A less elementary approach with much more material can be found in [L2].

7. Preliminaries about $\mathbf{Q}_p$. We first recall a few basic facts about $\mathbf{Q}_p$. It is defined as the completion of $\mathbf{Q}$ under the p-adic valuation $v = v_p$ and its norm is denoted by $|\ |_p$. When p is fixed and no confusion will arise we will denote the corresponding norm simply by $|\ |$. There is a maximal compact (open) subring

$$\mathbf{Z}_p = \{x \in \mathbf{Q}_p \mid |x|_p \leq 1\} \tag{40}$$

which is called the ring of integers in $\mathbf{Q}_p$. The unique maximal ideal of $\mathbf{Z}_p$

will be denoted by $p\mathbf{Z}_p$, or simply by (p). We have a canonical homomorphism of *rings*

$$\begin{aligned}\pi : \mathbf{Z}_p &\mapsto (\mathbf{Z}_p/p\mathbf{Z}_p) \simeq \mathbb{F}_p\\ x &\mapsto x \pmod{p}.\end{aligned}$$

The ideals $p^n\mathbf{Z}_p = (p^n)$, n in $\mathbb{Z}$, are compact open sets in $\mathbf{Q}_p$, shrinking to 0 as $n \to \infty$, and form what is called a fundamental system of neighbourhoods of the identity in $\mathbf{Q}_p$. Any open neighbourhood of 0 contains some member of the family p^n.

We give an alternative definition of $\mathbf{Z}_p$ as a *projective limit* of the finite groups $\{(\mathbb{Z}/p^n\mathbb{Z})\}_{n=1}^{\infty}$. Notice that if $m < n$, we have the natural projection maps

$$\begin{aligned}\phi_{n,m} :\ (\mathbb{Z}/p^n\mathbb{Z}) &\mapsto (\mathbb{Z}/p^m\mathbb{Z})\\ a + p^n\mathbb{Z} &\mapsto a + p^m\mathbb{Z}.\end{aligned}$$

We define $\mathbf{Z}_p$ by

$$\mathbf{Z}_p = \left\{(x_1, x_2, \dots) \in \prod_{n=1}^{\infty} (\mathbb{Z}/p^n\mathbb{Z}) \ \mid \phi_{ji}(x_j) = x_i \text{ if } i < j\right\}. \tag{41}$$

From (41) it is clear that $\mathbf{Z}_p$ is compact since it is defined as a closed subset of a product of compact sets (Tychonoff's theorem - remember that finite sets are compact!). It follows easily that $p^n\mathbf{Z}_p$ is also compact for all values of n.

Exercise 17. Show that the two definitions of $\mathbf{Z}_p$ given above are equivalent.

Exercise 18. Check that we have natural projection maps $\phi_n : \mathbf{Z}_p \mapsto (\mathbb{Z}/p^n\mathbb{Z})$ for $n \geq 1$, with $p^n\mathbf{Z}_p$ as the kernel and that these maps commute with the $\phi_{n,m}$ defined above.

The group of units in $\mathbf{Z}_p$ will be denoted by $\mathbf{Z}_p^{\times}$, U_v or $U_v^{(0)}$. It can be characterised as those elements x in $\mathbf{Z}_p$ such that $|x| = 1$. Notice that every element in $\mathbf{Q}_p^{\times}$ can be written uniquely in the form $p^n x$ for some integer n and some unit x. We thus have the decomposition

$$\mathbf{Q}_p^{\times} = \{p^{\mathbb{Z}}\} \times U_v^{(0)}. \tag{42}$$

We will now study the multiplicative structure of $U_v^{(0)}$ more closely. For $m \geq 1$, $U_v^{(0)}$ contains the compact open subgroups $U_v^{(m)}$ defined by

$$U_v^{(m)} = \{u \in U_v^{(0)} \,|\, u \equiv 1 \pmod{(p^m)}\}.$$

Notice that if $m_1 > m_2$, then $U_v^{(m_1)} \subset U_v^{(m_2)}$. For u in $U_v^{(m)}$ we can write $u = 1 + x$, where x is in (p^m). Then $x \mapsto 1 + x$ defines a map ψ from $(p^m) \mapsto U_v^{(m)}$.

Exercise 19. Check that ψ induces an isomorphism

$$(\mathbb{Z}/p\mathbb{Z}) \simeq (p^m)/(p^{m+1}) \simeq U_v^{(m)}/U_v^{(m+1)}. \tag{43}$$

You will need to prove that $(1 + x + y + xy)(1 + x + y)^{-1}$ is in $U_v^{(m+1)}$ if x and y are in (p^m), which can be done by expanding the second factor as a geometric series.

We deduce inductively from (43) that $U_v^{(m)}/U_v^{(m+i)} \simeq \mathbb{Z}/p^i\mathbb{Z}$ for all i in $\mathbb{N}$, when $m > 0$. In particular, this shows that the index of $U_v^{(m+i)}$ in $U_v^{(m)}$ is p^i if $m > 0$. What about the case $m = 0$? As we shall see this requires a slightly more subtle approach. In order to analyse this case we first need the following lemma.

Lemma 7.1. *(Hensel's Lemma) Let $f(X)$ be a monic polynomial with coefficients in $\mathbf{Z}_p$. Suppose there exist $a \in \mathbf{Z}_p$ such that $f(a) \equiv 0 \pmod p$ but $f'(a) \not\equiv 0 \pmod p$, then there exist $\alpha \equiv a \pmod p$ in $\mathbf{Z}_p$ such that $f(\alpha) = 0$.*

The proof of the lemma involves essentially Newton-Raphson iteration. For a slightly more general formulation and proof we refer to p. 42 of [L1]. Hensel's lemma says that if we can find a *simple* root modulo p we can lift it to get a root in $\mathbf{Z}_p$. Moreover, such a lift is unique.

We apply the lemma to the polynomial $f(X) = X^{p-1} - 1$, where p is an odd prime. Clearly $f(x)$ has $p - 1$ distinct roots modulo p (for instance, $1, 2, \dots, p - 1$ by Fermat's little theorem!). Hence, each such root a_i, $i = 1, \dots, p - 1$ modulo p can be lifted to a root α_i of $f(x)$ in $\mathbf{Z}_p$, and all these α_i are distinct since they are pairwise inequivalent modulo p. One checks easily that the set μ_{p-1} of these *roots of unity* form a group isomorphic to $\mathbb{F}_p^\times \simeq (\mathbb{Z}/p\mathbb{Z})^\times$, where $\mathbb{F}_p^\times$ denotes the multiplicative group of the finite field $\mathbb{F}_p \simeq (\mathbb{Z}/p\mathbb{Z})$. If we now restrict the map π to $U_v^{(0)}$ we obtain a homomorphism

$$\begin{aligned} \pi : \quad & U_v^{(0)} \mapsto \mathbb{F}_p^\times \\ & u \mapsto u \pmod p \end{aligned}$$

with kernel $U_v^{(1)}$. In fact, we have the split exact sequence

$$0 \longrightarrow U_v^{(1)} \longrightarrow U_v^{(0)} \xrightarrow{\pi} \mathbb{F}_p^\times \longrightarrow 0, \tag{44}$$

where the splitting is given by lifting the solutions of $X^{p-1}-1$ to $U_v^{(0)}$ using Hensel's lemma. As a result, we can now decompose $U_v^{(0)}$ as $\mu_{p-1} \times U_v^{(1)}$ and hence, by (42), we obtain the decomposition

$$\mathbf{Q}_p = \{p^{\mathbb{Z}}\} \times \mu_{p-1} \times U_v^{(1)}. \tag{45}$$

Thus, for every a in $U_v^{(0)}$ we can write

$$a = \omega(a)\langle a\rangle, \tag{46}$$

where $\omega(a) \equiv a \pmod{p}$ is in μ_{p-1} and $\langle a\rangle$ is in $U_v^{(1)}$. This decomposition is obviously unique. ω thus defines a character from $\mathbf{Z}_p$ to μ_{p-1} called the *Teichmüller character*. We also remark that from (44) it is clear that

$$U_v^{(0)}/U_v^{(1)} \simeq (\mathbb{Z}/p\mathbb{Z})^{\times},$$

so the index of $U_v^{(1)}$ in $U_v^{(0)}$ is exactly $p-1$ for odd primes p. From this it follows that the index of $U_v^{(m)}$ in $U_v^{(0)}$ is $(p-1)p^{m-1}$ when p is odd (notice that these numbers appeared in the Kummer congruences of Theorem 6.2!).

When $p=2$, we need to modify the arguments very slightly. We have a natural projection map $\pi_1 : \mathbb{Z}_2 \to \mathbb{Z}_2/4\mathbb{Z}_2$. $U_v^{(0)} = U_v^{(1)}$ is precisely the inverse image $\pi_1^{-1}\left((\mathbb{Z}_2/4\mathbb{Z}_2)^{\times}\right)$ while $U_v^{(2)}$ is the kernel of π_1. The equation x^2-1 has solutions ± 1 in $\mathbb{Z}_2$ and we call this group μ_2. As before we thus have the direct sum decomposition

$$\mathbf{Q}_2 = \left\{2^{\mathbb{Z}}\right\} \times \mu_2 \times U_v^{(2)}.$$

The Teichmüller character ω is given in this case by the unique decomposition

$$a = \omega(a)\langle a\rangle,$$

where $\omega(a) \equiv a (\mathrm{mod} 4)$ and $\langle a\rangle$ is in $U_v^{(2)}$.

We can do everything that we have discussed so far in this section for a finite extension of $\mathbf{Q}_p$. As we know, any finite extension of $\mathbf{Q}_p$ arises as the completion K_w of a finite extension K of $\mathbf{Q}$ with respect to a valuation w lying above v. That is, the valuation w is given by a prime ideal $\mathfrak{P}$ in the ring of integers $\mathcal{O}_K$ of K lying above the prime (p) (i.e., $\mathfrak{P} \cap \mathbf{Q} = (p)$). We will assume that $\mathfrak{P}$ does not lie above the prime ideal (2) in $\mathbf{Q}$. If $\mathfrak{P}$ does lie above (2), the arguments below have to be modified slightly, as before. We let $|\ |_w$ denote the norm arising from the valuation w. The ring of integers $\mathcal{O}_w$ in K_w consists of x in K_w such that $|x|_w \leq 1$. The units $\mathcal{U}_w$ (or $\mathcal{U}_w^{(0)}$)

are defined by the condition $|x|_v = 1$. For $m \geq 1$, we can also define the subgroups $\mathcal{U}_w^{(m)}$ by the condition

$$\mathcal{U}_w^{(m)} = \{u \in \mathcal{U}_w^{(0)} \mid u \equiv 1(\mathrm{mod} \mathfrak{P}^m)\}$$

as before. Hensel's lemma is valid with $\mathcal{O}_w$ in place of $\mathbf{Z}_p$. We also know that $\mathcal{O}_w/\mathfrak{P} \simeq \mathbb{F}_q$, where $q = p^f$ for some f, and $\mathbb{F}_q$ denotes the finite field with q elements. If μ_{q-1} is the group of roots of unity in $\mathcal{O}_w$ and $\mathbb{F}_q^\times$ is the multiplicative group of $\mathbb{F}_q$, we obtain the split exact sequence

$$0 \longrightarrow \mathcal{U}_w^{(1)} \longrightarrow \mathcal{U}_w^{(0)} \xrightarrow{\pi} \mathbb{F}_q^\times \longrightarrow 0, \tag{47}$$

and from this follows the decomposition

$$K_w = \{(\pi)^{\mathbb{Z}}\} \times \mu_{q-1} \times \mathcal{U}_w^{(1)}, \tag{48}$$

where π is an element of K_w of order 1 in $\mathfrak{P}$ (i.e., $\mathfrak{P} = (\pi)$).

8. Distributions and measures. Much of the exposition in this section comes from [Ko]. Those familiar with real or complex valued measures and distributions, or even Riemann integration, will find most of the facts and proofs mentioned below quite obvious.

Let X be a locally compact open subset of $\mathbf{Q}_p$ and $\mathbb{F}$ be one of the fields $\mathbf{Q}_p$, $\mathbf{R}$ or $\mathbf{C}$. Typically, X will be $\mathbf{Z}_p$ or $\mathbf{Z}_p^\times$. By $\mathcal{S}_\mathbb{F}(X)$ we denote the space of locally constant functions on X with *compact support* with values in $\mathbb{F}$. A function f is said to have compact support in X if it vanishes outside a compact subset K of X. A locally constant function is a function for which every point x has a neighbourhood U in X such that $f(u) = f(x)$ for all u in U. In the p-adic setting $\mathcal{S}_\mathbb{F}(X)$ turns out to be the right analogue of the Schwartz space on $\mathbf{R}$ that we introduced in Exercise 9. Any function f in $\mathcal{S}_\mathbb{F}(X)$ can be written as a finite sum

$$f(x) = \sum_{i=1}^{k} c_i X_{(a_i+V)}(x), \tag{49}$$

where V is a fixed open compact subgroup of $\mathbf{Z}_p$, c_i is in $\mathbf{Q}_p$ and $X_{(a_i+V)}$ denotes the characteristic function of the set $a_i + V$.

Exercise 20. Prove that every $f(x)$ in $\mathcal{S}(\mathbf{Z}_p)$ can be written in the form (49).

A *p-adic distribution* λ is a linear functional from the vector space $\mathcal{S}_\mathbb{F}(X)$ to $\mathbb{F}$. For f in $\mathcal{S}_\mathbb{F}(X)$, we write the value of λ at f as $\lambda(f)$.

Let μ be an additive set function which associates an element in $\mathbb{F}$ to each compact open set in X, i.e., if $U = \bigcup_{i=1}^{n} U_i$, is a disjoint union of compact open sets, then

$$\mu(U) = \sum_{i=1}^{n} \mu(U_i) \tag{50}$$

holds. Any such map μ defines a map on characteristic functions

$$\lambda(X_U) = \mu(U), \tag{51}$$

where X_U is the characteristic function of U. By (49) we see that we can extend λ by linearity to all functions f in $\mathcal{S}_{\mathbb{F}}(X)$. Conversely, any distribution λ defines such a map μ by

$$\mu(U) = \lambda(X_U). \tag{52}$$

It is customary to write the value of μ (thought of as a distribution) on f as $\mu(f)$ or more commonly as $\int f\mu$. Sometimes when we wish to stress the variable on which f depends we also write $\int f(x)\mu(x)$.

Exercise 21. Show that μ defined by (52) satisfies (50).

A set U of the form $a + (p^N)$ with a in X and $N \in \mathbb{Z}$, is called an interval. Intervals form a *basis* of open sets in $\mathbf{Q}_p$.

Lemma 8.1. *Every map μ from the set of intervals in X to $\mathbb{F}$ such that*

$$\mu(a + (p^N)) = \sum_{b=0}^{p-1} \mu(a + bp^N + (p^{N+1})) \tag{53}$$

extends uniquely to a distribution on X.

Exercise 22. Prove Lemma 8.1. as follows. First show that every compact open set U can be written as a finite disjoint union of intervals I_i. Then define $\mu(U) = \sum \mu(I_i)$ and check that the resulting definition does not depend on the choice of partition of U into intervals.

From now on we actually assume that X is *compact.* The space $\mathcal{S}_{\mathbb{F}}(X)$ is thus just the space of locally constant functions. By (52), we may regard any distribution λ as giving rise to a set function μ. If $\mu(U)$ is bounded for every compact open set $U \subset X$, i.e., $|\mu(U)| < B$ for some fixed $B > 0$, we say that λ is a bounded distribution and that the corresponding μ is a *measure.* Once we have a measure μ we can define the notion of a *Riemann sum* as follows. Let

$$X = \bigcup_{a+(p^N)\subset X} (a + (p^N)) \tag{54}$$

be a (disjoint) partition P_N of X into intervals, and let $Y_N = \{x_a\}$ be a set of points such that each x_a is in the a^{th} interval. The N^{th} Riemann sum of a function f is defined by

$$S_{P_N,Y_N} = \sum_{a+(p^N)\subset X} f(x_a)\mu(a+(p^N)). \tag{55}$$

Naturally S_{P_N,Y_N} depends on the choice of partition P_N and the choice of the points Y_N in each interval. However, in the limit for continuous functions we have the following

Theorem 8.2. *Let f be a continuous function from X to $\mathbf{Q}_p$. Then the limit of the Riemann sums*

$$\lim_{N\to\infty} S_{P_N,Y_N}$$

exists in $\mathbb{F}$ independent of the choice of the sets Y_N. We call this limit $\int f\mu$ (Note that if f is locally constant this agrees with the previous definition of $\int f\mu$).

Theorem 8.2 is quite easily proved and we leave the proof as an exercise. Alternatively, we refer the reader to pp. 36-40 of [Ko]. The basic point is that locally constant functions are dense in the space of continuous functions so the measure μ, being a *bounded* linear functional on $\mathcal{S}_{\mathbb{F}}(X)$ extends to a linear functional on the space of continuous functions. All the “usual” facts about integration hold in this context. For instance, if $|f(x)| < M$ for all x in X and $|\mu(U)| < B$ for all compact opens sets $U \subset X$, then

$$\left| \int f\mu \right| < MB. \tag{56}$$

If f and g are two continuous functions such that $|f(x) - g(x)| < \epsilon$ for all x in X, then

$$\left| \int f\mu - \int g\mu \right| < \epsilon B. \tag{57}$$

Lastly, for an open subset V of X, by $\int_V f\mu$ we mean $\int f X_V(x)\mu(x)$, where X_V denotes the characteristic function of V.

9. Examples of p-adic distributions and measures. We continue to rely heavily on [Ko] for our exposition. We give a number of examples of distributions and measures below. Unless otherwise specified we will assume that $X \subset \mathbf{Z}_p$. Of special interest are the Bernoulli distributions and measures which will allow us to define the p-adic Mellin transform and thence the p-adic L-functions.

Example 1. The Haar distribution μ_{Haar} is defined by

$$\mu_{Haar}(a + (p^N)) = \frac{1}{p^N}.$$

It is enough to verify (53) to show that this extends to a distribution. Clearly

$$\mu_{Haar}(a + (p^N)) = \frac{1}{p^N} = \sum_{b=0}^{p-1} \frac{1}{p^{N+1}} = \sum_{b=0}^{p-1} \mu_{Haar}(a + bp^N + (p^{N+1})),$$

so condition (53) is verified. Note that we may think of μ_{Haar} as taking values in any of the fields $\mathbf{Q}_p$, $\mathbf{R}$ or $\mathbf{C}$. If we wish to make it clear that the Haar distribution is taking values in a field $\mathbb{F}$ we will write $\mu_{Haar,\mathbb{F}}$. The very definition of μ_{Haar} shows that it is *translation invariant* (under addition). Note that if $\mathbb{F} = \mathbf{Q}_p$, the Haar distribution is not bounded and hence is not a measure. For this reason it is not a very interesting distribution from the point of view of p-adic L-functions.

Exercise 23. Calculate $\mu_{Haar,\mathbf{R}}(\mathbf{Z}_p^\times)$.

Exercise 24. Let $X = \mathbf{Z}_p \setminus \{0\}$ and let $\mu = \mu_{Haar,\mathbf{R}}$ denote the *real valued* Haar measure as above. For $\mathbb{F} = \mathbf{R}$, we define a distribution ν on $\mathcal{S}_\mathbb{F}(X)$ by setting

$$\nu(f) = \frac{p}{p-1} \int_X f(x) \frac{\mu(x)}{|x|}$$

for f in $\mathcal{S}_\mathbb{F}(X)$. Here, $|x|$ denotes the p-adic absolute value of x. Show that ν is a measure on X. Also show that ν is invariant under multiplication by elements in X, i.e., $\nu(xU) = \nu(U)$ for all x in X and compact open sets U (ν is called the *multiplicative Haar measure*).

Exercise 25. Calculate $\nu(\mathbf{Z}_p^\times)$.

Exercise 26. For any $s \in \mathbf{C}$, with $Re(s) > 0$ and X as in Exercise 24, calculate the value (as a function of s) of

$$\int_X f(x)|x|^s \nu(x),$$

where $f(x)$ denotes the characteristic function of $\mathbf{Z}_p$ viewed as a function of X simply by restriction (Caution: $f(x)$ is not in $\mathcal{S}_\mathbb{F}(X)$. However, $f(x)$ is a continuous function on X.). The above integral is called the local L-function at the prime p. Perhaps it is now clear why the Gamma function enters the functional equation for the Riemann ζ-function. We need to keep track of the information from *all* primes - both finite and infinite. The local L-functions are nothing but Mellin transforms of suitable functions in the respective Schwartz spaces.

Exercise 27. Try and generalise Exercises 23-26 for a finite extension K_v of $\mathbf{Q}_p$.

Example 2. The Dirac Measure δ_a concentrated at a point a in $\mathbf{Z}_p$ is given by

$$\delta_a = \begin{cases} 1 & \text{if } a \in U; \\ 0 & \text{otherwise.} \end{cases}$$

It is trivial to check that δ is a measure.

Example 3. Let $X = \mathbf{Z}_p$ and $\mathbb{F} = \mathbf{Q}_p$. The Bernoulli distributions are given by

$$\mu_{B,k}(a + (p^N)) = p^{N(k-1)} B_k \left(\frac{a}{p^N} \right). \tag{58}$$

Once again, we verify condition (53). We have

$$\sum_{b=0}^{p-1} \mu_{B,k}(a + bp^N + (p^{N+1})) = \sum_{b=0}^{p-1} \mu_{B,k} p^{(N+1)(k-1)} B_k \left(\frac{a + bp^N}{p^{N+1}} \right). \tag{59}$$

Set $\frac{a}{p^N} = c$. The right-hand side of (59) then becomes

$$p^{(N+1)(k-1)} \sum_{b=0}^{p-1} B_k \left(c + \frac{b}{p} \right) = A. \tag{60}$$

On the other hand, we have

$$\begin{aligned} p^{k-1} \sum_{b=0}^{p-1} \frac{te^{(c+b/p)t}}{e^t - 1} &= \frac{p^{k-1} te^{ct}}{e^t - 1} \sum_{b=0}^{p-1} e^{bt/p} = \frac{p^{k-1} te^{ct}}{e^t - 1} \frac{e^t - 1}{e^{t/p} - 1} \\ &= \frac{p^k (t/p) e^{(pc)t/p}}{e^{t/p} - 1} = p^k \sum_{i=0}^{\infty} B_i(pc) \frac{(t/p)^i}{t!}. \end{aligned} \tag{61}$$

Comparing the coefficient of t^k in the first and last expressions of (60) we get

$$p^{-N(k-1)} \frac{A}{k!} = \frac{B_k(pc)}{k!}. \tag{62}$$

Hence we can show that

$$\sum_{b=0}^{p-1} \mu_{B,k}(a + bp^N + (p^{N+1})) = p^{N(k-1)} B_k(pc)$$

$$= p^{N(k-1)} B_k \left(\frac{a}{p^N} \right) = \mu_{B,k}(a + (p^N)).$$

This shows that our definition does give a distribution. We record the following easy observations for future reference.

$$\mu_{B,k}(\mathbf{Z}_p) = B_k \tag{63a}$$

$$\mu_{B,k}((p)) = p^{k-1}B_k \tag{63b}$$

From (63a) and (63b) we conclude that

$$\mu_{B,k}\left(\mathbf{Z}_p^\times\right) = (1-p^{k-1})B_k. \tag{63c}$$

In particular, $\mu_{B,0}$ is nothing but the Haar measure (recall that $B_0(X) \equiv 1$) and $\mu_{B,1}\left(\mathbf{Z}_p^\times\right) = 0$.

As can be checked easily, $\mu_{B,k}$ is not bounded. In order to obtain a measure we must *regularise* the Bernoulli distributions as follows.

For any rational integer $r \neq 1$, and any compact open subset U of $\mathbf{Z}_p$, we define

$$\mu_{k,r}(U) = \mu_{B,k}(U) - r^{-k}\mu_{B,k}(rU). \tag{64}$$

Exercise 28. Show that $\mu_{1,r}$ is a measure and that $|\mu_{1,r}(U)| \leq 1$ for every compact open subset of $\mathbf{Z}_p$.

We have the following relation between $\mu_{k,r}$ and $\mu_{1,r}$.

Theorem 9.1. *Let d_k be the least common denominator of the coefficients of $B_k(x)$. Then*

$$d_k\mu_{k,r}(a+(p^N)) \equiv d_k k a^{k-1}\mu_{1,r}(a+(p^N)) \pmod{p^N}. \tag{65}$$

The proof of Theorem 9.1 is again a fairly straightforward computation which is not particularly illuminating. Once again, we refer the reader to [Ko] or leave it to be attempted as an exercise. For those who understand the language, Theorem 9.1 says that $\mu_{k,r}$ is absolutely continuous with respect to $\mu_{1,r}$ and the Radon-Nikodym derivative is precisely the function kx^{k-1}. From Theorem 9.1 and Exercise 23 follows.

Corollary 9.2. *$\mu_{k,r}$ is a measure for all k.*

Proof of Corollary 9.2. If V is any compact open subset of $\mathbf{Z}_p$, then V is contained in (p^m) for some m (possibly negative). Hence, we see that $|x|^{k-1} < p^{-m(k-1)}$ on V. Let us set $\mathrm{ord}_p\,(kd_k) = l$. Then

$$\begin{aligned}\int_V \mu_{k,r} &= \sum_{a+(p^N)\subset V} \mu_{k,r}(a+(p^N)) \\ &\equiv \sum_{a+(p^N))\subset V} \mu_{1,r}(a+(p^N)) \pmod{p^{N-m(k-1)-l}}.\end{aligned}$$

Let $B = \mu_{1,r}(V)$. Letting $N \to \infty$ we see that

$$\int_V \mu_{k,r}(V) = k \int_V x^{k-1}\mu_{1,r} \tag{66}$$

which proves the Corollary.

Note that (66) not only gives the boundedness of the regularised Bernoulli distributions but also explicitly calculates the derivative of $\mu_{k,r}$ with respect to $\mu_{1,r}$. Now that we have the measures $\mu_{B,k}$, we record the following volume computations below.

$$\frac{1}{k}\mu_{k,r}(\mathbf{Z}_p) = \mu_{B,k}(\mathbf{Z}_p) - r^{-k}\mu_{B,k}(\mathbf{Z}_p) = (1 - r^{-k})B_k,$$
$$\frac{1}{k}\mu_{k,r}((p)) = \mu_{B,k}((p)) - r^{-k}\mu_{B,k}((p)) = (1 - r^{-k})p^{k-1}B_k.$$

Thus,

$$\frac{1}{k}\mu_{k,r}(\mathbf{Z}_p^\times) = \mu_{B,k}(\mathbf{Z}_p^\times) - r^{-k}(\mathbf{Z}_p^\times) = (r^{-k} - 1)(1 - p^{k-1})\left(-\frac{B_k}{k}\right).$$

Hence, we see that

$$\frac{1}{k}\mu_{k,r}(\mathbf{Z}_p^\times) = (r^{-k} - 1)(1 - p^{k-1})\zeta(1-k). \tag{67}$$

One usually sets

$$\zeta_p(k) = (1 - p^{-k})\zeta(k). \tag{68}$$

Note that $\zeta_p(k)$ is just the usual ζ-function with the Euler factor at p removed. By (66) and (68) we see that we can rewrite (67) as

$$\int_{\mathbf{Z}_p^\times} x^{k-1}\mu_{1,r} = (r^{-k} - 1)\zeta_p(1-k).$$

This can be expressed by

$$\zeta_p(1-k) = \frac{1}{(r^{-k} - 1)} \int_{\mathbf{Z}_p^\times} x^{k-1}\mu_{1,r}. \tag{69}$$

The importance of (69) lies in the following: Although the right-hand side is *a priori* defined only for positive integers k, we will be able to give it meaning for any p-adic integer s - in fact, it is a *p-adic Mellin transform* (see Exercise 14). This will enable us to define the p-adic ζ-function for any p-adic integer.

Warning. The reader may be bothered by the fact that because of (63c), $\mu_{1,r}\left(\mathbf{Z}_p^\times\right) = 0$, and, hence, in (69) we seem to be integrating on a set of measure 0. Notice, however, that $\mathbf{Z}_p^\times$ has subsets with non-zero measure. Measures with values in $\mathbf{Q}_p$ thus behave very differently from the more familiar measures which take only positive real values. This is why we need the condition $|\mu(U)| < B$ for *all* compact open sets U in X in order for (56) to hold. Were we dealing with the usual Lebesgue measure, for example, we would need the condition only for $U = X$.

10. The exponential function. The purpose of this section is to give meaning to the expression x^s when x is in $\mathbf{Z}_p^\times$ and s is an arbitrary p-adic integer. This will help us make sense of the integral in (69) for all p-adic integers s. We will first obtain x^s as a continuous function, and later, also prove that it is analytic.

First, we consider some u in $U_v^{(1)}$. We can write $u = 1+z$, where $z \in (p)$. If n_1 and n_2 are such that $n_1 - n_2 \equiv 0(\text{mod } p^N)$, i.e., $n_1 - n_2 = bp^N$, for some integer b, we can write

$$\begin{aligned} |u^{n_1} - u^{n_2}| &= |(1+z)^{n_1} - (1+z)^{n_2}| = |1+z|^{n_1}|1-(1+z)^{bp^N}| \\ &= |bp^N z + \tbinom{bp^N}{2} z^2 + \dots|_p \le |p^{N+1}| = \frac{1}{p^{N+1}}. \end{aligned} \tag{70}$$

Thus, we see that if n_1 and n_2 are close p-adically, then so are u^{n_1} and u^{n_2}. This shows that the exponential function $u \mapsto u^s$, $s \in \mathbb{Z}$, is a continuous function on $\mathbb{Z}$ endowed with the p-adic topology. It can thus be extended as a continuous function to the completion of $\mathbb{Z}$ under the p-adic norm and this is precisely $\mathbf{Z}_p$.

We have thus succeeded in defining a continuous function $u \mapsto u^s$ for any p-adic integer s, with u in $U_v^{(1)}$. However, we would like to give meaning to the function x^s when x belongs to $U_v^{(0)}$. To do this, we first remark that we can write any s in $U_v^{(0)}$ as $s = s_0 + (p-1)s_1$, where s_0 can be chosen to lie in the set $\{0, 1, \dots, p-2\}$ and s_1 lies in $U_v^{(0)}$.

Having done this, we define x^s by

$$x^s = x^{s_0+(p-1)s_1} = x^{s_0}x^{(p-1)s_1} = x^{s_0}(x^{p-1})^{s_1}. \tag{71}$$

This last expression makes sense since $y = x^{p-1}$ lies in $U_v^{(1)}$ and by remarks of the previous paragraphs y^s is meaningful for all s in $\mathbf{Z}_p$. We have thus succeeded in *p-adically interpolating* the function x^s for all x in $U_v^{(0)}$.

If s is a p-adic integer which is not a rational integer, then there will be arbitrarily large integers in any sequence s_n of integers approaching s in the

p-adic topology. Hence, if x actually lies in (p) the only way to define x^s as a limit will be as the zero function. One thus has to abandon the idea of p-adically interpolating x^s when x is in p.

So far the exponential function x^s has only been obtained as a continuous function. From the point of view of p-adically interpolating the right-hand side of (69) this is entirely adequate, and the reader may skip straight to Section 11 where this is undertaken. However, it turns out that x^s is actually a p-adic analytic function, i.e., it possesses a power series expansion in the neighbourhood of any point s in $\mathbf{Z}_p$. We can see this directly from (70) where we write u in $U_v^{(1)}$ as $1+z$ and make a binomial expansion. It would remain only to show that the binomial series converges for all s in $\mathbf{Z}_p$. A more indirect approach is through the logarithm and exponential functions discussed below.

The best setting for the logarithm and exponential functions is the field of *p-adic complex numbers* $\mathbf{C}_p$ which we will introduce in Section 12. For the moment we will suppose that our setting is $\bar{\mathbf{Q}}_p$, the algebraic closure of $\mathbf{Q}_p$.

We define the functions $\log_p(1+x)$ and $\exp(x)$ for x in a finite extension of $\mathbf{Q}_p$ as follows.

$$\log_p(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}.$$

$$\exp(x) = 1 + x + \frac{x^2}{2!} + \ldots = \sum_{n=0}^{\infty}\frac{x^n}{n!}. \tag{72}$$

We have to determine the radius of convergence of these series. This is easily done since we have only to verify when the n^{th} term goes to zero. Note that

$$v_p\left(\frac{x^n}{n}\right) = nv_p(x) - v_p(n) \geq nv_p(x) - \log n/\log p. \tag{73}$$

(Note that $\log n$ and $\log p$ in (73) denote the usual natural logarithms!) Hence, if $v_p(x) > 0$, the left-hand side of (73) tends to infinity as $n \to \infty$ and we see that the series for $\log_p(1+x)$ converges. $\log_p(1+x)$ thus has 1 as its radius of convergence. From now on, since no confusion will occur, we denote $\log_p$ simply by log.

Exercise 29. Show that the radius of convergence of $\exp(x)$ is $p^{-\frac{1}{p-1}}$.

Let p be an odd prime. Let K_w be a finite extension of $\mathbf{Q}_p$ $\mathfrak{P}$ be the prime ideal of $\mathcal{O}_w$ lying above (p) and e be its index of ramification. Then if x is in $\mathfrak{P}$ but not in $\mathfrak{P}^2$, we see that $v_p(x) = 1/e$ and hence, $|x| = p^{-\frac{1}{e}}$. By Exercise 29 we see that $\exp(x)$ converges for x in $\mathfrak{P}^n$, if $n > e/(p-1)$ (When $p = 2$, one can easily make the necessary modifications.).

The identities $\log(xy) = \log(x) + \log(y)$ and $\exp(x + y) = \exp(x)\exp(y)$ are identities of power series, and are hence valid whenever x, y , xy and $x+y$ are in the relevant domains of convergence. The identities $\log(\exp(x)) = x$ and $\exp(\log(1 + x)) = 1 + x$ are also identities for power series, and hence valid for any value of x for which both series converge. In particular, this says that for $n > e/(p - 1)$ the function exp defines an isomorphism

$$\exp : \mathcal{U}_w^{(n)} \mapsto \mathfrak{P}^n$$

whose inverse is obviously given by log. Notice that these maps induce the isomorphisms given by (43) (this is obvious if one truncates the power series for log and exp).

Using the power series for the exponential and logarithmic functions we can obtain the power series for the function x^s, when x is in $U_v^{(0)}$. Indeed, setting $s = s_0 + (p - 1)s_1$, as before, and $y = x^{p-1}$, we see that

$$x^s = x^{s_0} y^{s_1} = x^{s_0}\exp(s_1 \log y),$$

where y is in $U_v^{(1)}$. Hence, $\log y$ is in (p) and the power series for $\exp(s_1 \log y)$ is valid, so we find

$$x^s = x^{s_0}\left(1 + s_1 \log y + (s_1 \log y)^2/2! + \dots\right).$$

If we substitute $s_1 = \frac{s-s_0}{p-1}$ back into the equation above we can obtain a power series in $s - s_0$. This shows that the function x^s is actually a p-adic analytic function of s for any value of s.

11. Interpolation and Congruences. We now proceed to give meaning to the expression

$$f(s) = \int_{\mathbf{Z}_p^\times} x^{s-1}\mu_{1,r}$$

for any s in $\mathbf{Z}_p$. We have already seen in the last section that x^{s-1} is a continuous function. It remains to be shown that $f(s)$ is also a continuous function. Using the notation of the last section we write $x^s = x^{s_0} y^{s_1}$ with y in $U_v^{(1)}$. If $n_1 \equiv n_2 \pmod{p^{N+1}}$, (70) tells us that

$$|y^{n_1} - y^{n_2}| \leq \frac{1}{p^{N+1}}.$$

By Exercise 28 and (57), we see immediately that

$$|f(n_1) - f(n_2)| \leq \frac{1}{p^{N+1}}.$$

This shows that f is a continuous function on $\mathbb{N}$ endowed with the p-adic topology and hence, it extends to a continuous function on the closure $\mathbf{Z}_p$. We have already made sense of the function r^{-s} when r is a rational integer not divisible by p. Hence, we can now define

$$\zeta_p(1-s) = \frac{1}{r^{-s}-1}\int_{\mathbf{Z}_p^\times} x^{s-1}\mu_{1,r}\ , \tag{74}$$

or, equivalently, we may write

$$\zeta_p(s) = \frac{1}{r^{-(1-s)}-1}\int_{\mathbf{Z}_p^\times} x^{-s}\mu_{1,r}. \tag{75}$$

This last formula is obviously valid as long as $s \neq 1$. If $s \neq 1$, we see that $\zeta_p(s)$ is obviously a p-adic analytic function of s. To see this one simply expands x^{-s} in power series and integrates the resulting series term by term to get the power series for $\zeta_p(s)$. One checks easily that $(s-1)\zeta_p(s)$ is analytic and one concludes that $\zeta_p(s)$ has a simple pole at $s = 1$.

Warning. The definition of $\zeta_p(s)$ given above is not the one given in the lectures (which was from [Ko]). The roles of s and $1-s$ have been switched so $\zeta_p(s)$ now has a pole at 1 and not at 0 as before. In general, both definitions of $\zeta_p(s)$ crop up in the literature.

The definition of $\zeta_p(s)$ seems to depend on the choice of integer r that we made initially. However, we notice that the values of $\zeta_p(s)$ at the negative integers $1-k$ are simply $-\frac{B_k}{k}$ and these are independent of the choice of r. These values determine the continuous function $\zeta_p(s)$ uniquely since they determine it on a dense subset. Hence, $\zeta_p(s)$ does not depend on r. Notice also that the value of $\zeta_p(s)$ is identically 0 whenever s_0 is odd, so we are interested only in even s_0.

We have thus succeeded in our main goal which was to interpolate the values of the usual Riemann ζ-function at the integers to obtain a p-adic analytic function $\zeta_p(s)$. Now that we have achieved this we would like to examine the values of $\zeta_p(s)$ at the rational integers and read off their number theoretic properties, just as we did in the complex valued case. Theorems 6.1 and 6.2 now simply reduce to estimating the integral in (75) with appropriate choices of integers r. We first prove Theorem 6.2. For this we choose $r \in \{2, 3, \ldots, p-1\}$ such that $r \bmod p$ generates $\mathbb{F}_p^\times$, i.e., the order of $r \bmod p$ is exactly $p-1$. If $m \equiv n \pmod{(p-1)p^a}$, then by (70), we know that

$$(r^{-n}-1) \equiv (r^{-m}-1) \pmod{p^{a+1}}. \tag{76}$$

We write

$$\Big|\int_{\mathbf{Z}_p^\times} x^{m-1}\mu_{1,r} - \int_{\mathbf{Z}_p^\times} x^{n-1}\mu_{1,r}\Big| = \Big|\int_{\mathbf{Z}_p^\times} x^{m-1} - x^{n-1}\mu_{1,r}\Big|.$$

By using (70) again, and also Exercise 28 and (57), we see that this last expression can be estimated by

$$|\int_{\mathbf{Z}_p^\times} x^{m-1} - x^{n-1}\mu_{1,r}\Big| \leq \frac{1}{p^{a+1}}.$$

Hence, we conclude that

$$\int_{\mathbf{Z}_p^\times} x^{m-1}\mu_{1,r} \equiv \int_{\mathbf{Z}_p^\times} x^{n-1}\mu_{1,r} \pmod{p^{a+1}}. \tag{77}$$

Multiplying the congruences (76) and (77) together, we see that

$$\frac{1}{r^{-m}-1}\int_{\mathbf{Z}_p^\times} x^{m-1}\mu_{1,r} \equiv \frac{1}{r^{-n}-1}\int_{\mathbf{Z}_p^\times} x^{n-1}\mu_{1,r} \pmod{p^{a+1}},$$

and this says exactly that

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

proving Theorem 6.2 .

To prove Theorem 6.1 we proceed in two stages. First, we consider the case when $p-1 \not| n$. We choose r as above. We then have the estimate

$$\Big|\frac{B_n}{n}\Big| = \Big|\frac{1}{r^{-n}-1}\Big|\Big|\frac{1}{1-p^{n-1}}\Big|\Big|\int_{\mathbf{Z}_p^\times} x^{n-1}\mu_{1,r}\Big| = \Big|\int_{\mathbf{Z}_p^\times} x^{n-1}\mu_{1,r}\Big| \leq 1.$$

This shows that $B_n \in \mathbf{Z}_p$ for all p such that $(p-1) \not| n$.

Exercise 30. If $p-1$ does divide n, prove that $pB_n \equiv -1 \pmod{p}$ and hence, that $B_n + 1/p \in \mathbf{Z}_p$. (Hint: If $p > 2$, choose $r = p+1$. If $p = 2$ choose $r = p^2 + 1 = 5$.)

Let q be any prime and consider the expression

$$B_n + \sum_{(p-1)|n} \frac{1}{p}. \tag{78}$$

If $(q-1) \not| n$, then $q \neq p$, for any of the p occurring in the sum in (78), so $1/p \in \mathbb{Z}_q$ for all these p. Since we already know that $B_n \in \mathbb{Z}_q$, we see that $B_n + \sum_{(p-1)|n} \frac{1}{p} \in \mathbb{Z}_q$. On the other hand, if $q-1$ divides n, $1/q$ appears in the sum in (78) and Exercise 30 tells us that $B_n + 1/q \in \mathbb{Z}_q$. For $p \neq q$ occurring in the sum, $1/p$ is a unit, so all of these terms are also in $\mathbb{Z}_q$. We thus see that the expression (78) lies in $\mathbb{Z}_q$ for all primes q, and hence, in $\mathbb{Z}$ itself. This proves Theorem 6.2.

12. Some extensions. In this section we discuss generalisations of some of the notions we have studied so far. No proofs will be given. The idea is to introduce various new objects and discuss a few of their properties so that the reader has some familiarity with them when they are encountered elsewhere.

The only example of p-adic interpolation that we have considered in these notes is that of the Riemann ζ-function. This was done so that one conveyed the basic idea without bothering too much about technical details. It should be clear to the reader that by interpolating the values of the Hurwitz ζ-function and the Dirichlet L-functions we can likewise obtain p-adic analogues. We already have all the machinery set up should we want to define such p-adic L-functions using measures and and Mellin transforms and it may be a useful exercise (to see if your understanding is complete) to try and do so. Using formula (6), one can then get the p-adic Dedekind ζ-function, at least for cyclotomic fields. Studying such generalisations of L-functions naturally gives generalisations of the Kummer congruences. For instance, if $\chi \not\equiv 1$ we have

Theorem 12.1. *If* $(p-1) \not| m$ *and* $m \equiv n \pmod{(p-1)p^a}$ *and* $n \not\equiv 0 \pmod{(p-1)p^a}$*, then*

$$\frac{B_{m,\chi}}{m} = \frac{B_{n,\chi}}{n} \pmod{p^{a+1}}.$$

The values of p-adic L-functions at $s = 1$ are also extremely important from the point of view of studying class numbers and regulators of number fields. We refer to section 5.5 of [W] for this material including a discussion of Leopoldt's conjecture.

We now discuss the p-adic complex numbers. The field of complex numbers is both complete as well as algebraically closed. In this section we introduce the fields $\mathbf{C}_p$, which also share these two properties. Starting with $\mathbf{Q}_p$ we construct $\mathbf{C}_p$ as follows. First take the algebraic closure $\bar{\mathbf{Q}}_p$ of $\mathbf{Q}_p$. This is easily seen to be not complete (Proposition 5.1 of [W]). Let $\mathbf{C}_p$ denote the completion of $\bar{\mathbf{Q}}_p$ with respect to the p-adic valuation. $|\ |_p$ thus extends to a norm on $\mathbf{C}_p$. It is a fact that $\mathbf{C}_p$ is algebraically closed (Proposition 5.2 of [W]).

The multiplicative structure of $\mathbf{C}_p$ is very similar to that of $\mathbf{Q}_p$. We let W be the roots of unity of order prime to p and U be the subgroup defined by

$$U = \{u \in \mathbf{C}_p \mid |u-1|_p < 1\}.$$

Then we have the decomposition

$$\mathbf{C}_p^\times = \{p^{\mathbf{Q}}\} \times W \times U,$$

where $p^{\mathbf{Q}}$ denotes the group of all rational powers of p, and the embedding of $p^{\mathbf{Q}}$ in $\mathbf{C}_p$ is chosen so that $p^r p^s = p^{r+s}$.

One can see easily that we can define the function log on the set U by means of power series as before, and that this function has radius 1. Similarly, exp can also be defined, and it has radius of convergence $p^{-1/(p-1)}$ in $\mathbf{C}_p$. If x is in $\mathbf{C}_p$ we can write $x = p^r \omega u$, with $\omega \in W$ and $u \in U$, and, moreover, this decomposition is unique. We can extend the function log to all of $\mathbf{C}_p$ so that $\log(xy) = \log x + \log y$ by setting $\log p^r \omega u = \log u$. This extension is the unique one with the additional property that $\log p = 0$.

It turns out that we can define p-adic L-functions as functions from $\mathbf{C}_p$ to $\mathbf{C}_p$. There are several ways of doing this but we simply write down a formula for the p-adic Hurwitz ζ-function without proof and refer the reader to section 5.2 of [W] for details. From the expression for Hurwitz ζ-function it will be easy for the reader to write down the definitions for the p-adic analogues of the other L-functions discussed in these notes. Let us set $H(s,a,F) = F^{-s} H(s, \frac{a}{f})$. By the expression $\binom{s}{n}$, $s \in \mathbf{C}_p$ and $n \in \mathbb{N}$, we shall simply mean $\frac{s(s-1)\dots(s-(n-1))}{n!}$. With this notation we have

Theorem 12.2. *The function*

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

is analytic except for a simple pole at $s = 1$ (with residue $1/F$) with the property that $H_p(1-k,a,F) = \omega^{-k}(a)H(1-k,a,F)$, if $n \geq 1$ and ω is the Teichmüller character.

References

[Ka] S. A. Katre, *The Stickelberger's Theorem*, This volume.

[Ko] N. Koblitz, *p-adic numbers, p-adic analysis and zeta-functions*, Graduate Texts in Mathematics, Springer Verlag (1984) **58**.

[L1] S. Lang, *Algebraic Number Theory*, Graduate Texts in Mathematics, Springer Verlag (1994) **110**.

[L2] ———, *Cyclotomic Fields I and II*, Graduate Texts in Mathematics, Springer Verlag (1990) **121**.

[S] J.-P. Serre, *A Course in Arithmetic*, Graduate Texts in Mathematics, Springer Verlag (1973) **7**.

[W] L. Washington, *Introduction to Cyclotomic Fields*, Graduate Texts in Mathematics, Springer Verlag (1997) **83**.

Ravi Raghunathan
School of Mathematics
Tata Institute of Fundamental Research
Mumbai 400-005
e-mail: ravir@math.tifr.res.in

On the Theorem[1] of Hasse-Minkowski

R. Sridharan [2]

Let K be a field (which, in what follows, will always be assumed to be of characteristic different from two). A *quadratic form* over K in n variables is a homogeneous polynomial f with coefficients in the variables $X_1, X_2, \cdots, X_n$:

$$f(X_1, X_2, \cdots, X_n) = \sum_{1 \leq i \leq n} a_i {X_i}^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} X_i X_j \qquad (*)$$

An *isotropy* of a quadratic form f over K is a non-zero vector, $(x_1, x_2, \cdots, x_n) \in K^n$, such that $f(x_1, x_2, \cdots, x_n) = 0$. We say that a quadratic form f over K is isotropic, if it has an isotropy.

Example. The quadratic form $X_1^2 - X_2^2$ (called the *hyperbolic plane*) is obviously isotropic. But the form $X_1^2 + X_2^2$ over the field $\mathbb{Q}$ of rational numbers is not isotropic.

Recall that an *algebraic number field* is by definition, a field extension of $\mathbb{Q}$ of finite degree. The classical theorem whose proof we wish to discuss here is due to Hasse-Minkowski, which gives a necessary and sufficient condition for a quadratic form over such a field to be isotropic in terms of “local” conditions. To state and prove this theorem we need some preliminary notation and results.

Let K be any field. Recall that a *valuation* of K is a map $v : K \to \mathbb{R}^+$, $\mathbb{R}^+$ denoting the non-negative real numbers, such that for $x, y \in K$,

1. $v(x) = 0$ if and only if $x = 0$,

2. $v(x + y) \leq v(x) + v(y)$,

3. $v(xy) = v(x)v(y)$, for $x, y \in K$.

[1] For some interesting historical comments on Hasse's proof of this theorem see the article of Hasse entitled ”Kurt Hensels entscheidender Anstoss zur Entdeckung des Lokal-Global-Prinzips” in Crelle's Journal, 209, $3-4$, 1962, where Hasse discusses the important role played by Hensel in his proof of this theorem. Minkowski's proof was never published.

[2] I thank S.A.Katre and Dinesh Thakur for inviting me to give a few lectures on this theorem in the summer school on Cyclotomic Fields held at Pune in June 1999. I want take this opportunity to thank S.D.Adhikari who was mainly responsible for my writing up these notes. I am grateful to Preeti Raman who among other things, made this Tex version possible. I also thank Raja Sridharan who went through the writeup and suggested useful changes. As ever, I am deeply indebted to Parimala for all her help.

The map $v : K \to \mathbb{R}^+$ defined by $v(0) = 0$ and $v(x) = 1$ for $x \neq 0$ is obviously a valuation of K called the *trivial valuation.* From now on by a valuation we mean a non-trivial valuation. We have a notion of equivalence of valuations defined by: $v \sim v'$, if and only if there exists a positive real number c such that $v' = v^c$. It is easy to see that v and v' are equivalent if and only if for any $x \in K$, $v(x) < 1$ implies $v'(x) < 1$. We shall often confuse between a valuation and its equivalence class. For any valuation v of K, we denote by K_v, the completion of K, with respect to v: K_v is in fact the completion of K with respect to the metric defined for $x, y \in K$ by $d(x, y) = v(x - y)$. The topological space K_v has an obvious field structure and K sits in K_v as a dense subfield.

We start with the following lemma on valuations.

Lemma 1. *Let $v_1, v_2, \cdots, v_n$ be pairwise inequivalent non-trivial valuations of any field K. Then the image of K under the canonical injection $K \to \prod_{1\leq i\leq n} K_{v_i}$ is dense. In other words, given $(x_1, x_2, \cdots, x_n) \in \prod_{1\leq i\leq n} K_{v_i}$ and an $\epsilon > 0$ in $\mathbb{R}$, there exists $x \in K$, such that $v_i(x - x_i) < \epsilon$.*

Proof. We note first that it is enough to find for each r, $1 \leq r \leq n$, $\theta_r \in K$ such that $v_r(\theta_r) > 1$, $v_m(\theta_r) < 1$, for $m \neq r$, $1 \leq m \leq n$. For, then as $s \to +\infty$, we have $\frac{\theta_r^s}{1+\theta_r^s} = \frac{1}{1+\theta_r^{-s}} \to 1$ with respect to v_r and $\frac{\theta_r^s}{1+\theta_r^s} = \frac{1}{1+\theta_r^{-s}} \to 0$ with respect to v_m, for $m \neq r$.Then it is enough to take $\xi = \sum_{r=1}^{n} \frac{\theta_r^s}{1+\theta_r^s} x_r$, for a sufficiently large s.

We show the existence of $\theta = \theta_1$, with $v_1(\theta) > 1$ and $v_r(\theta) < 1$ for $2 \leq r \leq n$. To do this we use induction on n.

Let $n = 2$. Since v_1 and v_2 are inequivalent, there exist α, β such that $v_1(\alpha) < 1$ and $v_2(\alpha) \geq 1$ and $v_1(\beta) \geq 1$ and $v_2(\beta) < 1$. Then $\theta = \beta\alpha^{-1}$ will do.

Let $n \geq 3$. By induction, there is a $\phi \in K$, such that $v_1(\phi) > 1$ and $v_r(\phi) < 1$, for $2 \leq r \leq n - 1$. By the case $n = 2$, there is a $\psi \in K$, such that $v_1(\psi) > 1$ and $v_n(\psi) < 1$. Then put

$$\theta = \begin{cases} \phi, & \text{if } v_n(\phi) < 1 \\ \phi^s\psi, & \text{if } v_n(\phi) = 1 \\ \frac{\phi^s}{1+\phi^s}\psi, & \text{if } v_n(\phi) > 1 \end{cases}$$

where $s \in \mathbb{N}$ is sufficiently large. This completes the proof.

We recall that the valuations of the field $\mathbb{Q}$ are given up to equivalence by the usual absolute value (called the *archimedean valuation* of $\mathbb{Q}$), or by the normalised p-adic valuation, corresponding to any prime p, defined for $a \in \mathbb{Q}$, $a \neq 0$ by $v_p(a) = p^{-w_p(a)}$ and $v_p(0) = 0$, $p^{w_p(a)}$ being the maximum power of p which divides a. Let K be an algebraic number field. It is well known that any valuation of $\mathbb{Q}$ extends to finitely many inequivalent valuations of K. We note that if v is any valuation of K which is an extension of a p-adic valuation of $\mathbb{Q}$, then v satisfies the stronger condition $2'$: $v(x+y) \leq max(v(x), v(y))$, for $x, y \in K$.

If f is a quadratic form over K, which is isotropic, then obviously for any v, f is isotropic over K_v. The theorem of Hasse-Minkowski is indeed the converse of this statement, namely,

Theorem 2. *Let f be a quadratic form over an algebraic number field K. If f is isotropic over K_v for all v, then f is isotropic over K.*

In order to prove the theorem, we begin with some general facts on quadratic forms. Let f be a quadratic form, given by $f = \sum_{1 \leq i \leq n} a_i X_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} X_i X_j$. If we denote by $\bar{X}$, the row vector $(X_1, X_2, ..., X_n)$ and by A_f, the symmetric matrix whose diagonal entries are a_i and the off-diagonal $(i,j)^{th}$ entries are a_{ij}, then we have $f(\bar{X}) = \bar{X} A_f \bar{X}^t$. Let f and g be quadratic forms, given by $f = \sum_{1 \leq i \leq n} a_i X_i^2 + \sum_{1 \leq i < j \leq n} 2a_{ij} X_i X_j$ and $g = \sum_{1 \leq i \leq n} b_i X_i^2 + \sum_{1 \leq i < j \leq n} 2b_{ij} X_i X_j$. We say that f and g are *equivalent*, written $f \sim g$ if there exists a $u \in Gl_n(K)$, such that $g(\bar{X}) = f(\bar{X}u)$. In other words, $\bar{X} A_g \bar{X}^t = \bar{X} u A_f (\bar{X}u)^t = \bar{X} u A_f u^t \bar{X}^t$. This implies that $A_g = u A_f u^t$ and conversely, if there exists a $u \in Gl_n(K)$, with $A_g = u A_f u^t$, by reversing the steps we see that $f \sim g$. If f and g are equivalent and $A_g = u A_f u^t$, then $det(A_g) = det(A_f).det(u)^2$, so that the class of $det(A_f)$ modulo K^{*2} depends only on the equivalence class of f, and is called the *discriminant* of the quadratic form f, denoted by $disc(f)$. We say that a quadratic form f is *diagonal* if $a_{ij} = 0$ for $i \neq j$ or what is the same, A_f is a diagonal matrix. If the diagonal entries are $a_1, a_2, \cdots, a_n$, we shall denote in what follows, such a form by $< a_1, a_2, \cdots, a_n >$.

Proposition 3. *Any quadratic form (over a field of characteristic different from* 2*), is equivalent to a diagonal form.*

Proof. Let f be a quadratic form. Since the characteristic of K is not 2 and f is not identically zero, there exists $u = (u_1, u_2, \cdots, u_n) \in K^n$ such that $a = f(u_1, u_2, \cdots, u_n) = u A_f u^t \neq 0$. Let $W \subset K^n$ be defined by $W = \{w \in K^n \mid u A_f w^t = 0\}$. Obviously W is a subspace of K^n and $W \cap K.u = 0$, since $u \notin W$. We have that $K^n = W \oplus K.u$, since any $z \in K^n$ can be written as $(z - \lambda u) + \lambda u$ and for $\lambda = a^{-1} u A z^t$, $z - \lambda u \in W$. We now choose a basis of K^n which consists of u and a basis of W. For this choice of a basis of K^n, A_f has the form

$$\begin{pmatrix} a & 0 \\ 0 & B_{(n-1)\times(n-1)} \end{pmatrix}.$$

The proof now follows by induction on n.

In view of the above proposition, from now on, we assume that f is a diagonal form. The number of non-zero a_i (which is simply the rank of the matrix A_f), is called the *rank of f*.

We now prove the following fact, which will be used in the proof of the next proposition.

Lemma 4. *Any quadratic form of rank greater than or equal to 3, over a finite field of characteristic different from 2, is isotropic.*
Proof. [3] Obviously it is enough to show that the equation $aX^2+bY^2=1$ has a solution over any finite field $\mathbb{F}_q$ of q elements. The number of elements of the set $S=\{a\lambda^2 \mid \lambda\in\mathbb{F}_q\}$ has cardinality $\frac{q+1}{2}$, which is also the cardinality of the set $S'=\{1-b\mu^2 \mid \mu\in\mathbb{F}_q\}$. Since the number of elements of $\mathbb{F}_q$ is q, S and S' must intersect, which proves the lemma.

Let K be an algebraic number field, v a valuation of K and let $F=K_v$ denote the completion of K at v. If v lies over the unique archimedean valuation of $\mathbb{Q}$, then K_v is isomorphic to either the real number field $\mathbb{R}$, or the field $\mathbb{C}$ of complex numbers. Assume now that v is non-archimedean. The set $\mathcal{O}_F=\{x\in F \mid v(x)\leq 1\}$ is easily checked to be a subring of K and has a unique non-zero prime ideal, i.e., $\{x\in\mathcal{O}_F \mid v(x)<1\}$, which is principal. Any generator $\pi=\pi_v$ of this ideal is called a *uniformising parameter for v*. The field $\bar{F}=\mathcal{O}_F/(\pi)$ (called the *residue field at v*) is a finite extension of the prime field $\mathbb{Z}/p\mathbb{Z}$, (where v is an extension of the p-adic valuation of $\mathbb{Q}$) whose degree is denoted by f_v *(called the residue class field degree at v)*, so that $\bar{F}$ is a finite field with $q=p^{f_v}$ elements. By a *unit* of F, we mean an invertible element of $\mathcal{O}_F$, i.e., an element not in $\pi\mathcal{O}_F$. We note that an element $u\in F$ is a unit if and only if $v(u)=1$. We record the next proposition which is needed in the proof of theorem 2.

Proposition 5. *Let v be a non-archimedean, non-dyadic, valuation of K, i.e., v is not an extension of the 2-adic valuation of $\mathbb{Q}$. Let $f=<u_1,u_2,u_3>$ be a rank 3 quadratic form over $F=K_v$, where u_i for $1\leq i\leq 3$, are units of F. Then f is isotropic over F.*

[3] I thank Dinesh Thakur for bringing to my attention the slick proof given here, which is different from the one I had given in the lecture.

Proof. In fact proving the proposition is equivalent to showing that there exist $\lambda_i \in \mathcal{O}_F$, for $1 \leq i \leq 3$, not all zero such that $\sum_{1\leq i\leq 3} u_i\lambda_i^2 = 0$. Since $\mathcal{O}_F$ is a subring of F, which is closed in F, it is complete for the topology of F. In fact $\mathcal{O}_F$ is a topological ring for which a fundamental system of neighbourhoods of 0 are (π^n) for $n \geq 0$.

Let $^-$ denote reduction modulo (π). Since $\bar{F}$ is a finite field, $\bar{f}$ is isotropic, by Lemma 4. Therefore there exist $\lambda_1, \lambda_2, \lambda_3 \in \mathcal{O}_F$, such that $\sum_{1\leq i\leq 3} \bar{u_i}\bar{\lambda_i^2} = 0$. We can assume that one of the λ_i is a unit, let it be λ_1. We assume by induction that there exists an integer $n \geq 1$ and $\lambda_1, \lambda_2, \lambda_3$ in $\mathcal{O}_F$, with λ_1 a unit such that $\sum_{1\leq i\leq 3} u_i\lambda_i^2 \in (\pi^n)$, so that its image in $\mathcal{O}_F/(\pi^n)$ is zero. Certainly this is true for $n = 1$, as we have just now remarked. We wish to solve for an $h \in (\pi^n)$, such that $u_1(\lambda_1 + h)^2 + \ u_2\lambda_2^2 + u_3\lambda_3^2 \in (\pi^{n+1})$. Since $h \in (\pi^n)$, $h^2 \in (\pi^{2n}) \subset (\pi^{n+1})$, the choice $h = \frac{-(u_1\lambda_1^2+u_2\lambda_2^2+u_3\lambda_3^2)}{2u_1\lambda_1}$ would do. Since $\mathcal{O}_F$ is complete, an iteration of this procedure leads to a Cauchy sequence, which, since $\mathcal{O}_F$ is complete, converges and yields a solution of the equation $\sum_{1\leq i\leq 3} u_iX_i^2 = 0$ in $\mathcal{O}_F$. This proves the proposition.

We record another result, on the completion of a number field K, with respect to a non-archimedean valuation, which we shall use in the proof of Thoerem 2.

Lemma 6. *Let K be a number field and $F = K_v$ be the completion of K at a non-archimedean valuation v of K, (possibly dyadic). Let $\mathcal{O}_F$ be the ring of integers of F and $\pi = \pi_v$ be a parameter. Then any element in $1 + (4\pi)$ is the square of a unit of $\mathcal{O}_F$.*

Proof. Let $u = 1 + 4\lambda\pi$, for some $\lambda \in \mathcal{O}_F$. Setting $z_0 = 1$, we assume by induction that there exists an integer $n \geq 1$ and an unit $z_n \in \mathcal{O}_F$, such that $z_n^2 \equiv u \bmod (4\pi^{n+1})$ and $z_n - z_{n-1} \in (2\pi^n)$. For $n = 1$, we choose $z_1 = 1 + 2\lambda\pi$, which obviously satisfies the above conditions. We want to solve for an $h \in (2\pi^{n+1})$, such that setting $z_{n+1} = z_n + h$, we have $z_{n+1}^2 \equiv u \bmod (4\pi^{n+2})$. Since z_n is a unit and $u - z_n^2 \in (4\pi^{n+1})$, the choice $h = \frac{u-z_n^2}{2z_n}$ would do. This procedure leads to a Cauchy sequence in $\mathcal{O}_F$, which converges to a unit z, such that $z^2 = u$.

The proof of Theorem 2 is achieved by first proving it for quadratic forms of low ranks and then appealing to induction. To prove the theorem, we assume that $f =< a_1, a_2, \cdots, a_r >$, $r \geq 2$, and $a_i \neq 0$, for $1 \leq i \leq r$.

Let $r = 2$. Since $< a_1, a_2 >= a_1 < 1, a_1^{-1}a_2 >$, we may assume that $f =< 1, -\lambda >$, $\lambda \in K^*$, which is isotropic if and only if λ is a square in K. Theorem 2 in this case follows from the following theorem, which is a consequence of the first inequality of Class field theory.

Theorem 7. *Let K be a number field, $\lambda \in K$ and $L = K(\sqrt{\lambda})$. If for every valuation v of K, $K_v(\sqrt{\lambda}) = K_v$, then $K(\sqrt{\lambda}) = K$.*

Let $r = 3$. As above, we may assume that $f = \ < -1, \lambda, \mu >$. This form is isotropic if and only if there exist $\alpha_1, \alpha_2, \alpha_3 \in K$ such that $\alpha_1^2 = \lambda\alpha_2^2 + \mu\alpha_3^2$. Obviously α_2 and α_3 cannot both be zero. We may assume without loss in generality that α_2 is not zero, so that $\lambda = (\frac{\alpha_1}{\alpha_2})^2 - \mu(\frac{\alpha_3}{\alpha_2})^2$, i.e., λ is a norm in the extension $K(\sqrt{\mu})$ over K. In this case, the theorem of Hasse-Minkowski follows from the following more general,

Theorem 8. *(Hasse norm theorem) Let K be an algebraic number field and L over K a cyclic extension of K. Then an element $\lambda \in K$ is a norm from K if and only if $\lambda \in K_v$ is a norm from L_w over K_v, for all v, where w is some valuation of L, extending v.*

The theorem applied to the case where L over K is a quadratic extension yields the Hasse-Minkowski theorem, for $r = 3$.

Let $r = 4$. We may assume, as before by scaling that $f =< 1, -a, -b, c >$. We first consider the case where the discriminant of f is 1, i.e., $abc \in K^{*2}$, so that f can be replaced by the equivalent quadratic form $< 1, -a, -b, ab >$. The condition that f is isotropic is equivalent to saying that there exist $\lambda_1, \lambda_2, \lambda_3, \lambda_4 \in K$ such that ${\lambda_1}^2 - a{\lambda_2}^2 - b{\lambda_3}^2 + ab{\lambda_4}^2 = 0$. Let v be any valuation of K. If ${\lambda_3}^2 - a{\lambda_4}^2 = 0$, for some $\lambda_3, \lambda_4 \in K_v^*$, then $K_v(\sqrt{a}) = K_v$, so that any element of K_v, in particular b is a norm from $K_v(\sqrt{a}) = K_v$. If on the other hand there is a solution in K_v, with ${\lambda_3}^2 - a{\lambda_4}^2 \neq 0$, then $b = \frac{{\lambda_1}^2 - a{\lambda_2}^2}{{\lambda_3}^2 - a{\lambda_4}^2}$ is a norm in $K_v(\sqrt{a})$ over K_v. The Hasse norm theorem quoted above shows then that $b \in K$ is a norm from $K(\sqrt{a})$. This implies that $b = {\lambda_1}^2 - a{\lambda_2}^2$, which shows that f is isotropic over K.

Let now $f =< 1, -a, -b, c >$ and $\mathrm{disc}(f) = abc \notin K^{*2}$. To prove the theorem in this case, we need the following lemma.

Lemma 9. *Let $q =< a_1, a_2, \cdots, a_r >$ be a quadratic form over a field K (of characteristic not 2) and $L = K(\sqrt{d})$ be a quadratic extension of K. If q is not isotropic over K and is isotropic over L, then q is equivalent to $\lambda < 1, -d, b_1, b_2, \cdots, b_{r-2} >$, with $\lambda, b_1, \cdots, b_{r-2} \in K$.*

Proof. Let $\lambda_i + \mu_i\sqrt{d} \in L$, for $1 \leq i \leq r$, such that $\sum_{1 \leq i \leq r} a_i(\lambda_i + \mu_i\sqrt{d})^2 = 0$, so that $\sum_{1 \leq i \leq r} a_i{\lambda_i}^2 + d\sum_{1 \leq i \leq r} a_i{\mu_i}^2 = 0$ and $\sum_{1 \leq i \leq r} a_i\lambda_i\mu_i = 0$. Since q is not isotropic over K, $\sum_{1 \leq i \leq r} a_i{\lambda_i}^2 \neq 0$, $\sum_{1 \leq i \leq r} a_i{\mu_i}^2 \neq 0$ and $d = -\frac{\sum_{1 \leq i \leq r} a_i{\lambda_i}^2}{\sum_{1 \leq i \leq r} a_i{\mu_i}^2} = \frac{-q(\bar{\lambda})}{q(\bar{\mu})}$, where $\bar{\lambda} = (\lambda_1, \cdots, \lambda_r)$ and $\bar{\mu} = (\mu_1, \cdots, \mu_r)$. The equation $\sum_{1 \leq i \leq r} a_i\lambda_i\mu_i = 0$ shows first that $\bar{\lambda}$ and $\bar{\mu}$ are linearly independent vectors in K^r and that if we extend $\bar{\lambda}, \bar{\mu}$, to a suitable basis of K^r, then q has the form $< -dq(\mu), q(\mu), b_1, \cdots, b_{r-2} >$. This proves the lemma.

Let $L = K(\sqrt{d})$. Let w be a valuation of L, which extends a valuation v of K and let L_w and K_v be their respective completions. Then L_w contains $K_v(\sqrt{d})$. Since f is equivalent to $< 1, -a, -b, ab >$ over L and f is isotropic over K_v, f is isotropic over L_w. Therefore by our previous consideration, f is isotropic over L. By the above lemma, f is equivalent to $< \lambda, -\lambda d, \mu, \mu' >$. Therefore comparing discriminants, we have $d = abc = -\lambda^2 d\mu\mu'$, so that f is equivalent to $< \lambda, -\lambda d, \mu, -\mu >$. Hence obviously, f is isotropic, as it contains $< \mu, -\mu >$.

Let $r \geq 5$ and $f =< a_1, a_2, a_3, a_4, a_5, \cdots, a_r >$, $a_i \in K$, for $1 \leq i \leq r$. Let S be a finite set of valuations of K such that S contains the 2-adic valuations i.e., the valuations of K lying over the prime 2 of $\mathbb{Q}$, the archimedean valuations and such that a_3, a_4, a_5 are units in K_v for $v \notin S$. We can choose such a finite set, since for any element $a \in K$, $v(a) \leq 1$ for all but a finite number of valuations v of K and if $a \neq 0$, applying the above remark to a^{-1} too, we have that $v(a) = 1$ for all but a finite number of valuations v of K, i.e., a is a unit in K_v for all but a finite set of valuations v of K. By Proposition 5, it follows that the quadratic form $< a_3, a_4, a_5 >$ and hence $< a_3, a_4, a_5, \cdots, a_r >$ is isotropic over K_v for $v \notin S$. We now claim that for any $v \in S$, there exists a $\mu_v \in K_v^*$, which is a value of $< a_1, a_2 >$ and such that $-\mu_v$ is a value of $< a_3, a_4, a_5, \cdots, a_r >$. To prove this claim, we consider two cases.

Case 1 Suppose $< a_1, a_2 >$ is anisotropic over K_v. Since f is isotropic over K_v, there exists $(\lambda_1, \lambda_2, \cdots, \lambda_r) \in K_v^r$, such that $\sum_{1 \leq i \leq r} a_i \lambda_i^2 = 0$. We then choose μ_v to be $a_1\lambda_1^2 + a_2\lambda_2^2$, which obviously cannot be zero.

Case 2 Suppose $< a_1, a_2 >$ is isotropic over K_v. Then choose μ_v, to be any non-zero element such that, $-\mu_v$ is represented by $< a_3, a_4, \cdots, a_r >$. Since $< a_1, a_2 >$ is isotropic over K_v, it represents all elements of K_v, in particular μ_v. Let $\mu_v = a_1 x_v^2 + a_2 y_v^2$, for some $x_v, y_v \in K_v$.

By Lemma 1, there exist $x, y \in K$ such that x is close to x_v and y is close to y_v for $v \in S$, so that $\mu = a_1x^2 + a_2y^2$ is close to μ_v and in fact belongs to the same square class as μ_v for every $v \in S$. If v non-archimedean, this is guaranteed by Lemma 6. If $v \in S$ is archimedean and real, this simply means that μ and μ_v should have the same sign. If $v \in S$ is archimedean and complex, there is nothing to check since every element of $\mathbb{C}^*$ is a square. Thus the form $< a_1, a_2 >$ represents μ over K, so that $< a_1, a_2 >$ is equivalent to $< \lambda, \mu >$ for $\lambda \in K$ and μ and μ_v are in the same square class for $v \in S$. Since for $v \in S$, $< a_3, a_4, \cdots, a_r >$ represents $-\mu_v$, it also represents $-\mu$ over K_v (since these elements have the same square class). Thus the form $< \mu, a_3, a_4, \cdots, a_r >$ is isotropic over K_v for $v \in S$. For $v \notin S$, by Proposition 5, $< a_3, a_4, a_5 >$ and hence $< a_3, a_4, \cdots, a_r >$ is isotropic over K_v and a fortiori, $< \mu, a_3, a_4, \cdots, a_r >$ is isotropic over K_v, so that $< \mu, a_3, \cdots, a_r >$ is isotropic over K_v, for all v. By induction on r, it follows that $< \mu, a_3, \cdots, a_r >$ is isotropic over K, so that $< a_3, a_4, \cdots, a_r >$ represents $-\mu$ over K and $f =< a_1, a_2, a_3, \cdots, a_r >=< \lambda, \mu, -\mu, \cdots >$. Hence f contains $\mu < 1, -1 >$ and therefore it is isotropic, which proves the theorem.

R. Sridharan
School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India.
e-mail: sridhar@math.tifr.res.in

The Kronecker-Weber Theorem

Eknath Ghate

1. Introduction

These are some brief notes on the famous Kronecker-Weber theorem, which says that cyclotomic extensions of $\mathbb{Q}$ capture *all* abelian extension of $\mathbb{Q}$. Kronecker stated this theorem in 1853, but his proof was incomplete. Weber gave a proof in 1886, but apparently there was still a gap in it. Correct proofs were given soon after by Hilbert and Speiser.

In these notes we shall derive the theorem as a consequence of the theorems of (global) class field theory. The main reference we use is Janusz' book [1]. This is a good first introduction to class field theory - it derives most of the main theorems with minimal use of heavy machinery, and I recommend it to you for further study.

If time permits, I will give another proof of the Kronecker-Weber theorem: namely the one given in Chapter 14 of Washington's book [7]. In this approach, the theorem is deduced from the corresponding statement for local fields, which, in turn, is proved using only 'elementary' facts about the structure of local fields and their extensions. Since the exposition in Washington is good, I will not reproduce this proof in these notes. A word of warning though: one needs to be fairly well acquainted with local fields to enjoy Washington's proof. As an excellent background builder for this, and for many other things, I recommend reading Serre's book [2].

2. Cyclotomic extensions of $\mathbb{Q}$

Let us start by describing what cyclotomic fields, the objects of study of this summer school, look and smell like.

Let ζ_n denote a fixed primitive n^{th} root of unity, and let $\mathbb{Q}(\zeta_n)$ be the number field generated by all the n^{th} roots of unity. The field $\mathbb{Q}(\zeta_n)$ is called the n^{th} cyclotomic field. Most of you have probably already met these fields in the course of working out the proof of the following theorem, which I suggest you now (re-)try and prove for yourself as a warm-up exercise:

Theorem 1 *Let $\phi(n)$ denote the cardinality of $(\mathbb{Z}/n)^\times$. Then $\mathbb{Q}(\zeta_n)$ is an abelian extension of $\mathbb{Q}$ of degree $\phi(n)$. More precisely, there is an isomorphism:*

$$\begin{array}{rcl}(\mathbb{Z}/n)^\times & \xrightarrow{\sim} & \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \\ a\,(\mathrm{mod}\ n) & \mapsto & \sigma_a,\end{array}$$

where $\sigma_a(\zeta_n) = \zeta_n^a$.

Since a sub-extension of an abelian extension is also abelian, cyclotomic fields and their sub-fields already give us an abundant supply of abelian extensions of $\mathbb{Q}$. The obvious question that is now begging to be asked is whether or not there are any more. The answer is a resounding NO!!! More formally, we have the

Theorem 2 (Kronecker-Weber) *Every finite abelian extension of* $\mathbb{Q}$ *is contained in a cyclotomic field.*

The rest of these notes will sketch a proof of this beautiful fact using class field theory. But, before we embark on this, let us make a small diversion. At this stage, you may be wondering as to whether every finite abelian group actually occurs as a Galois group of some Galois extension of $\mathbb{Q}$. This is in fact true, and to give you an idea of how one might prove it, let us work out an example.

Let us construct a Galois extension of $\mathbb{Q}$ with Galois group $G = \mathbb{Z}/7 \times \mathbb{Z}/13 \times \mathbb{Z}/13$. The idea is to construct Galois extensions which realize each of the cyclic factors of G. If we can do this in such a way so that these extensions are linearly disjoint, then we have won the game, because we can then just take the compositum of these extensions. So let us first construct an extension with Galois group $\mathbb{Z}/7$. The trick is to choose a prime p such that $p \equiv 1 (\text{mod } 7)$. The first prime that works is $p = 29$. Now consider the extension $\mathbb{Q}(\zeta_{29})/\mathbb{Q}$. By Theorem **??**, it is an abelian extension with Galois group $\mathbb{Z}/28$. Clearly, since $7|28$, it has a sub-field K_1 whose Galois group is $\mathbb{Z}/7$.

So far so good. We now can similarly construct another extension $K_2/\mathbb{Q}$ with Galois group $\mathbb{Z}/13$. This time we note that $53 \equiv 1 (\text{mod } 13)$, that $\text{Gal}(\mathbb{Q}(\zeta_{53})/\mathbb{Q}) = \mathbb{Z}/52$, and that $13|52$. So the field K_2 with $\text{Gal}(K_2/\mathbb{Q}) = \mathbb{Z}/13$ exists.

Now we have only one more factor to worry about, the 'second' factor of $\mathbb{Z}/13$ in G. This time we choose a different prime congruent to 1 (mod 13). In fact 79 seems to work. As above, in $\mathbb{Q}(\zeta_{79})$ there is a sub-field K_3 with $\text{Gal}(K_3/\mathbb{Q}) = \mathbb{Z}/13$.

Now note that K_1, K_2 and K_3 are linearly disjoint, that is, the intersection of any two of these fields is $\mathbb{Q}$. This is because they each lie in cyclotomic fields $\mathbb{Q}(\zeta_p)$, for different p, which themselves are linearly disjoint. (You could try and use ramification theory to prove this - as a hint note that only p ramifies in $\mathbb{Q}(\zeta_p)$). We now choose $K = K_1K_2K_3$. Then some Galois theory shows $\text{Gal}(K/\mathbb{Q}) = G$, and we are done.

By now, you probably know what to do in general. So why not now try and prove the following theorem:

Theorem 3 *Every finite abelian group is the Galois group of some Galois extension of $\mathbb{Q}$.*

The following interesting fact may come in handy in the course of your proof:

Theorem 4 (Dirichlet) *There are infinitely many primes in every arithmetic progression.*

Now don't be asking whether all finite groups can be realized as Galois groups...! This is one of the hardest problems in mathematics, and is an active area of current research. Let us state it as

Question 1 (Inverse Galois Problem) *Is every finite group the Galois group of a finite Galois extension of $\mathbb{Q}$?*

3. Class field theory

Let us now give a short 'proof' of the Kronecker-Weber theorem using class field theory. That this theory should yield a proof at all is hardly surprising, because CLASS FIELD THEORY FOR $\mathbb{Q}$ = THE THEORY OF ABELIAN EXTENSIONS OF $\mathbb{Q}$. However, I should mention up front that class field theory is a rather broad subject, one that has undergone many re-formulations in terms of both the language and tools it has used to state and prove its main results. To do it justice would require the better part of a year of course work - a time frame somewhat beyond the scope of our five lectures!

Nonetheless, rather than despair, let us be brave, and try and at least get a flavour of some of the statements of the more important theorems of the theory. We shall derive the Kronecker-Weber theorem as an easy consequence of these theorems.

3.1 The Artin Map

The basic object around which most of the statements of class field theory revolve is the Artin map, which has already been introduced by Sury in his lectures. Let us recall its definition.

Let L/K be an abelian extension of number fields. Let I_K denote the group of fractional ideals of K. Let S denote a finite set of prime ideals of K, including all the primes that ramify in L, and let I_K^S denote the subgroup of I_K generated by all the prime ideals outside S. For each fractional ideal $\mathfrak{A}$ in I_K^S, write $\mathfrak{A} = \prod_{\mathfrak{p}} \mathfrak{p}^{a(\mathfrak{p})}$, and set

$$\phi_{L/K}(\mathfrak{A}) = \prod_{\mathfrak{p}} \left[\frac{L/K}{\mathfrak{p}}\right]^{a(\mathfrak{p})}.$$

Here $\left[\frac{L/K}{\mathfrak{p}}\right] \in \mathrm{Gal}(L/K)$ is the Frobenius element at $\mathfrak{p}$. That is, if $\mathfrak{P}$ is a prime of L lying over $\mathfrak{p}$, then $\left[\frac{L/K}{\mathfrak{p}}\right]$ is the element σ in $\mathrm{Gal}(L/K)$ characterized by the property

$$\sigma(x) \equiv x^{N(\mathfrak{p})} \pmod{\mathfrak{P}}, \text{ for all } x \text{ in } \mathcal{O}_L \ , \tag{1}$$

where $\mathcal{O}_L$ is the ring of integers of L, and $N(\mathfrak{p})$, the norm of $\mathfrak{p}$, is the cardinality of the residue field of $\mathfrak{p}$.

The homomorphism $\phi_{L/K} : I_K^S \to \mathrm{Gal}(L/K)$ is called the Artin map for the extension L/K. The first deep theorem about it is:

Theorem 5 *The Artin map $\phi_{L/K}$ is surjective.*

We shall not say anything about the proof of this theorem, except that one possible approach to it is, funnily enough, via analysis (L-Series and Density Theorems are catchwords here).

Another important theorem that we shall need, that can also be established by analytic methods, is the following:

Theorem 6 *Let L_1 and L_2 be two finite Galois (not necessarily abelian) extensions of K, and let S_1 and S_2 denote the sets of primes of K which split completely in L_1 and L_2 respectively. Then $S_1 \subset S_2$ (except for a set of density 0) if and only if $L_2 \subset L_1$.*

Again, we shall not define what it means for a set of primes to have density 0; suffice it to say that sets of finite cardinality have density 0, and it is only such exceptional sets that will appear in the application we have in mind below.

3.2 The kernel of $\phi_{L/K}$

One of the aims of class field theory is to describe the kernel of the Artin map explicitly. Note that a prime ideal $\mathfrak{p} \in \ker \phi_{L/K}$ if and only if the Frobenius element at $\mathfrak{p}$ is trivial, that is:

$$\left[\frac{L/K}{\mathfrak{p}}\right] = 1.$$

Since you know that the Frobenius element has order $f(\mathfrak{P}/\mathfrak{p})$, the residue degree of $\mathfrak{P}|\mathfrak{p}$, we see that, in this case, both $e(\mathfrak{P}/\mathfrak{p}) = f(\mathfrak{P}/\mathfrak{p}) = 1$. This forces $g(\mathfrak{P}/\mathfrak{p}) = [L : K]$, which is to say that $\mathfrak{p}$ splits completely in the extension L/K. Thus, apart from a finite set of primes, the primes in the kernel of the Artin map are exactly the primes that split completely.

Definition 1 *A* modulus *for K is a formal product*

$$\mathfrak{m} = \prod_{\mathfrak{p}} \mathfrak{p}^{n(\mathfrak{p})}$$

taken over all primes (including the infinite primes) of K. The exponents $n(\mathfrak{p})$ are non-negative integers, and are positive for only a finite number of $\mathfrak{p}$. Furthermore $n(\mathfrak{p}) = 0$ or 1 when $\mathfrak{p}$ is real, and $n(\mathfrak{p}) = 0$ when $\mathfrak{p}$ is complex.

A modulus $\mathfrak{m}$ may be written as $\mathfrak{m}_f\mathfrak{m}_\infty$, where the first (resp. second) factor is divisible only by the finite (resp. infinite) places. Now let $\mathcal{O}_K$ denote the ring of integers of K. Set

$$\begin{aligned} K_{\mathfrak{m}} &= \{a/b \mid a, b \in \mathcal{O}_K, (a), (b) \text{ relatively prime to } \mathfrak{m}_f\}, \\ K_{\mathfrak{m},1} &= \{\alpha \in K_{\mathfrak{m}} \mid \alpha \equiv 1 \pmod{\mathfrak{m}}\} \end{aligned}$$

Here the condition $\alpha \equiv 1 \pmod{\mathfrak{m}}$ means the following: for each finite $\mathfrak{p}$ dividing $\mathfrak{m}_f$, we require that $v_{\mathfrak{p}}(\alpha - 1) \geq n(\mathfrak{p})$, and for each real prime $\mathfrak{p}$ dividing $\mathfrak{m}_\infty$, we require α to be positive at this place.

Write $I_K^{\mathfrak{m}}$ for the group I_K^S where S is the set of primes dividing $\mathfrak{m}_f$. We assume that $\mathfrak{m}_f$ is divisible by all the finite primes that ramify in L. Note that, via the map $x \mapsto (x)$, $K_{\mathfrak{m},1}$ may be thought of as a subgroup of $I_K^{\mathfrak{m}}$.

Definition 2 *The quotient*

$$\frac{I_K^{\mathfrak{m}}}{K_{\mathfrak{m},1}}$$

is called the ray class group *modulo* $\mathfrak{m}$. *Note that when $\mathfrak{m} = 1$ this is just the usual* class group *of K.*

Each prime in K may also be viewed as a product of primes in L. In this way $\mathfrak{m}$ may also be considered as a modulus for L, and so it makes sense to speak of the group $I_L^{\mathfrak{m}}$. Moreover, there is a natural norm map

$$\begin{aligned} N_{L/K} : I_L^{\mathfrak{m}} &\to I_K^{\mathfrak{m}}, \\ \mathfrak{P} &\mapsto \mathfrak{p}^f, \end{aligned}$$

where $f = f(\mathfrak{P}/\mathfrak{p})$. The first approximation to the kernel of the Artin map is given by the following proposition:

Proposition 1 *Let L/K be a finite abelian extension, and let $\mathfrak{m}$ be any modulus of K such that $\mathfrak{m}_f$ is divisible by all the primes of K which ramify in L. Then $N_{L/K}(I_L^{\mathfrak{m}}) \subset \ker \phi_{L/K}$.*

Proof: This follows immediately from the fact that the Artin map maps $\mathfrak{p}^f$ to $\left[\frac{L/K}{\mathfrak{p}}\right]^f = 1$.

The following key theorem now tells us exactly what the 'missing part' of the kernel of the Artin map is.

Theorem 7 (Artin Reciprocity Theorem) *Let L/K be a finite abelian extension. Then there exists a modulus $\mathfrak{m}$ divisible by at least the primes of K which ramify in L such that the kernel of the Artin map is given by:*

$$\ker \phi_{L/K} = N_{L/K}(I_L^{\mathfrak{m}}) \cdot K_{\mathfrak{m},1}. \tag{2}$$

Let us say that the modulus $\mathfrak{m}$ *divides* $\mathfrak{m}'$ (and write $\mathfrak{m}|\mathfrak{m}'$) if each place that occurs in $\mathfrak{m}$ occurs in $\mathfrak{m}'$ with equal or larger exponent. It is a fact that if (**??**) holds for $\mathfrak{m}$ then it holds for all moduli $\mathfrak{m}'$ which are divisible by $\mathfrak{m}$. So the following definition is rather natural:

Definition 3 *The greatest common divisor of all the moduli $\mathfrak{m}$ such that* (**??**) *holds is called the* conductor *of L/K.*

3.3 An example

It is about time that we gave an example to illustrate the above concepts. What better place to start than with cyclotomic fields. So let us set $K = \mathbb{Q}$, and $L = \mathbb{Q}(\zeta_n)$, for some fixed n.

We will show that the Artin reciprocity theorem (Theorem **??**) holds for $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ with the modulus $\mathfrak{m} = n \cdot \infty$, where ∞ is the unique real infinite place of $\mathbb{Q}$. We must include ∞ in the modulus since ∞ ramifies in the totally imaginary field $\mathbb{Q}(\zeta_n)$.

Let p be a prime not dividing n. Then the ideal (p) is unramified in L. Moreover, σ_p (see Theorem **??** for notation) satisfies the condition (**??**) characterizing the Frobenius at p, so we see that $\phi_{\mathbb{Q}(\zeta_n)/\mathbb{Q}}(p) = \sigma_p$. This shows that for any two positive integers a and b relatively prime to n,

$$\phi_{\mathbb{Q}(\zeta_n)/\mathbb{Q}}(a/b) = \sigma_{ab^*}, \tag{3}$$

where b^* is a positive integer prime to n with $bb^* \equiv 1(\text{mod } n)$. This formula allows us to compute the kernel of the Artin map. Indeed, we may easily compute that

$$\ker \phi_{\mathbb{Q}(\zeta_n)/\mathbb{Q}} = \{(a/b) \mid a, b \text{ positive}, a \equiv b \ (\text{mod } n)\} = \mathbb{Q}_{\mathfrak{m},1}. \tag{4}$$

The surjectivity of the Artin map is of course something we are assuming (see Theorem **??**). But note that in this case the surjectivity is essentially equivalent to Dirichlet's theorem that there are infinitely many primes in every arithmetic progression (see Theorem **??**).

Putting things together, we see that the Artin map induces an isomorphism between the ray class group modulo $\mathfrak{m} = n \cdot \infty$, and $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q}) = (\mathbb{Z}/n)^\times$.

As it turns out $\mathfrak{m} = n \cdot \infty$ is in fact the greatest common divisor of all the moduli such that (**??**) above holds, and so it is also the conductor of the extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$.

3.4 The Kronecker-Weber theorem

We can now derive the Kronecker-Weber theorem. We need one last result:

Proposition 2 *Let L/K be an abelian extension and $\mathfrak{m}$ a modulus so that (**??**) above holds. Let E/K be an arbitrary Galois extension such that*

$$N_{E/K}(I_E^{\mathfrak{m}}) \subset N_{L/K}(I_L^{\mathfrak{m}}) \cdot K_{\mathfrak{m},1}. \tag{5}$$

Then $L \subset E$.

Proof: Say $\mathfrak{p}$ is a prime of K that does not divide $\mathfrak{m}_f$. Then, if $\mathfrak{p}$ splits completely in E, it is trivially the norm of a prime of E, and so $\mathfrak{p} \in N_{E/K}(I_E^{\mathfrak{m}})$. By (**??**) and Theorem **??** above, we see that $\mathfrak{p} \in \ker \phi_{L/K}$, so that $\mathfrak{p}$ splits completely in L. We now apply Theorem **??** to see that $L \subset E$.

Proof of Kronecker-Weber: Say that L is an arbitrary abelian extension of $\mathbb{Q}$. We want to show that $L \subset \mathbb{Q}(\zeta_n)$, for some n. Now Theorem **??**, applied to L, yields a modulus $\mathfrak{m}$ over $\mathbb{Q}$ such that (**??**) holds. We may suppose that $\mathfrak{m} = n \cdot \infty$, for some n. Now set $E = \mathbb{Q}(\zeta_n)$. Then, by (**??**), $\ker \phi_{E/\mathbb{Q}} = \mathbb{Q}_{\mathfrak{m},1}$ and so

$$N_{E/\mathbb{Q}}(I_E^{\mathfrak{m}}) \subset N_{E/\mathbb{Q}}(I_E^{\mathfrak{m}}) \cdot \mathbb{Q}_{\mathfrak{m},1} = \mathbb{Q}_{\mathfrak{m},1} \subset N_{L/\mathbb{Q}}(I_L^{\mathfrak{m}}) \cdot \mathbb{Q}_{\mathfrak{m},1} = \ker \phi_{L/\mathbb{Q}}.$$

By Proposition **??**, we obtain $L \subset E = \mathbb{Q}(\zeta_n)$ as desired.

3.5 Existence theorem

Though we have now 'proved' the Kronecker-Weber theorem, let us pick up some loose ends and round off our whirlwind survey of class field theory. We return to the general situation: L/K will denote an abelian extension of number fields.

Note that if $K \neq \mathbb{Q}$, then so far, no part of the discussion above guarantees the existence of even one abelian extension L of K! This is remedied by the following:

Theorem 8 (Existence Theorem) *Let $\mathfrak{m}$ be a modulus. Then there exists a finite abelian extension L/K, such that every prime of K that ramifies in L occurs in $\mathfrak{m}_f$, and such that* (**??**) *above holds.*

This modulus $\mathfrak{m}$ may not be the conductor of the extension L/K, but by definition, the conductor certainly divides it. Also, curiously, some moduli may never be conductors at all (example: $\mathfrak{m} = \infty$ is never the conductor of a finite abelian extension of $\mathbb{Q}$). However, once a modulus $\mathfrak{m}$ occurs as a conductor, it is a fact that there is a maximal finite abelian extension, $L_{\mathfrak{m}}$, having $\mathfrak{m}$ as its conductor. It turns out that in this case $\ker \phi_{L_{\mathfrak{m}}/K} = K_{\mathfrak{m},1}$.

Definition 4 *$L_{\mathfrak{m}}$ is called the* ray class field *of conductor $\mathfrak{m}$.*

Note that via the Artin map, the Galois group of the ray class field of conductor $\mathfrak{m}$ is just the ray class group modulo $\mathfrak{m}$. Also note that every abelian extension L of K sits inside a ray class field, namely the one whose conductor is the conductor of L. Finally, in the case when $K = \mathbb{Q}$ (see Section 3.3), we see that the cyclotomic fields are the ray class fields (of conductor $\mathfrak{m} = n \cdot \infty$).

3.6 Classification theorem

We now wish to state the climactic theorem of class field theory - the Classification theorem - which says roughly that the abelian extensions L of K are parameterized by gadgets constructed purely out of K! Let us make some preliminary definitions:

Definition 5 *A group H is said to be a* congruence subgroup of level $\mathfrak{m}$ *if it satisfies*

$$K_{\mathfrak{m},1} \subset H \subset I_K^{\mathfrak{m}},$$

for some modulus $\mathfrak{m}$.

The key example of a congruence subgroup of course is the following: if L/K is a finite abelian extension of K, then the Artin reciprocity theorem says that $H = \ker \phi_{L/K}$, is a congruence subgroup of level $\mathfrak{m}$ for some modulus $\mathfrak{m}$.

To rid us of the somewhat unpleasant dependence on the modulus $\mathfrak{m}$, we now put an equivalence relation $\sim$ on the set of congruence subgroups.

But first let us make a remark. Let $\mathfrak{m}$ and $\mathfrak{m}'$ be two moduli, with $\mathfrak{m}'|\mathfrak{m}$. Then $I_K^{\mathfrak{m}}$ is a subgroup of $I_K^{\mathfrak{m}'}$. If H' is a congruence subgroup of level $\mathfrak{m}'$ then there may or may not be a congruence subgroup H of level $\mathfrak{m}$ such that $H = I_K^{\mathfrak{m}} \cap H'$. If this does happen then we say that the congruence subgroup H is the restriction of the congruence subgroup H'.

Now say $(H_1, \mathfrak{m}_1)$ and $(H_2, \mathfrak{m}_2)$ are two congruence subgroups. We set $H_1 \sim H_2$, if there exists a modulus $\mathfrak{m}$, with $\mathfrak{m}_i \bigm| \mathfrak{m}$, for $i = 1, 2$, and so that $I_K^{\mathfrak{m}} \cap H_1 = I_K^{\mathfrak{m}} \cap H_2$ as restricted congruence subgroups of level $\mathfrak{m}$.

Definition 6 *An* ideal group $[H]$ *is an equivalence class of congruence subgroups* $(H, \mathfrak{m})$ *with respect to the equivalence relation* $\sim$.

The ideal groups are the 'gadgets' referred to above which parameterize abelian extensions of K. In fact we have:

Theorem 9 (Classification Theorem) *The map*

$$L/K \quad \mapsto \quad [\ker \phi_{L/K}]$$

is an inclusion reversing bijection between the set of abelian extensions L of K and the set of ideal groups of K.

Here 'inclusion reversing' means that if the abelian extensions L_1 and L_2 correspond to the ideal groups $[H_1]$ and $[H_2]$ respectively, then

$$L_1 \subset L_2 \iff [H_2] \subset [H_1].$$

(Note: $[H_2] \subset [H_1]$ simply means that there are congruence subgroups $H \in [H_2]$ and $H' \in [H_1]$ of the same level such that $H \subset H'$; one needs to check that this is well defined).

3.7 Hilbert class field

There is one particular ray class field that is the simplest and most important. This is the Hilbert class field. It is defined to be the ray class field of conductor $\mathfrak{m} = 1$. We shall denote it by U. The following theorem now follows easily, after the discussion above.

Theorem 10 *The Hilbert class field U is the maximal finite everywhere unramified abelian extension of K. Moreover, the Artin map establishes an isomorphism between the class group of K and* $\mathrm{Gal}(U/K)$. *In particular the class number of K is just* $[U : K]$.

As a consequence of this theorem we see that a prime $\mathfrak{p}$ of K splits completely in U if and only if it is a principal ideal of K. This is not to be confused with the following theorem, which was proved by Furtwangler.

Theorem 11 (Principal Ideal Theorem) *Every ideal of K becomes principal in U.*

Obviously, the Hilbert class field of $\mathbb{Q}$ is just $\mathbb{Q}$ itself, since $\mathbb{Q}$ has class number 1. But the above theorems show that the Hilbert class field for a number field with non-trivial class number is a very interesting object. We shall say a little more about the Hilbert class field of an imaginary quadratic situation in the next section.

3.8 Complex multiplication

We have seen that the ray class field of $\mathbb{Q}$ of conductor $\mathfrak{m} = n \cdot \infty$ is exactly the cyclotomic field $\mathbb{Q}(\zeta_n)$, and that every abelian extension of $\mathbb{Q}$ sits in one of these ray class fields. This is indeed a very satisfying result since we can generate *explicitly* all the abelian extension of $\mathbb{Q}$ by values of the exponential function $e^{2\pi iz}$ at certain division points $z \in \mathbb{Q}/\mathbb{Z}$.

A central problem in class field theory is to be able to similarly generate the abelian extensions of an arbitrary number field by values of transcendental functions. In fact this problem has its origins in Kronecker's famous 'Jugendtraum' (= youthful dream, in German).

When K is an imaginary quadratic field, this problem has been completely solved by the so called theory of 'complex multiplication'. Essentially the idea is that the ray class fields are generated by values of the famous j function at points in the imaginary quadratic field K, as well as by values of the Weber function w, at division points of an elliptic curve with complex multiplication by K. It would take us too far afield from the purpose of these notes to make this any more precise. However, to get our toes wet, let us at least describe how to generate the Hilbert class field of K.

For each $z \in \mathbb{C}$ with positive imaginary part, let $j(z)$ be the corresponding value of the elliptic modular function, defined by

$$j(z) = 1728 \cdot \frac{g_2(z)^3}{g_2(z)^3 - 27g_3(z)^2},$$

where

$$g_2(z) = 60 \sum_{(0,0)\neq(m,n)\in\mathbb{Z}^2} \frac{1}{(mz+n)^4}, \; g_3(z) = 140 \sum_{(0,0)\neq(m,n)\in\mathbb{Z}^2} \frac{1}{(mz+n)^6}.$$

Then we have the beautiful theorem:

Theorem 12 *Let Let K be an imaginary quadratic field, with ring of integers $\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\tau$. Then $j(\tau)$ is an algebraic integer, and $U = K(j(\tau))$ is the Hilbert class field of K.*

3.9 Hilbert's twelfth problem

The generation of ray class fields by the values of transcendental functions was emphasized by Hilbert in his 'twelfth problem' presented at the Paris International Congress of Mathematicians in 1900. He wrote

> "The extension of Kronecker's theorem to the case that in place of the realm of rational numbers or of the imaginary quadratic field any algebraic field whatever is laid down as realm of rationality, seems to me of the greatest importance. I regard this problem as one of the most profound and far reaching in the theory of numbers and of functions."

So far very little progress has been made on the problem in general. However, in closing these notes, let us at least mention some additional special cases that have been partially treated:

3.9.1 CM fields

From the point of view of 'complex multiplication' the most natural way to to generalize the results obtained for imaginary quadratic fields is to replace elliptic curves by higher dimensional abelian varieties. This was done by Shimura and Taniyama, who managed to generate class fields of 'CM fields'. A CM field is the higher analog of an imaginary quadratic field: it is a totally imaginary quadratic extensions of a totally real field. It must be pointed out that unfortunately not all abelian extension of CM fields can be generated by this method.

Shimura and Taniyama's theory is exposed in their book [5]. There is also a new edition, [4], now on the market.[1]

See also Wafa Wei's (unpublished) thesis, where she gives some information about the maximal abelian extension of a CM field that can be generated by the values of automorphic functions [8].

3.9.2 Real quadratic fields

Some partial results have been obtained by Shimura in this case using abelian varieties with real multiplication. For more details see the last chapter of his book [3].

[1] In the preface to [5] it was claimed that Hecke had shown how to generate class fields of certain CM bi-quadratic extensions of $\mathbb{Q}$ by values of Hilbert modular functions. Apparently this work of Hecke was not complete (see the preface to [4]), but in any case, has since been corrected and subsumed by the work [4].

3.9.3 Stark's method

Another approach to Hilbert's twelfth problem has been proposed by Stark, who has shown that certain abelian extensions of arbitrary number fields can be generated by the values of Artin L-functions at $s = 0$. You could look at Tate's efficient monograph [6] for more details.

REFERENCES

1. G. Janusz, *Algebraic number fields*, Academic Press, New York-London, 1973.

2. J.-P. Serre, *Local fields*, GTM **67**, Springer-Verlag, Berlin-New York, 1979.

3. G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton Univ. Press, Princeton, 1971.

4. G. Shimura, *Abelian varieties with complex multiplication and modular functions*, Princeton Mathematical Series, **46**, Princeton Univ. Press, Princeton, 1999.

5. G. Shimura and Y. Taniyama, *Complex multiplication of abelian varieties and its applications to number theory*, Publications of the Mathematical Society of Japan, **6**, The Mathematical Society of Japan, Tokyo, 1961.

6. J. Tate, *Les conjectures de Stark sur les fonctions L d'Artin en* $s = 0$, Birkhäuser, 1984.

7. L. Washington, *Introduction to cyclotomic fields, Second edition*, Springer - Verlag, Berlin-New York, 1996.

8. W. Wei, *Moduli fields of CM-motives applied to Hilbert's 12th problem*, Preprint, Available on the world wide web at http://www.mathematik.uni-bielefeld.de/sfb343/preprints/pr94070.ps.gz, 1994.

Eknath Ghate
School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India.
e-mail: eghate@math.tifr.res.in

Kronecker-Weber via Ramification Theory

Sharad V. Kanetkar

In this note we prove the well known theorem of Kronecker-Weber using only ramification theory. The following steps are described in a series of exercises in [1, pp. 125-127].

Kronecker-Weber Theorem.

Theorem : Every finite abelian extension of $\mathbf{Q}$ (field of rational numbers) is contained in a cyclotomic field.
Proof : Let K be a finite abelian extension of $\mathbf{Q}$ with $G = \mathrm{Gal}(K/\mathbf{Q})$.

Step 1 : It is enough to assume that K is of degree p^m over $\mathbf{Q}$ for some prime p. For if G is expressed as a direct product of its Sylow subgroups :

$$G \cong S_{p_1} \times \cdots \times S_{p_r},$$

then fixed subfields k_i (of K) with groups S_{p_i} will generate K. If k_i belongs to a cyclotomic field F_i, for $i = 1, 2, \cdots, r$; then $K \subseteq F_1 F_2 \cdots F_r \subseteq$ some cyclotomic field. Hence we assume $K = k_1$ and $[K : \mathbf{Q}] = p^m$.
Step 2 : It is enough to assume that p is the only prime ramified in K. Suppose $q \in \mathbf{Z}$ is a prime (other that p) which is ramified in K. Let $E(\cdot|\cdot)$ and $e(\cdot|\cdot)$ denote the inertia group and the ramification index respectively. Let U be a prime of K lying above q with $e(U \mid q) = e$. Now the higher ramification group $V_1(U \mid q)$ is a $q-$subgroup of a $p-$group G [1, page 121]. Hence $|V_1(U \mid q)| = 1$ and $|V_0/V_1| = e$. Since G is abelian $|V_0/V_1| \mid (q-1)$ [1, page 124, Ex. 26(c)]. This gives $e|(q-1)$. Now there is a (unique) subfield $K_1 \subseteq \mathbf{Q}(\zeta_q)$ (where $\mathbf{Q}(\zeta_m)$ denotes the m-th cyclotomic field, i.e. ζ_m is a primitive m-th root of unity) with $[K_1 : \mathbf{Q}] = e$. Since $e \mid p^m$ and $q \neq p$, q is tamely ramified in both K_1 and K. Now q is totally ramified in $\mathbf{Q}(\zeta_q)$ and hence in K_1. This gives that the ramification index of q in K_1 is also e. Let U_1 be a prime of L lying above U in K. Now, $\mathrm{Gal}(K/\mathbf{Q})$ and $\mathrm{Gal}(K_1/\mathbf{Q})$ are both p-groups and since $\mathrm{Gal}(L/\mathbf{Q})$ injects into $\mathrm{Gal}(K/\mathbf{Q})\times$ $\mathrm{Gal}(K_1/\mathbf{Q})$, it is also a p-group. This shows that $V_1(U_1 \mid q)$ is both a p-group and a q-group implying that it is trivial. Thus $E(U_1 \mid q)$ is cyclic. Let W be the (unique) prime of K_1 lying below U_1. Hence, by restriction, $E(U_1 \mid q)$ injects into $E(U \mid q) \times E(W \mid q)$. All these three groups are cyclic and the last two have

order e each. This shows that $E(U_1 \mid q)$ is of order e. Thus the ramification index of q in L is also e. Since $e(U_1 \mid q) = e(U \mid q) = e, e(U_1 \mid U) = 1$. Let L_1 be the inertia field of U_1, i.e., L_1 is the fixed field of $E(U_1 \mid q)$. Then for any field F containing $L_1, U_1 \cap F$ is totally ramified in L. Thus for $F = L_1K_1, (U_1 \cap F)$ is totally ramified in L. But $F \supset K_1$ and therefore $e(U_1 \mid (U_1 \cap F)) \mid e(U_1 \mid U)$. This implies $e(U_1 \mid U_1 \cap F) = 1$. Thus $U_1 \cap F$ is totally ramified as well as unramified in L implying $F = L$. Hence if L_1 belongs to a cyclotomic field then since $K_1 \subset \mathbf{Q}(\zeta_q)$, L will be a subfield of a cyclotomic field. But $K \subset L$ and hence K will be a subfield of some cyclotomic field proving the theorem. Thus it is enough to replace K by L_1. But it is easy to see that all unramified primes of K are unramified in L_1 and, in addition, q is also unramified in L_1 (but ramified in K). Thus continuing this process of reduction we can assume that there are no primes other that p which are ramified in K. This finishes the proof of step 2.

Step 3 : **Case(i)** $p = 2$, $[K : \mathbf{Q}] = 2^m$.
In this case 2 is totally ramified in K since otherwise no prime will be ramified in the fixed field of $E(U \mid 2)$ and this will imply, by [1,page 137, Cor.3], that $E(U \mid 2) = G$. Thus 2 is totally ramified in K. Thus $e(U \mid 2) = 2^m$. If $m = 1$ then $[K : \mathbf{Q}] = 2$ and $K = \mathbf{Q}[\sqrt{d}]$ for some square-free integer d. But the $\text{Disc}(K/\mathbf{Q}) = d$ *or* $4d$. Since 2 is the only ramified prime of K, 2 is the only possible divisor of d. Hence

$$K = \mathbf{Q}[\sqrt{2}] \text{ or } \mathbf{Q}[\sqrt{-2}] \text{ or } \mathbf{Q}[\sqrt{1}].$$

All these fields are subfields of $\mathbf{Q}[\zeta_8]$. Hence the theorem is proved in this case. If $m > 1$ then consider $L = \mathbf{Q}(\zeta_{2^{m+2}}) \cap \mathbf{R}$, where $\mathbf{R}$ is the field of real numbers. Then $[L : \mathbf{Q}] = 2^m$ and $L \subset \mathbf{R}$. Hence L contains a unique quadratic subfield, namely $\mathbf{Q}[\sqrt{2}]$. Hence $\text{Gal}(L/\mathbf{Q})$ contains unique subgroup of index 2. Thus L is a cyclic extension. Now consider the field LK. Let μ be the extension of σ (where $< \sigma >= \text{Gal}(L \mid \mathbf{Q})$ to LK. Let F be the fixed field of μ. Since μ restricted to L generates $\text{Gal}(L/\mathbf{Q}), F \cap L = \mathbf{Q}$. If $[F : \mathbf{Q}] > 2$ then $F \cap \mathbf{R} \neq \mathbf{Q}$ and it will contain $\mathbf{Q}[\sqrt{2}] \subset L$ but $F \cap L = \mathbf{Q}$. Hence $[F : \mathbf{Q}] \leq 2$. If $[F : \mathbf{Q}] = 2$ then $F = \mathbf{Q}[\sqrt{-2}]$ or $\mathbf{Q}[i]$ and both are contained in $\mathbf{Q}[\zeta_8]$. Thus $K \subseteq LK = FL \subseteq \mathbf{Q}(\zeta_{2^{m+2}})$ and the theorem is proved. If $F = \mathbf{Q}$ then $< \mu >= \text{Gal}(LK/\mathbf{Q})$ and since

$$\text{Gal}(LK/\mathbf{Q}) \hookrightarrow \text{Gal}(L/\mathbf{Q}) \times \text{Gal}(K/\mathbf{Q}),$$

order of any element of $\text{Gal}(LK/\mathbf{Q}) \leq \text{lcm}\ (\mid \text{Gal}\ (L/\mathbf{Q}) \mid, \mid \text{Gal}\ (K/\mathbf{Q}) \mid) = 2^m$. Thus $2^m \leq [LK : \mathbf{Q}] \leq 2^m$. Hence $L = LK$ implying $K \subseteq L \subseteq \mathbf{Q}[\zeta_{2^{m+2}}]$. Thus the theorem is proved in this case also.

Case(ii) p is odd and $[K:\mathbf{Q}] = p^m$.
Consider the case $m = 1$. Hence K is of degree p over $\mathbf{Q}$ and p is the only ramified prime in K. Thus if U is the prime of K lying above p then $e(U \mid p) = p$.
Claim : $\text{diff}(R/Z) = U^{2(p-2)}$, where R is the ring of integers of K.
Proof : Let $\pi \in U - U^2$ then π satisfies a monic irreducible polynomial over $\mathbf{Z}$, say,

$$f(x) = x^p + a_{p-1}x^{p-1} + \cdots + a_0.$$

Let ϑ_U be the valuation corresponding to the DVR R_U. Then $\vartheta_U(\pi) = 1$ and since $U^p = pR$, $\vartheta_U(p) = p$. Now the coefficients a_i are symmetric polynomials in $\sigma\pi$, $\sigma \in \text{Gal}(K/\mathbf{Q})$ and $\vartheta_U(\sigma\pi) = 1$, $\forall \sigma \in \text{Gal}(K/\mathbf{Q})$. Hence $\vartheta_U(a_i) \geq 1$ and hence $p \mid a_i$. But $a_0 = \pm \prod(\sigma\pi)$ and hence $\vartheta_U(a_0) = p$. Now in the expression

$$f'(\pi) = p\pi^{p-1} + (p-1)a_{p-1}\pi^{p-2} + \cdots + a_1,$$

all terms have valuations distinct *mod p*. Therefore

$$\begin{aligned} \vartheta_U(f'(\pi)) &= \min\{\vartheta_U(p\pi^{p-1}), \vartheta_U((p-1)a_{p-1}\pi^{p-2} \cdots \vartheta_U(a_1)\}. \\ \text{Hence}, 2p-1 &\geq \vartheta_U(f'(\pi)) \geq p. \end{aligned}$$

But by Hilbert's formula [1, page 124, Exc. 27],

$$\vartheta_U(f'(\pi)) = \vartheta_U(\text{diff}(R/Z)) = \sum_{i=0}^{\infty}(\mid V_i \mid -1)$$

Since $\mid V_i \mid$ is a power of p, $(p-1) \mid \vartheta_U(f'(\pi))$. Hence $\vartheta_U(f'(\pi)) = 2p-2$. And $\text{diff}(R/\mathbf{Z}) = U^{2p-2}$ (because no other prime is ramified in k). Thus the claim is proved.
Now let $m = 2$.
Claim : G is cyclic.
Proof : Consider the inertia field corresponding to the prime p. In this field p is unramified. Hence no prime is ramified in this inertia field. Hence it must be equal to $\mathbf{Q}$. Thus K is totally ramified with $e(U/p) = p^2$. Since V_1 is Sylow$-p$ subgroup of $\text{Gal}(K/\mathbf{Q})$, $\mid V_1 \mid = p^2 = \mid V_0 \mid$. Let $V_r = V_r(U/p)$ be the least r for which $\mid V_r \mid < p^2$. But $V_{r-1}/V_r \hookrightarrow R/U \cong \mathbf{Z}/p\mathbf{Z}$ and hence $\mid V_r \mid = p$. Let H be any subgroup of G having order p. Let K_H be the fixed field of H. Then $[K_H : K] = p$ and $\text{diff}(R_H/\mathbf{Q}]) = U^{2p-2}$. Hence from the transitivity of different,

$$\text{diff} R/\mathbf{Z}) = \text{diff}(R/R_H).U^{(2p-2)p}, \quad [1, \text{page 96, Ex.38}].$$

Hence $\mathrm{diff}(R/R_H)$ is independent of H as long as $[H : \mathbf{Q}] = p$. Now by Hilbert's formula the power of U dividing $\mathrm{diff}(R/R_H)$ is given by

$$\alpha = \sum_{i=0}^{\infty} \mid V_i \cap H \mid -1.$$

Hence α is strictly maximized when $H = V_r$. Since α is independent of H, V_r is the only subgroup of order p in G. Thus G is cyclic, proving the claim. Thus in case $m = 1$, k is unique, otherwise KK_1 will be of degree p^2 containing two distinct subfields of degree p. Hence K is the unique subfield of $\mathbf{Q}[\zeta_{p^2}]$. Thus the theorem is true for the case $m = 1$.

Now let $m > 1$. Let L denote the unique subfield of $Q[\zeta_{p^{m+1}}]$ of degree p^m over $\mathbf{Q}$. Then $\mathrm{Gal}(L/\mathbf{Q})$ is cyclic of order p^m. Then LK is cyclic by the claim. But

$$\mathrm{Gal}(LK/\mathbf{Q}) \hookrightarrow \mathrm{Gal}(L/\mathbf{Q}) \times \mathrm{Gal}(k/\mathbf{Q}),$$

hence,

$$\begin{aligned} \mid \mathrm{Gal}(LK/\mathbf{Q}) \mid \; &\leq \; lcm(\mid \mathrm{Gal}(L/\mathbf{Q}) \mid, \mid \mathrm{Gal}(K/\mathbf{Q}) \mid) \\ &= \; p^m. \end{aligned}$$

Therefore $L \subseteq LK \subseteq L$ and hence $K \subseteq L \subseteq \mathbf{Q}(\zeta_{p^{m+1}})$, and the theorem is proved in this case also.

REFERENCE

1. D. A. Marcus, *Number Fields*, Springer Verlag, 1977.

Sharad V. Kanetkar
Bhaskaracharya Pratishthana
56/14, Erandavane, Damle Path
Off Law College Road
Pune-411 004
e-mail : bhaskara_p@vsnl.com

Stickelberger Revisited

C S Yogananda

In the present article, we give the proof of the Stickelberger's Theorem in the spirit of Kummer (rediscovered by Thaine) given by Washington [4].

Theorem. *Let $G = Gal(\mathbf{Q}(\zeta_m)/\mathbf{Q})$ and for $(a,m) = 1, \sigma_a : \zeta_m \mapsto \zeta_m^a$. Let*

$$\theta = \frac{1}{m} \sum_{\substack{a=1 \\ (a,m)=1}} a\sigma_a^{-1}.$$

If $\beta \in \mathbf{Z}[G]$ is such that $\beta\theta \in \mathbf{Z}[G]$ then $\beta\theta$ annihilates the ideal class group of $\mathbf{Q}(\zeta_m)$.

Proof: The proof is by looking at the factorisation of certain Gauss sums. Let $\mathcal{C}$ be an ideal class in $\mathbf{Q}(\zeta_m)$. There exist infinitely many unramified primes of degree 1 in $\mathcal{C}$ (follows from Dirichlet's theorem on primes in arithmetic progression and Chebotarev density theorem). Let λ be such an ideal which is above the rational prime l; since l splits completely in $\mathbf{Q}(\zeta_m)$ we have that $l \equiv 1 \pmod{m}$. Fix a primitive root s modulo l and define a Dirichlet character mod l, $\chi : (\mathbf{Z}/l\mathbf{Z})^* \to \mathbf{C}^*$ by $\chi(s) = \zeta_m$. Consider the Gauss sum

$$g(\chi) = -\sum_{b=1}^{l-1} \chi(b)\zeta_l^b.$$

It is easy to see which are the primes dividing $g(\chi)$ in $\mathbf{Q}(\zeta_m, \zeta_l)$. First of all, since $g(\chi)g(\bar{\chi}) = l$, only primes above l occur in the factorisation of $g(\chi)$. Since l splits completely in $\mathbf{Q}(\zeta_m)$, the Galois conjugates of λ, $\sigma_a^{-1}(\lambda)$, $1 \leq a \leq m, (a,m) = 1$, are all the factors of l; if $\mathfrak{L}$ is the prime above λ in $\mathbf{Q}(\zeta_m, \zeta_l)$, (remember $\mathbf{Q}(\zeta_m, \zeta_l)/\mathbf{Q}(\zeta_m)$ is fully ramified above l) the prime factorisation of the principal ideal $(g(\chi))$ would look like

$$(g(\chi)) = \prod_{\substack{a=1 \\ (a,m)=1}}^{m} \sigma_a^{-1}(\mathfrak{L})^{r_a}$$

where $0 \leq r_a \leq l-1$ (σ_a being extended to $\mathbf{Q}(\zeta_m, \zeta_l)$.

As $g(\chi)^{l-1}$ is in $\mathbf{Q}(\zeta_m)$ and $\mathfrak{L}^{l-1} = \lambda$ we have the following factorisation in $\mathbf{Q}(\zeta_m)$:

$$(g(\chi)^{l-1}) = \prod_{\substack{a=1 \\ (a,m)=1}}^{m} \sigma_a^{-1}(\lambda)^{r_a}.$$

In other words, $\sum r_a \sigma_a^{-1}$ *annihilates the class* $\mathcal{C}$ in the ideal class group of $\mathbf{Q}(\zeta_m)$.

We now use the Galois action on $g(\chi)$ to determine the integers r_a. Consider

$$\tau \in \mathrm{Gal}(\mathbf{Q}(\zeta_m, \zeta_l)/\mathbf{Q}(\zeta_m)), \quad \tau(\zeta_l) = \zeta_l^s.$$

We have $g(\chi)^\tau = \chi(s)^{-1} g(\chi)$ and

$$(\zeta_l^s - 1)/(\zeta_l - 1) \equiv 1 + \zeta_l + \cdots + \zeta_l^{s-1} \equiv s \pmod{\sigma_a^{-1}(\mathfrak{L})}.$$

As $\mathbf{Q}(\zeta_m, \zeta_l)/\mathbf{Q}(\zeta_m)$ is totally ramified above $\sigma_a^{-1}(\lambda)$, $1 \leq a \leq m, (a, m) = 1$, the inertia group of $\sigma_a^{-1}(\mathfrak{L})$ coincides with the full Galois group and hence τ acts trivially mod $\sigma_a^{-1}(\mathfrak{L})$. Therefore we have

$$\frac{g(\chi)}{(\zeta_l - 1)^{r_a}} \equiv \frac{g(\chi)^\tau}{(\zeta_l^s - 1)^{r_a}} \equiv \frac{g(\chi)}{(\zeta_l - 1)^{r_a}} \frac{\chi(s)^{-1}}{s^{r_a}} \pmod{\sigma_a^{-1}(\mathfrak{L})}.$$

Since $\sigma_a^{-1}(\mathfrak{L})$ occurs to first power in $(\zeta_l - 1)$ we have $g(\chi)/(\zeta_l - 1)^{r_a}$ relatively prime to $\sigma_a^{-1}(\mathfrak{L})$ and hence we get

$$\zeta_m = \chi(s) \equiv s^{-r_a} \pmod{\sigma_a^{-1}(\mathfrak{L})}.$$

Since $\zeta_m \in \mathbf{Q}(\zeta_m)$ this congruence holds modulo $\sigma_a^{-1}\lambda$ and applying σ_a we obtain

$$\zeta_m^a \equiv s^{-r_a} \pmod{\lambda}.$$

Now, since the mth roots of unity are distinct mod λ, the order of ζ_m mod λ is exactly m and so,

$$\zeta_m \equiv s^{-(l-1)c/m} \pmod{\lambda}$$

where c is an integer prime to m.

Therefore

$$r_a \equiv \frac{(l-1)ac}{m} \pmod{l-1}$$

which implies that $l - 1$ does not divide r_a as a and c are prime to m. Furthermore, we have $0 \leq r_a \leq l - 1$ and so

$$r_a = (l-1)\left\{\frac{ac}{m}\right\}$$

where $\{\cdot\}$ denotes the fractional part. We have

$$\sum_{(a,m)=1} (l-1)\left\{\frac{ac}{m}\right\} \sigma_a^{-1} = (l-1)\sigma_c \theta.$$

Thus we get that the ideal $\lambda^{(l-1)\sigma_c\theta}$ is a principal ideal generated by $g(\chi)^{l-1}$.

Let $\beta \in \mathbf{Z}[G]$ be such that $\beta\theta \in \mathbf{Z}[G]$ and $\gamma = g(\chi)^{\sigma_c^{-1}\beta}$. Then $\gamma^{l-1} \in \mathbf{Q}(\zeta_m)$ and $\lambda^{\beta\theta(l-1)} = (\gamma^{l-1})$ which implies that (γ^{l-1}) is the $(l-1)$st power of an ideal in $\mathbf{Q}(\zeta_m)$. Hence the extension $\mathbf{Q}(\zeta_m, \gamma)/\mathbf{Q}(\zeta_m)$ can be ramified only at the primes dividing $l-1$. But since $\mathbf{Q}(\zeta_m) \subseteq \mathbf{Q}(\zeta_m, \gamma) \subseteq \mathbf{Q}(\zeta_m, \zeta_l)$ it follows that $\mathbf{Q}(\zeta_m, \gamma)/\mathbf{Q}(\zeta_m)$ is totally ramified at primes above l. Therefore $\mathbf{Q}(\zeta_m, \gamma) = \mathbf{Q}(\zeta_m)$ which implies that $\gamma \in \mathbf{Q}(\zeta_m)$. We can therefore take $(l-1)$st root and obtain the following equality of ideals in $\mathbf{Q}(\zeta_m)$

$$\lambda^{\beta\theta} = (\gamma).$$

This completes the proof of the theorem. □

Actually, as we have seen in Katre's article [1], the Stickelberger's theorem holds for subfields of $\mathbf{Q}(\zeta_m)$ as well. (It would be interesting to see if the above methods can be used to give a proof of this general case.)

Thaine's Theorem

We shall state a simpler case of Thaine's theorem and give an outline of the proof to illustrate the main ideas. This account is based on Thaine's paper [3] and Washington's book [4], Chapter 15, §2.

Theorem. *Let $F = \mathbf{Q}(\zeta_p)^+$ and $\Delta = Gal(F/\mathbf{Q})$. Let E be the group of units of F, $C = C_F$, the group of cyclotomic units, A, the class group of F; put $B = E/C$. If $\theta \in \mathbf{Z}[G]$ annihilates the p-part of B then θ annihilates the p-part of A.*

An outline of the proof: Choose n large enough such that $p^n > |A|$ and $p^n > |B|$. Then the p-Sylow subgroups of A and B are, respectively, isomorphic to A/A^{p^n} and $E/E^{p^n}C$. Let $q \equiv 1 \pmod{p^n}$ be a prime; note that q splits completely in $\mathbf{Q}(\zeta_p)$. Suppose $\delta = \prod_b (\zeta_p^b - 1)^{y_b}$ is a cyclotomic unit in $\mathbf{Q}(\zeta_p)$. Let Q denote a prime above q in $\mathbf{Q}(\zeta_p)$ and $\tilde{Q}$ the unique prime in $\mathbf{Q}(\zeta_{pq})$ above Q. Put $\eta = \prod_b (\zeta_p^b\zeta_q - 1)^{y_b}$. It turns out that η is a unit in $\mathbf{Q}(\zeta_{pq})$ with some special properties: (i) η has norm 1 to $\mathbf{Q}(\zeta_p)$ (since $q \equiv 1 \pmod{p}$) and (ii) $\eta \equiv \delta$ mod primes above q. Property (1) in conjunction with Hilbert's Theorem 90 implies the existence of an $\alpha \in \mathbf{Q}(\zeta_{pq})$ such that $\eta = \alpha^\tau/\alpha$ where τ is a generator of $\mathrm{Gal}(\mathbf{Q}(\zeta_{pq})/\mathbf{Q}(\zeta_p))$. Since η is a unit the principal ideal (α) satisfies: $(\alpha)^\tau = (\alpha)$ which implies (since τ is a generator of the Galois group) that (α) is the product of an ideal I from $\mathbf{Q}(\zeta_p)$ and ramified primes:

$$(\alpha) = I \cdot \prod_\sigma \sigma(\tilde{Q})^{r_\sigma}$$

where the product is over $\sigma \in \mathrm{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q})$. Taking norm to F we get

$$(\text{Norm } \alpha) = (I\bar{I})^{q-1} \cdot \prod_{\sigma} \sigma(Q\bar{Q})^{r_\sigma} = (I\bar{I})^{q-1} \cdot (Q\bar{Q})^{\sum \sigma r_\sigma}.$$

Thus we have that $\sum \sigma r_\sigma$ annihilates the idealclass above q in the quotient A/A^{p^n}. If s is a primitive root mod q working exactly as in section 2 we get that $s^{r_\sigma} \equiv \epsilon \equiv \delta \pmod{\sigma^{-1}\tilde{Q}}$. Since s^{r_σ} and δ are in F this congruence gives us: $s^{r_\sigma} \equiv \sigma(\delta) \pmod{Q\bar{Q}}$ which determines r_σ modulo $q-1$ and hence modulo p^n. A careful choice of the unit we start with, δ, gives us the necessary information on r_σ to conclude.

This is the essential idea behind Thaine's proof.

Generalisation to imaginary quadratic fields: There had been generalisations of Stickelberger's theorem to the case of totally real fields (see the notes at the end of the Chapter 6 in [4]) and the Thaine's theorem can also be deduced from the results of Mazur-Wiles. But the point about Thaine's work was its simplicity and adaptability to other situations. Most famously, Rubin [2] was able to obtain a generalisation to the case of abelian extensions of imaginary quadratic fields which he later used to obtain, for the first time, examples of finite Shafarevich-Tate groups of elliptic curves.

REFERENCES

1. S. A. Katre, Gauss-Jacobi sums and Stickelberger's theorem, These proceedings.

2. K. Rubin, Global units and ideal class groups, *Invent. Math.* 89 (1987), 511–526.

3. F. Thaine, On the ideal class groups of real abelian number fields, *Ann. of Math.* 128 (1988), 1–18.

4. L. C. Washington, *Introduction to Cyclotomic Fields*, GTM, Springer-Verlag.

C. S. Yogananda
MO-Cell (DAE), Dept. of Maths, IISc.
Bangalore - 560012.
e-mail: yoga@math.iisc.ernet.in

Index of the Stickelberger Ideal and a Reflection Theorem

R. Sujatha

These are the notes of two talks given at the Instructional Conference, on the index of the Stickelberger ideal and a Reflection theorem. We mainly follow the proofs given by Washington in his book "Cyclotomic fields".

Index of the Stickelberger ideal

Notation: p an odd prime, $n \geq 1$ an integer.
$G = \mathrm{Gal}(\mathbb{Q}(\zeta_{p^n})/\mathbb{Q})$
$R = \mathbb{Z}[G]$, the group ring.
θ=Stickelberger element; recall that it is defined by

$$\theta = 1/p^n \Sigma' a \sigma_a^{-1} \ \in \ \mathbb{Q}[G],$$

where Σ' indicates summation over all integers a from 1 to p^n such that $(a, p) = 1$.
$I := R\theta \cap R$ is the Stickelberger ideal.
$J = \sigma_{-1}$ is the complex conjugation.
$R^- := \{x \in R \mid Jx = -x\}$
$I^- := I \cap R^- = R\theta \cap R^-$.
h=class number of $\mathbb{Q}(\zeta_{p^n})/\mathbb{Q}$.
$\mathbb{Q}(\zeta_{p^n})^+$ = Maximal totally real subfield of $\mathbb{Q}(\zeta_{p^n})$.
h^+ = class number of $\mathbb{Q}(\zeta_{p^n})^+$.
Recall that $h^+ \mid h$.

Definition. The *relative class number* h^- is the quotient h/h^+.

The main theorem that we want to prove, which is due to Iwasawa, is the following:

Theorem I. $[R^- : I^-] = h^-$.

We will first establish a few useful results that will be needed.

Lemma 1. $R^- = (1-J)R$.

Proof. The inclusion $\supseteq$ is obvious. To prove the other inclusion, consider an element $\alpha = \Sigma z_a \sigma_a^{-1} \in R^-,\ z_a \in \mathbb{Z}$. Then $J\alpha = \sigma_{-1}\alpha = -\alpha$, hence we get $z_{-a} = -z_a$. Now suppose that P is a set of representatives in $(\mathbb{Z}/m\mathbb{Z})^*$ for $(\mathbb{Z}/m\mathbb{Z})^*/(\pm 1)$. Then $\alpha = (1-J)\beta$, where $\beta = \sum\limits_{c\in P} z_c \sigma_c^{-1}$. □

We will need some local equalities, as the proof of the theorem will proceed prime by prime, by working over completions. Let q be a prime, $R_q = \mathbb{Z}_q[G]$, i.e. $R_q = \mathbb{Z}[G] \otimes \mathbb{Z}_q$ and $I_q = R_q I$. Then I is dense in I_q and $R_q^- = (1-J)R_q$, with $I_q^- = I_q \cap R_q^-$.

Lemma 2. *We have*
(a) $I_q = R_q\theta \cap R_q$
(b) $I_q^- = R_q\theta \cap R_q^-$
(c) $I_q^- = I^-.\mathbb{Z}_q$
(d) If $p \neq q$, *then* $I_q = R_q\theta$.

Proof. We will need the following fact: Suppose I' is the ideal of $\mathbb{Z}[G]$ generated by elements of the form $(c - \sigma_c)$, with $(c,p) = 1$. Let $\beta \in \mathbb{Z}[G]$; then $\beta\theta \in \mathbb{Z}[G]$ if and only if $\beta \in I'$. Now $I'_q = I'\mathbb{Z}_q$, and we have

$$R_q\theta \cap R_q = I'_q\theta = I'\mathbb{Z}_q\theta = I'\theta\mathbb{Z}_q = I\mathbb{Z}_q = I_q$$

and (a) follows.

Clearly (b) follows from (a). Also, (d) follows from (a). For, if $p \neq q$, then $\theta \in R_q$, hence $I_q = R_q\theta \cap R_q = R_q\theta$.

Finally, we prove (c). To prove (c), first note that $(1+J)\theta = N$, where $N = \sum\limits_{\sigma\in G} \sigma$. Let $x \in I'$, then $x\theta \in I$, and we have

$$x\theta \in I^- \iff (1+J)x\theta = 0 \iff x(1+J)\theta = 0 \iff xN = 0.$$

Similarly, suppose $y \in I_q^-$. Then $y\theta \in I'_q\theta = I_q$ (by (a)) and we have

$$y\theta \in I_q^- \iff (1+J)y\theta = 0 \iff yN = 0.$$

Clearly $I^-\mathbb{Z}_q \subseteq I_q^-$. To prove the other inclusion, suppose that $y\theta \in I_q^-$, with $y \in I'_q$. Then we can write

$$y = \sum_c \sum_\sigma a_\sigma^c \sigma(c - \sigma_c),\ a_\sigma^c \in \mathbb{Z}_q.$$

Note that

$$yN = 0 \implies \sum_c \sum_\sigma a_\sigma^c (c-1) = 0.$$

The idea is to approximate y by an element $x \in I'$ such that $xN = 0$; this will give us an element $x\theta \in I^-$ such that $x\theta$ is close to $y\theta$. This in turn will imply that I_q^- is contained in the closure of I^- which is $I^-.\mathbb{Z}_q$. The approximation uses the following principle:

Suppose $b_i \in \mathbb{Z},\ s_i \in \mathbb{Z}_q$, and suppose $\sum\limits_{i=1}^{m} b_i s_i = 0$. Then there is a sequence $(t_1^{(n)}, \cdots, t_m^{(n)}) \in \mathbb{Z}^m$, whose limit is $(s_1, \cdots, s_m)$ and such that $\Sigma b_i t_i^{(n)} = 0$.

The details are left to the reader. □

We have an isomorphism $R_q \simeq R \otimes \mathbb{Z}_q$, under which R_q^- maps isomorphically onto $R^- \otimes \mathbb{Z}_q$ and I_q^- maps isomorphically onto $I^- \otimes \mathbb{Z}_q$. Thus $(R_q^-/I_q^-) \simeq (R^-/I^-) \otimes \mathbb{Z}_q$ and this is isomorphic to the q-part of (R^-/I^-). Hence Theorem I will follow from the more general one below:

Theorem 3. *$[R_q^- : I_q^-] =$ q-part of $h^-(\mathbb{Q}(\zeta_{p^n}))$.*

Proof. Observe that $x \in R^-$ if and only if $[(1 - J/2)]x = x$. We consider the three different possibilities: $q \notin \{2, p\},\ q = 2$ and $q = p$.

(a)$q \notin \{2, p\}$: In this case $(1 \pm J/2) \in R_q$ and $R_q = R_q^+ \oplus R_q^-,\ I_q = I_q^+ \oplus I_q^-$, as is seen by writing $x = (x + Jx/2) + (x - Jx/2)$. Hence

$$I_q^- = \frac{1-J}{2} I_q = \frac{1-J}{2} R_q \theta = R_q^- \theta.$$

Consider the linear map

$$\begin{array}{rl} A & : R_q^- \to R_q^- \\ & \;\; x \ \mapsto x\theta. \end{array}$$

From a basic linear algebra result,

$$[R_q^- : AR_q^-] = [R_q^- : R_q^- \theta] = q\text{-}part\ of\ \det A.$$

Over $\bar{\mathbb{Q}}_q[G]^-$, we have a decomposition

$$\bar{\mathbb{Q}}_q[G]^- = \bigoplus_{\chi\ odd} \epsilon_\chi \bar{\mathbb{Q}}_q[G],$$

where each summand is one dimensional and

$$\epsilon_\chi = 1/p^n \Sigma_{a=1,\ (a,p)=1}^{p^n}\ \chi(a)\sigma_a^{-1}.$$

Recall from one of the previous lectures in the conference that

$$\epsilon_\chi \theta = B_{1,\bar{\chi}} \epsilon_\chi.$$

Thus

$$\begin{aligned}[R_q^- : I_q^-] &= q\text{- } part\ of \prod_{\chi\ odd} B_{1,\bar{\chi}} \\ &= q\text{- } part\ of\ 2p^n \prod_{\chi\ odd} (-1/2B_{1,\bar{\chi}}) \\ &= q\text{- } part\ of\ h^-(\mathbb{Q}(\zeta_{p^n})).\end{aligned}$$

This proves the theorem in this case.

To prove the theorem in the case (b), we note that $(\frac{1+J}{2}) \notin R_2$. We modify θ to obtain an element already in $\mathbb{Q}_2[G]^-$. Let

$$\tilde{\theta} = \Sigma'(a/p^n - 1/2)\sigma_a^{-1} = \theta - N/2,$$

where Σ', as before indicates summation over integers from 1 to p^n which are prime to p. We have $J\tilde{\theta} = -\tilde{\theta}$; hence $\tilde{\theta}$ is in the "$-$" component. We shall need the following lemma:

Lemma 4. *(i) $I_2^- \subseteq R_2\tilde{\theta}$, (ii) $[R_2\tilde{\theta} : I_2^-] = 2$.*

Proof. (i): We have $I_2^- = R_2\theta \cap R_2^-$. Let $x \in R_2$ and $x\theta \in I_2^-$. Then

$$x\theta = [(1-J)/2]x\theta = x[(1-J)/2](\tilde{\theta} + N/2) = x\tilde{\theta}$$

as $(\frac{1-J}{2})N/2 = 0$ and $(\frac{1-J}{2})\tilde{\theta} = \tilde{\theta}$.

(ii): First, we claim that if $x \in R_2$, then either $x\tilde{\theta} \in R_2$ or $(x-1)\tilde{\theta} \in R_2$. Note that

$$x\tilde{\theta} = x\theta - xN/2 \in R_2 \Longrightarrow xN/2 \in R_2,$$

and that a similar statement holds for $(x-1)\tilde{\theta}$.

Let $x = \Sigma x_\sigma \sigma$; then

$$xN = (\Sigma x_\sigma N)\ and\ (x-1)N = (-1 + \Sigma x_\sigma)N.$$

Either (Σx_σ) or $(-1+\Sigma x_\sigma)$ is even. Therefore either $xN/2$ or $(x-1)N/2 \in R_2$ and the claim is proved. We therefore have $[R_2\tilde{\theta} : R_2\tilde{\theta} \cap R_2] = 2$ (note that the index is not 1 since $\tilde{\theta} \notin R_2$; $x\tilde{\theta} \in R_2\tilde{\theta} \Longrightarrow 2x\tilde{\theta} = 2x\theta - xN \in R_2\tilde{\theta} \cap R_2$). Therefore assertion (ii) will be true if we can show that

$$R_2\tilde{\theta} \cap R_2 = R_2\theta \cap R_2^- = I_2^-.$$

By (i), $I_2^- \subseteq R_2\tilde{\theta} \cap R_2$. To show the other inclusion, suppose $x\tilde{\theta} \in R_2\tilde{\theta} \cap R_2$, where $x = \Sigma x_\sigma \sigma \in R_2$. As remarked above,

$$x\tilde{\theta} \in R_2 \Longrightarrow xN/2 \in R_2 \Longrightarrow \Sigma x_\sigma \equiv 0 \mod 2.$$

Let $y_\sigma = x_\sigma$ for $\sigma \neq 1,\ J$ and $y_1 = x_1 - \frac{\Sigma x_\sigma}{2},\ y_J = x_J - \frac{\Sigma x_\sigma}{2}$. Checking that $\Sigma y_\sigma = 0$, we have therefore $y = \Sigma y_\sigma \sigma \in R_2$ satisfies $yN = 0$. Further

$$x - y = (\frac{\Sigma x_\sigma}{2})(1 + J) \Longrightarrow (x - y)\tilde{\theta} = 0.$$

Therefore

$$x\tilde{\theta} = y\tilde{\theta} = y\theta - (yN/2) = y\theta \in R_2\theta.$$

Since $x\tilde{\theta} \in R_2$ and $[(1 - J)/2]x\tilde{\theta} = x\tilde{\theta}$, we have $x\tilde{\theta} \in R_2\theta \cap R_2^- = I_2^-$, so $R_2\tilde{\theta} \cap R_2 = I_2^-$. □

We continue with the proof of Theorem 3 in the next case.

(b) $p = 2$: As before, we look at the linear map

$$\begin{aligned} A \ :\ R_2^- &\to R_2^- \\ x &\mapsto \tilde{\theta}x \end{aligned}$$

Note that

$$x \in R_2^- \Longrightarrow \frac{1}{2}xN = 0, \text{ hence } \tilde{\theta}x \in R_2^-.$$

We have

$$\begin{aligned} [R_2^- : R_2^-\tilde{\theta}] =&2\text{-part of } \det\ A \\ =&2\text{-part of } \prod_{\chi\ odd} B_{1,\bar{\chi}} \qquad (\text{as } \chi\ odd \Longrightarrow \epsilon_\chi\tilde{\theta} = \epsilon_\chi\theta) \\ =&2^{\frac{1}{2}|G|}.\tfrac{1}{2}.(2\text{-part of } h^-). \end{aligned}$$

Since this index is finite, we should have

$$\frac{1}{2} \mid G \mid = \mathbb{Z}_2\text{-rank of } R_2^- = \mathbb{Z}_2\text{-rank of } R_2^-\tilde{\theta}.$$

But

$$R_2^-\tilde{\theta} = (1 - J)R_2\tilde{\theta} = R_2(2\tilde{\theta}) = 2R_2\tilde{\theta}.$$

Therefore $[R_2\tilde{\theta} : R_2^-\tilde{\theta}] = 2^{\frac{1}{2}|G|}$.

Finally,

$$\begin{aligned} [R_2^- : I_2^-] =&([R_2^- : R_2^-\tilde{\theta}]/[R_2\tilde{\theta} : R_2^-\tilde{\theta}]).[R_2\tilde{\theta} : I_2^-] \\ =&\tfrac{1}{2}.(2 - \text{part of } h^-).2 \\ =&2 - \text{part of } h^-. \end{aligned}$$

(c) $q = p$: In this case the problem is that θ has p^n in its denominator. As before, we consider the element $\tilde{\theta} - \frac{1}{2}N$. Let $x = \Sigma_{(b,p)=1}\ x_b\sigma_b \in R_p$. Note that

$$x\tilde{\theta} \in R_p^- \Longleftrightarrow x\theta \in R_p.$$

Now

$$x\theta = 1/p^n \sum_c \sum_a a x_{ac} \sigma_c;$$

and hence

$$\begin{aligned} x\theta \in R_p &\Longleftrightarrow \sum_a a x_{ac} \equiv 0 \mod p^n \ \forall\ c \text{ such that}(c,p) = 1 \\ &\Longleftrightarrow \Sigma a x_a \equiv 0 \mod p^n. \end{aligned}$$

We use this condition to see that $(x-b)\theta \in R_p$ for exactly one integer b mod p^n and that

$$[R_p\tilde{\theta} : R_p\tilde{\theta} \cap R_p^-] = p^n.$$

Since $(\frac{1-J}{2})N = 0$, and $R_p^+\tilde{\theta} = 0$, we see that

$$R_p\tilde{\theta} = R_p^-\tilde{\theta} = R_p^-(\theta - \frac{1}{2}N) = R_p^-\theta$$

and

$$R_p\tilde{\theta} \cap R_p^- = R_p^-\theta \cap R_p^- \subseteq R_p\theta \cap R_p^- = I_p^-.$$

Further, $R_p\theta \cap R_p^- \subseteq R_p^-\theta \cap R_p^-$. For, if $x\theta \in R_p^-$, then $x\theta = [(1-J)/2]x\theta \in R_p^-\theta$. Therefore $I_p^- = R_p^-\theta \cap R_p^-$ and this implies that

$$[R_p^-\theta : I_p^-] = p^n.$$

As before, we consider

$$\begin{aligned} A \ : \ R_p^- &\to R_p^- \\ x &\mapsto p^n\theta x, \end{aligned}$$

and we have

$$\begin{aligned} [R_p^- : p^n R_p^- \theta] =& p - part\ of \ \det\ A \\ =& p - part\ of\ p^{(\frac{n}{2})|G|} \prod_{\bar{\chi}\ odd} B_{1,\bar{\chi}} \\ =& p^{(\frac{n}{2})|G|}(1/p^n)(p - part\ of\ h^-). \end{aligned}$$

As $[R_p^-\theta : p^n R_p^-\theta] = p^{(\frac{n}{2})|G|}$, we have

$$[R_p^- : I_p^-] = p - \text{part of } h^-.$$

□

This completes the proof of Theorem 3 and hence also of Theorem I. □

Reflection Theorem

We will now prove a Reflection theorem due to Leopoldt; we begin by setting up notation.

Notation and background: p an odd prime, L/K is a Galois extension with $\mathrm{Gal}(L/K) = G$. We shall assume that $\zeta_p \in L$.
For any abelian group Y, Y^p and Y_p will denote respectively the image and kernel of multiplication by p in Y. We let p-rank $Y = \dim_{\mathbb{F}_p} Y/Y^p$.

Let L' be the maximal unramified abelian p-extension of L such that $\mathrm{Gal}(L'/L)$ is isomorphic to $\mathbb{Z}/p\mathbb{Z} \times \cdots \times \mathbb{Z}/p\mathbb{Z}$. This extension is in fact Galois over K, by virtue of its maximality. Let $H = \mathrm{Gal}(L'/L)$; then H is a normal subgroup of $\mathrm{Gal}(L'/K)$ with quotient G. As H is abelian, G acts on H by inner automorphisms, i.e. if $\bar{g} \in G$ and $h \in H$, the action is given by

$$\bar{g}.h = ghg^{-1}$$

where $g \in \mathrm{Gal}(L'/K)$ is a lift of $\bar{g}$. This action is well-defined as H is abelian, and thus H has a natural structure as a $\mathbb{Z}[G]$-module.
Let A be the p-Sylow subgroup of the ideal class group of L. The Artin map defines an isomorphism

$$H \simeq A/A^p$$

which is in fact a G-isomorphism, A/A^p being considered as a G-module with the canonical action. Now as $\zeta_p \in L$, abelian extensions of L of exponent p are obtained by extracting p-th roots of elements in L. In particular, L'/L is a Kummer extension. We let $B \subseteq L^*/(L^*)^p$ be the associated Kummer group i.e.

$$L' = L\{(b^{1/p}) \mid b \in B\}.$$

Let W_p be the group of p-th roots of unity. There is a pairing

$$\begin{aligned} H \times B \ &\to W_p \\ < h, b > &\mapsto h(b^{1/p})/b^{1/p}. \end{aligned}$$

This pairing is non-degenerate (i.e. if $< h, B >= 1$ for some $h \in H$, then $h = 1$) and bilinear. It induces an isomorphism of G-modules

$$B \simeq \mathrm{Hom}(A/A^p, W_p) \tag{1}$$

where the action of G on the group of homomorphisms is the natural one derived from the action of G on A/A^p and W_p.

Let $b \in B,\ L(b^{1/p}) \subseteq L'$. As $L(b^{1/p})$ is unramified, we must have $(b) = I^p$ for some ideal I of L. We thus have a homomorphism

$$\begin{aligned} \phi \ : \ B &\to A_p \\ b &\mapsto I. \end{aligned}$$

Note that ϕ is a G-homomorphism. We want to describe Ker ϕ.

Suppose $\phi(b) = 1$, then $(b) = (a)^p,\ a \in L$. Thus $b = \epsilon a^p$ for

some $\epsilon \in E :=$ units of L. Further, changing b by an element in $(L^*)^p$ doesn't change the ideal class of I, hence we see that

$$\text{Ker } \phi \subseteq (EL^{*^p})/L^{*^p} \simeq E/E^p.$$

The inclusion Ker $\phi \subseteq E/E^p$ is again a G-morphism. Thus we have

$$\begin{array}{rll} B & \simeq A/A^p & \text{(not a } G\text{-map)} \\ \phi : B \to & A_p & (G\text{-map}) \\ \text{Ker } \phi & \subseteq E/E^p & (G\text{-map}). \end{array}$$

These will be used in the reflection theorem proved below.

What does the reflection theorem say? We let $L = \mathbb{Q}(\zeta_p)$ and A be the p-Sylow subgroup of the ideal class group of L. The group A is a $\mathbb{Z}_p[\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})]$ module and can be decomposed as usual, as $A = A^+ \oplus A^-$. One of the consequences of the reflection theorem, which is due to Kummer, is that

$$A^- = 0 \Longrightarrow A^+ = 0.$$

We shall now state and prove the theorem. Recall that A has a direct sum decomposition corresponding to the idempotents of the group ring $\mathbb{Z}_p[\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})]$, *viz.*

$$A = \overset{p-1}{\underset{i=0}{\oplus}} \epsilon_i A.$$

Theorem II. *With notation as above, suppose i is even, j is odd and $i+j \equiv 1 \mod (p-1)$. Then*

$$p\text{-}rank\ \epsilon_i A \leq p\text{-}rank\ \epsilon_j A \leq 1 + p\text{-}rank\ \epsilon_i A.$$

Corollary.

$$p \mid h^+ \Longrightarrow p \mid h^-.$$

Proof of Corollary. We have $A = A^+ \oplus A^-$, and

$$A^+ = \underset{i\ even}{\oplus} \epsilon_i A, \ \ A^- = \underset{i\ odd}{\oplus} \epsilon_i A.$$

Suppose $p \not| h^-$. Then

$$A^- = 0.$$

Theorem II implies that $A^+ = 0$, hence $p \not| h^+$. □

Proof of Theorem II. . Let $G = \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. There is an isomorphism $H \simeq A/A^p$ as G-modules, hence $\epsilon_i H \simeq \epsilon_i(A/A^p)$ for all i. Let $h \in \epsilon_i H$; then

$$\sigma_a h = h^{\omega^i(a)} \ \forall \ a \in (\mathbb{Z}/p\mathbb{Z})^*,$$

where ω denotes the Teichmüller character. Let $b \in \epsilon_k B$. We have

$$\begin{aligned} < h, b >^{\omega(a)} &= < h, b >^{\sigma_a} \ \ (as \ \ < h, b >^p = 1 \ and \ \omega(a) \equiv a \ mod \ p) \\ &= < h^{\sigma_a}, b^{\sigma_a} > \\ &= < h^{\omega^i(a)}, b^{\omega^k(a)} > \\ &= < h, b >^{\omega^{i+k}(a)} \ \ \forall \ a \ (as \ h \in \epsilon_i H, \ b \in \epsilon_k B). \end{aligned}$$

We therefore get that if $(i + k) \not\equiv 1 \mod (p-1)$, then $< h, b >= 1$.

Using this along with the fact that the pairing between $B = \oplus \epsilon_k B$ and $H = \oplus \epsilon_i H$ is non-degenerate (see the discussion preceding (1)), we see that the induced pairing,

$$\epsilon_i H \times \epsilon_j B \to W_p$$

is non-degenerate when $i + j \equiv \ 1 \mod (p-1)$. Therefore, as H is G-isomorphic to A/A^p, we have

$$\epsilon_j B \simeq \epsilon_i H \simeq \epsilon_i(A/A^p).$$

Similarly, we have

$$\phi : \epsilon_j B \to \epsilon_j A_p,$$

and (Ker $\phi) \cap \epsilon_j B$ is isomorphic to a subgroup of $\epsilon_j(E/E^p)$. Grant now the following statement:
(2)

$$\epsilon_j(E/E^p) \simeq \begin{cases} \mathbb{Z}/p\mathbb{Z}, \ j \ even, \ j \not\equiv \ 0 \mod (p-1); \ or \ j \equiv 1 \mod (p-1) \\ 0 \text{ otherwise.} \end{cases}$$

This statement follows from an analysis of the units of $\mathbb{Q}(\zeta_p)$. We have p-rank $\epsilon_i A = \dim \ \epsilon_i(A/A^p)$ by definition.

Clearly, p-rank $\epsilon_j A = \dim \ \epsilon_j(A/A^p) = \dim \ \epsilon_j A_p$. Also,

$$\dim \ \epsilon_i(A/A^p) = \dim \ \epsilon_j B \leq \dim \ \epsilon_j(E/E^p) + \dim \ \epsilon_j A_p.$$

Using the statement (2) above, we see that

$$\begin{aligned} &p\text{-rank}(\epsilon_i A) \leq 1 + p\text{-rank}(\epsilon_j A) \ \textit{if } j \textit{ even and } j \not\equiv 0 \mod (p-1); \\ &p\text{-rank}(\epsilon_i A) \leq \quad p\text{-rank}(\epsilon_j A) \quad \textit{if } j \textit{ odd and } j \not\equiv 1 \mod (p-1). \end{aligned}$$

If $j \equiv 1 \mod (p-1)$, then we find

$$p\text{-rank } \epsilon_0 A \leq 1 + p\text{-rank } \epsilon_1 A.$$

We claim that $\epsilon_0 A = \epsilon_1 A = 0$. This follows essentially from

Stickelberger's theorem. Recall that Stickelberger's theorem implies that $(c-\sigma_c)\theta$ annihilates $\epsilon_i A$, where θ is the Stickelberger element, $c \in \mathbb{Z}$, $(c,p) = 1$. On the other hand, one can show that

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

and deduce that $(c - \omega^i(c))B_{1,\omega^{-i}}$ annihilates $\epsilon_i A$. For $i = 0$, this implies that $(c-1)/2$ annihilates A_0, and so $A_0 = 0$. For $i = 1$, take $c = 1 + p$, We have

$$(c - \omega(c))B_{1,\omega^{-1}} = pB_{1,\omega^{-1}} = \Sigma_{a=1}^{p-1} a\omega^{-1}a \equiv p - 1 \not\equiv 0 \mod p.$$

But $\epsilon_1 A$ is a p-group and $(c - \omega(c))B_{1,\omega^{-1}}$ annihilates $\epsilon_1 A$ implies that $A_1 = 0$, hence the claim. Therefore Theorem II is proved. □

R. SUJATHA
School of Mathematics
Tata Institute of Fundamental Research
Mumbai 400-005
e-mail: sujatha@math.tifr.res.in

Fermat's last theorem for regular primes

Dinesh S. Thakur

The so-called Fermat's last theorem (FLT for short) that 'there are no integral solutions to $x^n + y^n = z^n$, with $n > 2$ and $xyz \neq 0$ (a non-triviality condition)' is now proved by Wiles. The proof is outside the scope of this summer school, though we will sketch some ideas in the later lectures.

Now we will prove the so-called 'regular prime case' of FLT, following Kummer in essence, but connecting with later ideas and simplifications by other mathematicians.

It may be advisable to first learn the proof for $n = 3$ from Hardy and Wright or from Ireland-Rosen pp. 285-286, which prove even stronger result, using only the basics of quadratic fields that we have seen.

Since FLT for the exponent n implies FLT for multiples of n and since we proved it for $n = 4$, it is enough to prove FLT for $n = p$ an odd prime. The basic strategy is the same as before: try to get an infinite descent and thus a contradiction starting with an assumed non-trivial solution. We have a factorization $z^p = \prod(x + \zeta_p^i y)$ and we try to conclude that each factor (apart from GCD) is a p-th power. This conclusion is justified in the case of unique factorization domains. But since we only have unique factorization in ideals in general, we can conclude that each factor is a p-th power of an ideal, which can be further assumed to be principal, if p does not divide h, the class number of $\mathbf{Q}(\zeta_p)$. Such a (odd) prime is called a *regular prime*. But we want more. If a principal ideal is known to be a p-th power, its generator is of the form $u\alpha^p$, for some unit u and we would like to know when u itself can be concluded to be a p-th power. Since $(\sum a_i \zeta_p^i)^p \equiv \sum a_i^p (\zeta_p^p)^i \equiv \sum a_i$ modulo p, we know that the p-th powers are congruent to rational integers modulo p.

Kummer's lemma: If p is a regular prime and u is a unit of $\mathbf{Z}[\zeta_p]$ congruent to a rational integer modulo p, then u is a p-th power.

This is the hard part of the proof. This and characterization of regularity in terms of divisibility properties of Bernoulli numbers (essentially the special zeta values) are the great achievements of Kummer regarding the theorem. We will first prove FLT for regular p, assuming the Kummer lemma and then prove the Kummer's lemma.

What do we know about the regular primes? There are only three irregular primes less than hundred: 37, 59 and 67. Numerical data and heuristics (see Washington) suggest that approximately 61 % of the primes are regular, but it is not even proved that there are infinitely many. On the other hand, it is known that (see Katre's article) there are infinitely many irregular primes!

FLT for regular primes. If p is a regular prime, then $x^p + y^p = z^p$ has no non-trivial integral solutions.

In fact, there are no non-trivial solutions, even in $\mathbf{Z}[\zeta_p]$, but we will not prove this.

From what we have seen earlier, we might want to try to show there are no local solutions. But indeed there are local and global solutions, namely the trivial ones. So the next natural try is to see whether there are any local non-trivial solutions or even non-trivial solutions modulo a prime. If we look at the prime p itself, this leads naturally the so-called *first case*: p does not divide xyz. Indeed, $x^p + y^p = z^p$ has no non-trivial solutions modulo 9, if $p = 3$ and modulo 25, if $p = 5$. So the first case for $p = 3$ and $p = 5$ follows. For $p = 7$, there is even a 7-adic non-trivial solution (exercise). So we follow a different path. If p divides xyz, it is called the *second case* and then we need Kummer's lemma, which we do not need for the first case.

It should be noted though that (i) Terjanian gave a short, elementary proof of the first case for exponent $2p$ instead, (ii) Sophie Germain gave an elementary proof of the first case, for odd prime p such that $2p + 1$ is also a prime, and combining her ideas with sieve techniques, it was proved by Adleman, Heath-Brown and Fouvry that the first case holds for infinitely many primes p, (iii) Eichler proved the first case, under much weaker hypothesis that $p^{\lfloor\sqrt{p}\rfloor-2}$ does not divide h. There are even some easy, non-trivial conditions, for which no counter-example is known and under which the first case is proved. For these and more results, see Ribenboim's book on FLT.

Proof for the first case: For the rest of the chapter, let p be a prime greater than 3, $\zeta := \zeta_p$, $K := \mathbf{Q}(\zeta)$, $\mathcal{O} := \mathbf{Z}[\zeta]$, h be the class number of $\mathcal{O}$ and $\lambda := 1 - \zeta$.

Suppose $x^p + y^p = z^p$, with x, y and z relatively prime (without loss of generality, as we can just take out the common factors: note that if x, y , z were to be cyclotomic integers, we can not do this, as there may be non-principal ideal common factor) and with p not dividing xyz. We want to get a contradiction.

If $x \equiv y \equiv -z$ modulo p, then $-2z^p \equiv z^p$ modulo p which is a contradiction, since $p > 3$. Hence, changing signs and labelling of x, y and z if necessary, we can assume that $x \not\equiv y$ modulo p. We have $z^p = \prod(x + \zeta^i y)$.

Lemma 1: Ideals $(x + \zeta^i y)$ are relatively prime: Otherwise a prime $\wp$ divides two of them. By eliminating x, we see that $\wp$ divides λ or y, whereas eliminating y, we see that $\wp$ divides λ or x. So $\wp = \lambda$. But then $x + y \equiv$

$x+\zeta^i y \equiv 0$ modulo λ, hence $z^p \equiv x+y \equiv 0$ modulo λ, which leads to the contradiction p divides z.

So by the unique factorization of ideals, each factor is a p-th power of an ideal which is principal by regularity. Hence $x+\zeta^i y = \epsilon_i \alpha_i^p$, where ϵ_i are units and $\alpha_i \in \mathcal{O}$.

Lemma 2: Any unit ϵ of $\mathcal{O}$ is a power of ζ times a real unit: Since $\epsilon/\overline{\epsilon}$ is an algebraic integer with all its absolute values 1, it is a root of unity, so can be written as $\pm\zeta^{2r}$. If the sign is positive, $\zeta^{-r}\epsilon$ is invariant under the complex conjugation and is thus real and we are done. Now we have modulo λ, $\epsilon = \sum a_i \zeta^i \equiv \sum a_i \equiv \overline{\epsilon}$, whereas the negative sign would lead to $\epsilon \equiv -\overline{\epsilon}$, a contradiction. (Note that the first half of the argument is a special case of what we have seen before in much more general situation).

So, since a p-th power is congruent to a rational integer modulo p, we get $x+\zeta y = \zeta^r \epsilon_1 \alpha^p \equiv \zeta^r \epsilon_1 a$ modulo p, with rational integer a and a real unit ϵ_1. Conjugation gives $x+\zeta^{-1}y \equiv \zeta^{-r}\epsilon_1 a$ modulo p. Hence $\zeta^{-r}(x+\zeta y) \equiv \zeta^r(x+\zeta^{-1}y)$ i.e., $x+\zeta y - \zeta^{2r}x - \zeta^{2r-1}y \equiv 0$ modulo p. If the ζ powers occurring here are distinct, then since they form a part of a basis, we get p divides x and y, which is a contradiction. Otherwise, we break into the following cases, each leading to an easy contradiction listed: (i) $1=\zeta^{2r}$, then p divides y, (ii) $1=\zeta^{2r-1}$, then p divides $x-y$ and (iii) $\zeta = \zeta^{2r-1}$, then p divides x. This finishes the proof of the first case.

Proof of the second case: Suppose there is a second case non-trivial solution, then taking out the common λ factors, if any, we have $\alpha^p+\beta^p = \epsilon\delta^p\lambda^{mp}$, with m least positive such, $\alpha, \beta, \gamma \in \mathcal{O}$ not divisible by λ and ϵ a unit. Writing $\alpha^p+\beta^p = \prod(\alpha+\zeta^i\beta)$ we analyze the GCD:

Lemma 3: Changing β by $\zeta^i\beta$ if necessary, we have $\alpha+\beta = \lambda^{p(m-1)+1}I'J_0^p$ and $\alpha+\zeta^k\beta = \lambda I'J_k^p$, where $I' = (\alpha,\beta)$, J_i are pairwise prime and not multiples of λ and $m>1$.

Assuming this for now, we have $(\alpha+\zeta^k\beta)/(\alpha+\beta) = \lambda^{-p(m-1)}(J_k/J_0)^p$, so again by regularity assumption on p, we know that J_k/J_0 is principal generated by μ_k/n_k say, with $\mu_k \in \mathcal{O}$ and $n_k \in \mathbf{Z}$ both not divisible by λ. Hence $(\alpha+\zeta^k\beta)\lambda^{p(m-1)} = \epsilon_k(\alpha+\beta)(\mu_k/n_k)^p$. Subtracting the resulting equation for $k=2$ from $(1+\zeta)$ times the equation resulting for $k=1$, we get $\zeta(\alpha+\beta)\lambda^{p(m-1)} = (\alpha+\beta)[(\mu_1/n_1)^p\epsilon_1(1+\zeta) - (\mu_2/n_2)^p\epsilon_2]$ implying $(\mu_1 n_2)^p - (\epsilon_2(\mu_2 n_1)^p)/(\epsilon_1(1+\zeta)) = \lambda^{p(m-1)}(n_1n_2)^p\zeta/(\epsilon_1(1+\zeta))$.

Now $1+\zeta = (1-\zeta^2)/(1-\zeta)$ is a unit. Hence we get an equation of the form $\alpha_1^p + \epsilon'\beta_1^p = \epsilon''\delta_1^p\lambda^{p(m-1)}$, with α_1 and β_1 not multiples of λ. This would contradict the minimality of m above, if we can show that the unit ϵ' is a p-th power. This we do using Kummer's lemma: By Lemma 3, we have $m>1$, so λ^p divides $\alpha_1^p+\epsilon'\beta_1^p$. Now $\beta_1 = \mu_2 n_1$ is prime to λ, so ϵ' is a p-th power modulo λ^p and hence modulo p, so it is congruent to rational integer modulo p. Kummer lemma then provides the contradiction we need. This finishes the proof of the Theorem, modulo the proof of Lemma 3 and Kummer's lemma, which we now proceed to prove.

Proof of Lemma 3: Since α is not divisible by λ, we have $\alpha \equiv a + d\lambda$ modulo λ^2, with a not divisible by p. Let $ca \equiv d$ modulo p.

Then $(\zeta^c\alpha - a)/(1-\zeta) \equiv -a(1-\zeta^c)/(1-\zeta) + d\zeta^c \equiv -ac + d \equiv 0$ modulo λ, i.e., $\zeta^c\alpha \equiv a$, a rational integer, modulo λ^2.

So multiplying α and β by appropriate ζ powers, we can assume that $\alpha \equiv a$ and $\beta \equiv b$ modulo λ^2. Since $(p) = (\lambda^{p-1})$, we have $\alpha^p \equiv a^p$ and $\beta^p \equiv b^p$ modulo $\lambda^{p+1} = p\lambda^2$, as the p-th binomial coefficients are divisible by p. If m were 1, $a^p + b^p = \alpha^p + \beta^p + \rho\lambda^{p+1} = \epsilon\lambda^p(\delta^p + \epsilon^{-1}\rho\lambda)$. But if $a^p + b^p$ has valuation e at p, then it has valuation $e(p-1)$ at λ, whereas the right side has valuation p, a contradiction proving $m > 1$.

Now $\epsilon\delta^p\lambda^{mp} = \prod(\alpha + \zeta^k\beta)$. So λ divides some term of the product, but the difference in i-th and j-th term is $\zeta^j(\zeta^{i-j} - 1)\beta$ which is divisible by λ to the exactly first power. Hence λ divides each term and only one term, say one corresponding to j_0, can be divisible by higher power of λ. Hence the λ powers are as claimed in the lemma. Now $I' = (\alpha, \beta)$ divides each term and is not divisible by λ, so we see that J_i' are relatively prime. (If $\wp$ divides two of them, then $\lambda I'\wp$ divides the difference and hence divides $\lambda\beta$ and similarly divides $\lambda\alpha$. This implies $I'\wp$ divides $(\alpha, \beta) = I'$, a contradiction). Hence, each is a p-th power, as claimed. This proves lemma 3.

Proof of Kummer's lemma: First proof: The first proof is the simplest, assuming the following fact from the class field theory, for the special case $F = K := \mathbf{Q}(\zeta)$.

Fact: Given any number field F, its maximal abelian unramified extension H_F (called the Hilbert class field of F) has degree h_F, the class number of F.

Suppose u is as in the hypothesis, but not a p-th power, then $K(u^{1/p})$ is (a Kummer extension) an abelian extension of degree p of K. We will get a contradiction to the regularity of p, once we show that this extension is unramified everywhere:

Without loss of generality, we can assume that $u \equiv 1$ modulo p (replace u by u^{p-1}, if necessary). We then claim that, in fact, $u \equiv 1$ modulo λ^p: Write $u = 1 + \kappa p$. Now κ is congruent to some rational integer k modulo λ, as its residue field is $\mathbf{Z}/p\mathbf{Z}$. Hence $u \equiv 1 + kp$, a rational integer modulo λ^p. But conjugate of u is also congruent to the same integer modulo λ^p, as λ is its own conjugate. Hence $1 = Norm(u) \equiv (1+kp)^{p-1} \equiv 1 - kp$, which implies $k \equiv 0$ modulo p proving the claim.

Now let w be a p-th root of u. Then $(1-w)/\lambda$ is a root of $((\lambda x-1)^p+u)/\lambda^p$, which is a monic polynomial with algebraic integer coefficients by the claim above. Hence $(1-w)/\lambda$, which generates the same extension as w, is an algebraic integer. But the other roots are $(1-\zeta^i w)/\lambda$ and hence the differences of the roots are $(\zeta^i-\zeta^j)w/\lambda$ which are units. Hence the relative field discriminant is a unit, as it divides the product of these differences. Hence the extension is everywhere unramified (including at the infinite places, which are complex, so there is nothing to check there) as claimed. This finishes the first proof.

See also Washington pp. 80-81 for a couple of variations on this proof.

Kummer's lemma: Second proof: This proof again uses what we proved above that if the Kummer's lemma is false for u, then the Kummer extension we thus get is everywhere unramified, but rather than using the fact about Hilbert class field to get a contradiction, we get immediate contradiction by applying the following theorem to $K=\mathbf{Q}(\zeta)$ and $L=K(u^{1/p})$.

Hilbert theorem 94: Let K be a number field and let L be its cyclic extension of degree p, an odd prime and with relative discriminant 1. Then there is a non-principal ideal J of $\mathcal{O}_K$ such that $I=J\mathcal{O}_L$ is principal. Further, J^p is principal, so that p divides the class number of K.

Proof: We use the following theorem:

Theorem: Let L be a cyclic extension of a number field K of degree p, an odd prime. Let $Gal(L/K)=\langle\sigma\rangle$. Then there is $U\in\mathcal{O}_L^*$ of relative norm 1 which is not $(1-\sigma)$-th power of a unit (in $\mathcal{O}_L^*$).

This says that $H^1(Gal(L/K),\mathcal{O}_L^*)\neq 0$ and follows by easy Herbrand quotient argument, if you know that technology. We will give a classical proof (see Hilbert chapter 15) based on construction of a 'relative units basis'. But first let us see how it implies the theorem: By Hilbert 90, proved in Narlikar lectures, (i.e, $H^1(Gal,L^*)=0$), we can write $U=\alpha^{1-\sigma}$, with $\alpha\in L^*$. Multiplying by suitable rational integer, we can assume that $\alpha\in\mathcal{O}_L^*$.

Let $I=(\alpha)$. Then $I^\sigma=(U\alpha)=I$. Consider a prime factor $\wp$, if any, which does not lie in K (i.e., is not inert). If $\wp^\sigma\neq\wp$, so that $\wp$ splits, then the relative norm of $\wp$, which is a prime ideal in K divides I. If $\wp^\sigma=\wp$, then $\wp$ is ramified, which contradicts the discriminant hypothesis. Hence I comes from an ideal J of $\mathcal{O}_K$.

If J were principal, then $\alpha=\overline{U}\alpha_1$ for some unit $\overline{U}\in\mathcal{O}_L^*$ and $\alpha_1\in\mathcal{O}_K$. Now $\alpha^{1-\sigma}=\overline{U}^{1-\sigma}$, a contradiction. Finally, J^p is the relative norm of I, so is principal generated by relative norm of α. This finishes the proof of Hilbert theorem 94 modulo the proof of the Theorem.

Proof of the Theorem: To construct the U we want, we have to get a good control on the units with respect to the Galois action. Let $\epsilon_i \in \mathcal{O}_K^*$, $1 \leq i \leq r = r_K = r_1 + r_2 - 1$ be independent units. By the Dirichlet theorem, they form a finite index subgroup of $\mathcal{O}_K^*$. Since our extension is cyclic of odd degree p, the real primes cannot ramify and we have $r_L = p(r_1 + r_2) - 1$. If we choose a unit $U_1 \in \mathcal{O}_L^*$ independent of ϵ_i, then $r + p - 1$ units ϵ_j, $U_1^{\sigma^i}$ $(0 \leq i < p - 1)$ are independent: Otherwise we have a (non-zero) polynomial $F(\sigma)$ of degree $\leq p - 2$ and with integral coefficients such that $U_1^{F(\sigma)} \in \mathcal{O}_K^*$, with obvious meaning attached to this exponentiation. Since (the norm) $1 + \sigma + \cdots + \sigma^{p-1}$ is irreducible of degree $p - 1$, expressing the GCD as a linear combination and clearing the denominators, we get $Fg_1 + (1 + \sigma + \cdots + \sigma^{p-1})g_2 = a \in \mathbf{Z} - \{0\}$, with some polynomials g_1 and g_2 with integral coefficients. This would lead to a contradiction $U_1^a \in \mathcal{O}_K^*$ to the choice of U_1. Similarly, if we choose U_2 independent from the previous list, then $U_2^{\sigma^i}$, $0 \leq i \leq p-2$ are also independent and so on. Continuing this way, we get a set of $(r+1)(p-1) + r = r_L$ independent units consisting of ϵ_i's and $r + 1$ blocks of $U_i^{\sigma^j}$.

We can do even better.

Theorem: Units $\overline{U_1}, \cdots, \overline{U_{r+1}} \in \mathcal{O}_L^*$ exist such that if $\prod \overline{U_i}^{F_i(\sigma)} = U^{1-\sigma}\epsilon$, where $U \in \mathcal{O}_L^*$ and where $\epsilon \in \mathcal{O}_K^*$ or $\epsilon \in \mathcal{O}_L^*$ with $\epsilon^p \in \mathcal{O}_K^*$, then p divides $F_i(1)$, for all i.

Such a system $\overline{U_i}$ is called a *fundamental system of relative units*. Note that p divides $F_i(1)$ is equivalent to λ dividing $F_i(\zeta)$ or $1 - \sigma$ dividing $F_i(\sigma)$, since $\zeta^i \equiv 1$ modulo λ. Let us write $[\epsilon]$ as a short-hand for an arbitrary unit of K or a unit of L with its p-th power belonging to K.

Proof: We have found a nice subset of units generating a finite index subgroup of $\mathcal{O}_L^*$. Hence, for large enough p^m, $\prod U_i^{F_i(\sigma)}[\epsilon]$ cannot be a p^m-th power of a unit, unless all the coefficients of $F_i(\sigma)$ (of degree $\leq p - 2$) are divisible by p.

Now $(1 - \sigma)^p = 1 - \sigma^p + pg(\sigma)$, so that $(1 - \sigma)^{pm}$-th symbolic power is actually p^m-th power. So let e_1 be the largest non-negative integer such that $\prod_1^{r+1} U_i^{F_i(\sigma)}[\epsilon] = \overline{U_1}^{(1-\sigma)^{e_1}}$, with $\overline{U_1} \in \mathcal{O}_L^*$ and without having all $F_i(\zeta)$ divisible by λ, say $F_1(\zeta)$ is not divisible by λ. Next, let e_2 be the largest non-negative integer such that $\prod_2^{r+1} U_i^{F_i(\sigma)}[\epsilon] = \overline{U_2}^{(1-\sigma)^{e_2}}$, with $F_2(\zeta)$ not divisible by λ and so on.

Now suppose there are $g_i(\sigma)$ such that $\prod_1^{r+1} \overline{U_i}^{g_\sigma}[\epsilon]$ is $1 - \sigma$-th power of a unit, but with not all $g_i(\sigma)$ being divisible by $1 - \sigma$, say with g_h being first such, so that we can drop the first $h - 1$ terms in the product above without loosing the property. Raise the two sides to $(1 - \sigma)^{e_h}$-th power, so that it is now $(1 - \sigma)^{e_h + 1}$-th power, so substituting the definitions of $\overline{U_i}$ in terms of

U_i, we get that all the exponents are divisible by p, which is a contradiction as U_h has exponent $F_h(1)g_h(1)$ not divisible by p. This proves the Theorem.

Now we construct U as claimed: First note that given $\eta_1, \cdots, \eta_{r+2} \in \mathcal{O}_K^*$, there are integers a_i not all divisible by p such that $\prod \eta_i^{a_i} = 1$ (Since any $r+1$ units are dependent by the Dirichlet theorem, we have $\prod_1^{r+1} \eta_i^{d_i} = 1$. Taking out common p factors if any, we have $\prod_1^{r+1} \eta_i^{c_i} = \zeta^c$. (Here we make a simplifying assumption $\zeta_{p^2} \notin K$, which is true in our case $K = \mathbf{Q}(\zeta_p)$. Without it, we need an additional easy induction argument). Similarly, $\prod_2^{r+2} \eta_i^{b_i} = \zeta^b$, where without loss of generality p does not divide bc. Then the product of the b-th power of the first combination with $-c$-th power of the second is the combination we want).

Now relative norm of ζ is $\zeta^p = 1$, so by Hilbert 90, $\zeta = E^{1-\sigma}$, where without loss of generality E is a unit. (Otherwise, put $U = \zeta$). Now $(E^p)^{1-\sigma} = 1$, so $E^p =: \epsilon \in \mathcal{O}_K^*$ and clearly $E \notin K$. Note that by definition of E, its relative norm is $E^p = \epsilon$.

For $1 \leq i \leq r+1$, let η_i be the relative norm of $\overline{U_i}$ and let $\eta_{r+2} = \epsilon$. Find the integers a_i as above and put $U := \prod \overline{U_i}^{a_i} E^{a_{r+2}}$, so that its relative norm is 1. If it were a $1-\sigma$-th power, by the definition of the fundamental set, p divides a_i for $1 \leq i \leq r+1$. So we can express the norm of U as $1 = (\prod \eta_i^{a_i/p} E^{a_{r+2}})^p$ implying that the bracketed quantity is ζ^b and hence $E^{a_{r+2}} \in K$. This would give a contradiction $E \in K$, since p does not divide a_{r+2}. Hence U has the property we want and the Theorem and the second proof of Kummer's lemma is complete.

Since the class number is difficult to calculate, Kummer gave an easy *criterion for checking regularity.* He proved (1) If p divides h^+, the class number of $\mathbf{Q}(\zeta)^+$, then p divides $h^- := h/h^+$ and (2) The prime p divides h^- if and only if p divides one of the Bernoulli numbers B_j for some even j between 2 and $p-3$. By (1), this condition holds if and only if p is irregular.

For (1), we will see an algebraic proof in Sujatha's lecture, which is based on the Spiegelgungsatz or the reflection theorem (see also Washington 10.2 or Lang) obtained by comparing the class field theory information with Kummer theory information, giving switch of the parity (sign) of the characters. A proof involving p-adic class number formula and another proof (apparently close to Kummer's original proof) of Kummer's lemma based on similar considerations were presented in the workshop and can be found in Washington 5.6 (pp. 77-81).

For (2), again we have seen an algebraic proof as a corollary to Stickelbeger-Herbrand theorem in Katre's lectures. Another way is to note that the units of K and K^+ being closely related, the regulators are essentially the same so taking the ratio of analytic class number formulas for K and K^+, we get a formula for h^- in terms of special values of L-functions, which are Bernoulii numbers, as explained in Raghunathan's lectures. Thus we get $h^- = 2p\prod_{\chi \text{ odd}}(-B_{1,\chi}/2)$ which is congruent to $\prod(-B_j/2j)$ modulo p, where j runs through even integers between 2 and $p-3$. For details, see Washington Theorem 4.17.

We end by making some *miscellaneous comments*: For $n = 2$, we had a *parametric solution*, in fact. This is possible, since the corresponding curve is of genus zero. Since for $n > 2$, the equation represents a curve of *genus* $(n-1)(n-2)/2 > 0$, the parametric solutions are impossible. Hard part is to show that not a single non-trivial solution is possible.

More than 10 years before Wiles, Faltings proved *Mordell conjecture* that non-singular curves of genus more than one over a number field can have only finitely many solutions in a number field. This proves that for each $n > 3$, the Fermat equation with the exponent n can have only finitely many solutions up to scaling (as we have to dehomogenize), but in any number field. So in some respects, it is a stronger result.

In fact, even before Faltings proved this, Arizona undergraduate Filaseta noticed that as an easy consequence of Mordell conjecture, given any n, there is M_n such that Fermat equation for exponent nk with $k > M_n$ has no non-trivial integral solutions (Exercise: Note that a solution for exponent nk gives a solution for exponent n). This, coupled with careful counting, led Heath-Brown and Granville to conclude from Faltings result that for almost all (i.e., density 1) exponents n, FLT has no non-trivial integral solutions.

My colleague Bill McCallum has given another proof of the second case of Fermat for regular primes p, by using geometric and p-adic techniques (so-called Coleman-Chabauty method), which seem to have more potential. Interestingly, the techniques only imply existence of at most $p-3$ primitive solutions in the first case, the case which is supposed to be easier.

The *abc conjecture* states that given $\epsilon > 0$, there is $C_\epsilon > 0$, such that for any non-zero relatively prime integers a, b, c with $a + b = c$, we have $\max(|a|, |b|, |c|) \leq C_\epsilon(\prod_{p|abc} p)^{1+\epsilon}$. This is believed because of analogies with function fields (See Lang's Algebra) and is stronger than Mordell conjecture (this implication was proved by Elkies). It is easy to see that it implies FLT for large enough n.

Fermat's false proof?: It is well-known that Fermat wrote in the margin that he had a 'truly marvelous' proof, but in public only repeated weaker claims. So most probably he quickly found a mistake in his original false argument. Can it be the following infinite descent method attempt (the method that he was so fond of and used successfully for $n = 4$)?

Consider $a^n + b^n = c^n$, where $n \geq 3$ odd and $c \neq 0$ is even. Set $x + y = a^n$, $x - y = b^n$. Then $(x^2 - y^2)/(4x^2) = (ab/c^2)^n =: (z/x)^n$ and we get $x(x^n - 4z^n) = x^{n+1} - 4xz^n = x^{n-1}y^2$, which is a square.

Let d be a GCD of x and z, let $x' = x/d$, $z' = z/d$, so that $d^{n+1}x'(x'^n - 4z'^n)$ and so $x'(x'^n - 4z'^n)$ are squares. The last two factors are coprime, so $x' = p^2$, $x'^n - 4z'^n = q^2$ and now $p^{2n} - q^2 = (p^n + q)(p^n - q) = 4z'^n$.

The GCD of p^n+q and p^n-q is 2. Therefore $p^n+q = 2r^n$ and $p^n-q = 2s^n$, so that $p^n = r^n + s^n$ is another solution!

It seems we can apply descent (or even ascent, thanks to Faltings), but the problem is that the 'new' solution is one we started with!

This argument, attributed to Lexell, appears in Euler, Opera Postuma Math. et Physica (1862), vol. 1, 231-232.

Acknowledgements: I thank Andrew Granville for faxing me a copy of this argument from Euler dug out by Bombieri. I thank Pavlos Tzermias for carefully going through the manuscript. We have borrowed from many books which discuss Fermat for regular primes, e.g., those by Washington, Edwards, Lang, Landau, Hilbert, Borevich-Shafarevich, Ribenboim. See also Rosen's article in 'Modular forms and Fermat's last theorem' volume.

Dinesh Thakur
Department of Mathematics
University of Arizona
Tucson, AZ 85721
USA
e-mail: thakur@math.arizona.edu

Reciprocity Laws: Artin-Hilbert

Parvati Shastri

1. Introduction

In this volume, the early reciprocity laws including the quadratic, the cubic have been presented in Adhikari's article [1]. Also, in that article we have seen an exposition on the Eisenstein reciprocity without the use of class field theory. In the International Congress of Mathematicians, 1900, Hilbert asked for the most general reciprocity law, (Hilbert's problem 9) which would hold in any number field. In order to formulate and prove such a general reciprocity law, Hilbert introduced the *norm residue symbol* known after him as the Hilbert Symbol, in place of the power residue symbol and proved a reciprocity law for this symbol. From this law, one can derive all the earlier power reciprocity laws. Later on, Artin introduced a symbol named after him as the Artin Symbol, and proved a reciprocity law for his symbol. This is the crux of class field theory, today. In this article, we shall assume but recall the relevent theorems from class field theory, and deduce Hilbert's reciprocity law and show how this would imply the power reciprocity laws, that you have seen earlier. Further, we will also explicitly derive the quadratic reciprocity law, without the use of class field theory. We begin with Kummer Theory.

2. Kummer Theory

Theorem 1 *Let K be a field of characteristic 0, and $\mu_n \subset K$, be the group of all n^{th} roots of unity.*[1] *Let Δ be a subgroup of K^* such that $K^{*n} \subset \Delta \subset K^*$ and $L = K(\sqrt[n]{\Delta})$. Then L/K is a Galois extension (in fact, abelian of exponent n) and there exists a canonical isomorphism* [2]

$$\frac{\Delta}{K^{*n}} \longrightarrow Hom(G(L/K), \mu_n)\,.$$

Proof: (Sketch) Assume K^*/K^{*n} is finite. Let $G = \text{Gal}(L/K)$ and let $a \in \Delta$. Define

$$\chi_a : G \longrightarrow \mu_n$$

[1] This result holds also in characteristic $p > 0$ under the additional hypothesis that, $(char K, n) = 1$.

[2] If K^*/K^{*n} is infinite, then in the case when Δ/K^{*n} is also infinite, or equivalently, L/K is infinite, one needs to take continuous homomorphisms. However, we need to apply this result in the case when K^*/K^{*n} is finite, and we sketch a proof for this case only.

by

$$\chi_a(\sigma) = \frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}}.$$

We have a homomorphism

$$\theta : \Delta \longrightarrow \mathrm{Hom}(G, \mu_n)$$

defined by

$$\theta(a) = \chi_a.$$

It is clear that $\ker \theta = K^{*n}$. We claim that θ is surjective. Let $\chi \in \mathrm{Hom}(G, \mu_n)$. Then by Dedekind's theorem on the linear independence of characters, the sum $\sum_{\sigma\in G} \chi(\sigma)\sigma^{-1} \neq 0$. Let $\lambda \in L^*$ be such that $b = \sum_{\sigma\in G} \chi(\sigma)\sigma^{-1}(\lambda) \neq 0$. For $\tau \in G$, we have,

$$\begin{array}{rcl} \tau(b) & = & \tau\left(\sum_{\sigma\in G} \chi(\sigma)\sigma^{-1}\right)(\lambda) \\ & = & \sum_{\sigma\in G} \chi(\sigma)(\tau\sigma^{-1})(\lambda) \\ & = & \chi(\tau)\left(\sum_{\sigma\in G} \chi(\tau^{-1}\sigma)\tau\sigma^{-1}\right)(\lambda) \\ & = & \chi(\tau)b. \end{array}$$

Let $a = b^n$. Then $a \in K^n$ and we get $\theta(a) = \chi$. Therefore we get an induced isomorphism,

$$\frac{\Delta}{K^{*n}} \cong \mathrm{Hom}(G, \mu_n).$$

Remark 1 The above correspondence gives a bijection between subgroups Δ of K^* with $K^{*n} \subset \Delta \subset K^*$ and abelian extensions of exponent n. (For a proof of this statement as well as for generalization to infinite extensions, the reader can refer to [4], Chapter 8, or [3], p.15. See also [2], §4, this volume.)

3. Local Reciprocity Law

Let K be a local field of characteristic 0. By this we mean, a finite extension of $\mathbb{Q}_p$. For any local field K, we fix the following notations.

$\mathcal{O}_K$	ring of integers of K
m_K	the maximal ideal of K
π_K	a generator for m_K
U_K	group of units of K
$\kappa(K)$	the residue class field $\mathcal{O}_K/m\mathcal{O}_K$.

The following is the main theorem of local class field theory.

Theorem 2 *Let $L|K$ be a finite abelian extension of local fields with Galois group $G(L|K)$. Let $N_{L|K} : L \longrightarrow K$ denote the norm map. Then there exists a canonical isomorphism*

$$r_{L|K} : G(L|K) \longrightarrow K^*/N_{L|K}L^*$$

We will not prove this theorem, but briefly discuss how the isomorphism is defined. First assume that $L|K$ is unramified. Then $r_{L|K}$ can be described as follows. We know that $K^* = U_K \times (\pi_K)$, where (π_K) is the (multiplicative) cyclic group generated by π_K. Since $L|K$ is unramified, $N_{L|K}$ is surjective on the unit group, that is, $N_{L|K}(U_L) = U_K$. Clearly $N_{L|K}(\pi) = \pi^n$, where $n = [L : K]$. It follows that $K^*/N_{L|K}(L^*) \cong (\pi)/(\pi^n)$, since $K^{*n} \subset N_{L|K}(L^*)$. On the other hand, if $L|K$ is unramified, the Galois group $G(L|K)$ is isomorphic to the Galois group of the residue classfield extension $\kappa(L)|\kappa(K)$. This is a finite extension of a finite field. Recall that a finite extension of a finite field is cyclic and there is a distinguished generator for the Galois group, viz., the Frobenius. So there is a unique generator of $G(L|K)$ which corresponds to the Frobenius. Let us denote this generator by ϕ.[3] The reciprocity map $r_{L|K}$ is given by,

$$r_{L|K}(\phi) = \pi \bmod N_{L|K}(L^*).$$

In the general case, $r_{L|K}$ is defined, subject to the following two properties:

(i) (Functoriality) If $L|K$ and $L'|K'$ are finite Galois extensions of local fields with $K \subset K', L \subset L'$, then the diagram

$$\begin{tikzcd}
G(L'|K') \arrow[r, "r_{L'|K'}"] \arrow[d, "\mathrm{Res}"'] & K'^*/N_{L'|K'}L'^* \arrow[d, "N_{K'|K}"] \\
G(L|K) \arrow[r, "r_{L|K}"] & K^*/N_{L|K}L^*
\end{tikzcd}$$

is commutative, where Res denotes the restriction map $\mathrm{Res}(\sigma) = \sigma|L \ \forall \sigma \in G(L'|K')$.
(ii) If $L|K$ is a finite unramified extension, then $r_{L|K}$ is simply the map $r_{L|K}(\phi_{L|K}) = [\pi]$, where $[\pi]$ is the class of π mod $N_{L|K}(L^*)$.

Let $L|K$ be a totally ramified cyclic extension and σ be a generator for $G(L|K)$. Then one can show that there exists a finite abelian extension Σ

[3] ϕ is characterized by the property,

$$\phi(x) \equiv x^q \bmod \pi \ \forall \ x \in \mathcal{O}_K,$$

where q is the cardinality of $\kappa(K)$.

of K such that $L\Sigma|\Sigma$ is unramified and the restriction of the Frobenius of $L\Sigma|\Sigma$ to L is σ. Then $r_{L|K}(\sigma) = N_{\Sigma|K}(\pi_\Sigma)$. The general case is reduced to the cyclic case. Also this map is surjective and the kernel of this map is precisely the commutator subgroup $[G, G]$.

Thus, there is a canonical (i.e., satisfying (i) and (ii) above) isomorphism

$$r_{L|K} : G(L|K)^{\text{ab}} \longrightarrow K^*/N_{L|K}(L^*).$$

In particular, for finite *abelian extensions*, the Galois group $G(L|K)$ is isomorphic to the norm residue group $K^*/N_{L|K}(L^*)$.

4. Local Artin Symbol

Let the notation be as in section 2. Let $(*, L|K)$ be the inverse of the reciprocity map. By composing it with the natural map $K^* \longrightarrow N_{L|K}(L^*)$, we get for every $a \in K^*$, a symbol which we still denote by $(a, L|K)$ taking values in $G(L|K)$. This is called the Artin symbol; i.e., the local Artin symbol is induced by the inverse of the local reciprocity map.

Observe that we have the following simple description of the Artin symbol in the special cases $a = \pi,\ u$ where π is a parameter and u is a unit in K, viz.,

$$(\pi, L|K) \text{ is \quad the \ Frobenius} \in G(L|K)$$

and

$$(u, L|K) = 1.$$

5. Hilbert Symbol

We now define the Hilbert Symbol. Let μ_n be the group of n^{th} roots of unity. Assume that K is a local field containing μ_n. We have, by Kummer Theory,

$$K^*/K^{*n} \cong \operatorname{Hom}\left(G(L|K), \mu_n\right),$$

where $L = K(\sqrt[n]{K^*})$. (Note that K^*/K^{*n} is finite, since K is a local field.) On the other hand, local reciprocity law gives an isomorphism,

$$K^*/N_{L|K}(L^*) \cong G(L|K).$$

Since, $K^{*n} \subset N_{L|K}(L^*) \subset K^*$, it follows that $K^{*n} = N_{L|K}L^*$. Hence we get a pairing,

$$\langle\ ,\ \rangle_n : K^*/K^{*n} \times K^*/K^{*n} \longrightarrow \mu_n$$

given by

$$\langle a, b\rangle_n = \chi_b\left((a, L|K)\right),$$

where χ_b is the character associated to b by Kummer Theory, and $(a, L|K)$ is the local Artin symbol. By the properties of the Artin symbol, it follows that

$$\langle a, b\rangle_n = \chi_b\left((a, L|K)\right) = \frac{(a, L|K)\left(\sqrt[n]{b}\right)}{\sqrt[n]{b}} = \frac{\left(a, K(\sqrt[n]{b})|K\right)\left(\sqrt[n]{b}\right)}{\sqrt[n]{b}} \in \mu_n.$$

By composing it with the natural map $K^* \to K^*/K^{*n}$, we get a pairing

$$\langle\ ,\ \rangle_n : K^* \times K^* \longrightarrow \mu_n.$$

This is called the *Hilbert symbol* of degree n. In what follows, we will fix an n, and drop the suffix n.

Remark 2 It follows easily by the definition that the Hilbert symbol is non degenerate in the sense that,

$$\langle a, b\rangle = 1 \ \forall\ b \in K^* \Rightarrow a \in K^{*n}$$

and

$$\langle a, b\rangle = 1 \ \forall\ a \in K^* \Rightarrow b \in K^{*n}.$$

We now recall a few basic properties of the Hilbert symbol, which are needed in the sequel.

Lemma 1 The Hilbert symbol has the following properties:

(i) (Bimultiplicativity)
$\langle aa', b\rangle = \langle a, b\rangle.\langle a', b\rangle,\ \langle a, bb'\rangle = \langle a, b\rangle.\langle a, b'\rangle \ \forall\ a, b \in K^*.$

(ii) $\langle 1-a, a\rangle = 1 = \langle a, 1-a\rangle \ \forall\ a \in K^*.$

(iii) $\langle a, b^{-1}\rangle = \langle a, b\rangle^{-1} = \langle a, -a\rangle = 1 = \langle a, 1\rangle.$

(iv) (Skew symmetry) $\langle a, b\rangle = \langle b, a\rangle^{-1}.$

Proof: Part (i) follows by definition (easy to check). For proving Parts (ii), (iii) and (iv), observe that, for $a, b \in K^*$, $\langle a, b\rangle = 1$ if and only if a is a norm from $K(\sqrt[n]{b})$.
We have,

$$1 - a = \prod_{i=1}^{n}\left(1 - \zeta^i \sqrt[n]{a}\right),$$

where ζ is a primitive n^{th} root of unity, i.e., $1-a$ is a norm from $K(\sqrt[n]{a})$. So (ii) follows.

Next, by (i) we have, $\langle a,1\rangle.\langle a,1\rangle = \langle a,1\rangle$. Hence $\langle a,1\rangle = 1$. Similarly, $\langle a,b\rangle.\langle a,b^{-1}\rangle = \langle a,bb^{-1}\rangle = \langle a,1\rangle = 1$ Now, observe that $-a = \frac{1-a}{1-a^{-1}}$. Therefore, if we take $b=-a$, we get $\langle a,-a\rangle = 1$. This completes the proof of Part (iii).

(iv) By (iii) we have, $\langle ab,-ab\rangle = 1$. Now use bimultiplicativity and simplify using (iii) to get (iv).

6. Power Residue Symbol

We now assume that $(n,p)=1$ where p is the characteristic of the residue class field and compute the Hilbert symbols $\langle u,v\rangle$ and $\langle \pi,u\rangle$, where u,v are units and π is a parameter of K. It follows from standard facts of local theory, that $K(\sqrt[n]{v})$ is unramified, and that the norm function is surjective on the unit group. Therefore $\langle u,v\rangle = 1$ ∀ units $u,v\in K$. Also, by the discussions in Section 4, we know that the parameter corresponds to the Frobenius under the reciprocity map. Thus

$$(\pi, K(\sqrt[n]{u})|K)\,(x) \equiv x^q \bmod \pi \ \forall\ x\in\mathcal{O}_L,$$

where $L=K(\sqrt[n]{u})$. In particular, $(\pi,K(\sqrt[n]{u})|K)\,(\sqrt[n]{u}) \equiv \sqrt[n]{u}^q \bmod \pi$ $\equiv u^{\frac{q-1}{n}}.\sqrt[n]{u} \bmod \pi$. So $\langle\pi,u\rangle \equiv u^{\frac{q-1}{n}} \bmod \pi$. We define the n^{th} *power residue symbol*, by

$$\left(\frac{u}{\pi}\right) = \langle \pi,u\rangle.$$

Note that $\langle\pi,u\rangle$ is a root of unity in K and is independent of the parameter chosen.

7. Artin's Reciprocity Law

Let K be a number field. Let V_K be the set of all valuations of K including the archimedean ones. Let $L|K$ be a finite abelian extension of K. For every valuation $v\in V_K$, fix a valuation w of L, which extends v. Note that the archimedean completions of K_v are isomorphic to either $\mathbb{R}$ or $\mathbb{C}$. Therfore either $L_w\cong K_v\cong\mathbb{R}$ or $\mathbb{C}$ or $L_w\cong\mathbb{C}$ is a quadratic extension of $K_v\cong\mathbb{R}$, with Galois group isomorphic to $\mathbb{Z}/2\mathbb{Z}$. In order to state Artin's reciprocity law, we need to define Artin's symbol at the archimedean completions also. If $L_w|K_v$ is quadratic, let σ be the nontrivial automorphism of $L_w|K_v$. We define $(a,L_w|K_v)=1$, if $a>0$, $(a,L_w|K_v)=\sigma$ if $a<0$. If $L_w\cong K_v$ we define $(a,L_w|K_v)=1$. With this notation, we have,

Theorem 3 *For any $a \in K^*$,*

$$\prod_{v \in V_K} (a, L_w|K_v)) = 1.^4$$

8. Hilbert's Reciprocity Law

As in the case of Artin symbol, we also need to extend Hilbert's symbol at the archimedian completions, in an obvious manner. For $a, b \in \mathbb{R}$, define

$$\langle a, b \rangle := (-1)^{\frac{\operatorname{sgn} a - 1}{2} \cdot \frac{\operatorname{sgn} b - 1}{2}}.$$

With this definition, we have

Theorem 4 *Let K be a number field, $\mu_n \subset K$, and V_K be as in the previous section. Let $a, b \in K^*$. Then*

$$\prod_{v \in V_K} \langle a, b \rangle_v = 1,$$

where $\langle a, b \rangle_v$ is the n^{th} Hilbert symbol at the completion K_v.

Proof: This is immediate from Artin's reciprocity law. In fact, we have,

$$\begin{array}{rcl} \prod_{v \in V_K} \langle a, b \rangle_v & = & \prod_v \chi_b \left(a, K_v(\sqrt[n]{b}) \right) \\ & = & \frac{\left(\prod_v \left(a, K_v(\sqrt[n]{b})\right)\right)(\sqrt[n]{b})}{\sqrt[n]{b}} \\ & = & \frac{\mathrm{Id}(\sqrt[n]{b})}{\sqrt[n]{b}} \\ & = & 1. \end{array}$$

(The last but one equality is by Artin's reciprocity law.)

9. Power Reciprocity Law

We need to define global power residue symbol, in terms of the local symbols. Let K be a number field containing μ_n, as in the previous section. Let $a, b \in K$. Let $(b) = \prod_{\mathfrak{p}} \mathfrak{p}^{n_i}$ be the factorization of the principal ideal (b). Let $v_{\mathfrak{p}}$ be the valuation on K corresponding to the prime ideal $\mathfrak{p}$. We define the power residue symbol $\left(\frac{a}{b}\right)$ to be the product of the local power residue symbols; i.e.,

$$\left(\frac{a}{b}\right) = \prod_{\mathfrak{p}} \left(\frac{a}{b}\right)_{v_{\mathfrak{p}}}.$$

[4] Note that this would also mean that all but finitely many terms in this product are equal to 1.

This is well defined, since the local power residue symbol is independent of the parameter chosen.

Theorem 5 *Let K be a number field, $\mu_n \subset K,\ a, b \in K^*$. Let $(a), (b), (n)$ be relatively prime and let*

$$V_{n\infty} = \{v_{\mathfrak{p}} : \mathfrak{p}|(n)\} \cup \{v : v \textit{ is archimedian}\}.$$

Then

$$\left(\frac{a}{b}\right)\left(\frac{b}{a}\right)^{-1} = \prod_{v \in V_{n\infty}} \langle a, b\rangle_v.$$

Proof: Let us look at Hilbert's reciprocity law,

$$\prod_{v \in V_K} \langle a, b\rangle_v = 1.$$

Let $V_0 = \{v \in V_K : v|abn\infty\}$ and $V_1 = V_K - V_0$. The left hand side can be written as

$$\prod_{v|b}\langle a, b\rangle_v \prod_{v|b}\langle a, b\rangle_v \prod_{v|n\infty} \langle a, b\rangle_v \prod_{v \in V_1} \langle a, b\rangle_v.$$

Now, observe that

$$\prod_{v \in V_1} \langle a, b\rangle_v = 1,$$

since each of the symbols is trivial. Also,

$$\langle a, b\rangle_v = \left(\frac{b}{a}\right)_v \text{ if } v|a$$

and

$$\langle a, b\rangle_v = \langle b, a\rangle_v^{-1} = \left(\frac{a}{b}\right)_v^{-1} \text{ if } v|b.$$

Hence the theorem follows.

10. Quadratic Reciprocity Law

In order to derive the quadratic reciprocity law, we need to compute the Hilbert symbols in the special case, $n = 2, K = \mathbb{Q}$. First assume a, b are odd positive integers. Then we have,

$$\left(\frac{a}{b}\right)\left(\frac{b}{a}\right) = \langle a, b\rangle_2 \langle a, b\rangle_\infty.$$

Here $\langle a, b\rangle_2$ denotes the Hilbert symbol at the dyadic completion $\mathbb{Q}_2$. Since a, b are positive, $\langle a, b\rangle_\infty = 1$. So, we only need to compute $\langle a, b\rangle_2$. Let $U_{\mathbb{Q}_2}$

be the unit group of $\mathbb{Q}_2$. Note that $U_{\mathbb{Q}_2}/U^2_{\mathbb{Q}_2}$ is generated by $\{5, -1\}$.[5] By the bimultiplicative and skew symmetric properties of the Hilbert symbol, it is enough to compute $\langle 5, -1\rangle_2$, $\langle 5, -5\rangle_2$, $\langle -1, -1\rangle_2$. We have, $\langle 5, -1\rangle_2 = \langle 5, 5\rangle_2 = 1$ and $\langle -1, -1\rangle_2 = -1$. From this it follows that

$$\left(\frac{a}{b}\right) \cdot \left(\frac{b}{a}\right) = (-1)^{\frac{a-1}{2}}(-1)^{\frac{b-1}{2}},$$

for $a, b \in \{5, -1\}$. For arbitrary odd integers, the result follows from the multiplicativity of these symbols and the fact that for an odd integer a, $a^2 \equiv 1 \bmod 4$.

REFERENCES

1. S.D.Adhikari,*The Early Reciprocity Laws: From Gauss to Eisenstein*, This volume.
2. M. J. Narlikar, *Abelian Kummer Theory*, This volume.
3. J.Neukirch, Class Field Theory, Springer-Verlag, 1986.
4. V. Suresh, Chapter 8 in *Introduction to Class Field Theory*, (Lecture Notes of the Instructional School on Algebraic Number Theory, held in the Department of Mathematics, University of Mumbai, December 1994-January 1995).
5. J. Tate, Problem 9: The General Reciprocity Law, Proceedings of Symp. in Pure Math. Vol. 28, 1976, pp. 311-322.

Parvati Shastri
Department of Mathematics
University of Mumbai
Mumbai 400 098
e-mail: parvati@math.mu.ac.in

[5] This is a consequence of Hensel's lemma.

Main Conjecture of Iwasawa Theory

C. S. Rajan

Abstract. 1) Weil conjectures 2) Iwasawa's theorem on growth of class groups 3) Iwasawa's construction of p-adic L-functions via Stickelberger elements 4) Main conjecture.

1. Weil Conjectures

1.1. Zeta and L-functions. We recall the Riemann zeta function

$$\zeta(s) = \sum_{n \geq 1} n^{-s} \qquad (\mathrm{Re}(s) > 1).$$

The above series converges absolutely for $\mathrm{Re}(s) > 1$ and defines an analytic function there. $\zeta(s)$ admits an analytic continuation to the entire complex plane except for a pole at $s = 1$. The zeta function and its generalisations enjoy remarkable analytic properties, which have significant consequences to arithmetic. One of the outstanding problems concerned with the zeta function is the Riemann Hypothesis:

Riemann Hypothesis(RH): If ρ is any zero of $\zeta(s)$ with $\mathrm{Re}(\rho) \geq 0$ (such zeros are called the non-trivial zeros of $\zeta(s)$), then

$$\mathrm{Re}(\rho) = 1/2.$$

i.e., all the zeros to the right of the line $\mathrm{Re}(s) = 0$ lie on the line $\mathrm{Re}(s) = 1/2$.

How does one tackle this conjecture? One idea going back to Hilbert stems from the fact that the eigenvalues of a hermitian (skew-hermitian) matrix are real (purely imaginary). In other words the eigenvalues of a hermitian matrix all lie on a straight line. This leads to the following question:

Question 1.1.1. Is it possible to find a skew-hermitian operator acting on some space, such that the zeros of $\zeta(s+1/2)$ with non-negative real part, occur as eigenvalues of this operator?

Very little is known about this question. Recently some exciting numerical computations show that the known zeros of the $\zeta(s)$ behave like the eigenvalues of a random hermitian matrix. See works of Montogomery, Odlyzko, Katz, Sarnak and others.

The solution to many arithmetical problems lie in the analytic properties of the zeta function. It seems profitable then to look at analogous situations, also with a view to throw some light on the Riemann Hypothesis. One generalisation is to number fields. The Dedekind zeta function of a number field K is,

$$Z_K(s) = \sum_{\mathfrak{a} \neq 0} (N\mathfrak{a})^{-s},$$

where $\mathfrak{a}$ runs over the non-zero ideals of the ring of integers $\mathcal{O}_K$ of K, and $N\mathfrak{a} = |\mathcal{O}_K/\mathfrak{a}|$ is the number of elements in the ring $\mathcal{O}_K/\mathfrak{a}$. $Z_K(s)$ is defined as above for $\mathrm{Re}(s) > 1$, and has properties similar to the Riemann zeta function. The key fact that enables us to define the zeta function is that for a non-zero ideal $\mathfrak{a}$ in $\mathcal{O}_K$, $\mathcal{O}_K/\mathfrak{a}$ is a finite ring. It can be shown that $Z_K(s)$ converges absolutely, and hence defines an analytic function in the half plane, $\mathrm{Re}(s) > 1$.

In order to generalise the definition of the zeta function to a more general context, it is easier to work with the Euler product expansion of the zeta function. Recall that for $\mathrm{Re}(s) > 1$, the zeta function can be expressed as a product,

$$Z_K(s) = \prod_{\mathfrak{m}} (1 - (N\mathfrak{m})^{-s})^{-1},$$

where $\mathfrak{m}$ runs over the maximal ideals of $\mathcal{O}_K$. Here we remark, although it is not essential for what follows, the fact that the above expression defines the zeta function is another way of expressing the unique factorisation of an ideal into a product of prime ideals, generalising the unique factorisation property of integers.

More generally let $f_1(x_1, \cdots, x_n), \cdots, f_r(x_1, \cdots, x_n)$ be polynomials with integral coefficients in the ring $\mathbb{Z}[x_1, \cdots, x_n]$. Let I be the ideal in $\mathbb{Z}[x_1, \cdots, x_n]$ generated by $f_1, \cdots, f_n$ and let A be the ring,

$$A = \mathbb{Z}[x_1, \cdots, x_n]/I.$$

The ring A can be thought of as the ring of polynomial functions restricted to the variety X defined by the common zeros of the polynomials $f_1, \cdots, f_n$, or equivalently of the common zeros of polynomials in the ideal I:

$$X = \{(a_1, \cdots, a_n) \in \mathbb{C}^n \mid f((a_1, \cdots, a_n)) = 0, \ \forall f \in I\}.$$

It follows from the finite generation of A as an algebra over $\mathbb{Z}$, that the following are equivalent:

i) $\mathfrak{m}$ is a maximal ideal of A.

ii) $A/\mathfrak{m}$ is a finite field.

Let X' denote the collection of maximal ideals of A. One can define the zeta function,

$$Z(X, s) = \prod_{\mathfrak{m} \in X'} (1 - (N\mathfrak{m})^{-s})^{-1}.$$

Here we assume that I is not the unit ideal. It can be shown that $Z(X,s)$ converges absolutely in some half plane, and thus defines an analytic function there.

Example 1. Let L/K be a finite Galois extension of number (global) fields with Galois group G. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$ unramified in L, and let $\mathcal{P}$ be a prime ideal of $\mathcal{O}_L$ dividing $\mathfrak{p}$. Then the Frobenius element $\sigma_{\mathcal{P}}$ is given by,

$$\sigma_{\mathcal{P}}(x) \equiv x^{N\mathfrak{p}}(mod\ \mathcal{P}),$$

where $x \in \mathcal{O}_L$. For a fixed $\mathfrak{p}$, the Frobenius elements $\sigma_{\mathcal{P}}$ for $\mathcal{P}|\ \mathfrak{p}$ form a conjugacy class inside G. Let ρ be a finite dimensional representation of G into $GL(n,\mathbb{C})$. The incomplete Artin L-function associated to ρ is defined by,

$$L'(s,\rho) = \prod_{\mathfrak{p}} \det(1-\rho(\sigma_{\mathfrak{p}})N\mathfrak{p}^{-s})^{-1}, \mathrm{Re}(s) > 1.$$

Here the product is over the unramified primes of K with respect to L. It is possible to define the factors at the ramified primes, in order that the completed L-function has an analytic continuation to the entire plane and satisfies a suitable functional equation.

In particular, take $K=\mathbb{Q}$, and $L=\mathbb{Q}(\zeta_n)$, where ζ_n is a primitive n^{th} root of unity. Then $G \simeq (\mathbb{Z}/n\mathbb{Z})^*$. Let χ be a character of $(\mathbb{Z}/n\mathbb{Z})^*$, a Dirichlet character. Let

$$L(s,\chi) = \prod_{(p,n)=1} (1-\chi(p)p^{-s})^{-1}, \mathrm{Re(s)}> 1.$$

These are the L-functions considered by Dirichlet.

1.2. Arithemetical applications. a) *Dirichlet's theorem and Prime number theorem.* The non-vanishing of $L(s,\chi)$-functions at $s=1$ when χ is a Dirichlet character, imply Dirichlet's theorem on infinitely many primes of the form $an+b,\ (a,b)=1,\ n\in\mathbb{N}$. More generally, the non-vanishing of $L(s,\chi)$ on the line $\mathrm{Re}(s)=1$ for all characters $\chi : (\mathbb{Z}/n\mathbb{Z})^* \to \mathbb{C}^*$, imply for any $a,\ b,\ (a,b)=1$ that,

$$\#\{p \le x \mid p = am+b, \text{ for some } m\} \sim \frac{1}{\phi(a)}\frac{x}{\log x}, \text{ as } x\to\infty,$$

where ϕ is Euler ϕ-function. In particular, when we take χ to be the trivial character, we obtain the prime number theorem, viz., that the number of primes of size at most x, grows asymptotically like $x/\log x$ as $x\to\infty$.

b) *Class number formulas.* It can be seen that the Dedekind zeta function $Z_K(s)$ has a meromorphic continuation to the entire plane, which is holomorphic except at $s=1$. Let h denote the class number of K, R the regulator of K, D the absolute value of the discriminant of K, r_1 the number of real embeddings of K, $2r_2$ the number of complex embeddings, and w the number of roots of unity in K. Then the class number formula is,

$$\text{res}_{s=1} Z_K(s) = \frac{2^{r_1}(2\pi)^{r_2} hR}{w\sqrt{D}}.$$

Remark. The main theme in defining these zeta functions, is the local-global principle. The various Euler factors encode the arithmetical information at the various primes, and global arithmetical consequences arise from the analytical aspects of the global L-functions. The fact that these L-functions have suitable analytical properties, can be considered as a vast generalisation of the classical quadratic and higher reciprocity laws. However even the analytic properties similar to that enjoyed by $\zeta(s)$ are not known for these L-functions. Again these analytic properties have remarkable arithmetical applications. For example, take I to be generated by the polynomial $f(x, y) = y^2 - x^3 + ax + b,\ a,\ b \in \mathbb{Z}$, in the ring $\mathbb{Z}[x, y]$. X defines an elliptic curve minus the point at infinity. The analytic continuation of $Z(X, s)$ and related objects have been established by Wiles, and by Ribet's theorem imply Fermat's conjecture. Thus these generalisations seem to be even more difficult to handle than $\zeta(s)$!

1.3. Finite fields. However if we assume that the ideal I contains a prime number, the problem becomes tractable. In this case we are essentially replacing the base ring $\mathbb{Z}$ by a finite field $\mathbb{F}_q$ consisting of q elements. Again take I to be an ideal in the polynomial ring $\mathbb{F}_q[x_1, \cdots, x_n]$, and let $A = \mathbb{F}_q[x_1, \cdots, x_n]/I$ be the ring of polynomial functions on the variety $X = V(I)$ defined by I. As above, we can define the zeta function,

$$Z_X(s) = \prod_{\mathfrak{m} \in X'} (1 - (N\mathfrak{m})^{-s})^{-1}.$$

Again this can be shown to make sense.

We will now translate the above definition for the zeta function, into a form which counts number of common solutions of polynomials in I, over finite extensions of the base field $\mathbb{F}_q$. To do this, suppose that $(a_1, \cdots, a_n)$ is a common solution of the polynomials $f_1, \cdots, f_r$ in some field k containing $\mathbb{F}_q$. Define a ring homomorphism $A \to k$, by sending the generators x_i to the element a_i. This prescription allows us to define a bijection between the set of solutions of $f_1, \cdots, f_r$ (equivalently the set of common solutions of the polynomials in the ideal generated by $f_1, \cdots, f_r$) in k, and the collection of ring homomorphisms from $A \to k$.

Let X_m be the set of solutions with values in $\mathbb{F}_{q^m}$. Identifying as above with ring homomorphisms, we see that this is the same as giving a maximal ideal $\mathfrak{m} \in X'$, and an embedding f of the finite field $f : A/\mathfrak{m} \to \mathbb{F}_{q^m}$, which is identity on the base field $\mathbb{F}_q$.

Being a subset of $\mathbb{F}_{q^m}^n$, X_m is finite. Define

$$\nu_m = |X_m|.$$

Let μ_l be the number of maximal ideals $\mathfrak{m}$ in X' such that the cardinality of the residue field $N\mathfrak{m} = q^l$.

EXERCISE 1.3.1. Show that $\nu_m = \sum_{l|m} l\mu_l$.

An alternate expression for the zeta function in terms of the number of solutions of varieties over finite fields is the following:

$$\log Z(X, s) = \sum_{m \geq 1} \frac{\nu_m q^{-ms}}{m}.$$

PROOF. Exercise. *Hint:* The number of embeddings of $\mathbb{F}_{q^l}$ over $\mathbb{F}_q$ into $\mathbb{F}_{q^m}$ is $l = [\mathbb{F}_{q^l} : \mathbb{F}_q]$, where we assume that $l|m$. □

We can reformulate by substituting $t = q^{-s}$, to obtain a formal power series

$$\log Z(X, t) = \sum_{m \geq 1} \frac{\nu_m t^m}{m}.$$

We now look at some examples.

EXERCISE 1.3.2. 1) X a point. Then $\nu_n = 1$ for all n, and

$$Z(X, t) = \exp\left(\sum_{n \geq 1} \frac{t^n}{n}\right) = \frac{1}{1-t}.$$

2) Take X to be defined by the zero ideal in the polynomial ring $k[x_1, \cdots, x_m]$. Then $\nu_n = q^{mn}$ and

$$Z(X, t) = \exp\left(\sum_{n \geq 1} \frac{q^{mn} t^n}{n}\right) = \frac{1}{1-q^m t}.$$

Thus the above product converges absolutely to the right of $\mathrm{Re}(s) > n$.

3) Consider the variety defined by the zero ideal in $\mathbb{Z}[x_1, \cdots, x_n]$. Show that

$$X' = \cup_p X'_p,$$

where p is a rational prime and X'_p denotes the set of maximal ideals in X'_p with residue field a finite extension of $\mathbb{F}_p$. Hence conclude,

$$Z(X, s) = \prod_p (1 - q^{n-s})^{-1} = \zeta(s-n).$$

In particular this shows that the above product defining the zeta function converges absolutely to the right of $\mathrm{Re}(s) > n + 1$.

1.4. Affine varieties. Let I be an ideal in the ring $k[x_1, \cdots, x_n]$. To the ideal I, we can associate the set of solutions $X(R)$ over any k-algebra R,

$$\begin{aligned} X(R) &= \{s \in R \mid f(s) = 0,\ \forall f \in I\} \\ &= \{\phi : k[x_1, \cdots, x_n]/I \to R, \phi \text{ a ring homomorhism}\}. \end{aligned}$$

We will refer to the functor which assigns to any k-algebra R, the set of solutions $X(R)$ as the (affine) algebraic variety associated to the ideal I. For example, one can take I to be the zero ideal in $k[x_1, \cdots, x_n]$. Then the set of solutions associated to any k-algebra R is R^n. The algebraic variety associated to the zero ideal in $k[x_1, \cdots, x_n]$ will be referred to as the affine n-space $\mathbf{A}^n_k$ over the field k.

Note that if J is an ideal containing I, then there is a natural inclusion for any k-algebra R, of the set of common solutions of the polynomials in J, to the corresponding set for I. We will refer to this as the algebraic subsets of the variety X defined by I. A topology can be defined on the variety X, by taking algebraic subsets as closed subsets.

Remark. In particular one can take $R = \bar{k}$, an algebraic closure of the field k. Sometimes we will also refer to $X(\bar{k})$ as the algebraic variety associated to I. However the set $X(\bar{k})$ determines the radical

$$\begin{aligned} R(I) &= \{f \in k[x_1, \cdots, x_n] \mid f^n \in I, \text{ for some } n\} \\ &= \{f \in k[x_1, \cdots, x_n] \mid f(a) = 0,\ \forall a = (a_1, \cdots, a_n) \in X(\bar{k})\}, \end{aligned}$$

of the ideal I, and not necessarily the ideal I.

Remark. The Noether normalization theorem asserts that given a ring A as above, there is a finite morphism from a suitable polynomial ring on d generators onto A. d is then the dimension of the corresponding variety. Using this, it can be shown that the zeta function corresponding to the variety defined by the solutions of the ideal corresponding to A, converges absolutely in a suitable half plane.

1.5. Projective varieties. Before going onto further examples, we will now discuss a bit about projective varieties defined by homogeneous polynomials

$$f_1(x_0, \cdots, x_m), \cdots, f_r(x_0, \cdots, x_m).$$

Since the polynomials are homogeneous, if $a = (a_0, \cdots, a_m)$ is a common solution of f_i in some k-algebra R, then any multiple $\lambda a = (\lambda a_0, \cdots, \lambda a_m)$ for $\lambda \in R$ is also a solution. By the solutions in a k-algebra R of the projective variety defined by $f_1, \cdots, f_r$, we mean the equivalence classes of *non-zero* solutions, where two solutions are equivalent if they are scalar multiples of each other.

Remark. Let $\mathbf{P}^m_k$ denote the m-dimensional projective space defined by the zero ideal in $k[x_0, \cdots, x_m]$. As a set $\mathbf{P}^m_k$ can be identified with the set of lines passing through the origin. The set of solutions $a = (a_0, \cdots, a_m) \neq 0$ with $a_i \neq 0$ for some i, can be identified with $\mathbf{A}^m_k$ (as sets to start with), by sending $(a_0, \cdots, a_m) \mapsto (a_0/a_i, \cdots, a_m/a_i)$, with the i^{th} co-ordinate omitted. Given homogeneous polynomials $f_1, \cdots, f_r$ as above with degrees respectively $d_1, \cdots d_r$, then the set of common solutions of $f_1, \cdots f_r$ in $\mathbf{P}^m_k$ with $a_i \neq 0$, corresponds via the above correspondence to the set of common solutions of the 'dehomogenized' polynomials $x_i^{-d_j} f_j(x_0, \cdots, x_m)$ in the new variables $y_0 = x_0/x_i, \cdots, y_m = x_m/x_i$. This dehomogenization process is suitable for studying the local behaviour of a projective variety. Conversely given a polynomial, for example $f(x, y) = y^2 - x^3 - ax - b$, the solutions of the corresponding 'homogeneous' polynomial $f(x, y, z) = y^2z - x^3 - axz^2 - bz^3$ in $\mathbf{P}^2$, can be taken as the 'completion' or the projective analogue of the set of 'affine' solutions of the polynomial $f(x, y)$.

To distinguish from the affine case, we will denote the projective variety of solutions by $\bar{X}$, and by $\bar{\nu}_n$ the number of 'projective' solutions in $\mathbb{F}_{q^{n+1}}$. The zeta function $Z(\bar{X}, t)$ is defined by the formal power series,

$$\log Z(X, t) = \sum_{m \geq 1} \frac{\bar{\nu}_m t^m}{m}.$$

Example 2. Let $\mathbf{P}^m$ denote the m-dimensional projective space defined by the zero ideal in $\mathbb{F}_q[x_0, \cdots, x_m]$. Then

$$\bar{\nu}_n = \frac{q^{(m+1)n} - 1}{q^n - 1} = 1 + q^n + \cdots + q^{mn}, \text{ and so}$$

$$Z(\mathbf{P}^m, t) = \exp\left((1 + q^n + \cdots + q^{mn})t^n/n\right) = 1/(1-t)(1-qt)\cdots(1-q^mt).$$

Exercise 1.5.1. Let $\mathbf{P}^m_k$ denote the m-dimensional projective space defined by the zero ideal in $k[x_0, \cdots, x_m]$. Show the following:

i) The set of projective solutions of the zero ideal can be identified with $(k^{n+1} \backslash \{0\})/k^*$.

ii) $\mathbf{P}^1_{\mathbb{R}} \simeq S^1$ as a topological space, (or as a manifold) where S^1 denotes the circle in the complex plane.

iii) Show that $\mathbf{P}^m_{\mathbb{R}} \simeq S^m/\pm 1$, obtained by identifying pairs of antipodal points on the m-sphere. Hence these spaces are all compact.

iv) Show that $\mathbf{P}^1_{\mathbb{C}}$ can be identified with the Riemann sphere, the one point compactification of the complex plane. Show that in general $\mathbf{P}^m_{\mathbb{C}}$ are compact topological spaces.

v) Let $\mathbf{A}^m_k$ denote the affine m-space defined by the zero ideal in $k[x_1, \cdots, x_m]$. Show that the projective m-space admits a decomposition (for example, as sets of solutions-disjoint),

$$\mathbf{P}_k^m = \mathbf{A}_k^m \cup \mathbf{A}_k^{m-1} \cdots \cup \mathbf{A}_k^0.$$

1.6. Smoothness. Heuristically a variety X of dimension d defined over the complex numbers, is smooth at a point $p \in X(\mathbb{C})$, if locally near the point p, the set of solutions looks like a ball in $\mathbb{C}^d$. Based on the implicit function theorem, an affine variety X of dimension d, defined by an ideal $I = (f_1, \cdots, f_m) \subset \mathbb{C}[x_1, \cdots, x_n]$ is said to be smooth at a point $a = (a_1, \cdots, a_n) \in X(\mathbb{C})$ if the matrix of partial derivatives,

$$\left((\frac{\partial f_i}{\partial x_j})(a_1, \cdots, a_n) \right),$$

has rank $n - d$.

Note that this definition is algebraic in character, and can be carried over to arbitrary fields, and not necessarily the complex numbers. We now take this as the definition of smoothness for an algebraic variety over k, in a neighbourhood of a point $a \in X(\bar{k})$.

Example 3. Let X be a variety defined by a single equation $f \in k[x_1, \cdots, x_n]$. Then X is smooth at a point $a = (a_1, \cdots, a_n) \in k^n$, if on Taylor expansion at the point a,

$$f(x_1, \cdots, x_n) = f(a) + \sum_i \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \text{ higher degree terms},$$

at least one $\partial f / \partial x_i(a)$ is non-zero.

For projective varieties, the condition of smoothness at a point, can be defined by dehomoginizing (see Remark above), and working with the corresponding affine space of solutions. Suppose now that we are considering the projective variety defined by a single homogeneous polynomial$f$. It can be checked that the projective variety $\bar{X}_f$ defined by f is smooth at a point $(a_0, \cdots, a_n)$, if $(a_0, \cdots, a_n)$ is not a solution of the system of equations,

$$\frac{\partial f}{\partial x_1}(x_1, \cdots, x_n) = 0, \cdots, \frac{\partial f}{\partial x_m}(x_1, \cdots, x_n) = 0.$$

Example 4. Let $f(x_0, \cdots, x_n) = a_0 x_o^m + \cdots + a_n x_n^m$, $(m, p) = 1$ and $a_0 \cdots a_n \neq 0$, where p is the characteristic of the base field. Then it can be checked that the projective variety defined by f is smooth.

1.7. Weil conjectures. Let X be a smooth, projective variety over a finite field k of dimension d.

a) $Z(X, t)$ is a rational function in t, i.e., of the form $P(t)/Q(t)$, where $P(t)$, $Q(t)$ are polynomial functions of t with rational integral coefficients. It also satisfies a suitable functional equation of the form,

$$Z(X, 1/q^d t) = \pm t^\alpha q^\beta Z(X, t),$$

where α and β are determined by the geometry of X.

b) (Riemann Hypothesis) There is a factorisation of $P(t)$ and $Q(t)$ in the ring $\mathbb{Z}[t]$ as,

$$P(t) = P_1(t) \cdots P_{2d-1}(t),$$
$$Q(t) = P_0(t) \cdots P_{2d}(t),$$

where $P_0(t) = (1-t)$ and $P_{2d}(t) = (1-q^d t)$. There is a factorisation of $P_i(t)$ over $\mathbb{C}$ of the form,

$$P_i(t) = \prod_{j=1}^{b_i}(1 - \alpha_{ij}t).$$

α_{ij} are algebraic integers, and for any embedding $\iota : \bar{\mathbb{Q}} \to \mathbb{C}$, of an algebraic closure $\bar{\mathbb{Q}}$ of $\mathbb{Q}$, we have for all $i,\ j$ the analogue of the Riemann Hypothesis,

$$|\iota(\alpha_{ij})| = q^{i/2}.$$

c) (Topological interpretation for b_i) Suppose now that X is a smooth, projective variety over the integers (or after inverting a few primes), say defined by homogenous polynomials f_i. This amounts to saying that the set of solutions $X(\mathbb{C})$ of X over $\mathbb{C}$, admits a structure of a smooth, complex analytic manifold. In particular, the Betti numbers B_i are defined, where B_j is the $\mathbb{Z}$-rank of the cohomology group $H^j(X(\mathbb{C}), \mathbb{Z})$. For a prime p, it makes sense to consider the variety modulo p, i.e., the space of solutions defined by considering the polynomials f_i taken modulo modulo p. Then for any sufficiently large prime p, the conjecture is,

$$b_j = B_j.$$

In other words, the shape or the topology of the variety over $\mathbb{C}$, controls the growth of the number of solutions of the equations defining the variety over various finite fields.

Example 5. Suppose X is a smooth, projective curve defined over a finite field (analogous to that of number fields). For example, X can be the space of homogenous solutions in $\mathbf{P}^2$ of a homogeneous polynomial $f(x, y, z)$ in three variable of degree d, satisfying the smoothness condition:

there are no common solutions of the system of equations

$$\frac{\partial f}{\partial x} = 0,\ \frac{\partial f}{\partial y} = 0,\ \frac{\partial f}{\partial z} = 0.$$

In this case it can be seen that $B_0 = B_2 = 1$ and $B_1 = (d-1)(d-2)$. Further $P_1(X, s) = \prod(1 - \alpha_{1j}q^{-s})$. Hence to say that the absolute values of α_{1j} be $\sqrt{q}$ is equivalent to saying that the zeros of P_1 lie on the line $\mathrm{Re}(s) = 1/2$. This is analogous to the usual Riemann Hypothesis for the Dedekind zeta function.

1.8. Frobenius and Weil cohomology. Let $F_q : \mathbb{A}^n \to \mathbb{A}^n$ be the Frobenius map sending,

$$(a_1, \cdots, a_n) \mapsto (a_1^q, \cdots, a_n^q).$$

Then X_m is the set of fixed points of F_q^m in $\overline{\mathbb{F}}_q^n$ which lie in the variety X. Now recall the Lefschetz fixed point theorem in topology: suppose $f : X \to X$ is a continuous self map of a topological manifold of dimension d, with isolated fixed points. Then the number of fixed points $\nu(f)$ is given by the expression,

$$\nu(f) = \sum_{i=0}^{d} (-1)^i \mathrm{Tr}(f \mid H^i(X, \mathbb{Z})).$$

In view of c) of the Weil conjectures stated above, Weil was inspired to conjecture the existence of a suitable cohomology theory for varieties defined over abstract fields and finite fields in particular having the following properties:

- The analogue of the Lefschetz fixed point theorem for the Frobenius morphism should be true. Since we are counting the number of fixed points, this forces the cohomology groups to have values in a characteristic zero ring, which we can assume to be a characteristic zero field L. Then we should have,

$$\nu_n(X) = \sum_{i=0}^{2d} (-1)^i \mathrm{Tr}(F^n \mid H^i(\overline{X}, L),$$

  where F^n denotes the corresponding action of the Frobenius acting on $H^i(\overline{X}, L)$. From this it follows (exercise)

$$Z(X, t) = \prod_{i=0}^{2d} \mathrm{d}et(1 - tF \mid H^i(X, L))^{-1}.$$

- The cohomology groups should have the correct Betti numbers. This means the following: Suppose $X(\mathbb{C})$ is the set of projective solutions of a system of homogeneous polynomial equations with integral coefficients. Assume that the corresponding space $X(\mathbb{C})$ is smooth as a topological manifold. Note that it will be a compact manifold, being a closed, subspace of the (compact) projective space. Since the equations have integral coefficients, it makes sense to consider the equations modulo a prime p, and to consider the corresponding projective variety of solutions X_p. Then for large enough p, the dimension of the $H^i(X_p, L)$, should be the i^{th} Betti number of the manifold $X(\mathbb{C})$.
- The duality for the zeta function should correspond to Poincare duality for the cohomology theory.

After initial efforts of Serre, such a cohomology theory was developed by Grothendieck and Artin. Grothendieck proved the Weil conjectures except for the Riemann Hypothesis, as a corollary of the cohomological machinery he had developed. Notice that this answers one of the initial hopes we had about the Riemann Zeta function- that of expressing the zeros of the zeta function as the eigenvalues of the Frobenius operator acting on the cohomology groups! The Riemann Hypothesis was proved by Deligne.

1.9. Modular forms. Consider the Ramanujan τ function defined formally by,

$$\sum_{n=1}^{\infty} \tau(n) q^n = q \prod_{n=1}^{\infty} (1-q^n)^{24}.$$

The corresponding Dirichlet series,

$$\sum_{n=1}^{\infty} \tau(n) n^{-s} = \prod_{p} (1-\tau(p)p^{-s}+p^{11-2s})^{-1},$$

defined for $\mathrm{Re}(s)$ large enough, admits analytic properties similar to the zeta function above- analytic continuation to the entire plane, is entire, and satisfies a functional equation. Ramanujan conjectured,

$$|\tau(p)| \leq 2p^{11/2}.$$

This conjecture was shown to be true by Deligne, as a consequence of the Weil conjectures. We have thus profited by searching for analogies of the Riemann Hypothesis!

Standard estimates for modular forms give the inequality, $\tau(n)=O(n^6)$. This was later refined by Rankin and Selberg, who provided the estimate $\tau(n)=O(n^{6-\epsilon})$.

Rankin's idea was to consider the Dirichlet series $\sum_n |\tau(n)|^2 n^{-s}$, the 'convolution' of the above series for $\tau(n)$ with itself, and to show that this Dirichlet series has suitable analytic properties. The germ of the idea to prove the Riemann Hypothesis, lies in adapting the Rankin-Selberg method to varieties defined over finite fields, and utilise the cohomological machinery developed by Grothendieck-roughly this amounts to relating Artin type L-functions on a n-fold product of the curve X^n for a curve X, to similar L-functions on curves. The starting point for the induction is the analogue of the prime number theorem, equivalent to providing a non-vanishing region for the L-functions.

1.10. Curves, Picard varieties and class groups. We now try to examine some algebraic ways of defining cohomology groups in the context of curves. It is this example which has served as a prime motivation for much of Iwasawa theory.

Let E be an elliptic curve defined over $\mathbb{C}$, i.e., a smooth, projective curve of genus 1 with a distinguished base point serving as the origin of a group law on the curve. E can also be explicitly given by a Weirstrass equation of the form

$$y^2z = 4x^3 + axz^2 + bz^3,\ a, b \in \mathbb{C},$$

in $\mathbf{P}^2_{\mathbb{C}}$, where the polynomial $4x^3 + ax + b$ has distinct roots in $\mathbb{C}$. Here the origin is given by the point at infinity $(0, 1, 0) \in \mathbf{P}^2_{\mathbb{C}}$. Complex analytically E is isomorphic to the complex analytic manifold $\mathbb{C}/L$, where L is a lattice in $\mathbb{C}$, i.e., a closed, discrete subgroup of $\mathbb{C}$ of rank 2 over $\mathbb{Z}$. Let $\mathcal{P}_L(z)$ be the Weierstrass $\mathcal{P}$-function associated to the lattice L,

$$\mathcal{P}_L(z) = \frac{1}{z^2} + \sum_{\omega \in L\backslash 0} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right).$$

The corresponding Weierstrass equation as given above is the projective equivalent of the well-known identity

$$\mathcal{P}'_L(z)^2 = 4\mathcal{P}_L(z)^3 - g_4(L)\mathcal{P}_L(z) - g_6(L),$$

where $\mathcal{P}'_L(z) = \sum_{\omega \in L} -2/(z-\omega)^3$, and $g_4(L) = 60 \sum_{\omega \in L\backslash 0} 1/\omega^4$, $g_6 = 140 \sum_{\omega \in L\backslash 0} 1/\omega^6$.

The fundamental group $\pi_1(E, 0)$ can be identified with the lattice L. Since L is commutative, the first homology group of $H_1(E, \mathbb{Z})$ can again be identified with L. Given a finite covering $\mathbb{C}/L' \to \mathbb{C}/L$, where $L' \supset L$ is a lattice in $\mathbb{C}$, there exists an integer n such that $nL \subset L' \subset L$ ($\mathbb{C}/L'$ is an algebraic object and is in fact again an elliptic curve). Thus the collection of covers given by the multiplication by n map, $n_L : \mathbb{C}/L \to \mathbb{C}/L$ is cofinal in the collection of (algebraic) coverings of the curve E. The profinite completion $\hat{\pi}_1(E)$ of the fundamental group (analogous to the definition of the Galois group for an infinite Galois extension), can thus be identified with the profinite limit of

$$\pi_1(E) \simeq \lim_{\leftarrow} L/nL \simeq \lim_{\leftarrow} (\mathbb{Z}/n\mathbb{Z})^2,$$

where L/nL can be identified with the Galois group of the covering map $n : \mathbb{C}/L \to \mathbb{C}/L$. Since $(\mathbb{C}/L)/\mathrm{Ker}\ n_L \simeq \mathbb{C}/L$, we have a natural identification,

$$\hat{\pi}_1(E) \simeq \lim_{\leftarrow} \mathrm{Ker}\ n_L.$$

The key point to be observed in the above isomorphism, is that the left hand side is defined topologically as the profinite completion of the fundamental group, whereas the right hand side can be interpreted algebraically as the inverse limit over n of the n-division points $E_{[n]}$ on E. We can thus take the right hand side as a substitute for the first homology of an elliptic curve, defined over an arbitrary field k. To make a suitable theory, and in order to have values in a characteristic zero field, one should fix a prime l coprime to the characteristic of k, and consider $\lim_{\leftarrow} E_{[l^m]}$ as a substitute for the first homology.

For curves C of higher genus, we have the Jacobian variety $J(C)$, which complex analytically can be identified with $H^1(X,\mathbb{R})/H^1(X,\mathbb{Z})$, with $H^1(X,\mathbb{R})$ equipped with an almost complex structure coming from the complex structure on X. We can again imitate the same construction, provided there is an algebraic interpretation of the Jacobian of a curve. Such a construction is obtained by constructing the Picard group of a curve. We will now tabulate some of the well known analogies between number fields and algebraic curves:

NUMBER FIELDS	CURVES
number	meromorphic function
non-zero prime ideal $\mathfrak{p}$	Point P
valuation corresponding to $\mathfrak{p}$	order of zero at P
ideal $\mathfrak{a} = \prod_i \mathfrak{p}_i^{a_i}$	divisor $D = \sum_i a_i P_i$
principal ideal $(a) = \prod_i \mathfrak{p}_i^{a_i}$	divisor of a meromorphic function $(f) = \sum_i \mathrm{ord}_{P_i}(f) P_i$
ideal class group	Picard group=divisors of degree zero modulo principal divisors

In analogy with the Picard group of a curve, one is led to considering the ideal class group of a number field. If X is a curve defined over $\mathbb{F}_q$, then the group of divisors made up of points on the curve over $\mathbb{F}_q$ modulo principal rational divisors is finite. The relevant Picard group is obtained by considering these rational Picard groups over all the finite extensions $\mathbb{F}_{q^n}$, or equivalently that of the Picard group of the curve over $\bar{\mathbb{F}}_q$.

2. Diagonal Fermat hypersurfaces

We will now verify part of the Weil conjectures for the zeta function of a diagonal Fermat hypersurface of the form

$$f(x_0, \cdots, x_m) = x_0^k + \cdots x_m^k.$$

This example was worked out by Weil. The idea behind the proof is to count the number of solutions using character sums. The rationality of the zeta function then amounts the Davenport-Hasse theorem relating Gauss sums over extension fields. The Riemann Hypothesis comes from the fact that the absolute values of a Gauss sum $G(\chi)$ is $\sqrt{q}$, for a non-trivial character χ of $\mathbb{F}_q^*$.

Fix q a power of a prime p and let $k_s = \mathbb{F}_{q^s}$. Assume that $n|(q-1)$. Consider the projective variety defined by the homogeneous equation

$$f(x_o, \cdots, x_r) = x_0^n + \cdots + x_r^n = 0$$

in projective r-space $\mathbf{P}^r$. Let

$$\overline{N}_s = \#\{(a_0, \cdots, a_r) \in (k_s^{r+1} \backslash 0)/k_s^* \mid a_0^n + \cdots + a_r^n = 0\}$$

be the set of projective solutions of f in $\mathbf{P}^r_{k_s}$. We first concentrate on a single field, say $k = k_0$, and count the number of solutions over k. Let

$$N = \#\{(a_0, \cdots, a_r) \in k_s^{r+1} \mid a_0^n + \cdots + a_r^n = 0\}$$

be the number of affine solutions. Then

$$\overline{N} = \frac{N-1}{q-1}.$$

For each $u \in k$, let

$$N(u) = \#\{x \in k \mid x^n = u\}.$$

We will indicate the steps involved in calculation N.

Step 1: Expressing N as a character sum.

i) For each $u \in k$, let

$$N(u) = \#\{x \in k \mid x^n = u\} = \begin{cases} 1 & \text{if } u = 0, \\ d & \text{if } u \text{ is a n}th \text{ power}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.0.1}$$

$$\text{Then } N = \sum_{u \mid L(u)=0} N(u_0) \cdots N(u_r),$$

where $L(u) = u_0 + \cdots + u_r$.

ii) Choose a generator ω of k^* (a primitive root), and for

$$\alpha \in S = \left\{0, \frac{1}{n}, \cdots, \frac{n-1}{n}\right\},$$

$$\text{define } \chi_\alpha(\omega) = e^{2\pi i \alpha}.$$

Extend to 0, by defining

$$\chi_\alpha(0) = \begin{cases} 1 & \text{if } \alpha = 0, \\ 0 & \alpha \neq 0. \end{cases} \tag{2.0.2}$$

$$\text{Then } N(u) = \sum_{\alpha \in S} \chi_\alpha(u), \tag{2.0.3}$$

$$\text{and so } N = \sum_{\{u|\ L(u)=0\}} \chi_{\alpha_0}(u_0) \cdots \chi_{\alpha_r}(u_r). \tag{2.0.4}$$

Step 2: Relating N to Jacobi sums.

The contribution from the term corresponding to $\alpha_0 = \cdots = \alpha_r = 0$ is q^r as $L(u) = 0$.

iii) If there exists $0 < s \leq r$ such that $\alpha_0, \cdots, \alpha_{s-1}$ are non-zero, and $\alpha_s = \cdots = \alpha_r = 0$, then the contribution of the corresponding term is zero. Hence

$$N = q^r + \sum_{L(u)=0,\ \alpha_i \neq 0} \chi_{\alpha_0}(u_0) \cdots \chi_{\alpha_r}(u_r). \tag{2.0.5}$$

iv) If $u_0 = 0$, then contribution is zero. Assume now $u_0 \neq 0$. Write for $i \geq 1$, $u_i = u_o v_i$. Then v_i satisfy the equation

$$\sum v_i = -1.$$

Now $\#\{u \mid L(u) = 0, u_0 \neq 0\} = (q-1)\#\{v = (v_1, \cdots, v_r) \mid \sum v_i = -1\}$. Let $T = \{\alpha = (\alpha_1, \cdots, \alpha_r) \in S^r \mid \alpha_i \neq 0,\ \sum_{i=0}^r \alpha_i = 0\}$. For $\alpha \in T$, define $\alpha_0 = -(\sum_{i=1}^r \alpha_i \neq 0$. Define the Jacobi sum for $\alpha \in T$, as

$$J_k(\alpha) = \sum_{\sum v_i = -1} \chi_{\alpha_1}(v_1) \cdots \chi_{\alpha_r}(v_r). \tag{2.0.6}$$

$$\text{Then } N = q^r + (q-1) \sum_{\alpha \in T} J_k(\alpha). \tag{2.0.7}$$

Step 3: Jacobi and Gauss sums.

v) Having related the number of solutions to Jacobi sums, we now relate this to Gauss sums. This would enable us to use the Davenport-Hasse theorem, relating Gauss sums over extension fields, in order to compute the zeta function.

Let $\psi : (k, +) \to S^1$ be a fixed additive character of the field k. Define the Gauss sum,

$$G_k(\chi, \psi) = \sum_{x \in \mathbb{F}_q} \chi(x)\psi(x).$$

The Gauss sum is the Fourier transform of the multiplicative character χ with respect to the additive character ψ, and can be thought of as the analogue of the Gamma function. For $\alpha \in T$ show that,

$$J_k(\alpha) = \frac{1}{q} G_k(\chi_{\alpha_0}, \psi) \cdots G_k(\chi_{\alpha_r}, \psi),$$

where $\alpha_0 = -(\sum_{i=1}^{r} \chi_{\alpha_i})$.

Hint: Can assume that $x_0 \neq 0$. Write $x_i = x_0 y_i$.

vi) Now substitute for $k = k_s$, with the obvious notation. Then,

$$\overline{N}_s = (N_s - 1)/(q^s - 1) = 1 + q^s + \cdots q^{s(r-1)} + \sum_{\alpha \in T_s} J_s(\alpha).$$

Step 4: Properties of Gauss sums.

vii) In order to compute the zeta function, we need to know the behaviour of Gauss sums by field extensions. The main result that is needed to show the rationality of the zeta function is the theorem of Davenport-Hasse,

$$G_s(\chi \circ N_s, \psi \circ T_s) = (-1)^{s-1} G(\chi, \psi)^s,$$

where $N_s : k_s^* \to k_0 = k$ denotes the norm map and $T_s : k_s \to k$ is the trace map. Our assumption that $n|(q-1)$ implies as χ runs over the characters of order n of k_0^*, then $\chi \circ N_s$ runs over the characters of order n of k_s^*. Hence we get,

$$J_s(\alpha \circ N_s) = (-1)^{(r-1)(s-1)} J(\alpha)^s,$$

with the evident notation.

Step 5. Rationality and the Riemann Hypothesis.

viii) The logarithmic derivative of $Z(X, t)$ becomes

$$\sum_{s=1}^{\infty} \overline{N}_s t^{s-1} = -\sum_{h=0}^{r-1} \frac{d}{dt}(1 - q^h t) + (-1)^r \sum_{\alpha \in T} \frac{d}{dt} \log(1 - (-1)^{r-1} J(\alpha) t).$$

Hence $Z(X, t) = \dfrac{1}{(1-t)\cdots(1-q^{r-1}t)} \left(\prod_{\alpha \in T} (1 - (-1)^{r-1} J(\alpha) t) \right)^{(-1)^{r-1}}$.

ix) Riemann Hypothesis follows from the fact,

$$|G_{\mathbb{F}_q}(\chi, \psi)| = \sqrt{q}.$$

Step 6. Betti numbers

x) We will leave it to the reader to check that the Betti numbers of the corresponding complex points of the projective variety are the same as that indicated by the Weil conjectures.

3. Growth of class groups along cyclotomic towers

Let us look at some of the aspects of zeta functions of varieties defined over finite fields, with a view to carry them over to number fields.

1) The first aspect that comes to our mind, is that the Galois group $G(\bar{\mathbb{F}}_q/\mathbb{F})$ of an algebraic closure $\bar{\mathbb{F}}_q$ over $\mathbb{F}_q$ is isomorphic to the profinite completion $\hat{\mathbb{Z}} = \lim_{\leftarrow}\mathbb{Z}/n\mathbb{Z}$. A topological generator is given by the Frobenius automorphism $\phi(x) = x^q$ of $\bar{\mathbb{F}}_q$. This gives a distinguished generator, with respect to which it is possible to talk about characteristic polynomials. Thus one type of extensions in number fields we would like to look at are extensions which are topologically generated by a single element.

2) A second aspect is that $\bar{\mathbb{F}}_q$ is generated by roots of unity over the base field. This property is satisfied by the above example.

3) In analogy with curves over finite fields, the ideal class groups of number fields can be considered as a substitute for the Picard group. The analogue of working over the algebraic closure, lies in considering the collection of the ideal class groups along a tower of compatible cyclic extensions.

Example 6. Let p be a prime, and let $q = 4$ if $p = 2$, and $q = p$ if p is odd. Let d be a positive integer coprime to p. Denote by K_n be field $\mathbb{Q}(\zeta_{dqp^n})$. Then $\mathrm{Gal}(K_n/K_0) \simeq \mathbb{Z}/p^n\mathbb{Z}$.

The important observation of Iwasawa, is that instead of considering the full ideal class group, if we consider the p-primary component X_n of the ideal class group of K_n, these patch together to give a module over the group ring $\mathbb{Z}_p[\Gamma]$, where Γ is the Galois group of $(\cup_n K_n)/K_0$. The structure theory of modules over this ring can be worked out, and applying class field theory provides us with information about the growth of these groups X_n.

3.1. Inverse limits.

We recall the notion of inverse limits. The reader who is familiar with inverse limits, can skip this section. We work in a general category $\mathcal{C}$- for example the category of sets, topological spaces, groups,...Let $(I, \leq)$ be a partially ordered set, and an inverse system consists of the following: a) for each $i \in I$ an object X_i in $\mathcal{C}$, b) for $i \leq j$, a map $\phi_{ij} : X_j \to X_i$. This satisfies the compatibility condition that for $i \leq j \leq k$, then $\phi_{jk}\phi_{ij} = \phi_{ik}$. Let X be an object equipped with a collection of maps $\pi_i : X \to X_i$ satisfying the compatibility condition $\phi_{ij}\pi_j = \pi_i$ for pairs $i, j \in I$, $i \leq j$. We say that X is a inverse limit of the of the inverse system (X_i, ϕ_{ij}, I) if it satisfies the following universal property:
given a family of maps $\psi_i : Z \to X_i$ from an object Z in $\mathcal{C}$, satisfying the compatibility condition $\phi_{ij}\psi_j = \psi_i$, for $i \leq j$, $i, j \in I$, then there exists a unique map $\Psi : Z \to X$, such that $\pi_i\Psi = \psi_i$ for all $i \in I$. The inverse limit if it exists, is unique upto a unique isomorphism.

If $\mathcal{C}$ is a full subcategory of the category of sets, then in terms of elements X can be described as,

$$X = \{(x_i) \in \prod_i X_i \mid \phi_{ij}(x_j) = x_i, \text{ for } i \leq j\}.$$

Example 7. Let A be a ring with an ideal J. Let the indexing set be the natural numbers $\mathbb{N}$ with the usual order. For $m \leq n$, we have natural ring homomorphisms $A/J^n \to A/J^m$. The inverse limit $\hat{A}_J$ taken in the category of rings, is called the J-adic completion of A. For example, take A to be the integers $\mathbb{Z}$, and $J = (p)$ the ideal generated by a prime number p. Then the p-adic completion of $\mathbb{Z}$ along (p) gives the ring of p-adic integers $\mathbb{Z}_p$.

Another similar example can be obtained by taking $A = \mathbb{Z}[T]$ (or even $\mathbb{Z}_p[T]$), and J to be the ideal (p, T) generated by p and T. Then the completion gives the ring of formal power series $\mathbb{Z}_p[[T]]$, with coefficients in $\mathbb{Z}_p$. In this case, it is possible to consider the completion in the category of topological rings, by equipping the finite ring quotients $\mathbb{Z}[T]/(p,T)^n$ with the discrete topology. The profinite completion $\mathbb{Z}_p[[T]]$ has the structure of a compact ring.

Example 8. Let L/K be an algebraic extension of fields. Since each element in L is algebraic over K, L can be written as the union of finite extension fields M of K. Assume not that L can be written as a union of finite Galois extensions over K. We will then say that L is Galois over K. Moreover the Galois group $G(L/K)$ is the projective(inverse) limit of the finite Galois groups $G(M/K)$, where M is a finite Galois extension of K. The inverse limit is taken over the indexing set of all finite Galois extensions M of K, ordered by inclusion. $G(L/K)$ has thus the structure of a profinite group, and is compact in particular.

Suppose K is a number field, and v a place of K. v is ramified (unramified) if v is ramified in some finite extension (unramified in any finite extension) $M \subset L$ of K. A valuation belonging to the place v can be extended to any finite extension, and thus extended to L. If w is a valuation on L extending v, then we say that $w|v$. Define the decomposition group D_w and inertia group I_w at w as,

$$D_w = \varprojlim_M D_{w|M} \text{ and } I_w = \varprojlim_M I_{w|M}$$

3.2. Our aim now is to prove the following theorem of Iwasawa concerning of growth of ideal class groups along a $\mathbb{Z}_p$ tower. Let K_0 be a finite extension of $\mathbb{Q}$, and let K_∞/K_0 be a $\mathbb{Z}_p$ extension, i.e., $\Gamma := \mathrm{Gal}(K_\infty/K_0) \simeq \mathbb{Z}_p$.

Write

$$K_\infty = \cup_{n\geq 0} K_n$$
$$\text{with } \Gamma_n := \mathrm{Gal}(K_n/K_0) \simeq \mathbb{Z}/p^n\mathbb{Z}.$$

Denote by $\gamma = (\gamma_n)$ a topological generator for Γ. γ_n is a compatible collection of generators for the groups Γ_n.

Example 9. Let p, q and d be as in the example above, and let $K_n = \mathbb{Q}(\zeta_{dqp^n})$. Write $K_\infty = \cup_n K_n$. Then $G(K_\infty/K_0) \simeq \mathbb{Z}_p$. These extensions will be referred to as cyclotomic $\mathbb{Z}_p$-extensions, and will constitute our primary examples.

Let X_n be the p-Sylow subgroup of the ideal class group of K_n. Write

$$|X_n| = p^{e_n}.$$

The theorem of Iwasawa is the following:

Theorem 3.2.1. *Let K_∞/K_0 be a $\mathbb{Z}_p$ extension. There exists integers $\lambda \geq 0, \mu \geq 0, \nu$ and a positive n_0, such that for $n \geq n_0$,*

$$e_n = \lambda p^n + \mu n + \nu.$$

For the proof, let

$$X = \lim_{\leftarrow} X_n,$$

where the inverse limit is taken with respect to the norm maps $N : K_m \to K_n$, $m \geq n$. We have a compatible action of $\mathrm{Gal}(K_n/K_0) = \Gamma_n$ on X_n,

$$N(\gamma_m x_m) = r(\gamma_m) N(x_m),$$

where $\gamma_m \in \Gamma_m$, and r denotes the natural restriction map $\Gamma_m \to \Gamma_n$. Hence we get a continuous action of Γ on X.

Further since each X_n is a p-primary abelian group of order p^{e_n}, there is an action of $\mathbb{Z}_p$ on $X(n)$, via the projection $\mathbb{Z}_p \to Z_p/p^{e_n}\mathbb{Z}_p \to \mathbb{Z}/p^{e_n}\mathbb{Z}$. This action is compatible with the norm maps. Thus we get an action of $\mathbb{Z}/p^{e_n}\mathbb{Z}[\Gamma_n]$, the group ring of Γ_n with coefficients in $\mathbb{Z}/p^{e_n}\mathbb{Z}$ on X_n. Let

$$\Lambda := \mathbb{Z}_p[\Gamma] = \lim_{\overleftarrow{n}} \mathbb{Z}/p^{e_n}\mathbb{Z}[\Gamma_n].$$

This is a profinite ring. Note that both $\mathbb{Z}_p[\Gamma]$ are profinite spaces, and thus carry a topology with respect to which they are compact. It can be checked that we obtain a continuous action of Λ on X.

The proof of Iwasawa's theorem follows from:

1) Structure theory of $\mathbb{Z}_p[\Gamma]$-modules.

2) Using class field theory to obtain information about the finite layers X_n from X.

3.3. Application of Class Field Theory. We will begin with 2) first. First some preliminaries:

Lemma 3.3.1. *a) Let K_∞/K_0 be a $\mathbb{Z}_p$-extension. l a prime of K_∞ not dividing p. Then K_∞ is unramified at l. In particular there are only finitely many primes of K_0 which ramify in K_∞.*

b) At least one prime ramifies.

c) There is a finite extension K_n such that every prime of K_n which ramifies is totally ramified.

We will not present a proof of the lemma, which follows quite easily from class field theory, but instead will be content upon remarking that for the cyclotomic $\mathbb{Z}_p$-extension the properties stated in the lemma are seen to be satisfied.

Henceforth we will assume the following, and it can be seen that the cyclotomic $\mathbb{Z}_p$-extension satisfy the following:

1) K_0 is such that all primes of K_0 which ramify in K_∞ are totally ramified.

2) There is only one prime of K_0 which is totally ramified.

These assumptions imply that there is a unique place of K_∞, lying over the ramified prime of K_0, and we have an identification of the inertia group I at this prime with Γ.

3.3.1. *Hilbert class fields.* The main aim is to recover X_n from X. For this we use class field theory. Recall that the Hilbert class field H_K of K, is the maximal abelian unramified extension of K. By class field theory, we know that H_K/K is a finite, abelian extension and there is canonical identification,

$$F: \ C(K) \to G(H_K/K),$$

where $C(K)$ denotes the class group of K. The identification is obtained by sending a prime ideal $\mathfrak{p}$ of K, to the corresponding Frobenius element, and extending it multiplicatively to the class group. Recall that the Frobenius element for an abelian extension of number fields L/K, corresponding to an unramified prime ideal $\mathcal{P}$ of L lying over a prime ideal $\mathfrak{p}$ of K, is the unique element $F(\mathfrak{p})$ in $G(L/K)$ satisfying,

$$F(\mathfrak{p})(x) \equiv x^{N\mathfrak{p}} \pmod{\mathcal{P}}, \tag{3.3.1}$$

where x is an integral element in L. Since L/K is abelian, $F(\mathfrak{p})$ is independent of the choice of $\mathcal{P}$ dividing $\mathfrak{p}$.

Suppose K is a Galois extension of E, with Galois group Γ. Let G denote the Galois group of H_K/E. We have the exact sequence,

$$1 \to C(K) \simeq G(H_K/K) \to G \to \Gamma \to 1. \tag{3.3.2}$$

We have now two naturally defined actions of Γ on $C(K)$. The first action, is via the interpretation of Γ as automorphisms of K, and $C(K)$ in terms of fractional ideals in K. Namely, an element $\gamma \in \Gamma$, acts on an ideal I of K, by $(\gamma, I) \to \gamma(I)$. This descends down to an action of Γ on $C(K)$. The second action is via the exact sequence 3.3.2 above, and the interpretation of $C(K)$ as $G(H_K/K)$. For $\gamma \in \Gamma$, choose a lift to G, denoted again by γ. γ acts by sending an element $x \in C(K)$ to $x^\gamma := \gamma^{-1} x \gamma$. Via the identification $F : C(K) \to G(H_K/K)$, these two actions are the same:

Lemma 3.3.2. *Let $\mathfrak{p}$ be an unramified prime ideal of K over E. Then for $\gamma \in \Gamma$,*

$$\gamma^{-1} F(\mathfrak{p}) \gamma = F(\gamma^{-1}\mathfrak{p}).$$

Proof. Let x be an algebraic integer in E, and $\mathcal{P}$ be a prime ideal in the ring of integers of E lying over $\mathfrak{p}$. We have,

$$F(\mathfrak{p})(\gamma x) \equiv (\gamma x)^{N\mathfrak{p}} \pmod{\mathcal{P}}.$$

We apply γ^{-1} to both sides. Then $\gamma^{-1}\mathcal{P} | \gamma^{-1}\mathfrak{p}$, and $N(\gamma^{-1}\mathfrak{p}) = N(\mathfrak{p})$. Hence,

$$\begin{aligned} \gamma^{-1} F(\mathfrak{p}) \gamma(x) &\equiv \gamma^{-1}((\gamma x)^{N\mathfrak{p}}) \pmod{\mathcal{P}} \\ &\equiv x^{N\gamma^{-1}\mathfrak{p}} \pmod{\gamma^{-1}\mathcal{P}}. \end{aligned}$$

□

Let M_n be the maximal unramified abelian p-extension of K_n. From the identification of the Galois group with the Hilbert class field as in the above paragraph, we have the following identification, still denoted by F:

$$F : \ X_n \to G(M_n/K_n). \tag{3.3.3}$$

Let G_n denote the Galois group of M_n over K_0. Since K_{n+1} is a ramified extension of K_n and M_n an unramified extension of K_n, the fields M_n and K_{n+1} are linearly disjoint. The extension $M_n K_{n+1}$ is unramified over K_{n+1}, and the Galois group $G(M_n K_{n+1}/K_{n+1})$ can be identified with $G(M_n/K_n)$. There is a natural restriction map $r : G(M_n K_{n+1}/K_{n+1}) \to G(M_n/K_n)$. We have,

Lemma 3.3.3. *a) The following diagram is commutative:*

$$\begin{tikzcd} X_{n+1} \arrow[r, "F"] \arrow[d, "N"] & G(M_{n+1}/K_{n+1}) \arrow[d, "r"] \\ X_n \arrow[r, "F"] & G(M_n/K_n) \end{tikzcd}$$

Proof. Choose a prime ideal $\mathcal{P}$ of $\mathcal{O}_{M_{n+1}}$ lying over a prime $\mathfrak{p}$ of $\mathcal{O}_{K_{n+1}}$. Assume that $\mathfrak{p}$ is unramified over K_n and divides the prime p of $\mathcal{O}_{K_n}$. Then the norm from K_{n+1} to K_n of the ideal $\mathfrak{p}$ is p^f, where f is the degree of the residue field extensions, satisfying $N\mathfrak{p} = Np^f$. Let x be an integral element in M_n. The restriction $rF(\mathfrak{p})(x)$ satisfies,

$$\begin{aligned} rF(\mathfrak{p})(x) &\equiv x^{N\mathfrak{p}} \pmod{\mathcal{P}} \\ i.e.,\ F(\mathfrak{p})(x) &\equiv x^{Np^f} \pmod{\mathcal{P}} \\ &\equiv F(p)(x)^f \pmod{\mathcal{P}} \\ &\equiv F(p)^f(x) \pmod{\mathcal{P}} \\ &\equiv F(N\mathfrak{p}) \pmod{\mathcal{P}} \end{aligned}$$

□

3.3.2. Let M_∞ be the compositum of the fields M_n. M_∞/K_∞ is Galois with Galois group isomorphic to X. It can be shown that M_∞ is the maximal unramified p-abelian extension of K_∞. Denote by G the Galois group of M_∞ over K_∞. We have the exact sequence,

$$1 \to X \to G \to \Gamma \to 1. \tag{3.3.4}$$

As before, let γ denote a topological generator of Γ. Note that the element $\gamma - 1$ in the Iwasawa algebra acts by sending $(\gamma-1) : x \mapsto x^{\gamma-1} = \gamma x \gamma^{-1} x^{-1}$.

Proposition 3.3.1. *X_0 is the group of coinvariants for the Γ action on X, i.e., we have*

$$X/X^{\gamma-1} \simeq X_0 \simeq G(M_0/K_0).$$

More generally for $n \geq 0$,

$$X/X^{\gamma^{p^n}-1} \simeq X_n \simeq G(M_n/K_n).$$

Before we prove this proposition, we recall a bit about semi-direct products. Given an exact sequence of groups,

$$1 \to X \to G \xrightarrow{\pi} \Gamma \to 1,$$

we say that the exact sequence splits, or that G is a semi-direct product of Γ by X, written $G = X \times \Gamma$, if there exists a (continuous) section $s : \Gamma \to G$, such that $\pi \circ s = \mathrm{id}_\Gamma$. This amounts to giving a subgroup $H \subset G$, such that $H \cap X = \{1\}$, and H maps onto Γ. The splitting provides an action of Γ on X. It can be checked that the following are equivalent for a semi-direct product G: i) $s(\Gamma)$ is a normal subgroup of G. ii) the induced action of Γ by a splitting section s on X is trivial. iii) G is isomorphic to the direct product of the groups $X \times s(\Gamma)$.

Lemma 3.3.4. *The exact sequence 3.3.4 splits.*

Proof. Let I denote the inertia group of a prime of M_∞ lying over the unique prime of K_0 ramifying inside K_∞. Since K_∞ is totally ramified over K_0, I restricted to K_∞ is surjective. Further M_∞ is an unramified extension of K_∞. Hence $I \cap X = 1$. This implies that G is the semi-direct product of I and X. □

LEMMA 3.3.5. *The closure of the commutator group* $\overline{[G,G]}$ *is isomorphic to* $X^{\gamma-1}$.

PROOF. Take $x \in X$. Then

$$x^{\gamma-1} = \tilde{\gamma} x \tilde{\gamma}^{-1} x^{-1} \in [G,G].$$

This implies $X^{\gamma-1} \subset [G,G]$.

Conversely look at the exact sequence,

$$1 \to X/X^{\gamma-1} \to G/X^{\gamma-1} \to \Gamma \to 1.$$

Note that $X^{\gamma-1}$ is a closed normal subgroup of G (why?- because I normalises, and it is closed since it is the image of the compact group X by the map $\gamma - 1$).

By definition, the Γ which is topologically generated by γ, acts trivially on the extension. Hence the semi-direct product becomes a product. Hence $G/X^{\gamma-1}$ is a commutative group. This implies that $X^{\gamma-1} \supset \overline{[G,G]}$.

□

PROOF OF PROPOSITION 3.3.1 M_0 is the maximal p-primary abelian unramified extension of K_0. Since M_0 is an abelian extension of K_0, we have $G(M_\infty/M_0) \subset [G,G]$. Since M_0/K_0 is unramified implies that $G(M_\infty/M_0) \subset I$.

3.4. Structure theory of modules over the Iwasawa algebra.

In this section, we follow Serre in understanding the structure theory of modules over the Iwasawa algebra $\Lambda = \mathbb{Z}_p[\Gamma]$. The key point is to identify Λ with the power series ring $\mathbb{Z}_p[[T]]$.

PROPOSITION 3.4.1.

$$\Lambda \simeq \mathbb{Z}_p[[T]].$$

Note that if we work at the finite level, then we have an obvious isomorphism of the group ring

$$\mathbb{Z}_p[\Gamma_n] \simeq \mathbb{Z}_p[U]/(U^{p^n} - 1),$$

by sending a generator γ_n of Γ_n to the element U. The problem with this isomorphism is that the isomorphism is not compatible with the inverse system formed by $\mathbb{Z}_p[\Gamma_n]$. In order to build a compatible system, we need to work with a different generator for the ring $\mathbb{Z}_p[U]/(U^{p^n} - 1)$. Substitute $U = T + 1$. We then have

$$\mathbb{Z}_p[\Gamma_n] \simeq \mathbb{Z}_p[U]/(U^{p^n} - 1) \simeq \mathbb{Z}_p[T]/((1+T)^{p^n} - 1).$$

The polynomial $h_n(T) = (1+T)^{p^n} - 1$ is an example of a *distinguished polynomial*:

Definition 1. A polynomial of the form $T^n + a_1 T^{n-1} + \cdots + a_n$ is distinguished if all the coefficients a_i are divisible by p.

Lemma 3.4.1.

$$\lim_{\overleftarrow{n}} \mathbb{Z}_p[[T]]/(h_n) \simeq \mathbb{Z}_p[[T]].$$

Proof.

$$\begin{aligned} h_{n+1} = (1+T)^{p^{n+1}} - 1 &= ((1+T)^{p^n} - 1 + 1)^p - 1 \\ &= (h_n + 1)^p - 1 = h_n^p + p h_n^{p-1} + \cdots + p h_n. \end{aligned}$$

Thus by induction we see that h_{n+1} is in the ideal $(p, T)^{n+1}$. □

Lemma 3.4.2.

$$\mathbb{Z}_p[[T]]/(h_n) \simeq \mathbb{Z}_p[T]/(h_n) \simeq \mathbb{Z}_p[\Gamma_n].$$

Proof. The proof of this rests on the Euclidean algorithm for the power series ring: if f is a distinguished polynomial of degree n, and $g \in \mathbb{Z}_p[[T]]$, then there exists unique elements $q \in \mathbb{Z}_p[[T]]$ and a polynomial $r \in \mathbb{Z}_p[T]$ of degree less than n. □

From the above lemmas, we have a proof of the proposition. (Make a remark that the h_n are compatible with taking the inverse limit).

Remark. Note that if we are working over $\mathbb{Q}_p$ instead of $\mathbb{Z}_p$, then we will not have the above lemma, and the inverse limits of the group rings with coefficients in $\mathbb{Q}_p$ will not have the nice description in terms of formal power series as given above.

In terms of the identification of Λ with $\mathbb{Z}_p[[T]]$, we have by Proposition 3.3.1 (iii),

$$X_n = X/h_n(T)X.$$

In order to use the structure theorem, we need to know that X is finitely generated as a module over Λ, and this is provided by the following topological version of Nakayama's lemma. We begin with a topological version of Nakayama's lemma, which asserts the finite generation of X as a Λ-module.

Lemma 3.4.3. *Let $\mathcal{O}$ be a local ring with maximal ideal $\mathfrak{m}$. Let V be a compact topological module over $\mathcal{O}$ with respect to the $\mathfrak{m}$-adic topology.*

i) if $\mathfrak{m}V = V$, then $V = 0$.

ii) if $\mathcal{O}$ is compact, and $V/\mathfrak{m}V$ is finitely generated, then V is finitely generated.

Proof. Exercise. □

Going back to the proof of Iwasawa's theorem on the growth of class groups along $\mathbb{Z}_p$-extensions, we see that it follows from Nakayama's lemma, and the identification of $X/h_0X \simeq X_0$, that X is a finitely generated Λ-module.

Corollary 3.4.1. *X is finitely generated as a Λ-module.*

Remark. Note that this lemma is different from the usual Nakayama lemma, which assumes a priori that V is finitely generated. Here by imposing a topological notion, we are able to conclude the finite generation of V, from the finite generation of the covariants of V with respect to $\mathfrak{m}$.

We state without proof the following structure theorem for finitely generated modules over $\mathbb{Z}_p[[T]]$. For a proof, we refer to Washington's or Lang's book.

Theorem 3.4.2. *Let V be a finitely generated module over $\mathbb{Z}_p[[T]]$. Then there is morphism*

$$V \to \Lambda^r \oplus \prod_i \Lambda/(p^{n_i}) \oplus \prod_j \Lambda/(f_j^{m_j})$$

with finite kernel and cokernel, and where f_j are distinguished and irreducible polynomials.

As a corollary we obtain the following theorem:

Theorem 3.4.3. *There is an injective morphism from X to a Λ-module of the form $\prod_i \Lambda/(p^{n_i}) \oplus \prod_j \Lambda/(f_j^{m_j})$ with finite cokernel, and where f_j are distinguished polynomials.*

This follows from the fact that if X contains a copy of Λ, then X_0 will be infinite, contradicting the finiteness of the class number. In order to establish the growth of class numbers along a $\mathbb{Z}_p$-extension, we have to examine the growth of these modules. Note that

$$X_n = X/X^{\gamma^{p^n}-1},$$

a) $V = \Lambda/(p^m)$. Then,

$$V_n \simeq (\mathbb{Z}/p^m\mathbb{Z})[[T]]/(h_n) \simeq (\mathbb{Z}/p^m\mathbb{Z})[T]/(T^{p^n}-1).$$

Thus V_n is a free module of rank p^n over $\mathbb{Z}/p^m\mathbb{Z}$, and is of cardinality p^{mn}.

b) $V = \Lambda/(f)$, f a distinguished polynomial of degree d. We assume that V_n is finite for all n, an assumption satisfied for the module X in our situation. Note that since f is distinguished,

$$f \cong T^d \pmod{p}.$$

Hence there exists n_0, such that for $n > n_0$,

$$T^{p^{n-1}} \cong 0 \ (\mathrm{mod}\ (f, p)).$$

For an element $P \in \Lambda$, denote by $L(P)$ the left multiplication by P on V. Then for $n > n_0$,

$$L(T)^{p^{n-1}} \cong 0 \ (\mathrm{mod}\ p).$$

It follows that,

$$L((1+X)^{p^{n-1}}) \cong 1 \ (\mathrm{mod}\ p).$$

$$L((1+X)^{p^n}) \cong (L((1+X)^{p^{n-1}})^p \cong 1 \ (\mathrm{mod}\ p^2).$$

We have,

$$(1+T)^{p^{n+1}} - 1$$
$$= \{(1+T)^{p^n} + (1+T)^{2p^n} + \cdots + (1+T)^{(p-1)p^n}\}\{(1+T)^{p^n} - 1\}.$$

From this it can be seen by induction that for some unit element u,

$$(\gamma_n - 1)V = p^{n-n_0}u(\gamma_{n_0} - 1)V, \ n > n_0 \gg 0.$$

Since $(\gamma_{n_0} - 1)V$ is of finite index in $\mathbb{Z}_p^d$, it is isomorphic to $\mathbb{Z}_p^d$, and so upto a constant error term the growth as n varies, is given by dn, where d is the degree of the distinguished polynomial f. Thus μ is the sum of the degrees of the polynomials $(f_j^{m_j})$ occurring in the decomposition of the module V.

4. Construction of p-adic L-function via Stickelberger

Let χ be a Dirichlet character, and fix a prime p. Our objective in this section is to outline a construction of the p-adic L-function $L_p(s, \chi)$ associated to the character χ, following a method of Iwasawa using Stickelberger elements. In the last section we saw that the profinite limit X of the p-primary part of the class group along a $\mathbb{Z}_p$-extension, is upto a finite cokernel isomorphic to,

$$X \to \prod_i \Lambda/(p^{n_i}) \oplus \prod_j \Lambda/(f_j^{m_j}).$$

Essentially this amounts to saying that $\prod_i p^{n_i} \prod_j f_j^{m_j}$ annihilates X. Recall that the Stickelberger element for the field K_n annihilates the class group. The question naturally arises about the relationship of the Stickelberger elements along a $\mathbb{Z}_p$-tower of fields, and the annihilator obtained above. The limit of the Stickelberger elements taken along a $\mathbb{Z}_p$-extension can be seen to belong to the Iwasawa algebra, and the analytic function associated to this formal power series turns out to be the p-adic L-function constructed by Kubota and Leopoldt. Roughly, the Main conjecture of Iwasawa theory says that these two methods give raise to same element in the Iwasawa algebra upto a unit, thus expressing the characteristic function of X as a p-adic L-function, in analogy with the Weil conjectures for curves.

4.1. Preliminaries. Let $\bar{\mathbb{Q}}_p$ be an algebraic closure of the field $\mathbb{Q}_p$. The natural valuation $|\ |$ on $\mathbb{Q}_p$ extends to a valuation on $\bar{\mathbb{Q}}_p$, which is no longer discrete. We normalise the valuation by fixing $|p| = 1/p$. The field $\bar{\mathbb{Q}}_p$ is no longer complete with respect to this valuation. Denote by $\mathbb{C}_p$ the completion of $\bar{\mathbb{Q}}_p$ with respect to this valuation. The absolute value extends to $\mathbb{C}_p$, and $\mathbb{C}_p$ is complete, algebraically closed. This is the analogue of the complex numbers. In fact, it can be seen that as abstract fields $\mathbb{C}_p$ and $\mathbb{C}$ are isomorphic. Fix an embedding of $\bar{Q}$ into $\mathbb{C}_p$, where $\mathbb{C}_p$ is an algebraic closure of the field $\mathbb{Q}_p$. With this we can think of Dirichlet characters as having values in $\mathbb{C}_p$. We remark however that the use of $\mathbb{C}_p$ is more for convenience, and we could have as well worked with the locally compact field $\mathbb{Q}_p(\chi)$, obtained by adjoining the values of χ.

We recall the definition of the Teichmuller character ω defined on $\mathbb{Z}_p^*$. Let $q = p$ if p is odd, and $q = 4$ if $p = 2$. ω is the unique character ω on $\mathbb{Z}_p^*$ with values in the $(p-1)$th roots of unity if p is odd, and values in $\{\pm 1\}$ if $p = 2$, satisfying the following congruence:

$$\omega(a) \equiv a \pmod{p}.$$

It can be seen that $\omega(a)$ is also equal to $\lim_{n\to\infty} a^{p^n}$. Define

$$< a >= \omega(a)^{-1}a.$$

Let N be the conductor of χ. The generalised Bernoulli numbers are defined by the series,

$$\sum_{a=1}^{N} \frac{\chi(a)te^{at}}{e^{Nt}-1} = \sum_{n=0}^{\infty} B_{n,\chi}\frac{t^n}{n}.$$

The Dirichlet L-functions associated $L(s,\chi)$ admit an analytic continuation to the entire complex plane, and the values at the negative integers are given by the Bernoulli numbers,

$$L(1-n,\chi) = -\frac{B_{n,\chi}}{n}, \quad n \geq 1.$$

For any integer $k,\ n \geq 0$, define

$$S_{n,\chi}(k) = \sum_{a=1}^{k} \chi(a)a^n. \tag{4.1.1}$$

An important property of the Bernoulli numbers we will need is the following:

LEMMA 4.1.1. *In $\mathbb{C}_p$,*

$$B_{n,\chi} = \mathrm{lim}_{k\to\infty}\frac{1}{p^kN}S_{n,\chi}(p^kN) = \mathrm{lim}_{k\to\infty}\frac{1}{p^kN}\sum_{a=1}^{p^kN}\chi(a)a^n.$$

For a proof of this lemma, we refer to Iwasawa's book. Let $q = p$ if p is odd, and $q = 4$ if $p = 2$. We recall the following theorem:

THEOREM 4.1.1. *There exists a p-adic meromorphic function $L_p(s, \chi)$ defined in the domain $D = \{s \in \mathbb{C}_p \mid |s| < qp^{-1/(p-1)}\}$ satisfying the following:*

a) $L_p(s, \chi)$ is given by

$$L_p(s, \chi) = \frac{a_{-1}}{s-1} + \sum_{n=0}^{\infty} a_n (s-1)^n, \quad a_n \in \mathbb{Q}_p(\chi),$$

where $a_{-1} = 1 - 1/p$ if χ is the trivial character, and 0 otherwise.

$$\begin{aligned}(b)\ L_p(1-n, \chi) &= -(1 - \chi\omega^{-n}(p)p^{n-1})\frac{B_{n,\chi\omega^{-n}}}{n} \\ &= (1 - \chi\omega^{-n}(p)p^{n-1})L(1-n, \chi\omega^{-n}), \quad n \geq 1.\end{aligned}$$

Further as a p-adic meromorphic function on the domain D, $L_p(s, \chi)$ is uniquely characterised by the above properties.

In particular, let χ be an odd Dirichlet character, i.e., $\chi(-1) = -1$. It can be seen from properties of Bernoulli numbers, that for $n \equiv \pmod{\phi(q)}$, $B_{n,\chi} = 0$. Hence it follows that $L_p(s, \chi)$ is identically 0 if χ is an odd Dirichlet character.

4.2. Write the conductor of χ as dp^j, $(d, p) = 1$, for some $j \geq 0$. Let $q_n = dqp^n$, $n \geq 0$. Let

$$K_n = \mathbb{Q}(\zeta_{q_n}), \quad n \geq 0.$$

By the Kronecker-Weber theorem, we see that χ can be considered as a character of the Galois group $G(K_n/\mathbb{Q})$ of the field K_n over $\mathbb{Q}$. We have the exact sequence,

$$1 \to G(K_n/K_0) \to G(K_n/\mathbb{Q}) \to G(K_0/\mathbb{Q}) \to 1.$$

This exact sequence splits giving an idenitification

$$G(K_n/\mathbb{Q}) \simeq \Delta \times \Gamma_n,$$

where $\Delta \simeq G(K_0/\mathbb{Q})$ and $\Gamma_n \simeq G(K_n/K_0)$. Accordingly for $a \in G(K_n/\mathbb{Q})$, let $\delta(a) \in \Delta$ and $\gamma_n(a) \in \Gamma_n$ denote respectively the components with respect to the above decomposition. Note that $G(K_n/\mathbb{Q})$ can be identified with $(\mathbb{Z}/q_n\mathbb{Z})^* \simeq (\mathbb{Z}/q_0\mathbb{Z})^* \times \mathbb{Z}/p^n\mathbb{Z}$, where $q_n = m_0qp^n$, $(m_0, p) = 1$, $q = p$ if p is odd and $q = 4$ if $p = 2$. Then $\Delta \simeq (\mathbb{Z}/q_0\mathbb{Z})^*$ and $\Gamma_n \simeq \mathbb{Z}/p^n\mathbb{Z}$.

Write $\chi = \theta\psi$ accordingly, where θ is a character of Δ and ψ is a character of Γ_n. θ and ψ is called a character of the first kind, and ψ a character of the second kind. Note that θ is an unramified character, and that ψ is a even character of p-power order. Let $\mathcal{O}_\theta$ be the ring of integers in the field $\mathbb{Q}_p(\theta)$. The main theorem is the following:

THEOREM 4.2.1. *Let χ be an even character. There exists formal power series $f(T,\theta)$(if $\theta \neq 1$), $g(T,\theta)$, $h(T,\theta) \in \mathcal{O}_\theta[[T]]$, with*

$$h(T,\theta) = 1 - (1+q_0)/(1+T), \text{ and}$$

$$f(T,\theta) = \frac{g(T,\theta)}{h(T,\theta)}, \quad \theta \neq 1.$$

If $\theta = 1$, we take the above as the definition of $f(T,\theta)$ formally in the quotient field of $\mathcal{O}_\theta[[T]]$. Moreover in the domain D defined above,

$$L_p(s,\chi) = f(\psi(1+q_0)^{-1}(1+q_0)^s - 1, \theta).$$

There are a number of remarks to be made to clarify the above theorem.

REMARK. For $x \in \mathbb{C}_p$, define the exponential function,

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

It can be shown that this series converges absolutely in the region $|x| < p^{-1/(p-1)}$. Similarly define the logarithm function in the domain $\{x \in \mathbb{C}_p \mid |1-x| < 1\}$ by the series,

$$\log_p(1-x) = \sum_{n=1}^{\infty} \frac{x^n}{n}.$$

It can be shown that the above definition of $\log_p(x)$ can be extended uniquely to a continuous function on $\mathbb{C}_p^*$ such that $\log_p(p) = 0$. With this define,

$$(1+q_0)^s = \exp(s\log_p(1+q_0)).$$

REMARK. We have $|\log_p(1+q_0)| = |q_0|$. Thus for $s \in D$,

$$|(1+q_0)^s - 1| = |\exp(s\log_p(1+q_0)| < 1.$$

Since ψ of order a power of p, $\psi(1+q_0)$ is of p-power order, and so $|\psi(1+q_0)^{-1}(1+q_0)^s - 1| < 1$. Hence on substituting for $T = \psi(1+q_0)^{-1}(1+q_0)^s - 1$, in a power series with coefficients in $\mathcal{O}_\theta$, the series converges and defines an analytic function on D. Moreover the only possible zero for $h(T,\theta)$ is when $T = q_0$, equivalently when $s = 1$ and $\psi(1+q_0) = 1$. As $(1+q_0)$ is a generator for Γ, this implies ψ is trivial, and χ is an even character of the first kind. $L_p(s,\chi)$ is analytic except at $s = 1$, where it has a simple pole.

4.3. We will give an outline of the construction of the formal power series occurring in the theorem. In virtue of Theorem 4.1.1, and the above remark, it suffices to check that the values at the points $s = 1 - n,\ n$ a rational integer satisfy (b) of Theorem 4.1.1.

Let χ be an even character, and let $\chi = \theta\psi$ be the decomposition as characters of the first and second kind respectively. Let $\theta^* = \theta\omega^{-1}$. Since ψ is an even character, θ will be even, and θ^* an odd character. Let

$$S_n = \{a \mid 0 < a < q_n, (a, q_0) = 1\}.$$

Define $\xi_n(\theta),\ \eta_n(\theta) \in K_\theta[\Gamma_n]$, elements in the group ring of Γ_n as,

$$\xi_n(\theta) = -\frac{1}{q_n}\sum_{a\in S_n} a\theta^*(a)\gamma_n(a)^{-1}$$
$$\text{and } \eta_n(\theta) = (1 - (1+q_0)\gamma_n(1+q_0)^{-1})\xi_n(\theta).$$

Remark. Recall that upto a sign, the Stickelberger element ψ_n of K_n is,

$$\psi_n = -\frac{1}{q_n}\sum_{a\in S_n} a\delta(a)^{-1}\gamma_n(a)^{-1}.$$

$\xi_n(\theta)$ is then the projection of the Stickelberger element to the group ring $K_\theta[\Gamma_n]$ of Γ_n, with respect to the idempotent $\frac{1}{|\Delta|}\sum_{\delta\in\Delta}\theta^*(\delta)^{-1}\delta^{-1}$.

Proposition 4.3.1. *a) If $m \geq n \geq 0$, then $\xi_m(\theta) \mapsto \xi_n(\theta)$ and $\eta_m(\theta) \mapsto \eta_n(\theta)$, with respect to the map $K_\theta[\Gamma_m] \to K_\theta[\Gamma_n]$, induced by the projection map from Γ_m to Γ_n.*

b) $\frac{1}{2}\eta_n(\theta) \in \mathcal{O}_\theta[\Gamma_n]$.

c) If $\theta \neq 1$, then $\frac{1}{2}\xi_n(\theta) \in \mathcal{O}_\theta[\Gamma_n]$.

Proof. We outline a proof when p is assumed to be a odd prime.

a) Write $a \in S_{n+1}$ as $a = b + iq_n,\ b \in S_n,\ 0 < i < p$. Denote by ξ_n' the image of ξ_{n+1} in $K_\theta[\Gamma_n]$, with respect to the projection map $\Gamma_{n+1} \to \Gamma_n$. Then $\gamma_{n+1}(a) \mapsto \gamma_n(b)$ and $\theta^*(a) = \theta^*(b)$. We have,

$$\begin{aligned}\xi_n(\theta)' &= \xi_n(\theta) - \frac{1}{q_{n+1}}\sum_{0<i<p}\sum_{b\in S_n} iq_n\theta^*(b)\gamma_n(b)^{-1}\\ &= \xi_n(\theta) - \frac{p-1}{2}\sum_{b\in S_n}\theta^*(b)\gamma_n(b)^{-1}.\end{aligned} \tag{4.3.1}$$

Let $S_n' = \{a \mid a \in S_n,\ a < q_n/2\}$. Since θ^* is odd and γ_n is even, we get

$$\begin{aligned}\sum_{b\in S_n}\theta^*(b)\gamma_n(b)^{-1} &= \sum_{b\in S_n'}\theta^*(b)\gamma_n(b)^{-1} + \sum_{b\in S_n'}\theta^*(q_n - b)\gamma_n(q_n - b)^{-1}\\ &= 0.\end{aligned} \tag{4.3.2}$$

b)

$$\eta_n(\theta) = \xi_n(\theta) + \frac{1}{q_n} \sum_{a \in S_n} (1+q_0)a\theta^*(a)\gamma_n((1+q_0)a)^{-1}.$$

Write $(1+q_0)a = a' + a''q_n,\ 0 \leq a' \leq q_n$. Then $\omega((1+q_0)a) = \omega(a'),\ \theta((1+q_0)a) = \theta(a')$ and $\gamma_n((1+q_0)a) = \gamma_n(a')$. Further as a ranges over elements of S_n, so does a'. Hence it follows,

$$\eta_n(\theta) = \sum_{a \in S_n} a''\theta^*(a')\gamma_n(a')^{-1}. \tag{4.3.3}$$

c)

$$\xi_n(\theta) = -\frac{1}{q_n} \sum_{a \in S'_n} a\theta^*(a)\gamma_n(a)^{-1} + -\frac{1}{q_n} \sum_{a \in S'_n} (q_n - a)\theta^*(q_n - a)\gamma_n(q_n - a)^{-1}.$$

Since $\theta^*(q_n - a) = -\theta^*(a)$ and $\gamma_n(q_n - a) = \gamma_n(a)$, we have,

$$\xi_n(\theta) = -\frac{2}{q_n} \sum_{a \in S'_n} a\theta^*(a)\gamma_n(a)^{-1} + \sum_{a \in S'_n} \theta^*(a)\gamma_n(a)^{-1}.$$

Fix a_0 coprime to q_0. It can be from the definition of ω and of γ_n, that $\gamma_n(a) = \gamma_n(b)$ if and only if $\omega(a)^{-1}a \equiv \omega(b)^{-1}b \mod q_n$. Hence $\mod q_n$ we have,

(4.3.4)

$$\sum_{\substack{a \in S'_n \\ \gamma_n(a) = \gamma_n(a_0)}} a\omega(a)^{-1}\theta(a)\gamma_n(a)^{-1} = a_0\omega(a_0)^{-1}\gamma_n(a_0) \sum_{\substack{a \in S'_n \\ \gamma_n(a) = \gamma_n(a_0)}} \theta(a).$$

Now as a ranges over the indexing set in the above equation, it can be seen that the projection of a to Δ ranges over all the elements of Δ. Since θ is a non-principal character, this sum vanishes. □

Proof of the theorem. A slight extension of the isomorphism of the last section provides an identification of the group ring $\mathcal{O}_\theta[\Gamma] \simeq T\mathcal{O}_\theta[[T]]$. Let u be a generator of Γ. We recall that this identification is obtained by sending the generator u to $1+T$. Note that in the particular example we had, we could have taken $\gamma(1+q_0)$ as an explicit generator for Γ.

Via this identification, the proposition implies the existence of formal power series $f(T,\theta)$ (if $\theta \neq 1$), $g(T,\theta) \in \mathcal{O}_\theta[[T]]$ as,

$$f(T,\theta) = \mathrm{l}im_{\overleftarrow{n}}\, \xi_n(\theta), \text{if } \theta \neq 1, \tag{4.3.5}$$

$$\text{and } g(T,\theta) = \mathrm{l}im_{\leftarrow n} \xi_n(\theta), \text{if } \theta \neq 1. \tag{4.3.6}$$

In view of the theorem 4.1.1, we need to understand at the level of group rings, the effect on substituting $T = \zeta_\psi(1+q_0)^{1-m} - 1$, where m is a natural number. Define for n sufficiently large depending on the conductor of ψ and for a fixed integer t,

$$\phi^{\psi}_{m,n} : \mathcal{O}_\theta[\Gamma_n] \to \mathcal{O}_\theta/q_n\mathcal{O}_\theta$$

such that

$$\gamma_n(a) \mapsto \psi(a)^{-1} < a >^{-t},\ a \in \mathbb{Z},\ (a, q_0) = 1.$$

For n varying the various maps patch together to give a morphism of $\mathcal{O}_\theta$-algebras $\phi_t : \mathcal{O}_\theta[\Gamma] \to \mathcal{O}_\theta$.

Lemma 4.3.1. *Let $\xi \in \mathcal{O}_\theta[\Gamma]$ correspond to $f(T) \in \mathcal{O}_\theta[[T]]$ via the above isomorphism. Then*

$$\phi^{\psi}_t(\xi) = f(\psi(1+q_0)^{-1}(1+q_0)^{-t} - 1).$$

Proof. Assume first that $f(T) = 1+T$. Then $\xi = \gamma(1+q_0)$ and hence,

$$\phi_t(\gamma(1+q_0)) = \psi(1+q_0)^{-1}(1+q_0)^{-t} = f(\psi(1+q_0)^{-1}(1+q_0)^{-t} - 1).$$

The lemma now follows for all polynomials and hence for any f. □

In view of the above remarks, it suffices in order to prove the theorem to show that for any integer $m \geq 1$

$$g(\zeta_\psi(1+q_0)^{1-m} - 1, \theta)$$
$$= -h(\zeta_\psi(1+q_0)^{1-m} - 1, \theta)(1 - \chi\omega^{-m}(p)p^{m-1})B_{m,\chi\omega^{-m}}.$$

Applying $\phi_{t,n}$ to both sides of 4.3.3,

$$\phi_{t,n}(\eta_n) = \sum a''\chi_{t+1}(a')(a')^t \pmod{q)_n\mathcal{O}_\theta},$$

where $\chi_{t+1} = \chi\omega^{-t-1}$. From this it follows that,

$$(t+1)\phi_{t,n}(\eta_n) = -(1-\chi(1+q_0)(1+q_0)^{t+1})\frac{1}{q_n}\sum_a \chi_{t+1}(a)a^{t+1} \pmod{q)_n\mathcal{O}_\theta}.$$

Hence,

$$(t+1)\phi_t(\eta_n) = -(1-\chi(1+q_0)(1+q_0)^{t+1})\lim \frac{1}{q_n}\sum_a \chi_{t+1}(a)a^{t+1}.$$

It follows from the identification of $g(T,\theta)$ with $\lim \phi_t(\eta_n)$ that,

$$f(\chi(1+q_0)(1+q_0)^{-t} - 1, \theta) = -\frac{1}{t+1}\lim \frac{1}{q_n}\sum_a \chi_{t+1}(a)a^{t+1}.$$

We break the above sum into two parts, one indexed by a coprime to p, and the other by $p|a$, to obtain

$$\sum_a \chi_{t+1}(a)a^{t+1} = S_{t+1,\chi_{t+1}}(q_n) - \chi_{t+1}(p)p^{t+1}S_{t+1,\chi_{t+1}}(q_{n-1}),$$

with $S_{n,\chi}$ defined as in 4.1.1. It follows from Lemma 4.1.1 upon substituting $m = t + 1 \geq 1$,

$$f(\chi(1+q_0)(1+q_0)^{1-m} - 1, \theta) = -(1 - \chi_m(p)p^{m-1}\frac{B_{m,\chi\omega^{-m}}}{m}.$$

This gives an outline of the proof of the theorem, and for more details we refer the reader to the books by Iwasawa and Washington.

5. Main conjecture

In analogy with the Weil conjectures, it is natural to ask whether the characteristic function of the pro p-part of the class group X of a $\mathbb{Z}_p$ extension, considered in Section 2, is of zeta type. Note that this characteristic function is well defined upto a unit in the Iwasawa algebra. We have also seen, that the Stickelberger elements patch together along a cyclotomic $\mathbb{Z}_p$-extension, and gives rise to a power series, which interpolates the special values of Dirichlet L-series. The Main conjecture is indeed the expectation that these two power series should be equal upto a unit in the Iwasawa algebra.

Let $K_n = \mathbb{Q}(\zeta_{p^{n+1}})$, and θ be an odd character of $\Delta \simeq \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. Let X_n denote the Galois group of the maximal abelian p-extension Ω_n of K_n, unramified outside of the prime above p, and X be the profinite limit $X = \lim_{\leftarrow [n]} X_n$. The Iwasawa algebra,

$$\Lambda = \mathbb{Z}_p[\Gamma] = \lim_{\leftarrow} \mathbb{Z}_p[\Gamma_n] \simeq \mathbb{Z}_p[[T]].$$

For any character χ of Δ, denote by ϵ_χ the projector in the group ring $\mathbb{Z}_p[\Delta]$,

$$\epsilon_\chi = \frac{1}{(p-1)} \sum_{\delta \in \Delta} \chi^{-1}(\delta)\delta.$$

As seen in section 2, there is a map of Λ-modules,

$$\epsilon_\chi X \to \oplus_i \left(\Lambda/(p^{n_i(\chi)})\right) \oplus \left(\oplus_j \Lambda/(g_j(T,\chi)^{m_j})\right),$$

with finite cokernel. We denote by,

$$\mathrm{char}(X(\chi)) = \prod_i p^{n_i} \oplus \prod_j g_j(T,\chi)^{m_j},$$

the characteristic function of X as a Λ-module. $g(T,\chi)$ is well defined only upto a unit in Λ.

In the last section we had constructed $f(T,\chi) \in \Lambda$ satisfying for any natural number $m \geq 1$,

$$f(\psi(1+q_0)^{-1}(1+q_0)^{1-m} - 1, \chi) = L(1-m,\chi) =$$
$$-(1 - \chi\omega^{-n}(p)p^{n-1})\frac{B_{n,\chi\omega^{-n}}}{n}.$$

We can now state a form of Iwasawa's main conjecture, which has been proved by Mazur and Wiles.

THEOREM 5.0.2 (Main conjecture of Iwasawa theory). *For all nontrivial even characters χ of Δ,*

$$\mathrm{char}(X(\chi)) = f(T, \chi)\Lambda.$$

REFERENCES

(1) Ireland, K. and Rosen, M., A classical introduction to modern number theory, Graduate Texts in Math., Springer-Verlag.
(2) Iwasawa, K., Lectures on p-adic L-functions, Annals of Math. Studies **74**, Princeton University Press, Princeton.
(3) Lang, S., Cyclotomic fields, Graduate Texts in Math., Springer-Verlag.
(4) Washington, L.C., Introduction to Cyclotomic Fields, Graduate Texts in Math., **83**, Springer-Verlag.
(5) Weil, A., Number of solutions of equations in finite fields, Collected Papers I, 399-410.

C. S. Rajan
School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India.
e-mail: rajan@math.tifr.res.in

A Brief Introduction to Modular Forms

B. Ramakrishnan

1. Introduction

In this article, a brief introduction to the subject of modular forms is being given. For most of the statements I have not given the proofs. There are several good books written on this subject and I have listed some of them in the references. The interested reader can look at the books for the details of proofs. I have also given some articles related to the theory of newforms (which has been presented at the end). I hope that the reader will have some feeling for the subject from these notes. I would like to thank S. D. Adhikari and S. A. Katre for a careful reading of the manuscript.

2. Modular Forms over $SL_2(\mathbb{Z})$

Let k be a rational integer and let $\mathcal{H}$ denote the Poincaré upper half-plane, consisting of complex numbers z with positive imaginary part. Let $M_2(R)$ denote the set of all 2×2 matrices whose entries lie in the ring R. The *modular group*, denoted by $SL_2(\mathbb{Z})$ is the group defined as follows.

$$SL_2(\mathbb{Z}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in M_2(\mathbb{Z}) \middle| ad - bc = 1 \right\}.$$

It is a discrete subgroup of

$$GL_2^+(\mathbb{R}) = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in M_2(\mathbb{R}) \middle| ad - bc > 0 \right\}.$$

$GL_2^+(\mathbb{R})$ acts on $\mathcal{H}$ as follows. For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in GL_2^+(\mathbb{R})$ and $z \in \mathcal{H}$,

$$\gamma \cdot z = \frac{az+b}{cz+d}.$$

This is called the Möbius transformation. It is an easy exercise to check that

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

So, if $z \in \mathcal{H}$ and $\gamma \in GL_2^+(\mathbb{R})$, then $\gamma z \in \mathcal{H}$. i.e., $\mathcal{H}$ is preserved by the Möbius transformation. We also extend the action of $\gamma \in GL_2^+(\mathbb{R})$ to $\mathbb{Q} \setminus \{-d/c\}$ by the same rule.

Further, for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in GL_2^+(\mathbb{R})$ we define

$$\gamma \cdot \infty = \frac{a}{c} \quad \text{and } \gamma \cdot (\frac{-d}{c}) = \infty,$$

So, we have the action of $GL_2^+(\mathbb{R})$ on $\mathcal{H} \cup \{\infty\} \cup \mathbb{Q}$.

We first give the formal definition of a modular forms below and the explanations shall follow the definition.

Definition 2.1. *A function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is said to be a modular function (resp. form) of weight k for the full modular group $SL_2(\mathbb{Z})$, if it satisfies the following:*

(i) *f is a meromorphic (resp. analytic) function.*
(ii) $f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z),$ for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$.
(iii) *$f(z)$ has the following Fourier expansion.*

$$f(z) = \sum_{n=-m_0}^{\infty} a(n)q^n \qquad (q = e^{2\pi iz}), m_0 > 0$$

$$\left(\text{resp.} \quad f(z) = \sum_{n=0}^{\infty} a(n)q^n\right).$$

A modular form is said to be a cusp form if further $a(0) = 0$ in (iii) *above.*

Notation. The set of all modular forms (resp. cusp forms) of weight k for the group $SL_2(\mathbb{Z})$ is denoted by M_k (resp. S_k).

The condition (ii) above can be written as:

$$(cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right) =: f\Big|_k \gamma(z) = f(z) \quad \text{for all } \gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}).$$

In terms of this new stroke notation, we want the following.

$$f\Big|_k \gamma_1\Big|_k \gamma_2(z) = f\Big|_k \gamma_1\gamma_2(z).$$

In other words, defining $j(\gamma, z) = cz + d$ for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, we want the equality

$$j(\gamma_1\gamma_2, z) = j(\gamma_1, \gamma_2 \; z) \cdot j(\gamma_2, z),$$

which can easily be verified. The factor $j(\gamma, z)$ is called the *automorphy factor*. Let us now analyze the condition (iii). Consider the change of variable

$$z \longmapsto e^{2\pi iz} =: q.$$

This takes the Poincaré upper half-plane to the punctured open disk centred at the origin. We agree to take the point at ∞ to the origin under this map. The real line maps onto the boundary of the disk. Since we have the condition that f is meromorphic on $\mathcal{H}$, under this transformation, f has a Laurent expansion around the origin, say

$$f(z) = \sum_{-\infty}^{\infty} b(n)q^n.$$

We say that f is *meromorphic* at ∞ (resp. *holomorphic* at ∞) if $b(n) = 0$ for $n \ll 0$ (resp. for $n < 0$). Here $n \ll 0$ means that only finitely many coefficients $b(n)$ for $n < 0$ can be non-zero. So the condition (iii) is equivalent to saying that f is *meromorphic* (resp. *holomorphic*) at ∞.

Cusps. The points $\mathbb{Q} \cup \{\infty\}$ are called the *cusps*. Let $s \in \mathbb{Q}$. Then, s can be written in the reduced form $s = \frac{a}{c}$, where $\gcd(a, c) = 1$. Now complete a, c to get the matrix $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$. Then it is clear that $s = \gamma \cdot \infty$. But one has a natural equivalence (modulo the group $SL_2(\mathbb{Z})$) among the cusps, namely, the cusps s_1 and s_2 are said to be $SL_2(\mathbb{Z})$ equivalent if there exists a matrix $\gamma \in SL_2(\mathbb{Z})$ satisfying $s_1 = \gamma \cdot s_2$. From the above remarks. it is clear that all rational numbers are $SL_2(\mathbb{Z})$ equivalent to ∞. This is the reason why we have only one Fourier expansion for f.

Fundamental domain. The action of $SL_2(\mathbb{Z})$ on $\mathcal{H}$ divides $\mathcal{H}$ into equivalence classes, called *orbits*. Selecting one point from each orbit one gets a *fundamental set* for $SL_2(\mathbb{Z})$. Modifying the concept slightly for having nice topological properties, one obtains what are called *fundamental domains.* A fundamental domain for $SL_2(\mathbb{Z})$ is given as follows.

$$\mathcal{F} = \left\{ z \in \mathcal{H} \;\middle|\; |z| \geq 1, \frac{-1}{2} \leq \mathrm{Re}(z) \leq \frac{1}{2} \right\}. \tag{1}$$

The equivalence is defined as usual: two points z_1 and z_2 are said to be $SL_2(\mathbb{Z})$- equivalent, if there exists a matrix γ belonging to $SL_2(\mathbb{Z})$ which takes one to the other. The fact that $\mathcal{F}$ is a fundamental domain for $SL_2(\mathbb{Z})$ is equivalent to the following:

(i) For any $z \in \mathcal{H}$, there exists a unique z_1 belonging to $\mathcal{F}$ such that $z = \gamma z_1$, for some $\gamma \in SL_2(\mathbb{Z})$.

(ii) No two interior points of $\mathcal{F}$ are $SL_2(\mathbb{Z})$ equivalent. (If we take the fundamental domain suitably, one can omit the condition "interior".)

Let us briefly indicate the proof of the fact that $\mathcal{F}$ is a fundamental domain for $SL_2(\mathbb{Z})$. The modular group $SL_2(\mathbb{Z})$ contains the following two special matrices.

$$T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right); \qquad Tz = z + 1 \quad \text{(translation.)}$$
$$S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right); \qquad Sz = -1/z \quad \text{(inversion.)}$$

For a given point $z \in \mathcal{H}$, the idea is to apply the translations T^m ($m \in \mathbb{Z}$ to get a point, say z_1, inside the strip $-1/2 \leq \mathrm{Re}(z_1) \leq 1/2$. If this point z_1 is inside $\mathcal{F}$, then we are through. Otherwise, apply the inversion map S to z_1 to get another point z_2, which will lie outside the unit circle. The process is continued till we get a point inside our region $\mathcal{F}$. (It is a fact that this process ends after a finite number of steps!) This proof implies an interesting fact that the modular group $SL_2(\mathbb{Z})$ is generated by the two elements T and S.

Weight formula. Let f be a complex valued meromorphic function defined on $\mathcal{H}$. For a point $P \in \mathcal{H}$, let $v_P(f)$ denote the order of zero of f at P and $v_\infty(f)$ denote the least integer n for which $a(n)$ is non-zero in the expansion (iii) of Definition 2.1. Then we have the following theorem.

Theorem 2.1. *Let f be a non-zero modular function of weight k for the group $SL_2(\mathbb{Z})$. Then*

$$v_\infty(f) + \frac{1}{3}v_\rho(f) + \frac{1}{2}v_i(f) + \sum_{\substack{P \in \mathcal{H}\backslash SL_2(\mathbb{Z}) \\ P \neq i,\rho}} v_P(f) = \frac{k}{12}, \tag{2}$$

where $\rho = -1/2 + i\sqrt{3}/2$.

The above theorem is proved by integrating the function f'/f along a specific contour in $\mathcal{H}$ and using the residue theorem. We omit the details here. This theorem has many applications as we shall see in the sequel.

Remark 2.1. (1) There is no modular form of weight $k < 0$. (This is clear from the formula (2) as the left hand side is always non-negative.)

(2) There is no non-zero modular form of odd weight. This will follow directly using Definition 2.1. Take the matrix $\gamma = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. By the condition (ii) of Definition 2.1, we must have, $(-1)^k f(z) = f\Big|_k \gamma(z) = f(z)$. If k is odd, it follows that $f(z) = 0$.

(3) If f is a modular form of weight $k = 0$, then $f \in \mathbb{C}$. Assume that $f \neq 0$ is a modular form of weight 0. Then by weight formula, f does not vanish on $\mathcal{H} \cup \{\infty\}$. Put $c = v_\infty(f)$. Then, $g = f - c$ is a modular form of weight 0 and it vanishes at ∞ (by definition), This implies that $g = 0$. Therefore, $f = c \in \mathbb{C}$.

(4) Let $k = 2$. Then all the terms on the left hand side of (2) are non-negative. So, we must have $f = 0$.

From the above remark, it follows that the first non-trivial example of a modular form occurs only when $k \geq 4$. We shall present now an example of a modular form, which plays an important role in the theory of modular forms.

Eisenstein series. For $k \geq 4$, and $z \in \mathcal{H}$, put

$$G_k(z) = \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} (mz + n)^{-k} \tag{3}$$

We will show that $G_k(z) \in M_k$.
Since $k \geq 4$, the series on the right hand side of (3) converges absolutely and uniformly on any compact subset of $\mathcal{H}$. Therefore $G_k(z)$ defines a holomorphic function on $\mathcal{H}$. Now consider

$$\begin{aligned}\lim_{z\to i\infty} G_k(z) &= \lim_{z\to i\infty} \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} (mz+n)^{-k} \\ &= \sum_{\substack{n\in\mathbb{Z}\\n\neq 0}} n^{-k} + \lim_{z\to i\infty} \sum_{\substack{m,n\in\mathbb{Z}\\m\neq 0}} (mz+n)^{-k} \\ &= 2\sum_{n\geq 1} n^{-k} \\ &= 2\zeta(k),\end{aligned} \tag{4}$$

where $\zeta(s) = \sum_{n\geq 1} n^{-s}$ is the well known Riemann zeta function. Since the limit exists, it is clear that $G_k(z)$ has no negative term in its Fourier expansion and in fact, we have shown that the constant term in the Fourier expansion is $2\zeta(k)$. This implies that $G_k(z)$ is holomorphic at $i\infty$. It remains to prove that $G_k(z)$ is invariant under $SL_2(\mathbb{Z})$ with respect to the stroke operator $\Big|_k$.

$$\begin{aligned}G_k\Big|_k T(z) &= \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} (m(z+1)+n)^{-k} \\ &= \sum_{\substack{m,n'\in\mathbb{Z}\\(m,n')\neq(0,0)}} (mz+n')^{-k} \\ &= G_k(z).\end{aligned} \tag{5}$$

Next we consider the transformation with respect to S.

$$\begin{aligned}G_k\Big|_k S(z) &= z^{-k} \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} (m(-1/z)+n)^{-k} \\ &= \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} (-m+nz)^{-k} \\ &= G_k(z).\end{aligned} \tag{6}$$

(In the last line, we have used the fact that the series converges absolutely.) Since S and T generate $SL_2(\mathbb{Z})$, it follows that $G_k(z)$ is invariant under $SL_2(\mathbb{Z})$ with respect to the stroke operation. We have thus established the fact that $G_k(z)$ is a modular form of weight k for the group $SL_2(\mathbb{Z})$.

Let us now derive the Fourier expansion of $G_k(z)$.

First let us prove the following lemma which is needed in getting the Fourier expansion of $G_k(z)$.

Lemma 2.2.

$$\zeta(k) = \frac{-(2i\pi)^k}{2(k!)} B_k, \tag{7}$$

where B_k denotes the $k-$th Bernoulli number defined by

$$\frac{x}{e^x - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} x^k$$

Proof. The product formula for sine function is

$$\sin \pi z = \pi z \prod_{n=1}^{\infty} \left(1 - z^2/n^2\right).$$

Taking logarithmic derivative with respect to z,

$$\begin{aligned} \pi \cot \pi z &= \frac{1}{z} + \sum_{n=1}^{\infty} \left(\frac{1}{z+n} + \frac{1}{z-n} \right) \qquad (8) \\ \text{i. e.,} \quad \pi z \cot \pi z &= 1 + 2\sum_{n=1}^{\infty} \frac{z^2}{z^2 - n^2}. \end{aligned}$$

Substituting $2i\pi z = x$ on the right hand side of the above equation, we get,

$$\begin{aligned} \pi z \cot \pi z &= 1 + 2\sum_{n=1}^{\infty} \frac{x^2}{x^2 + (2\pi n)^2} \\ &= 1 + 2\sum_{n=1}^{\infty} \left(\frac{x}{2\pi n}\right)^2 \left(1 + \left(\frac{x}{2\pi n}\right)^2\right)^{-1} \\ &= 1 + 2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty} (-1)^{m-1} \left(\frac{x}{2\pi n}\right)^{2m} \\ &= 1 + 2\sum_{m=1}^{\infty} \zeta(2m) \left(\frac{x}{2\pi}\right)^{2m}. \end{aligned} \tag{9}$$

On the other hand,

$$\begin{aligned} \pi z \cot \pi z = \pi z \frac{\cos \pi z}{\sin \pi z} \quad &= i\pi z \frac{e^{i\pi z} + e^{-i\pi z}}{e^{i\pi z} - e^{-i\pi z}} \\ = i\pi z \frac{e^{2i\pi z} + 1}{e^{2i\pi z} - 1} \quad &= i\pi z + \frac{2i\pi z}{e^{2i\pi z} - 1} \\ = \frac{x}{2} + \frac{x}{e^x - 1} \quad &= \frac{x}{2} + \sum_{\ell=0}^{\infty} \frac{B_\ell}{\ell!} x^\ell. \end{aligned} \tag{10}$$

Comparing the k-th power of x in (9) and (10), we get the required result. □

Proposition 2.3.

$$G_k(z) = 2\zeta(k)\left[1 - \frac{2k}{B_k}\sum_{n\geq 1}\sigma_{k-1}(n)q^n\right]. \tag{11}$$

In the above, $\sigma_r(n) = \sum_{d|n} d^r$.

Proof. As in the proof of the above lemma, we have,

$$\begin{aligned}\pi\cot\pi z &= i\pi\left(1 - \frac{2}{1-e^{2i\pi z}}\right)\\ &= i\pi\left(1 - 2\sum_{n=0}^{\infty} e^{2i\pi nz}\right)\\ &= -i\pi\left(1 + 2\sum_{n=1}^{\infty} q^n\right).\end{aligned} \tag{12}$$

Recalling (8), we have,

$$\begin{aligned}\pi\cot\pi z &= \frac{1}{z} + \sum_{n=1}^{\infty}\left(\frac{1}{z+n} + \frac{1}{z-n}\right)\\ &= \frac{1}{z} + \sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\left(\frac{1}{z+n} - \frac{1}{n}\right)\end{aligned} \tag{13}$$

Comparing (12) and (13), we have,

$$\frac{1}{z} + \sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\left(\frac{1}{z+n} - \frac{1}{n}\right) = -i\pi\left(1 + 2\sum_{n=1}^{\infty} q^n\right). \tag{14}$$

Differentiating the above (with respect to z) $k-1$ times, we get,

$$(k-1)!\sum_{n=-\infty}^{\infty}(z+n)^{-k} = (2i\pi)^k\sum_{n=1}^{\infty} n^{k-1}q^n. \tag{15}$$

Using the definition of $G_k(z)$, we have,

$$\begin{aligned} G_k(z) &= 2\zeta(k) + 2\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty}(mz+n)^{-k} \\ &= 2\zeta(k) + 2\frac{(2i\pi)^k}{(k-1)!}\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} n^{k-1}q^{mn} \qquad \text{(using (15))} \\ &= 2\zeta(k) - \frac{4k}{B_k}\zeta(k)\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} n^{k-1}q^{mn} \text{ (using Lemma 2.2)} \\ &= 2\zeta(k)\left[1 - \frac{2k}{B_k}\sum_{n\geq 1}\sigma_{k-1}(n)q^n\right]. \end{aligned}$$

This completes the proof. □

We put

$$E_k(z) = \frac{1}{2\zeta(k)}G_k(z).$$

Then from the above Proposition, the Fourier expansion of E_k $(k \geq 4)$ can be written as

$$E_k(z) = 1 - \frac{2k}{B_k}\sum_{n\geq 1}\sigma_{k-1}(n)q^n. \tag{16}$$

The Eisenstein series $E_k(z)$ is called the *normalised* Eisenstein series.

Before we move onto another important example of a modular form (in fact, a cusp form), we remark about some properties of modular forms.

Remark 2.2.

(1) If $f \in M_k$, then λf also belongs to M_k, for every $\lambda \in \mathbb{C}$.
(2) If $f, g \in M_k$, then $f + g \in M_k$.
(3) If $f \in M_{k_1}$ and $g \in M_{k_2}$, then $fg \in M_{k_1+k_2}$.
(4) If f and g are the same as in the previous case with $g \neq 0$, then f/g is a modular function of weight $k_1 - k_2$.
(5) The above facts imply that M_k is a $\mathbb{C}$ -vector space.

Using the weight formula (2), one can arrive at the conclusion that the $\mathbb{C}$ -vector space M_k is one-dimensional for $4 \leq k \leq 10$, and in these cases M_k is generated by the Eisenstein series G_k. So, there is no cusp form of weight $k \leq 10$. The first example of a cusp form occurs when $k = 12$.

An example of a cusp form. Put

$$\Delta(z) = \frac{1}{1728}\left(E_4(z)^3 - E_6(z)^2\right). \tag{17}$$

Since $E_4^3(z)$ and $E_6^2(z)$ belong to M_{12}, $\Delta(z)$ is a modular form of weight 12 for $SL_2(\mathbb{Z})$. Using the Fourier expansion of $E_k(z)$ given by (16), we get

$$\begin{aligned}E_4(z) &= 1 + 240\sum_{n\geq 1}\sigma_3(n)q^n\\ E_6(z) &= 1 - 504\sum_{n\geq 1}\sigma_5(n)q^n\end{aligned}$$

and so, we have

$$\begin{aligned}E_4^3(z) &= 1 + 720\ q + 179280\ q^2 + \cdots\\ E_6^2(z) &= 1 - 1008\ q - 220752\ q^2 + \cdots .\end{aligned}$$

Therefore,

$$\begin{aligned}\Delta(z) &= \frac{1}{1728}\left(E_4(z)^3 - E_6(z)^2\right)\\ &= \frac{1}{1728}\left(1728q - 41472q^2 + \cdots\right)\\ &= \sum_{n=1}^{\infty}\tau(n)q^n,\end{aligned} \tag{18}$$

and hence $\Delta(z)$ is a cusp form of weight 12 for $SL_2(\mathbb{Z})$.
In the above, $\tau(n)$ is called the Ramanujan's tau function. As remarked above, $\Delta(z)$ is the first example of a cusp form. S. Ramanujan is the first mathematician to notice some nice arithmetical properties of the function $\tau(n)$. From the definition it is clear that $\tau(n) \in \mathbb{Z}\ \forall\ n$.

The first few values of $\tau(n)$ are: $\tau(1) = 1, \tau(2) = -24, \tau(3) = 252$, $\tau(4) = -1472$, $\tau(5) = 4830$, $\tau(6) = -6048$, $\tau(7) = -16744$, $\tau(8) = 84480$, $\tau(9) = -113643$, $\tau(10) = -115920$, $\tau(11) = 534612$, $\tau(12) = -370942, \cdots$. There is a table of $\tau(n)$'s for $n \leq 1000$ given by G. N. Watson. The interested reader can refer to [12]. There is a remarkable product formula for $\Delta(z)$ (due to Jacobi), which we shall give below without proof.

$$\Delta(z) = \sum_{n\geq 1}\tau(n)q^n = q\prod_{n\geq 1}(1-q^n)^{24} = (\eta(z))^{24}. \tag{19}$$

(In the above, $\eta(z)$ is the Dedekind eta-function.) The Ramanujan function $\tau(n)$ satisfies the following remarkable congruence:

$$\tau(n) \equiv \sigma_{11}(n) \pmod{691} \qquad \text{for all } n \geq 1. \tag{20}$$

There are some other congruences involving $\tau(n)$. But we shall not go into the details here.

REMARK 2.3. The space M_k is a finite dimensional vector space. Moreover, it is generated by the Eisenstein series $E_4(z)$ and $E_6(z)$. That is, given any modular form $f(z)$ in M_k, it can be expressed as a linear combination of E_4 and E_6. More precisely,

$$f(z) = \sum_{\substack{0\leq a,b\in\mathbb{Z}\\ 4a+6b=k}} c_{a,b}\ E_4(z)^a E_6(z)^b, \tag{21}$$

where $c_{a,b} \in \mathbb{C}$. Further, one has:

$$M_k = \mathbb{C}E_k \bigoplus S_k. \tag{22}$$

In fact, multiplication by the discriminant function $\Delta(z)$ gives an isomorphism between M_{k-12} and M_k. Using this one has the following dimension formula.

$$\dim_{\mathbb{C}} M_k = \begin{cases} \left[\frac{k}{12}\right] & if\ k \equiv 2 \pmod{12} \\ 1 + \left[\frac{k}{12}\right] & if\ k \not\equiv 2 \pmod{12} \end{cases} \tag{23}$$

The modular invariant $j(z)$.

The Klein's modular invariant is defined as follows:

$$j(z) = 1728\frac{E_4(z)^3}{E_4(z)^3 - E_6(z)^2} = \frac{E_4(z)^3}{\Delta(z)}. \tag{24}$$

Since $E_4(z)^3$ and $\Delta(z)$ are modular forms of the same weight (weight 12) for $SL_2(\mathbb{Z})$, by Remark 2.2, $j(z)$ is a modular function of weight 0 for $SL_2(\mathbb{Z})$. Since it is a modular function of weight 0 for $SL_2(\mathbb{Z})$, from the definition it follows that $j(z)$ is invariant under $SL_2(\mathbb{Z})$. That is

$$j\left(\frac{az+b}{cz+d}\right) = j(z) \text{ for all } \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}).$$

(That is why this function is called the *modular invariant*.) This function was studied extensively by F. Klein.

By the weight formula, it can be seen that

$$\Delta(z) \neq 0 \qquad \text{for all } z \in \mathcal{H}.$$

Further, $\Delta(z)$ vanishes only at ∞. Therefore, by the definition, $j(z)$ is a holomorphic function on $\mathcal{H}$ and it has a simple pole at ∞. It has the following Fourier expansion.

$$j(z) = q^{-1} + 744 + \sum_{n\geq 1} c(n)\ q^n. \tag{25}$$

The coefficients $c(n)$'s are integers. These coefficients also satisfy some nice congruence properties. Here again, we shall not go into the details. In the following remark, we shall mention some of the main properties of $j(z)$.

Remark 2.4.

1. The modular function j defines a bijection of $\mathcal{H}/SL_2(\mathbb{Z})$ onto $\mathbb{C}$.
2. The following are equivalent:
   (a) f is a modular function of weight 0 for $SL_2(\mathbb{Z})$.
   (b) f is a quotient of two modular forms of the same weight for $SL_2(\mathbb{Z})$.
   (c) f is a rational function of $j(z)$.

Hecke theory. In 1916, S. Ramanujan conjectured the following properties satisfied by the Ramanujan function $\tau(n)$.

(i) $\tau(n)$ is a multiplicative function. i.e.,

$$\tau(mn) = \tau(m)\tau(n) \quad \text{if } \gcd(m,n) = 1. \tag{26}$$

(ii) $|\tau(p)| \leq 2p^{11/2}$.

Both of these conjectures have been proved; the first one by L. J. Mordell in 1917 and the second one by P. Deligne in 1973. The proof of conjecture (i) is the starting point of the theory of Hecke operators. Here, we shall briefly explain the Hecke theory for modular forms of integral weight.

$$\text{Put} \qquad \Delta_n = \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2(\mathbb{Z}) \middle| ad - bc = n \right\}. \tag{27}$$

Now, define an equivalence relation on Δ_n as follows. Two elements γ_1, γ_2 of Δ_n are said to be equivalent (modulo $SL_2(\mathbb{Z})$) if and only if $\gamma_1\gamma_2^{-1} \in SL_2(\mathbb{Z})$. It can be easily seen that this is an equivalence relation. So, one has the following decomposition for Δ_n:

$$\Delta_n = \cup_i SL_2(\mathbb{Z})\gamma_i, \tag{28}$$

where γ_i's are a finite number of representatives for the equivalence classes. Define the n-th Hecke operator by

$$f\Big|T_n(z) := n^{\frac{k}{2}-1} \sum_i f\Big|_k \gamma_i \qquad (f \in M_k). \tag{29}$$

(Here $f\Big|_k \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)(z) := (ad-bc)^{k/2}(cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right)$.)

In our case, the exact set of representatives γ_i's are given as follows:

$$\left\{ \left(\begin{smallmatrix} a & b \\ 0 & d \end{smallmatrix}\right) \middle| a, b, d \in \mathbb{Z} \text{ such that } ad = n, 0 \leq b < d \right\}. \tag{30}$$

When $n = p$, a prime, and for $f \in M_k$, one has:

$$\begin{aligned} f\Big|T_p &= p^{\frac{k}{2}-1} \sum_{\substack{ad=p\\ 0\leq b<d}} f\Big|_k \left(\begin{smallmatrix} a & b\\ 0 & d\end{smallmatrix}\right) \\ &= p^{k-1} f(pz) + p^{-1} \sum_{0\leq b<d} f\left(\frac{z+b}{p}\right). \end{aligned} \tag{31}$$

If $f(z) = \sum_{n\geq 0} a_f(n)\ q^n$, then it is easily seen that

$$f\Big|T_p(z) = \sum_{n\geq 0} b(n)\ q^n, \tag{32}$$

where

$$b(n) = a_f(np) + p^{k-1} a_f(n/p) \qquad (n \geq 0). \tag{33}$$

(Notation: $a_f(m)$ denotes the m-th Fourier coefficient of f and it is defined to be zero when m is not an integer.)

From the work of H. Petersson, there is an inner product defined in the space M_k. Let $f, g \in M_k$ with f or g a cusp form. Then the Petersson inner product of f and g is defined by

$$\langle f, g\rangle = \int_{\mathcal{F}} f(z)\overline{g(z)} y^{k-2} dx dy \qquad (z = x + iy). \tag{34}$$

The requirement that f or g to be a cusp form is needed for the convergence of the integral.

The Hecke operators T_p for all primes p are hermitian with respect to the Petersson scalar product. Further, they form a commuting family of operators. Therefore, we have the following theorem from linear algebra.

Theorem 2.4. *The vector space S_k has a basis of simultaneous eigenforms with respect to all Hecke operators T_n.*

Remark 2.5. Note that the basis constructed in Theorem 2.4 is orthogonal.

Further, it can be shown that the Hecke operators satisfy the following commuting property in general.

$$T_m T_n = \sum_{d|\ \gcd(m,n)} d^{k-1} T_{mn/d^2}. \tag{35}$$

Therefore, for a Hecke eigenform (this means an eigenfunction with respect to all Hecke operators T_n) $f \in S_k$, we have,

$$a_f(m)a_f(n) = \sum_{d|\,\gcd(m,n)} d^{k-1} a_f(mn/d^2). \tag{36}$$

Hence, $a_f(m)a_f(n) = a_f(mn)$ if $\gcd(m,n) = 1$.
Since S_{12} is one-dimensional, $\Delta(z)$ is a Hecke eigenform and hence by the above observation, we get

$$\tau(m)\tau(n) = \tau(mn) \quad \text{if } \gcd(m,n) = 1, \tag{37}$$

which proves the conjecture (i) of Ramanujan mentioned above.

Theorem 2.5. *Let $k \geq 4$ and let $f = \sum_{n \geq 0} a_f(n) q^n \in M_k$ be a Hecke eigenform. Then $a_f(1) \neq 0$. In other words, the function f can be normalised.*

Proof. Since f is a Hecke eigenform, the following is true for all primes p and for all $n \geq 0$:

$$a_f(np) + p^{k-1} a_f(n/p) = \lambda_p a_f(n), \tag{38}$$

where λ_p is the eigenvalue.

If $a_f(1) = 0$, then it follows from the above expression that,

$$a_f(p) = 0 \quad \text{for all primes } p \tag{39}$$

which implies that $a_f(n) = 0$ for all $n \geq 1$.
Therefore, $f(z) = a_f(0) \in \mathbb{C}$. Since $k \geq 4$, $a_f(0) = 0$, and so we have, $f = 0$, a contradiction. Therefore, $a_f(1) \neq 0$. □

Definition 2.2. *A form $f \in M_k$ is said to be normalised, if the leading term in its Fourier expansion is equal to* 1*.*

Remark 2.6. We can find an orthogonal basis of S_k consisting of normalised Hecke eigenforms.

Theorem 2.6. *Let f be a normalised Hecke eigenform belonging to S_k. Then the eigenvalue of f for T_p is $a_f(p)$. Further, we have*

$$a_f(m)a_f(n) = \sum_{d|\,\gcd(m,n)} d^{k-1} a_f\left(mn/d^2\right) \tag{40}$$

if and only if f is a normalised Hecke eigenform in S_k.

Proof. If f is a Hecke eigenform with eigenvalue λ_p for T_p, then we have

$$a_f(np) + p^{k-1} a_f(n/p) = \lambda_p a_f(n) \qquad \text{for all } n \geq 0. \tag{41}$$

Substituting $n = 1$ in the above and using $a_f(1) = 1$, we have $\lambda_p = a_f(p)$.

We shall now prove the second statement. First assume that

$$a_f(m)a_f(n) = \sum_{d|\gcd(m,n)} d^{k-1} a_f\left(mn/d^2\right) \quad \forall\ m,n \geq 1. \tag{42}$$

Taking $m = p$ in the above equation we have,

$$a_f(np) + p^{k-1}a_f(n/p) = a_f(p)a_f(n) \qquad \forall\ p,n \geq 1$$

i.e.,

$$f\Big|T_p = a_f(p)f \qquad \forall\ p.$$

Since $a_f(1) = 1$ (by putting $n = 1$, $m = p$), we see that f is a normalised Hecke eigenform.
Conversely, let $f \in S_k$ be a normalised Hecke eigenform. Then, we have

$$a_f(np) + p^{k-1}a_f(n/p) = a_f(p)a_f(n) \qquad \forall\ p,n \geq 1$$

i.e.,

$$a_f(p^{n+1}) = a_f(p)a_f(p^n) - p^{k-1}a_f(p^{n-1}), \quad n \geq 1$$

and

$$a_f(np^\alpha) = a_f(p)a_f(np^{\alpha-1}) - p^{k-1}a_f(np^{\alpha-2}), \alpha \geq 2, \gcd(n,p) = 1.$$

From this we conclude that

$$a_f(m)a_f(n) = a_f(mn) \quad \text{if } \gcd(m,n) = 1.$$

The general identity can be easily proved and we leave it to the reader to verify. □

Remark 2.7. It can be easily seen that

$$\sigma_k(m)\sigma_k(n) = \sum_{d|\gcd(m,n)} d^k \sigma_k\left(mn/d^2\right).$$

To prove this one has to use the following.

$$\begin{aligned} \sigma_k(p^{n+1}) &= \sigma_k(p)\sigma_k(p^n) - p^k\sigma_k(p^{n-1}) \quad n \geq 1, \\ \sigma_k(mn) &= \sigma_k(m)\sigma_k(n) \quad \text{if } \gcd(m,n) = 1. \end{aligned} \tag{43}$$

Theorem 2.7. *Let $f, g \in S_k$ be two normalised Hecke eigenforms whose eigenvalues with respect to the Hecke operators T_p are equal. Then $f = g$.*

(Note: The above theorem is often referred to as the “multiplicity 1” theorem.)

Proof. We have proved that if f is a normalised Hecke eigenform, then the eigenvalue with respect to the Hecke operator T_p is the p-th Fourier coefficient $a_f(p)$ of f in the q-expansion. If f and g have the same eigenvalues for all T_p, then we have,

$$a_f(p) = a_g(p) \quad \forall\ p.$$

Since $a_f(1) = a_g(1) = 1$, we must have,

$$a_f(n) = a_g(n) \quad \forall\ n \geq 1.$$

This completes the proof. □

L-functions associated to modular forms. We have the following theorem of Hecke on the estimates for the Fourier coefficients of modular forms.

Theorem 2.8. *(Hecke) Let $f \in S_k$. Then, we have*

$$a_f(n) = O(n^{k/2}).$$

(In other words, the quotient $\frac{|a_f(n)|}{n^{k/2}}$ remains bounded when $n \to \infty$.)

As a consequence, we have the following corollary.

Corollary 2.9. *If $f \in M_k$ and f is not a cusp form, then $a_f(n) = O(n^{k-1})$.*

Remark 2.8. The exponent $k/2$ of the above theorem can be improved. In fact, the Ramanujan-Petersson conjecture says (Ramanujan for the weight $k = 12$ and Petersson for general k) that if $f \in S_k$, is a normalised Hecke eigenform, then

$$a_f(n) = O\left(n^{k/2-1/2}\sigma_0(n)\right).$$

This implies that $a_f(n) = O(n^{k/2-1/2+\epsilon})$ for every $\epsilon > 0$. As mentioned before, this conjecture was proved by P. Deligne as a consequences of the "Weil conjectures" about algebraic varieties over finite fields.

Let $f \in M_k$. Then the Dirichlet series associated to f is defined by

$$L_f(s) = \sum_{n \geq 1} a_f(n) n^{-s}. \tag{44}$$

We have seen that $a_f(n) = O(n^{k/2})$ if f is a cusp form and $a_f(n) = O(n^{k-1})$ if f is not a cusp form. Therefore, the Dirichlet series defined by (44) converges absolutely for Re $(s) > k/2 + 1$ if f is a cusp form and for Re $(s) > k$ if f is not a cusp form. E. Hecke found a remarkable connection between each modular form and its associated Dirichlet series.

Theorem 2.10. *(Hecke) If the Fourier coefficients $a_f(n)$ satisfy the multiplicative property*

$$a_f(m)a_f(n) = \sum_{d|\, \gcd(m,n)} d^{k-1}a_f(mn/d^2), \tag{45}$$

then the Dirichlet series $L_f(s)$ has an Euler product expansion of the form

$$L_f(s) = \prod_p \left(1 - a_f(p)p^{-s} + p^{k-1-2s}\right)^{-1}, \tag{46}$$

absolutely convergent with the Dirichlet series.

Remark 2.9. For the Ramanujan's discriminant function, we have the Euler product representation as follows.

$$\sum_{n=1}^{\infty} \tau(n)n^{-s} = \prod_p \left(1 - \tau(p)p^{-s} + p^{11-2s}\right)^{-1} \quad \text{for Re } (s) > 7.$$

Hecke also deduced the following analytic properties of $L_f(s)$.

Theorem 2.11. *(Hecke) Assume that $k \geq 4$. Let $L_f(s)$ be the Dirichlet series associated to the modular form $f \in M_k$, which is defined for* Re $(s) > k$. *Then, $L_f(s)$ can be continued analytically beyond the line* Re $(s) = k$ *with the following properties.*

(i) *If $a_f(0) = 0$, $L_f(s)$ is an analytic function of s.*

(ii) *If $a_f(0) \neq 0$, $L_f(s)$ is analytic for all s except for a simple pole at $s = k$ with residue*

$$\frac{(-1)^{k/2}a_f(0)(2\pi)^k}{\Gamma(s)}.$$

(iii) *The function $L_f(s)$ satisfies the functional equation*

$$(2\pi)^{-s}\Gamma(s)L_f(s) = (-1)^{k/2}(2\pi)^{s-k}\Gamma(k-s)L_f(k-s). \tag{47}$$

Remark 2.10. Hecke also proved a converse to the above theorem. He proved that every Dirichlet series which satisfies a functional equation of the type stated in the theorem together with some analytic and growth conditions necessarily arises from a modular form in M_k.

Remark 2.11. Let f be a modular form in M_k. Then f is a normalised Hecke eigenform if and only if the associated Dirichlet series $L_f(s)$ has an Euler product of the form

$$L_f(s) = \prod_p \left(1 - a_f(p)p^{-s} + p^{k-1-2s}\right)^{-1}.$$

3. Modular Forms of Higher Level

Let $N \in \mathbb{N}$. The subgroup

$$\Gamma(N) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}) \middle| \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \pmod N \right\} \tag{48}$$

of $SL_2(\mathbb{Z})$ is called the *principal* congruence subgroup of level N.

Definition 3.1. *A subgroup Γ' of $SL_2(\mathbb{Z})$ is called a congruence subgroup, if it contains $\Gamma(N)$ for some N. The least N for which $\Gamma(N) \subset \Gamma'$ is called the level of the congruence subgroup.*

Examples.

$$\Gamma_1(N) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}) \middle| \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \equiv \left(\begin{smallmatrix} 1 & * \\ 0 & 1 \end{smallmatrix}\right) \pmod N \right\}. \tag{49}$$

$$\Gamma_0(N) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z}) \middle| \; c \equiv 0 \pmod N \right\}. \tag{50}$$

Definition 3.2. *A function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is said to be a modular form of weight k for $\Gamma_0(N)$ with character χ (χ is a Dirichlet character modulo N), if*

(i) *f is an analytic function.*

(ii) $f\left(\frac{az+b}{cz+d}\right) = \chi(d)(cz+d)^k f(z)$, for all $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_0(N)$.

(iii) *for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in SL_2(\mathbb{Z})$, the function $(cz+d)^{-k} f\left(\frac{az+b}{cz+d}\right)$ has a Fourier expansion of the form*

$$\sum_{n \geq 0} b_\gamma(n) q_N^n \qquad (q_N = e^{2\pi i z/N}).$$

As before, (iii) is equivalent to saying that f is holomorphic at all cusps of $\Gamma_0(N)$. Let us elaborate a little bit more. $\Gamma_0(N)$ is a subgroup of finite index in $SL_2(\mathbb{Z})$. Let $[SL_2(\mathbb{Z}) : \Gamma_0(N)] = \mu$. Then $SL_2(\mathbb{Z})$ can be written as

$$SL_2(\mathbb{Z}) = \cup_{i=1}^{\mu} \Gamma_0(N)\gamma_i.$$

Then, the set $\left\{ \gamma_i \infty \middle| 1 \leq i \leq \mu \right\}$ contains the set of inequivalent cusps modulo $\Gamma_0(N)$. Let s be a cusp belonging to $\mathbb{Q}$. Let $\gamma_0 \in SL_2(\mathbb{Z})$ be such that $\gamma_0 \infty = s$. Let f be a function satisfying the condition (ii) stated in the above definition. Put $g = f\Big|_k \gamma_0$. Then, it can be checked that g satisfies the condition (ii) with respect to the group $\gamma_0^{-1}\Gamma_0(N)\gamma_0 \supset \Gamma(N) \ni \left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right)$. Therefore, $g(z+N) = g(z)$. This means that g has a Fourier expansion

$$g(z) = \sum_{n=-\infty}^{\infty} a_g(n) q_N^n. \tag{51}$$

Now, we say that f is meromorphic (resp. holomorphic) at the cusp s, if $a_g(n) = 0$ for $n \ll 0$ (resp. for $n < 0$).

Note. A modular form f defined by Definition 3.2 is said to be a cusp form if $b_\gamma(0) = 0$ for all $\gamma \in SL_2(\mathbb{Z})$ in condition (iii). The set of all modular (resp. cusp) forms of weight k for the group $\Gamma_0(N)$ with character χ is denoted as $M_k(N, \chi)$ (resp. $S_k(N, \chi)$). When χ is a trivial character, then the respective spaces are denoted as $M_k(N)$ and $S_k(N)$.

Facts.

(a) $M_k(N,\chi)$ is a finite dimensional $\mathbb{C}$ -vector space.

(b) One has the Petersson inner product defined as follows.

$$\langle f,g\rangle=\int_{\mathcal{F}_N} f(z)\overline{g(z)}y^{k-2}dxdy \qquad (z=x+iy), \tag{52}$$

where $f,g\in M_k(N,\chi)$ with f or g a cusp form and $\mathcal{F}_N$ is a fundamental domain for $\Gamma_0(N)$.

(c) One has the theory of Hecke operators. In this situation, the Hecke algebra is generated by the Hecke operators T_p for $p\not|N$ and U_p for $p|N$. These operators are defined by

$$\begin{aligned} f\Big|T_p(z)&=\frac{1}{p}\sum_{0\le b<p} f\left(\frac{z+b}{p}\right)\ +\ \chi(p)\ p^{k-1}f(pz) \quad (p\not|N),\\ f\Big|U_p(z)&=\frac{1}{p}\sum_{0\le b<p} f\left(\frac{z+b}{p}\right) \quad (p|N). \end{aligned} \tag{53}$$

Here, unlike before, only the operators T_p for $p\not|N$ are hermitian with respect to the Petersson norm and so we have the weaker form of the corresponding theorem of Hecke.

Theorem 3.1. *(Hecke-Petersson) The space $S_k(N)$ has a basis of eigenforms with respect to all Hecke operators T_p for $p\not|N$.*

Newform theory. Note that when $N_1|N_2$, $\Gamma_0(N_2)\subseteq\Gamma_0(N_1)$ and hence $S_k(N_1)\subseteq S_k(N_2)$. In fact, $S_k\subseteq S_k(N)$ for all $N\ge 1$. The reason for the Hecke operator U_p $(p|N)$, defined on the space $S_k(N)$, not being hermitian with respect to the Petersson norm is because of this duplication of forms in higher levels. In order to find a satisfactory theory as in the case of modular forms with respect to the full modular group $SL_2(\mathbb{Z})$, A. O. L. Atkin and J. Lehner [2] defined a certain subspace of $S_k(N)$ which has the required nice properties. More precisely, they defined the subspace containing all the duplicating forms which come from lower levels as

$$S_k^{old}(N):=\Big\{f(dz)\Big|f\in S_k(r), rd|N, r\neq N\Big\} \tag{54}$$

and defined the space $S_k^{new}(N)$ to be the orthogonal complement of $S_k^{old}(N)$ in $S_k(N)$ with respect to the Petersson scalar product. This space $S_k^{new}(N)$ has all the required nice properties.

Theorem 3.2. *(Atkin-Lehner)*

(a) *The space $S_k^{new}(N)$ has a basis of simultaneous eigenforms with respect to all Hecke operators.*

(b) *Let $f \in S_k^{new}(N)$ be a non-zero Hecke eigenform. Then, $a_f(1) \neq 0$. So, one can find a basis of normalised Hecke eigenforms. These are called* **newforms** *of level N.*

(c) *Let $f \in S_k^{new}(N_1)$ and $g \in S_k^{new}(N_2)$ be two newforms having the same eigenvalues for almost all Hecke operators, then $N_1 = N_2$ and $f = g$.*

(d) *Let $f \in S_k^{new}(N)$ be a newform. Then $a_f(n)$ (the n-th Fourier coefficient of f) is a multiplicative function. Further, it satisfies the following property.*

$$\begin{aligned} a_f(np) &= a_f(p)a_f(n) \quad \text{if } p|N, n \geq 1 \\ a_f(np) + p^{k-1}a_f(n/p) &= a_f(p)a_f(n) \quad \text{if } p \nmid N, n \geq 1. \end{aligned} \tag{55}$$

(e) *Let $f \in S_k^{new}(N)$ be a newform. Then the corresponding L- function $L_f(s)$ has an Euler product expansion (for $\mathrm{Re}(s) > \frac{k}{2} + 1$):*

$$L_f(s) = \prod_{p|N} \left(1 - a_f(p)p^{-s}\right)^{-1} \prod_{p \nmid N} \left(1 - a_f(p)p^{-s} + p^{k-1-2s}\right)^{-1}. \tag{56}$$

REFERENCES

(1) Tom M. Apostol, *Modular Functions and Dirichlet series in Number Theory*, Graduates Texts in Mathematics **41**, Springer-Verlag, Berlin - Heidelberg - New York, 1990.

(2) A. O. L. Atkin and J. Lehner, *Hecke operators on* $\Gamma_0(m)$, Math. Ann. **185** (1970), 134–160.

(3) H. Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics **17**, American Mathematical Society, Providence, RI, 1997.

(4) N. I. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Graduates Texts in Mathematics **97**, Springer-Verlag, Berlin - Heidelberg - New York, 1984.

(5) S. Lang, *Introduction to Modular Forms*, Grundl. Math. Wiss. **222**, Springer - Verlag, Berlin - Heidelberg - New York, 1976.

(6) W. -W. Li, *Newforms and functional equations*, Math. Ann. **212** (1975), 285–315.

(7) T. Miyake, *Modular Forms*, Springer–Verlag, 1989.

(8) Andrew P. Ogg, *Survey of modular functions of one variable.* Notes by F. van Oystaeyen. Modular functions of one variable, I, pp. 1–35. Lecture Notes in Math. Vol. **320**, Springer, Berlin, 1973. Corrections: Modular functions of one variable, IV, p. 145. Lecture Notes in Math., Vol. **476**, Springer, Berlin, 1975.

(9) B. Ramakrishnan, *Theory of Newforms*, Bull. Allahabad Math. Soc. **8/9** (1993/94), 69–89 (1997).

(10) J. -P. Serre, *A course in Arithmetic*, Graduate Texts in Mathematics **7**, Springer-Verlag, Berlin- Heidelberg - New York, 1977.

(11) G. Shimura, *Introduction to the theory of automorphic functions*, Princeton Univ. Press, Princeton 1971.

(12) G. N. Watson, *A table of Ramanujan's function* $\tau(n)$, Proc. London Math. Soc., (2) **51** (1949), 1–13.

B. Ramakrishnan
Mehta Research Institute
Chhatnag Road, Jhusi
Allahabad 211 019, India.
e-mail: ramki@mri.ernet.in

ELLIPTIC CURVES, SERRE'S CONJECTURE AND FERMAT'S LAST THEOREM

KIRTI JOSHI

Contents

Date: Version of Joshi 5/13/2000 9:35:17 AM.

1991 *Mathematics Subject Classification.* 14J20, 14C25.

Key words and phrases. elliptic curves, serre's conjecture, galois representations, fermat's last theorem, modular forms, modular representations, frey curves.

1. Introduction

In these lectures I want to explain a circle of ideas introduced about 15 years ago which led to the proof of Fermat's Last Theorem at the hands of Taylor and Wiles (see [49], [46]) and the subsequent refinement of these ideas at the hands of Breuil, Conrad, Diamond and Taylor led to a proof of the Shimura-Taniyama-Weil conjecture (see [3]). Needless to say that these two results rest on the work of a number of other mathematicians as well: H. Hida, B. Mazur, K. Ribet, F. Diamond, B. Edixhoven, P. Deligne, J.-P. Serre and many others (see [4], [5], [7], [8], [11], [17], [14], [15], [19], [24], [32], [33], [34], [37], [42], [46], [49], and references at the end of these articles)

The main conjecture which formed the backdrop of these developments is a conjecture of Serre (see [41]). Despite the developments which have taken place, this conjecture of Serre still remains intractable at the moment. In [41] Serre showed that his conjecture together with the observation of G. Frey (see [20], [21]) led to a proof of Fermat's Last Theorem.

In these lectures I will essentially outline a proof of this assertion that Serre's conjecture implies Fermat's last Theorem. This article, needless to say, is completely based on Serre's paper [41] and is meant to serve as an introduction to the circle of ideas introduced in Serre's paper and is, by no means, a substitute for it. I have attempted to keep this article as self contained as possible. However it is impossible to prove all the results or to develop the theory of Galois representations in any reasonable depth or detail in article of this length. A more comprehensive account of the subject can be found in [9] or [16] and Serre's book [38] is a classic introduction to the subject of Galois representations.

2. Some History

The term elliptic curves is of relatively recent vintage, but the fundamental objects which lead to these curves have been around for a long time. Euler and Legendre studied the following kind of complex integrals:

$$(2.1) \qquad \int_z^\infty \frac{dx}{((1-x^2)(1-k^2x^2))^{1/2}} \text{ and } \int_z^\infty \frac{dx}{(4x^3-g_2x-g_3)^{1/2}}$$

where k, g_2, g_3 are complex numbers. Integrals like these arise naturally while calculating the length of the arc of an ellipse and hence these integrals were called *elliptic integrals*. Functions which arise in inverting these integrals were called elliptic functions. These integrals naturally live on curves whose equations look like

$$y^2 = f(x) \tag{2.2}$$

where $f(x)$ is a polynomial in x of degree three or four with distinct roots and complex coefficients. Such curves are called *elliptic curves*.

3. Cubic Curves

Let K be a field. Consider curves defined by homogeneous polynomials of the form

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3 \tag{3.1}$$

where the coefficients $a_i \in K$. Observe that as the polynomial is homogeneous if (x_0, y_0, z_0) is a solution to (3.1) then so is $(\lambda x_0, \lambda y_0, \lambda z_0)$ for any constant $\lambda \neq 0$. So we discard the *trivial solution* $(0, 0, 0)$, and identify solutions which are non-zero multiples of one another. In other words we will work with *homogeneous* or *projective* coordinates X, Y, Z and hence forth identify the solution (x_0, y_0, z_0) with $(\lambda x_0, \lambda y_0, \lambda z_0)$ for any $\lambda \neq 0$.

Also note that (3.1) has another obvious solution $O = (0, 1, 0)$ and it is easy to check that this the only non-trivial solution with $z = 0$. We will call O the *point at infinity* on the curve.

If (x_0, y_0, z_0) is a solution with $z_0 \neq 0$ then we can scale the solution by $1/z_0$ to get another solution $(x_0/z_0, y_0/z_0, 1)$. Thus we see that, except O, any other solution of (3.1) can be taken to be of the form $(x_1, y_1, 1)$ for some (x_1, y_1) and that such solutions are points on the dehomogenised form of the equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{3.2}$$

where we have written $y = Y/Z, x = X/Z$. For simplicity we will always write the equation (3.1) in its dehomogenous form (3.2) remembering the extra point O. We define several important quantities associated to the curve:

$$b_2 = a_1^2 + 4a_2 \tag{3.3}$$

$$b_4 = 2a_4 + a_1a_3 \tag{3.4}$$

$$b_6 = a_3^2 + 4a_6 \tag{3.5}$$

$$b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^3 - a_4^2 \tag{3.6}$$

$$c_4 = b_2^2 - 24b_4 \tag{3.7}$$

$$c_6 = b_2^2 + 36b_2b_4 - 216b_6 \tag{3.8}$$

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_8^2 + 9b_2b_4b_6 \tag{3.9}$$

$$j = c_4^3/\Delta \tag{3.10}$$

We call Δ the *discriminant* of the curve.

Definition 3.11. An elliptic curve E/K is a curve given by

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{3.12}$$

where $a_i \in K$ and provided that the discriminant $\Delta = \Delta_E \neq 0$.

Definition 3.13. A solution (x_0, y_0, z_0) of (3.11) will be called a point on the curve defined by the equation and if $x_0, y_0, z_0 \in K$ we will say that (x_0, y_0, z_0) is a K-rational point, or when there is no cause for confusion, simply a rational point of E.

Example 3.14. For example $y^2 = x^3 - x$ is an elliptic curve over any field K in which $\Delta_E = -64 \neq 0$.

Example 3.15. Let E be defined by $y^2 = x^3 + x + t$. Then we can easily calculate the quantities defined above: $b_2 = 0$, $b_4 = 2$, $b_6 = 4t$, $b_8 = -1$, $c_4 = -48$, $c_6 = -864t$, $\Delta = -432t^2 - 64$, $j = -6192/(-27t^2 - 4)$. Hence this equation defines an elliptic curve if and only if $-432t^2 - 64 \neq 0$.

So far we have not placed any restrictions over K, but under additional assumptions on K we can simplify the equation of any elliptic curve considerably. For instance if $char(K) \neq 2$ we can replace y by $\frac{1}{2}(y - a_1x - a_3)$ to get

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6 \tag{3.16}$$

And if $char(K) \neq 2, 3$ then we can replace (x, y) by $(\frac{x-3b_2}{36}, \frac{y}{216})$ to get

$$y^2 = x^3 - 27c_4x - 54c_6 \tag{3.17}$$

One gets the additional relations $4b_8 = b_2b_6 - b_4^2$ and $1728\Delta = c_4^3 - c_6^2$.

Remark 3.18. The only transformations of (3.11) which preserve the form of the equation are of the following kind:

$$x = u^2x' + r \tag{3.19}$$

$$y = u^3y' + u^2sx' + t \tag{3.20}$$

with $u, r, s, t \in K$ and $u \neq 0$. I leave it to you to check that under these substitutions the new value of the new discriminant is $\Delta' = u^{-12}\Delta$ and $j' = j$. Thus the function j of the coefficients a_1, a_2, a_3, a_4, a_6 is an invariant of the curve and is called the *j-invariant of the elliptic curve.*

4. Singularities

Suppose E is given by

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{4.1}$$

but is not an elliptic curve, i.e., $\Delta_E = 0$. In this situation we will say that the curve is *singular* and we want to describe the geometric possibilities for E in this situation. Let

$$f(x,y) = y^2 + a_1xy + a_3y - (x^3 + a_2x^2 + a_4x + a_6) \tag{4.2}$$

Then it is not very difficult to show that $\Delta_E = 0$ if and only of

$$\frac{\partial}{\partial x}f(x,y) = \frac{\partial}{\partial y}f(x,y) = 0 \tag{4.3}$$

has a simultaneous solution, say (x_0, y_0). Then we can use Taylor series to write $f(x,y)$ as

$$f(x,y) - f(x_0,y_0) = [(y-y_0) - \alpha(x-x_0)][(y-y_0) - \beta(x-x_0)] - (x-x_0)^3 \tag{4.4}$$

where, $\alpha, \beta \in \bar{K}$, where $\bar{K}$ is an algebraic closure of K.

Remark 4.5. In general α, β may not be in K.

Definition 4.6. Let the notation be as above. If $\alpha \neq \beta$ we say that $P = (x_0, y_0)$ is a node and in this case the lines

$$y - y_0 = \alpha(x - x_0) \tag{4.7}$$

$$y - y_0 = \beta(x - x_0) \tag{4.8}$$

are two tangents to the curve E at (x_0, y_0).

If $\alpha = \beta$ we say that E has a cusp at $P = (x_0, y_0)$ and the line $y - y_0 = \alpha(x - x_0)$ is a multiple tangent to E at $P = (x_0, y_0)$.

Example 4.9. The curve $y^2 = x^3 + x^2$ is singular with a node at the point $(0, 0,)$.

Example 4.10. The curve $y^2 = x^3$ has a cusp at $(0, 0)$.

5. Group Law

Elliptic curves have many remarkable properties, and one of the important properties is that the set of points on an elliptic curve forms a group. I will briefly describe the group law. The group law is best described geometrically but it also admits an algebraic description.

Suppose E/K is an elliptic curve. Let L/K be any field extension. We will write $E(L)$ for the set of L-rational points of E. Note that $O \in E(L)$ for any L/K and so $E(L)$ is a non-empty set. If $P, Q \in E(L)$ are two points then the line joining P, Q intersects the curve in a third point. A little bit of algebra shows that this new point also has coordinates in L, i.e., it is also an L-rational point. This basic geometric fact underlies the group law on the elliptic curve. Let us denote this third point by R. Then we join O and R by a line. This line also intersects the curve again in a point $R' \in E(L)$. We declare that $P + Q = R'$. If $P = Q$ we take the line to be the tangent line to the curve at P. It is not hard to check that this makes the set $E(L)$, for any extension L/K, into an abelian group with O as its identity element.

We can also describe the group law algebraically and if you are not convinced that the above geometric description gives a group law, then you can carry out the tedious calculations required to verify that the algebraic formulas given below define a group structure on $E(L)$. Let $P_1 = (x_1, y_1), P_2 = (x_2, y_2) \in E(L)$. Then we define $-P_1 = (x_1, -y_1 - a_1x_1 - a_3)$. If $x_1 = x_2$ and $y_1 + y_2 + a_1x_2 + a_3 = 0$ then $P_1 + P_2 = O$ otherwise if $x_1 \neq x_2$ let

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1} \tag{5.1}$$

$$\nu = \frac{y_1x_2 - y_2x_1}{x_2 - x_1}. \tag{5.2}$$

If $x_1 = x_2$ then let

$$\lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} \tag{5.3}$$

$$\nu = \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3} \tag{5.4}$$

Then the line $y = \lambda x + \nu$ is a line through P_1, P_2 or is tangent to E at P_1 if $P_1 = P_2$. Then we have the following formula for $P_3 = (x_3, y_3) = P_1 + P_2$

$$x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2 \tag{5.5}$$

$$y_3 = -(\lambda + a_1)x_3 - \nu - a_3 \tag{5.6}$$

Example 5.7. Let $K = \mathbb{Q}$, and let E/K be defined by

$$y^2 = x^3 - 6x^2 + 11x + 3 \tag{5.8}$$

then $\Delta_E = -34928, j_E = 6912/2183$.

Further $P_1 = (1,3), P_2 = (2,3)$ are $\mathbb{Q}$-rational points on E and $P_1 + P_2 = (3,-3)$ and $2P_1 = P_1 + P_1 = (17/4, -29/8)$, $3P_1 = 2P_1 + P_1 = (633/169, 6046/2197)$ and $4P_1 = (17889/13456, -3262625/1560896)$ etc.

When $L = \mathbb{C}$ the structure of $E(L)$ is completely understood using the theory of elliptic functions. The following theorem is a key result in the classical theory of elliptic functions.

Theorem 5.9. *Let $E/\mathbb{C}$ be an elliptic curve. Then there exists a lattice $L \subset \mathbb{C}$ (so $L \cong \mathbb{Z} \oplus \mathbb{Z}$ as abelian group) and a homomorphism of groups*

$$\mathbb{C}/L \cong E(\mathbb{C}) \tag{5.10}$$

which is given by $0 \mapsto O \in E(\mathbb{C})$ and $0 \neq z \mapsto (\mathfrak{p}_L(z), \mathfrak{p}'_L(z), 1)$ where $\mathfrak{p}_L$ and $\mathfrak{p}'_L$ are the Weierstrass elliptic functions with period lattice L.

Remark 5.11. Thus the group law on $E(L)$ gives us a way of generating more solutions from the given ones.

Remark 5.12. Fermat's method of infinite descent is also a variant of the group law on a suitable elliptic curve over $K = \mathbb{Q}$.

The following theorem due to Mordell, which was later generalized by André Weil to higher dimensional analogues of elliptic curves (see [47]) gives us a fundamental insight into the structure of the group $E(K)$. For a proof see [43]

Theorem 5.13 (Mordell-Weil)**.** *Let K be a number field and let L/K be a finite extension. Then for any elliptic curve E/K, the group $E(L)$ is a finitely generated abelian group, i.e.,*

$$E(L) \cong \mathbb{Z}^r \oplus \text{Finite group} \tag{5.14}$$

Remark 5.15. The number of copies of $\mathbb{Z}$ which occur in the above description is called the rank of $E(L)$. Birch and Swinnerton-Dyer have made a fascinating conjecture about the rank of $E(K)$ and the order of zero of a complex analytic function associated to E (see [1], [2] and [43]).

6. Elliptic curves over finite fields

In our discussion so far we have not specified the field K except in the examples. The formulae we have written down for the group law are valid over any field K. In this section we will assume that $K = \mathbb{F}_q$ a finite field with q elements and characteristic $p > 0$.

Suppose $E/\mathbb{F}_q$ is an elliptic curve over $\mathbb{F}_q$. Then it is easy to see that $E(\mathbb{F}_q)$ is a finite set as there are only finite number of possible values for each coordinate of any $\mathbb{F}_q$-rational point. Hasse (see [23]) proved the following bound on the size of $E(\mathbb{F}_q)$ in terms of q. This estimate was later generalized by Weil in [48]. More precisely, Hasse proved the following:

Theorem 6.1 (Hasse-Weil Estimate)**.** *Let $E/\mathbb{F}_q$ be an elliptic curve over a finite field with q elements. Then we have*

$$|\#E(\mathbb{F}_q) - q - 1| \leq 2q^{1/2} \tag{6.2}$$

In other words if we write

$$\#E(\mathbb{F}_q) = q + 1 - a_q \tag{6.3}$$

then

$$|a_q| \leq 2q^{1/2}. \tag{6.4}$$

For a proof of the Hasse-Weil estimate, see [43].

This estimate was later generalized by Weil to arbitrary smooth projective curves and to abelian varieties and higher dimensional varieties. Weil's paper [48] is an excellent and elementary introduction to the subject. A more modern and elementary account can be found in [26].

Some of the Weil conjectures were proved by [18] and [31]. In 1974 Deligne proved the Weil conjectures (see [12]). For a reader with some basic background in algebraic geometry we recommend Katz's exposition of Deligne's proof (see [27]).

Remark 6.5. Elliptic curves over finite fields play an important role in primality testing and cryptography (see [28]).

7. Minimal Equations

Let E be an elliptic curve over $\mathbb{Q}$. We will say that the equation defining E is *minimal* at a prime $p|\Delta_E$ if the valuation $\nu_p(\Delta_E)$ is least amongst all possible choices of equations for E such that coefficients are p-adic integers. By formulas (3.18), we see that we can change the equation of E and we can get $\Delta' = u^{-12}\Delta$. Performing this change of coordinates as many times as possible we arrive at $\Delta_p(E) < 12$ and such that all the coefficients are p-adic integers. Thus we can conclude: if $v_p(a_i) \geq 0$ and $\nu_p(\Delta) < 12$ then this equation of E is minimal.

We note that $v_p(a_i) \geq 0$ and $\nu_p(\Delta_E) < 12$ is a sufficient condition for minimality of the equation.

Remark 7.1. Let E/K be an elliptic curve over a number field K and let $\mathcal{O}_K$ be the ring of integers and $\mathfrak{p} \subset \mathcal{O}_K$ a prime ideal. We can also formulate the notion of a minimal equation for E at $\mathfrak{p}$ exactly as above.

Example 7.2. Let $y^2 = x^3 + 16$, then this equation define an elliptic curve over $\mathbb{Q}$ and $\Delta_E = -2^{12}3^3$. The equation is not minimal at $p = 2$ and the transformation $x = 4x'$ $y = 8y' + 4$ transforms the equation of E to $(y')^2 + y' = (x')^3$. This equation has $\Delta = -27$, and is minimal at 2 and in fact at all primes.

It is not difficult to see using (3.18) that any equation E/K, where $K/\mathbb{Q}_p$ is a finite extension of $\mathbb{Q}_p$, has a minimal equation over K. The following proposition is easy to prove using the transformations (3.18) and the fact that every prime ideal in $\mathbb{Z}$ is principal.

Proposition 7.3. *Let $E/\mathbb{Q}$ be an elliptic curve over $\mathbb{Q}$. Then there exists an equation with integer coefficients for E which is minimal at all primes.*

From now on we will assume that all our elliptic curves are given by a minimal equation.

8. Reducing modulo primes

Let K be a number field, let $\mathcal{O}_K$ be the ring of integers of K and $\mathfrak{p}$ be a non-zero prime ideal in $\mathcal{O}_K$. Let E/K be an elliptic curve defined by an equation which is a *minimal equation* at $\mathfrak{p}$.

$$y^2 + a_1xy + a_3 = x^3 + a_2x^2 + a_4x + a_6 \tag{8.1}$$

We will assume for simplicity that $a_i \in \mathcal{O}_K$. Then we can reduce the equation defining E modulo $\mathfrak{p}$ and arrive a new equation which has coefficients in $k = \mathcal{O}_K/\mathfrak{p}$ which is a finite field:

$$y^2 + \bar{a}_1xy + \bar{a}_3 = x^3 + \bar{a}_2x^2 + \bar{a}_4x + \bar{a}_6 \tag{8.2}$$

where $\bar{a}_i = a_i \mod \mathfrak{p}$.

Observe that (8.2) represents an elliptic curve over $k = \mathcal{O}_K/\mathfrak{p}$ if and only if the discriminant $\bar{\Delta} = \Delta \mod \mathfrak{p}$ is not zero, i.e., provided that $\Delta \notin \mathfrak{p}$ or equivalently $\mathfrak{p} \not| \Delta$.

Thus this recipe of reducing the equation of an elliptic curve over a number field produces an elliptic curve over the finite field $\mathcal{O}_K/\mathfrak{p}$ provided $\mathfrak{p} \not| \Delta$. As Δ is divisible by only a finite number of primes we see that for all but finite number of primes in K we will get an elliptic curve over the corresponding finite field.

Definition 8.3. In the notations of the above paragraphs, we will say that $\mathfrak{p}$ is a prime of *good reduction if* $\mathfrak{p} \not| \Delta$.

Definition 8.4. If $\mathfrak{p}|\Delta$ then we will say that $\mathfrak{p}$ is a prime of *bad reduction* for E.

If $\mathfrak{p}$ is a prime of bad reduction then our recipe fails to produce an elliptic curve but produces a singular curve instead. By using our discussion of singular curves (see Section 4) we can further classify primes of bad reduction. We know from Section 4 that the reduction of E has either a node or a cusp.

Definition 8.5. In the notations as above, We will say that E has *semistable* or *multiplicative* reduction at $\mathfrak{p}$ if $\mathfrak{p}|\Delta$ and the reduction of E modulo $\mathfrak{p}$ has a node. If the two tangents at the node are defined over $\mathcal{O}_K/\mathfrak{p}$ then we say that E has *split multiplicative reduction* at $\mathfrak{p}$. If the two tangents at the node are not defined over $\mathcal{O}_K$ then we say that E has *non-split multiplicative reduction* at $\mathfrak{p}$. If the reduction of E at $\mathfrak{p}$ is a cusp then we say that E has *additive* or *unstable* reduction at $\mathfrak{p}$.

Proposition 8.6. *Let E/K be an elliptic curve defined by a minimal equation*

$$y^2 + a_1xy + a_3 = x^3 + a_2x^2 + a_4x + a_6 \tag{8.7}$$

with $a_i \in \mathcal{O}_K$. Let $\mathfrak{p}$ be a prime in $\mathcal{O}_K$.

(1) *E has good reduction at $\mathfrak{p}$ if and only if $\Delta_E \not\equiv 0 \mod \mathfrak{p}$,*
(2) *E has additive reduction modulo $\mathfrak{p}$ if and only if $\Delta_E \equiv c_4 \equiv 0 \mod \mathfrak{p}$,*
(3) *E has semistable reduction modulo $\mathfrak{p}$ if and only if $\Delta_E \equiv 0 \mod \mathfrak{p}$ and $c_4 \not\equiv 0 \mod \mathfrak{p}$.*

Remark 8.8. Additive reduction is the worse kind of bad reduction while semistable reduction is not too bad. The following may help illustrate the subtle difference between these two types of bad reduction. Let E/K be an elliptic curve and let $\mathfrak{p}$ be a prime of bad reduction for E. Let K'/K be a finite extension, $\mathfrak{p}'$ be any prime lying over $\mathfrak{p}$ in K'. If E has semistable reduction at $\mathfrak{p}$ then, E thought of as an elliptic curve over K', continues to have semistable reduction at $\mathfrak{p}'$. However, if E has additive reduction at $\mathfrak{p}$, then there exists a finite extension K'/K such that for any prime $\mathfrak{p}'$ lying over $\mathfrak{p}$, E has good or semistable reduction at $\mathfrak{p}'$. In other words, additive reduction may disappear or become semistable over a suitable finite extension, while semistable reduction persists. The following examples illustrate this point further.

Example 8.9. Let $K = \mathbb{Q}$, let E be defined by $y^2 = x^3 + x^2 + 17$. Then E has multiplicative reduction modulo 17.

Example 8.10. Let $K = \mathbb{Q}$ and let E be defined by $y^2 = x^3 + 17$. Then E has additive reduction modulo 17. Let $K' = \mathbb{Q}(17^{1/6})$, and then we can write the equation of E over K' as

$$(y/(17^{1/6})^3)^2 = (x/(17^{1/6})^2)^3 + 1 \tag{8.11}$$

Writing $y' = y/17^{1/2}, x' = x/17^{1/3}$ we get $y'^2 = x'^3 + 1$ and this new equation has good reduction modulo the unique prime lying over 17 in $\mathbb{Q}(17^{1/6})$.

9. The conductor of an elliptic curve

Let K be a number field. Let E/K be an elliptic curve. We can define an ideal, $N_E \subset \mathcal{O}_K$ of the ring of integers of K,called the conductor ideal or more simply the conductor of E. This ideal is defined as

$$N_E = \prod_{\mathfrak{p}|\Delta} \mathfrak{p}^{f_{\mathfrak{p}}(E/K)} \tag{9.1}$$

where the exponents $f_{\mathfrak{p}}(E/K)$ of $\mathfrak{p}$ are defined as follows.

$$f_{\mathfrak{p}}(E/K) = \begin{cases} 0 & \text{if } E \text{ has good reduction at } \mathfrak{p}, \\ 1 & \text{if } E \text{ has semistable reduction at } \mathfrak{p}, \\ 2 & \text{if } \mathfrak{p} \not| 6 \text{ and } E \\ & \quad\text{has additive reduction at } \mathfrak{p}, \\ 2 + \delta_{\mathfrak{p}}(E/K) & \text{if } \mathfrak{p}|6. \end{cases} \tag{9.2}$$

where the exact definition of $\delta_{\mathfrak{p}}(E/K)$ is irrelevant for our purposes. No exact formulas are known for $\delta_{\mathfrak{p}}(E/K)$, but Tate's algorithm (see [45]) gives a way of computing $f_{\mathfrak{p}}$ for all primes including those dividing 6. It is known, for instance, that

$$f_2(E/\mathbb{Q}) \leq 8 \tag{9.3}$$

$$f_3(E/\mathbb{Q}) \leq 5 \tag{9.4}$$

Observe that the conductor ideal is divisible by only those primes which divide Δ_E, and that it captures finer reduction information of E/K. The conductor is one of the fundamental arithmetical invariants of an elliptic curve.

Remark 9.5. Tate's algorithm [45], and other algorithms like [30] have now been implemented in many software packages (notably in the software package PARI-GP which is available on Internet; this package also contains an extensive elliptic curve computation package which computes many numerical invariants of elliptic curves).

Example 9.6. Let $y^2 = x^3 + 10x + 11$. Then $\Delta_E = -2^4 \cdot 13^2 \cdot 43$ and $N_E = 2^4 \cdot 13 \cdot 43$ and the equation is minimal.

10. Action of the Galois group

Let K be a number field or a finite extension of $\mathbb{Q}_p$ or a finite field. Let $\bar{K}$ be an algebraic closure of K. We write $G_K = \mathrm{Gal}(\bar{K}/K)$ for the Galois group of $\bar{K}/K$. Let E/K be an elliptic curve and suppose $P = (x_0, y_0) \in E(\bar{K})$. Thus P satisfies

$$y_0^2 + a_1x_0y_0 + a_3 = x_0^3 + a_2x_0^2 + a_4x_0 + a_6. \tag{10.1}$$

Suppose $\sigma \in G_K$. We apply σ to the above equation and get

$$\sigma(y_0)^2 + a_1\sigma(x_0)\sigma(y_0) + a_3 = \sigma(x_0)^3 + a_2\sigma(x_0)^2 + a_4\sigma(x_0) + a_6. \tag{10.2}$$

where we have used the fact that $\sigma(a_i) = a_i$ for any $a_i \in K$. Thus we see from the above equation that $P^\sigma = (\sigma(x_0), \sigma(y_0)) \in E(\bar{K})$ is also a point on the elliptic curve. In other words, G_K operates on $E(\bar{K})$.

We now study this action in some more detail. Let $n \geq 2$ be any integer and let

$$E[n] = \{P \in E(\bar{K}) | nP = O\} \tag{10.3}$$

Then it is easy to see that $E[n] \subset E(\bar{K})$ is a subgroup. We call $E[n]$ the group of *n-torsion points* on the curve E or simply the group of points of order dividing n.

The structure of $E[n]$ as an abelian group is completely understood.

Theorem 10.4. *Let K be a field and let E be an elliptic curve over K. If the characteristic of K does not divide n then one has*

$$E[n] \cong \mathbb{Z}/n \oplus \mathbb{Z}/n. \tag{10.5}$$

In particular for for a number field K, and when $n = p$ a prime, we see that $E[p]$ is a two dimensional vector space over $\mathbb{F}_p$.

When the characteristic of K divides n the result is slightly different but as we will not need it here we do not recall it and the reader is referred to [43] for more details.

Moreover, if $P \in E[n]$ and $\sigma \in G_K$ then

$$\sigma(nP) = n\sigma(P) = O \tag{10.6}$$

So that G_K acts on the $\mathbb{Z}/n$-module $E[n]$.

Example 10.7. Let $K = \mathbb{Q}$ and let E be defined by $y^2 = x^3 - x$. Then $E[2] = \{O, (0,0), (0,1), (0,-1)\}$ and so $G_{\mathbb{Q}}$ operates trivially on $E[2]$. For a more interesting example see Remark 11.20.

11. Tate Elliptic curves

Let $K = \mathbb{Q}_p$ be the field of p-adic numbers, where p is any prime, let $|\ |_p$ denote the p-adic absolute value, and let $\nu_p : \mathbb{Q}_p \to \mathbb{Z}$ be the normalized p-adic valuation, normalized so that $\nu_p(p) = 1$. As $\mathbb{Q} \subset \mathbb{Q}_p$, we can think of any elliptic curve over $\mathbb{Q}$ as an elliptic curve over $\mathbb{Q}_p$.

Tate discovered elliptic curves over $\mathbb{Q}_p$ with remarkable properties. These curves are called Tate elliptic curves (see [44]). These curves are sort of "universal models" for elliptic curves over $\mathbb{Q}_p$ with split multiplicative reduction modulo p. Tate curves are given by an explicit equation.

Fix $q \in \mathbb{Q}_p$ such that $|q|_p < 1$, so q is a p-adic integer divisible by p. Let E_q be defined by

$$y^2 + xy = x^3 + a_4(q)x + a_6(q) \tag{11.1}$$

where

$$a_4(q) = -5\sum_{n=1}^{\infty} \frac{n^3q^n}{1-q^n} \tag{11.2}$$

and

$$a_6(q) = -\frac{1}{12}\sum_{n=1}^{\infty} \frac{(7n^5+5n^3)q^n}{(1-q^n)} \tag{11.3}$$

Then the discriminant $\Delta_{E_q} = \Delta(q)$ is the famous Ramanujan function

$$\Delta(q) = q\prod_{n=1}^{\infty}(1-q^n)^{24} \tag{11.4}$$

and the j-invariant is

$$j(q) = \frac{1}{q} + 744 + 196884q + \cdots \tag{11.5}$$

Remark 11.6. It is clear from the definition of discriminant $\Delta(q)$ that

$$\nu_p(\Delta(q)) = \nu_p(q) \tag{11.7}$$

$$\nu_p(j(q)) = -\nu_p(q) \tag{11.8}$$

Further for $|q|_p < 1$, the series $a_4(q)$ and $a_6(q)$ converge and (11.1) has points in $\mathbb{Q}_p((u))$ given by convergent power series in u

$$x(q,u) = \sum_{n\in\mathbb{Z}} \frac{q^nu}{(1-q^nu)^2} - 2\sum_{n=1}^{\infty}\frac{nq^n}{1-q^n} \tag{11.9}$$

$$y(q,u) = \sum_{n\in\mathbb{Z}} \frac{q^{2n}u^2}{(1-q^nu)^3} + \sum_{n=1}^{\infty}\frac{nq^n}{1-q^n} \tag{11.10}$$

Then for any $|q|_p < 1$ and $u \in \bar{\mathbb{Q}}_p^*$ these power series converge and their values give a point $(x, y) \in E_q(\bar{\mathbb{Q}}_p)$. Moreover if we reduce the equation of E_q modulo p then we get

$$y^2 + xy = x^3 \tag{11.11}$$

and so one checks easily from this that E_q has split multiplicative reduction modulo p.

Observe that $x(q, u)$ and $y(q, u)$ as a function of u has the property that

$$x(q,u) = x(q,qu) \tag{11.12}$$

$$y(q,u) = y(q,qu). \tag{11.13}$$

Thus these functions are periodic with respect to the multiplicative group

$$q^{\mathbb{Z}} = \{q^m | m \in \mathbb{Z}\}. \tag{11.14}$$

Tate further showed that if we fix q with $|q|_p < 1$ then the mapping

$$\bar{\mathbb{Q}}_p^* \to E_q(\bar{\mathbb{Q}}_p) \tag{11.15}$$

given by $u \mapsto (x(q,u), y(q,u))$ is a surjective homomorphism of groups with kernel $q^{\mathbb{Z}}$ and is compatible with the action of Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ on both the sides. Thus one gets

$$\mathbb{Q}_p^*/q^{\mathbb{Z}} \cong E_q(\mathbb{Q}_p) \tag{11.16}$$

The crucial property of Tate curves is described in the following theorem.

Theorem 11.17 (Tate uniformization)**.** *Let $E/\mathbb{Q}$ be an elliptic curve and suppose that $p|\Delta_E$. Assume that E has split multiplicative reduction at p. Then there exists a $q \in \mathbb{Q}_p$ such that $E \cong E_q$ over $\mathbb{Q}_p$.*

Tate proved a result which is more general than the above statement. In particular, Tate's result is valid for curves over number fields and finite extensions of $\mathbb{Q}_p$ (see [43]). Moreover, one also get a similar assertion for non-split multiplicative reduction. But we will not need the general assertion here and we refer the reader to [43] for more details.

Example 11.18. Let $K = \mathbb{Q}_p$ and let E_q be a Tate curve over K. We can use Tate's theorem to describe $E[n]$ explicitly. The $u \in \bar{\mathbb{Q}}_p^*$ has the property that

$$nP = n(x(q,u), y(q,u)) = O \tag{11.19}$$

if and only if $u^n \in q^{\mathbb{Z}}$, i.e., $nP = O$ if and only if $u^n = q^m$ for some $m \in \mathbb{Z}$ and so $u = \zeta q^{m/n}$ for some n^{th} root of unity ζ. Thus we obtain an isomorphism

$$E[n] \cong \left\{ \zeta u^{m/n} | \zeta^n = 1, 0 \leq m \leq n \right\}. \tag{11.20}$$

12. Some Galois Theory

In this section we will recall a few facts about the structure of the Galois Groups of number fields or local fields. For proofs or details see [40].

Fix prime p. We would like to recall a few facts about the structure of the Galois group $G_{\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. We will do this by studying the structure of $G_p = \mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$. Let $\mathfrak{p}|p$ be any valuation lying over p in $\bar{\mathbb{Q}}$. The Galois group $G_{\mathbb{Q}}$ acts on the set of such $\mathfrak{p}$ transitively. Fix one such valuation $\mathfrak{p}|p$.

Definition 12.1. The decomposition group $D(\mathfrak{p}, p) \subset G_{\mathbb{Q}}$ at $(\mathfrak{p}, p)$ is defined to be the set of all $\sigma \in G_{\bar{\mathbb{Q}}}$ such that $\sigma(\mathfrak{p}) = \mathfrak{p}$.

In other words, $D(\mathfrak{p}, p)$ is the stabilizer of $\mathfrak{p}$. The decomposition group depends on $\mathfrak{p}$ and p. If we replace $\mathfrak{p}$ by another $\mathfrak{p}'|p$ then the decomposition group $D(\mathfrak{p}', p)$ is a conjugate of $D(\mathfrak{p}, p)$. Hence we will often suppress the dependence of $\mathfrak{p}$ and often call $D(\mathfrak{p}, p)$ the decomposition group at p.

The decomposition group encodes a lot of information about p, it contains interesting subgroups which encode information about the ramification of p in any extension. The following result identifies the decomposition group little more explicitly.

Proposition 12.2. *For every prime $\mathfrak{p}|p$ in $\bar{\mathbb{Q}}$, we have an isomorphism*

$$\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p) \cong D(\mathfrak{p}, p) \hookrightarrow \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}). \tag{12.3}$$

So for every prime $\mathfrak{p}|p$ in $\bar{\mathbb{Q}}$ we get an embedding of $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p) \hookrightarrow \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$.

Thus we reduce to the of study the Galois group of $\bar{\mathbb{Q}}_p/\mathbb{Q}_p$. From now on we study these groups.

We have a natural surjection

$$G_p \to \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p) \tag{12.4}$$

which is given by $\sigma \mapsto \sigma \mod p$.

Definition 12.5. The kernel $G_p \to \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$ of

$$I_p = \{\sigma \equiv 1 \mod p\} \tag{12.6}$$

is called the inertia subgroup at p.

Thus we have an isomorphism

$$G_p/I_p \cong \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p), \tag{12.7}$$

given by $\sigma \mapsto \sigma \mod p$.

The $\mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$ has natural element $\mathrm{Frob}_p \in \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F})$ which is given by raising to p^{th}-powers:

$$\mathrm{Frob}_p(x) = x^p \tag{12.8}$$

for all $x \in \bar{\mathbb{F}}_p$. We will call Frob_p the Frobenius (morphism) at p.

A proof of the following proposition can be found in [40].

Proposition 12.9. *Let L/K be any Galois extension of number fields. Let $\mathfrak{p}'\mathfrak{p}$ be a prime of L lying over a prime $\mathfrak{p}$ of K. Then L/K is unramified at $\mathfrak{p}$ if and only if the inertia subgroup $I_{\mathfrak{p}}$ at $\mathfrak{p}$ is trivial.*

Definition 12.10. Let $L/\mathbb{Q}$ be an arbitrary Galois extension. Assume that $L/\mathbb{Q}$ is unramified outside a finite set of primes. Let p be a prime at which L is unramified. Let $\mathfrak{p}$ be a prime lying over p in L. Then a Frobenius element at p is a conjugacy class of any element $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$ such that σ is in $D(\mathfrak{p}, p)$ and its image in $\mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$ under the isomorphism (12.7) is $\mathrm{Frob}_p \in \mathrm{Gal}(\bar{\mathbb{F}}_p/\mathbb{F}_p)$.

We will continue to use the notation Frob_p to denote Frobenius element at p, keep in mind that it is really a conjugacy class of elements which depends only on p and not on the choice of a prime lying over p in L.

The following variant of the Chebotarev density theorem (see [6] or [38]) will be used without proof.

Theorem 12.11. *Let $L/\mathbb{Q}$ be an arbitrary Galois extension which is unramified outside a finite set of primes. Then the the set of Frobenius elements of primes which are unramified in $L/\mathbb{Q}$ is dense in* $\mathrm{Gal}(L/\mathbb{Q})$.

The inertia subgroup I_p has finer structure as well. We can define a filtration

$$I_i = \left\{\sigma | \sigma \equiv 1 \mod p^{i+1}\right\} \tag{12.12}$$

for all $i \geq 0$, with $I_0 = I_p$. There is a natural surjection from

$$I_p \to \bar{\mathbb{F}}_p^* \tag{12.13}$$

given by the action of the inertia group on the roots of unity in $\bar{\mathbb{Q}}_p$ and it is standard that

$$I_1 = \ker(I_p \to \bar{\mathbb{F}}_p^*). \tag{12.14}$$

The quotient group

$$I_t = I_p/I_1 \tag{12.15}$$

is often called the *tame quotient of I_p and I_1 is called the* wild inertia subgroup.

13. Galois Representations

Let E/K be an elliptic curve and assume that K is a field of characteristic zero. Let $n \geq 2$ be an integer. Then as we have seen that G_K operates on $E[n] \cong \mathbb{Z}/n \times \mathbb{Z}/n$. This action is compatible with the group structure and so for each $\sigma \in G_K$, the mapping $P \mapsto \sigma(P)$ is an automorphism of the $\mathbb{Z}/n$-module $E[n]$. Thus we get a homomorphism

$$\rho : G_K \to \mathrm{Aut}(E[n]) = \mathrm{GL}_2(\mathbb{Z}/n) \tag{13.1}$$

where

$$\mathrm{GL}_2(\mathbb{Z}/n) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} | a, b, c, d \in \mathbb{Z}/n \text{ and } ad - bc \in (\mathbb{Z}/n)^* \right\} \tag{13.2}$$

is the group of invertible matrices with coefficients in $\mathbb{Z}/n$.

Elliptic curves are sources of such homomorphisms but there are other sources of such homomorphisms and the confluence of two such sources is essentially the content of Serre's conjecture.

Definition 13.3. Let R be any ring, and let K be either a number field or a finite extension of $\mathbb{Q}_\ell$ for a prime ℓ. A Galois representation of G_K is a continuous homomorphism

$$\rho : G_K \to \mathrm{GL}_n(R) \tag{13.4}$$

where $\mathrm{GL}_n(R)$ denotes the group of $n \times n$ invertible matrices with entries in R. If R has a natural topology then we give $\mathrm{GL}_n(R)$ the topology induced on it as an open subset of R^n, and otherwise we give R the discrete topology.

Example 13.5. We will be interested in the situation when R is either the field of p-adic numbers for some p or is $\mathbb{Z}/n$ for some integer n. In the first case we give $\mathrm{GL}_n(\mathbb{Q}_p)$ the topology induced from $\mathbb{Q}_p^n$ and then continuity condition is with respect to this topology.

Example 13.6. When $R = \mathbb{Z}/n$, then $\mathrm{GL}_n(R)$ is a finite group and we give it the discrete topology. Continuity of ρ then simply means that $\ker(\rho)$ is an open subgroup of G_K.

Remark 13.7. It is often convenient to think of $\mathrm{GL}_n(R)$ as the group invertible R-linear maps from an R-module $V = R^n$ to itself. Any representation $\rho : G_K \to \mathrm{GL}_n(R)$ thus gives rise to an R-linear action of G_K on R^n, i.e., for each $\sigma \in G_K$ we have an invertible R-linear mapping $R^n \to R^n$ which is continuous and which satisfies obvious conditions and conversely any such action gives rise to a representation ρ.

Example 13.8. Let $\mu_n = \left\{z \in \bar{K} | z^n = 1\right\}$ be the group of n^{th} roots of unity in $\bar{K}$. Then G_K acts on it. We have an isomorphism of abelian groups $\mu_n \cong \mathbb{Z}/n$. Thus this action gives rise to a homomorphism

$$G_K \to \operatorname{Aut}(\mu_n) = \mathrm{GL}_1(\mathbb{Z}/n) = (\mathbb{Z}/n)^*. \tag{13.9}$$

Example 13.10. We will need the following special case of the above example. Suppose $n = \ell$ for a prime ℓ and

$$\chi_\ell : G_K \to \operatorname{Aut}(\mu_\ell) = (\mathbb{Z}/\ell)^*. \tag{13.11}$$

This representation is called the *cyclotomic character* at ℓ.

Example 13.12. Let ℓ be a prime and let $K = \mathbb{Q}$. Let us examine $\chi_\ell : G_\mathbb{Q} \to (\mathbb{Z}/\ell)^*$ in a little more detail. Let ζ be any ℓ^{th}-root of unity; suppose that c is complex conjugation $c : \bar{\mathbb{Q}} \to \bar{\mathbb{Q}}$. Then we have $\chi(c)(\zeta) = \bar{\zeta} = \zeta^{-1}$ and in particular, we see that $\chi_\ell(c) = -1 \in (\mathbb{Z}/\ell)^*$.

Definition 13.13. We say that a representation $\rho : G_K \to \mathrm{GL}_n(R)$ is reducible if there exists a proper R-submodule $0 \neq W \subset V = R^n$ such that the action of ρ on V maps W into itself, i.e., for all $\sigma \in G_K$, we have $\sigma(W) \subset W$.

Definition 13.14. If $\rho : G_K \to \mathrm{GL}_n(R)$ is not reducible, then we say that ρ is an irreducible representation of G_K.

Example 13.15. Let E_q be a Tate curve over $\mathbb{Q}_p$. We had observed in Example 11.18 that $E_q[n] = \left\{\zeta q^{m/n} \middle| \zeta^n = 1, 0 \leq m \leq n\right\}$. We claim that the representation of $G_p = \operatorname{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ is reducible. To see this we define a surjective homomorphism $E_q[n] \to \mathbb{Z}/n$ which is given by $\zeta q^{m/n} \mapsto m \bmod n$. This homomorphism is clearly surjective and if we give the trivial action of G_p on $\mathbb{Z}/n$, then this homomorphism is also compatible with the action of G_p. Further we can also identify the kernel $\ker(E_q[n] \to \mathbb{Z}/n)$ with μ_n. Thus we have an exact sequence of abelian groups

$$0 \to \mu_n \to E_q[n] \to \mathbb{Z}/n \to 0 \tag{13.16}$$

and each map in the sequence is compatible with the action of G_p. Thus the representation of G_p on $E_q[n]$ is reducible as μ_n is a G_p-stable subspace of $E_q[n]$.

Example 13.17. Let E/K be an elliptic curve and let ℓ be a prime and $m \geq 1$ be any integer. Then we know that G_K operates on the group $E[\ell^m] \cong \mathbb{Z}/\ell^m \oplus \mathbb{Z}/\ell^m$. Thus we get a homomorphism $\rho_{\ell^m} : G_K \to \operatorname{Aut}(E[\ell^m]) \cong \mathrm{GL}_2(\mathbb{Z}/\ell^m)$.

It is not difficult to see that the composite map $G_K \to \mathrm{GL}_2(\mathbb{Z}/\ell^m) \to \mathrm{GL}_2(\mathbb{Z}/\ell^{m-1})$ is $\rho_{\ell^{m-1}}$. Thus we can put these representations together get a homomorphism $\rho : G_K \to \mathrm{GL}_2(\mathbb{Z}_\ell)$, and as $\mathrm{GL}_2(\mathbb{Z}_\ell) \hookrightarrow \mathrm{GL}_2(\mathbb{Q}_\ell)$ then further composition gives a representation $G_K \to \mathrm{GL}_2(\mathbb{Q}_\ell)$.

Proposition 13.18. *Let R be a field of characteristic zero and let $\rho_1, \rho_2 : G \to GL_n(R)$ be two continuous, finite dimensional irreducible representations of a topological group G. Suppose that the traces* $\mathrm{Trace}(\rho_1(g)) = \mathrm{Trace}(\rho_2(g))$ *for dense subset of $g \in G$. Then ρ_1 and ρ_2 are isomorphic representations.*

Proof. See [40] or [10]. □

14. Fine structure of Galois Representations

We want to use the structure of the Galois group $\mathrm{Gal}(\bar{\mathbb{Q}}_p/\mathbb{Q}_p)$ described in Section 12 to probe the structure of Galois representations.

Definition 14.1. Let $\rho : G_K \to \mathrm{GL}_n(R)$ be a continuous representation. Let $\mathfrak{p} \subset \mathcal{O}_K$ be a nonzero prime ideal in K. Then we say ρ is unramified at $\mathfrak{p}$ if the image, $\rho(I_{\mathfrak{p}})$, of the inertia subgroup, $I_{\mathfrak{p}}$ at $\mathfrak{p}$, under ρ is trivial.

Remark 14.2. The definition given above is independent of the choice of a prime lying over $\mathfrak{p}$ as the decomposition groups of all primes lying over $\mathfrak{p}$ in $\bar{K}$ are conjugate and so are the inertia subgroups.

Example 14.3. Let $\rho : G_K \to \mathrm{GL}_n(\mathbb{Z}/m)$ be a continuous representation. Let $H = \ker(\rho)$ be the kernel of ρ. Continuity of ρ shows that H is an open subgroup. Further as $\mathrm{GL}_n(\mathbb{Z}/m)$ is a finite group, we see that the image of ρ is a finite group as well. Let $K_\rho = \bar{K}^H$ be the fixed field of H. Galois Theory provides us an isomorphism:

$$\mathrm{image}(\rho) = \mathrm{Gal}(K_\rho/K) \tag{14.4}$$

Then $\rho : G_K \to \mathrm{GL}_m(\mathbb{Z}/n)$ is unramified at $\mathfrak{p}$ if and only if the extension K_ρ/K is unramified at $\mathfrak{p}$.

Remark 14.5. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_n(R)$ be any continuous Galois representation. Suppose ρ is unramified at p. Then $\rho(\mathrm{Frob}_p)$, where $frob_p$ is a Frobenius element at p, is conjugacy class of elements of $\mathrm{GL}_n(R)$. The characteristic polynomial $\det(1 - X\rho(\mathrm{Frob}_p))$ is well defined as it depends only on the conjugacy class of $\rho(\mathrm{Frob}_p)$. This characteristic polynomial plays a fundamental role in studying a Galois representations.

Example 14.6. Let $\chi_\ell : G_{\mathbb{Q}} \to (\mathbb{Z}/\ell)^*$ be the cyclotomic character. Then χ_ℓ is unramified at all primes $p \neq \ell$. We can calculate the Frobenius elements explicitly in this case. For all $p \neq \ell$, $\chi_\ell(\mathrm{Frob}_p) = p \mod \ell$ and hence $\det(1 - X\chi(\mathrm{Frob}_p)) = (1 - pX)$.

15. Modular forms

Let

$$\mathfrak{H} = \{z \in \mathbb{C} | \mathrm{Re}(z) > 0\} \tag{15.1}$$

be the upper half plane. Let

$$\mathrm{GL}_2^+(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \big| a, b, c, d \in \mathbb{R}, ad - bc > 0 \right\} \tag{15.2}$$

We write

$$\mathrm{GL}_2^+(\mathbb{Q}) = \mathrm{GL}_2^+(\mathbb{R}) \cap \mathrm{GL}_2(\mathbb{Q}) \tag{15.3}$$

We let $\mathrm{GL}_2^+(\mathbb{R})$ act on the topological space $\mathfrak{H}$ by the formula

$$gz = \frac{az+b}{cz+d} \tag{15.4}$$

for any $g \in \mathrm{GL}_2^+(\mathbb{R})$ and any $z \in \mathfrak{H}$.

In this and the subsequent sections we will be interested in subgroups of finite index in $\mathrm{SL}_2(\mathbb{Z})$. We fix a subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ of finite index.

Definition 15.5. A modular form of weight k on Γ is a holomorphic function $f : \mathfrak{H} \to \mathbb{C}$ such that

(1) $f(gz) = (cz+d)^k f(z)$ for all $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$

(2) for every $g \in \mathrm{GL}_2^+(\mathbb{Q})$ the function $\det(g)^{k/2}(cz+d)^{-k} f(gz)$, where $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has a Fourier expansion of the form

$$\det(g)^{-k}(cz+d)^{-k} f(gz) = \sum_{n=0}^{\infty} a_n e^{2\pi i n z/N} \tag{15.6}$$

for some integer $N \geq 1$.

Definition 15.7. We say that a modular form of weight k on Γ is a cusp form if in addition to the above two conditions we have

$$\det(g)^{-k}(cz+d)^{-k} f(gz) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z/N} \tag{15.8}$$

It is easy to verify that a linear combination of modular forms on Γ of weight k is again a modular form on Γ of weight k and product of two modular forms of weights k and m is a modular form of weight $k + m$ on Γ. Thus the set of modular forms of fixed weight on Γ form a complex vector space and the set of all cusp forms is a subspace of the space modular forms.

Example 15.9. Let $k > 2$ be an even integer. For any $z \in \mathfrak{H}$ consider the series

$$G_k(z) = \sum_{(m,n)\neq(0,0)}' \frac{1}{(mz+n)^k} \tag{15.10}$$

and the sum is over all integers (m, n) which are not simultaneously zero. Then it is easy to see that the series is absolutely convergent for all values of $z \in \mathfrak{H}$ and one has $G_k(gz) = (cz+d)^k G_k(z)$ for all $g \in \mathrm{SL}_2(\mathbb{Z})$. Its Fourier expansion is given by

$$G_k(z) = 2\zeta(k)\left(1 - \frac{2k}{B_k}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n\right) \tag{15.11}$$

and where $q = e^{2\pi iz}$. Thus $G_k(z)$ is a modular form of weight k on $\mathrm{SL}_2(\mathbb{Z})$. But $G_k(z)$ is not a cusp form.

Example 15.12. The Ramanujan $\Delta(q)$ function defined by

$$\Delta(q) = q\prod_{n=1}^{\infty}(1-q^n)^{24} = \sum_n = 1^{\infty}\tau(n)q^n.$$

where $q = e^{2\pi iz}$ is a cusp form of weight 1 on $\mathrm{SL}_2(\mathbb{Z})$.

Definition 15.13. Let Γ be a subgroup of finite index in $\mathrm{SL}_2(\mathbb{Z})$. Let $k \geq 1$ be an integer. Let $M_k(\Gamma)$ (resp. $S_k(\Gamma)$) be the space of modular forms of weight k (resp. cusp forms) on weight k on Γ.

Observe that $S_k(\Gamma) \subset M_k(\Gamma)$.

From now on we will be interested in the following kinds of subgroups of $\mathrm{SL}_2(\mathbb{Z})$.

Definition 15.14. Let $N \geq 1$ be any integer. Define $\Gamma_1(N)$ as follows

$$\Gamma_1(N) = \left\{ g = \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \mod N \middle| g \in \mathrm{SL}_2(\mathbb{Z}) \right\} \tag{15.15}$$

and

$$\Gamma_0(N) = \left\{ g = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \mod N \middle| g \in \mathrm{SL}_2(\mathbb{Z}) \right\} \tag{15.16}$$

I leave it as an exercise to check that these two subgroups are of finite index in $\mathrm{SL}_2(\mathbb{Z})$. Moreover one has a homomorphism

$$\Gamma_0(N) \to (\mathbb{Z}/N)^* \tag{15.17}$$

given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto d \mod N. \tag{15.18}$$

This defines a surjective homomorphism of groups and the kernel of this homomorphism is precisely the subgroup $\Gamma_1(N)$ and hence we see that $\Gamma_1(N) \subset \Gamma_0(N)$ is a normal subgroup.

We will be interested in studying modular forms on these two groups. It is standard (see [42]) that $M_k(\Gamma_1(N))$ and $M_k(\Gamma_0(N))$ are finite dimensional $\mathbb{C}$-vector spaces. In particular the spaces of cusp forms on these groups are finite dimensional as well.

Definition 15.19. A modular form (resp. cusp form) of level N, weight k is a form $f \in M_k(\Gamma_1(N))$ (resp. $f \in S_k(\Gamma_1(N))$).

Let $\chi : (\mathbb{Z}/N) \to \mathbb{C}^*$ be any homomorphism (i.e., a Dirichlet character). For $d|N$ we set $\chi(d) = 0$ and extend this function to $\mathbb{Z}/N$.

Definition 15.20. A modular form $f \in M_k(\Gamma_1(N))$ is said to be a modular form of level N, weight k and *nebentype* χ if f satisfies the following:

$$f(gz) = \chi(d)(cz+d)^k f(z) \tag{15.21}$$

and we write $M_k(\Gamma_0(N), \chi) \subset M_k(\Gamma_1(N))$ for the space of such forms on $\Gamma_1(N)$. We also define $S_k(\Gamma_0(N), \chi)$ in the obvious way.

One has the following decomposition of complex vector spaces

$$M_k(\Gamma_1(N)) = \oplus_{\chi \mod N} M_k(\Gamma_0(N), \chi), \tag{15.22}$$

and

$$S_k(\Gamma_1(N)) = \oplus_{\chi \mod N} S_k(\Gamma_0(N), \chi). \tag{15.23}$$

for space of cusp forms.

Definition 15.24. A modular form (resp. cusp form) of level N, weight k and nebentype χ is a form $f \in M_k(\Gamma_0(N), \chi)$ (resp. $f \in S_k(\Gamma_0(N), \chi)$).

Example 15.25. Let $N = 11$. Then the function $f(z) = (\Delta(q)\Delta(11q)^{1/2} = q\prod_{n=1}^{\infty}((1-q^n)(1-q^{11n})^2$ is a cusp form of level 11, weight 2 and nebentype $\chi = 1$.

Remark 15.26. Let $f(z) \in M_k(\Gamma_1(N))$ be a modular form of weight k and level N. Then as

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_1(N) \tag{15.27}$$

we see that $f(Tz) = f(z+1) = f(z)$ so $f(z)$ is a complex holomorphic function which is periodic with period 1 and so its Fourier expansion looks like

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z} \tag{15.28}$$

Definition 15.29. We will call the coefficients a_i the Fourier coefficients of f and write $K_f = \mathbb{Q}(a_0, a_1, \cdots)$ for the field generated by the Fourier coefficients of f. In general K_f is not even an algebraic extension of $\mathbb{Q}$. But under interesting circumstances it is a finite extension of $\mathbb{Q}$. In any case we will refer to K_f as the field of Fourier coefficients of f.

16. Hecke operators

We will restrict attention to the two special subgroups introduced earlier. For a general and comprehensive account of the theory see [42], [25], [29]. From now on we will concentrate on the space of cusp forms on $\Gamma_1(N)$ and keep a track of the nebentype as we go along.

The spaces of modular forms on $\Gamma_1(N)$ come equipped with a commutative family of operators, which were introduced by Hecke and are named after him. We will not recall the definition of Hecke operators but only recall the effect of Hecke operators on the Fourier coefficients of any modular form. For every $n \geq 1$ we have a $\mathbb{C}$-linear mapping

$$T_n : M_k(\Gamma_1(N)) \to M_k(\Gamma_1(N)) \tag{16.1}$$

which takes the subspace of cusp forms to cusp forms and forms withe Nebentype to forms with the same nebentype. We call T_n the n^{th} Hecke operators. First property of these operators is

$$T_{nm} = T_n T_m \qquad \text{for } (m,n) = 1, \tag{16.2}$$

Thus it is sufficient to define these operators when $n = \ell^m$ for any prime ℓ and $m \geq 1$. So fix a prime ℓ. Let $f \in S_k(\Gamma_0(N), \chi)$ be a form level N, weight k and nebentype χ.

$$T_\ell(f) = \sum_{n=1}^{\infty} a_{\ell n} q^n + \chi(\ell)\ell^{k-1} \sum_{n=1}^{\infty} a_n q^{\ell n} \tag{16.3}$$

recall that when $\ell | N$ we have set $\chi(\ell) = 0$.

We define T_{ℓ^m} recursively using the above definition for $m = 1$.

Definition 16.4. A Hecke eigenform $f \in S_k(\Gamma_0(N), \chi)$ is a common eigen function of all the Hecke operators T_ℓ, i.e., for all primes ℓ there exists a constant λ_ℓ such that

$$T_\ell(f) = \lambda_\ell f. \tag{16.5}$$

Then one gets relations

$$a_{\ell n} - \lambda_\ell a_n + \chi(\ell)\ell^{k-1}a_{n/\ell} = 0 \tag{16.6}$$

for all $n \geq 1$, and where we have set $a_{n/\ell} = 0$ if $\ell \not| n$. This relation gives $a_\ell = \lambda_\ell a_1$. Moreover if $a_1 = 0$ then one checks that $f = 0$. Thus we can assume a_1 is nonzero for any Hecke eigenform and we normalize any Hecke eigenform by setting $a_1 = 1$.

From the above discussion it is evident the Fourier coefficients of normalized Hecke eigenforms are the eigenvalues of the corresponding Hecke operators. This gives recursion relations between Fourier coefficients.

$$a_{\ell n} = a_\ell a_n - \chi(\ell)\ell^{k-1}a_{n/\ell} \tag{16.7}$$

and, in particular, if we take $n = \ell^m$ then we get

$$a_{\ell^{m+1}} = a_\ell a_{\ell^m} - \chi(\ell)\ell^{k-1}a_{\ell^{m-1}} \tag{16.8}$$

and $a_{rn} = a_r a_n$.

17. New forms

Let $M|N$ and let $d|(N/M)$ and assume $d > 1$ (so M divides N properly). Let $f \in S_k(\Gamma_1(M))$. Then the mapping $S_k(\Gamma_1(M)) \to S_k(\Gamma_1(N))$ defined by $f(z) \mapsto f(dz)$ is injective and this mapping takes forms with nebentype to forms with the same nebentype and it takes eigen vectors of Hecke operators T_ℓ for $\ell \not| N$ to eigenforms. We define $S_k(\Gamma_1(N))^{\text{old}}$ to be space of all forms which arise in this way from forms of lower level.

Definition 17.1. A form $f \in S_k(\Gamma_1(N))$ which is not in $S_k(\Gamma_1(N))^{\text{old}}$ is called a newform. The set of new forms is a subspace $S_k(\Gamma_1(N))$ and we denote it by $S_k(\Gamma_1(N))^{\text{new}}$.

It is a basic fact in the theory of new forms that $S_k(\Gamma_1(N))^{\text{new}}$ has a basis consisting of normalized new eigenforms. For a proof see [29].

Remark 17.2. Normalized Hecke newforms are uniquely identified by their Fourier coefficients and the Fourier coefficients are all algebraic numbers and that field K_f all the Fourier coefficients of a new eigen form f is a finite extension of $\mathbb{Q}$, i.e., K_f is a number field. For a proof see [42].

18. Galois representations associated to modular forms

Deligne's proof of the Ramanujan conjecture also produced the Galois representation associated to normalized new cusp eigenform. This representation was also constructed by Shimura for $k = 2$.

Theorem 18.1 (Deligne). *Let $f \in S_k(\Gamma_0(N), \chi)$ be a normalized new eigenform and let $f = \sum_1^\infty a_n q^n$ be its Fourier expansion. Let $K_f = \mathbb{Q}(a_0, a_1, \cdots)$ be the field of Fourier coefficients of f and let $\mathfrak{p}$ be any prime lying over p in K_f. Let $K_{f,\mathfrak{p}}$ be the p-adic field associated to K_f at $\mathfrak{p}$. Then there exists a two dimensional, irreducible representation,*

$$\rho_{f,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(K_{f,\mathfrak{p}}) \tag{18.2}$$

with the following properties:

1. *$\rho_{f,p}$ is unramified outside pN,*
2. *for all $\ell \not| (pN)$ the characteristic polynomial of $\rho(\mathrm{Frob}_\ell)$ is given by the formula:*

$$\det(1 - X\rho(\mathrm{Frob}_\ell)) = 1 - a_\ell X + \ell^{k-1}\chi(\ell)X^2 \tag{18.3}$$

The following consequence of Theorem 12.11 and Proposition 13.18 shows that the representation constructed in Theorem 18.1 is characterized, up to isomorphism, by the properties listed in Theorem 18.1.

Proposition 18.4. *Any continuous irreducible, finite dimensional representation of $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ is determined, up to isomorphism, by the traces of Frobenius elements at all the primes where the representation is unramified.*

Remark 18.5. Such a representation $\rho_{f,p}$ has the property that $\det(\rho(\mathrm{Frob}_\ell)) = -1$

19. The Ramanujan Estimate

The following estimate for the size of the Fourier coefficients of normalized Hecke eigen cusp forms of level N, weight k and nebentype χ was conjectured by Ramanujan (see [35]) and Deligne (see [11]) showed that it is a consequence of the Weil conjectures (see [12]).

The estimate is a generalization of the Hasse-Weil estimate for elliptic curves (see Theorem 6.1).

Theorem 19.1 (Deligne; Ramanujan). *Let $f = \sum_{n=1}^\infty a_n q^n$ be a normalized Hecke eigen form in $S_k(\Gamma_1(N))^{\mathrm{new}}$. Then for any prime ℓ we have*

$$|a_\ell| \leq 2\ell^{\frac{k-1}{2}}. \tag{19.2}$$

For $k = 2$, this estimate was also proved by Eichler and Shimura.

20. Reducing Galois representations modulo p

From now on we will assume that we have a normalized new Hecke eigen cusp form f such that $K_f = \mathbb{Q}$.

Most of theory outlined in this section works with out assumption. But we are restricting ourselves to this case as it keeps the notation fairly simple and transparent.

Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_n(\mathbb{Q}_p)$ be a continuous irreducible representation. As $G_{\mathbb{Q}}$ is a compact group, it is easy to see, using continuity of ρ, that the image of ρ is contained in a compact subgroup of the target. One can explicitly see this by observing that $G_{\mathbb{Q}}$ stabilizes a lattice $L \cong \mathbb{Z}_p^n \subset \mathbb{Q}_p^n$. Using this lattice we see that we ρ is the composite of $G_{\mathbb{Q}} \to \mathrm{GL}(L) \cong \mathrm{GL}_n(\mathbb{Z}_p) \hookrightarrow \mathrm{GL}_n(\mathbb{Q}_p)$. In other words we can arrange things in such a way that ρ has p-adic integer matrix entries. So we can then reduce these matrices modulo p. Thus we obtain, a representation $\bar{\rho} : G_{\mathbb{Q}} \to GL_n(\mathbb{Z}_p) \to \mathrm{GL}_n(\mathbb{Z}/p)$. Such a representation $\bar{\rho}$ is not in general unique and its construction depends in a rather strong way on the choice of the lattice which was used to carry out the reduction.

However, it is standard fact that such representation $\bar{\rho}$ is unique (up to isomorphism) if it is irreducible and its formation is independent of the choice of the lattice.

Example 20.1. Let $\Delta(q) = \sum_{n=1}^{\infty} \tau(n) q^n$ be the Ramanujan modular form of weight twelve and level 1. Take $p = 691$. Then Ramanujan observed that $\tau(n) \equiv \sigma_{11}(n) \mod 691$. This congruence implies that the mod 691 representation associated to Δ is reducible. For more on the connection between congruences between Fourier coefficients of modular forms and Galois representations see Serre's Seminar Bourbaki talk on the work of Swinnerton-Dyer [39].

21. Galois representations arising from elliptic curves

Let $E/\mathbb{Q}$ be an elliptic curve over $\mathbb{Q}$. Fix a prime p. One has a family of Galois representations associated to E, p, given by the action of the Galois group $G_{\mathbb{Q}}$ on the p^n-torsion points $E[p^n]$, which are denoted

$$\rho_{E,p^n} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/p^n) \tag{21.1}$$

It is not very difficult to verify that the composite homomorphism

$$\rho_{E,p^n} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/p^n) \to \mathrm{GL}_2(\mathbb{Z}/p^{n-1}) \tag{21.2}$$

is $\rho_{E,p^{n-1}}$. Such a system of representations can be assembled to give a continuous representation (note our abuse of notation)

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_p). \tag{21.3}$$

Theorem 21.4. *Let $E/\mathbb{Q}$ be an elliptic curve and let N_E be its conductor. Then*

(1) *$\rho_{E,p}$ is unramified for primes not dividing pN_E and*
(2) *let $\ell \not| pN_E$ be a prime. Then*

$$\det(1 - X\rho_{E,p}(\mathrm{Frob}_\ell)) = 1 - a_\ell X + \ell X^2. \tag{21.5}$$

where $a_\ell = \ell + 1 - \#E(\mathbb{F}_\ell)$.

Remark 21.6. It is clear from the above theorem and Deligne's theorem 18.1 that the representation associated to an elliptic curves looks like the representation associated to a modular form of level N_E, weight 2 and $\chi = 1$.

22. The Shimura-Taniyama-Weil Conjecture

There are many equivalent definition of modularity of an elliptic curve. We will restrict our attention to the line of thought which emerged in the discussion at the end of previous section.

Definition 22.1. Let $E/\mathbb{Q}$ be an elliptic curve. We will say that E is a modular elliptic curve if the representation

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}_p) \tag{22.2}$$

if there exists a normalized new cusp form f on $\Gamma_1(N)$ such that

$$a_\ell(f) = a_\ell(E) \tag{22.3}$$

for all but finite number of primes (not dividing the level), and where $a_\ell(f)$ is the ℓ^{th} Fourier coefficient of f.

Remark 22.4. By Proposition 18.4, we observe that $\rho_{E,p}$ is isomorphic to the Galois representation of f at p. Thus to say that E is modular is equivalent to saying that $a_\ell(E) = \ell + 1 - \#(E(\mathbb{F}_\ell))$ are the eigenvalues of the Hecke operators on a normalized Hecke newform of some level, of weight two and $\chi = 1$.

It is a standard fact that if $\rho_{E,p}$ arises from a modular form f for one prime p then it does so for all primes p. Thus in this sense the definition is independent of the prime p.

Conjecture 22.5 (Shimura-Taniyama-Weil)**.** Every elliptic curve $E/\mathbb{Q}$ is modular.

This conjecture is now a theorem of Breuil, Conrad, Diamond and Taylor, Wiles [3] and its proof uses methods of [49], [46].

23. Serre's Conjecture

Serre in [41] formulated a mod p version of a similar conjecture. This conjecture of Serre still remains intractable.

In this section we will consider representations $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$. Assume that ρ is irreducible and continuous. Observe that continuity of ρ implies that the image is finite and hence is contained in $\mathrm{GL}_2(\mathbb{F}_q)$ for some finite extension $\mathbb{F}_q/\mathbb{F}_p$.

Definition 23.1. Let ρ be as above. We will say ρ is odd if $\det(\rho(c)) = -1 \in \bar{\mathbb{F}}_p^*$ where $c : \bar{\mathbb{Q}} \to \bar{\mathbb{Q}}$ is complex conjugation.

We remark that if $p = 2$ then $-1 = 1$ so there is no restriction at $p = 2$.

Definition 23.2. We will say that $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ is modular if there exists a new eigen cusp form f on $\Gamma_1(N)$ for some $N \geq 1$ such that $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p^*)$ can be obtained by reduction the representation $\rho_{f,\mathfrak{p}} : G_{\mathbb{Q}} \to \mathrm{GL}_2(K_{f,\mathfrak{p}})$ modulo $\mathfrak{p}$ for some prime $\mathfrak{p}|p$ in K_f.

Conjecture 23.3 (Serre)**.** Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ be any continuous, irreducible, odd representation of $G_{\mathbb{Q}}$. Then ρ is modular.

Remark 23.4. The condition that ρ is odd is necessary by Remark 18.5.

Serre also gave a recipe for computing the level, $N(\rho)$ and the weight $k(\rho)$ and the nebentype ϵ_ρ of such a form. As a consequence of [5], [13], [19], [32], [37] that if ρ is modular of some level weight and nebentype, then it does indeed arise from modular form of weight and level predicted by Serre.

Remark 23.5. For our purposes it is sufficient to know that for a representation $\rho : G_{\bar{\mathbb{Q}}} \to \mathrm{GL}_2(\bar{\mathbb{F}}_p)$ the predicted level $N(\rho)$ is coprime to p and is divisible only by prime $\ell \neq p$ which for which ρ is ramified.

24. Frey Elliptic Curves

Let A, B, C be pair wise coprime integers such that $A + B + C = 0$. Then we can associate an elliptic curve to such a triple of integers. Let

$$E_{A,B,C} : y^2 = x(x - A)(x + B) \tag{24.1}$$

For this curve we have $\Delta_E = 16(ABC)^2$. And hence as one of A, B, C is even we have

Lemma 24.2. *For elliptic curve $E_{A,B,C}$ we have $p|\Delta \Leftrightarrow p|ABC$.*

We now study the bad reduction of $E_{A,B,C}$ using Proposition 8.4. We observe that $c_4 = 16(A^2 + AB + B^2)$

Lemma 24.3. *Let $E = E_{A,B,C}$ be an elliptic curve associated to a triple of integers A, B, C as above. Then $E_{A,B,C}$ has good reduction outside primes p not dividing ABC. For any $p|ABC$, the curve E has semistable reduction modulo p.*

Proof. It is not difficult to verify that this equation is in fact minimal for all primes $p \neq 2$. The proof is easy for $p \neq 2$. for instance if $p|A$ then the reduction looks like $y^2 = x^2(x+B) \mod p$. (Note that $c_4 \not\equiv 0 \mod p$ as any $p|\Delta_E$ divides exactly one of A, B, C). For $p = 2$ we have to do a little more work. One essentially reduces to the case when $A \equiv -1 \mod 4$ and $B \equiv 0 \mod 32$. Then we can substitute

$$x = 4X \tag{24.4}$$

$$y = 8Y + 4X. \tag{24.5}$$

Then the equation of E reduces to

$$Y^2 + XY = X^3 + cX^2 + dX \tag{24.6}$$

where

$$c = \frac{B-1-A}{4} \tag{24.7}$$

$$d = -\frac{AB}{16} \tag{24.8}$$

and then the reduction of E modulo 2 is given by

$$Y^2 + XY = \begin{cases} X^3 & \text{if } A \equiv 7 \mod 8, \\ X^3 + X^2 & \text{if } A \equiv 3 \mod 8. \end{cases} \tag{24.9}$$

This curve has distinct tangents over $\bar{\mathbb{F}}_2$ and hence the reduction of E modulo 2 is semistable at 2. □

Thus we get

$$N_E = \prod_{p|ABC} p \tag{24.10}$$

$$j_E = \frac{2^8(C^2 - AB)^3}{A^2B^2C^2} \tag{24.11}$$

Proposition 24.12. *For $p \geq 5$ the representation*

$$\rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p) \tag{24.13}$$

is irreducible and is unramified for primes not dividing ABC.

25. Frey Curves arising from FLT

Fix a prime $p \geq 5$. Let $a^p + b^p + c^p = 0$ be a nontrivial solution to Fermat's last theorem. We will assume that a, b, c are pairwise coprime and $abc \neq 0$. We will assume, without loss of generality, that $a \equiv -1 \mod 4$ and $b \equiv 0 \mod 4$. If we set $A = a^p, B = b^p, C = c^p$ then we get elliptic curves which were introduced and studied by G. Frey (see [20], [21], [22])

$$y^2 = x(x - a^p)(x + b^p) \tag{25.1}$$

26. Analysis of Ramification

In this section we analyze the ramification properties of the representations ρ obtained from the Frey elliptic curves $E = E_{A,B,C}$ for $A + B + C = 0$. Let $\rho = \rho_{E,p} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p)$ be the two dimensional representation of $G_{\mathbb{Q}}$ corresponding to its action on $E[p]$.

It suffices to concentrate on primes $\ell | (ABC)$ as ρ is unramified at primes p not dividing ABC. Fix a prime $\ell | ABC$. Then we know that E has semistable reduction at ℓ.

We will use Tate elliptic curves (see Section 11) to analyze the ramification at ℓ. For $\ell = 2$ we will need a refined version of Theorem 11.17 to take care of the possibility that the reduction type at 2 may be non-split semistable. But the difficulties are mostly technical and I will suppress this issue completely. The diligent reader is referred to [43].

By Tate's theorem, we can replace E by a Tate curve E_{q_ℓ} over $\mathbb{Q}_\ell$. Note that q_ℓ depends on E. In Remark 11.20 we had described $E_{q_\ell}[p]$ explicitly:

$$E_{q_\ell}[p] \cong \left\{ \zeta q_\ell^{m/p} | \zeta^p = 1 \text{ and } m \in \mathbb{Z}/p \right\} \tag{26.1}$$

We can read of the ramification properties of ρ at ℓ from this description.

Proposition 26.2. *Let $\ell | ABC$. Then the extension corresponding to $E_{q_\ell}[p]$ is given by $\mathbb{Q}_\ell(\zeta, q_\ell^{1/p})$. In particular this extension is unramified for $\ell \neq 2, p$ if and only if $\nu_\ell(q_\ell) \equiv 0 \mod p$.*

Proof. If $\ell \neq 2, p$ and $\nu_\ell(q_\ell) \equiv 0 \mod p$ then $q_\ell = \ell^p u$ where u is unit in $\mathbb{Z}_\ell$. Thus the extension is given by $\mathbb{Q}_\ell(\zeta, u^{1/p})$ where ζ is a p^{th}-root of unity and $\ell \neq p$ and u is a unit in $\mathbb{Z}_\ell$. Thus this extension is unramified at ℓ. Conversely, if $\mathbb{Q}_\ell(\zeta, q_\ell^{1/p})$ is unramified at ℓ then we see that q_ℓ is a p^{th} power up to a unit in $\mathbb{Z}_\ell$. □

One can also carry out the analysis at $\ell = p$ however, the extension is ramified at $\ell = p$. But the ramification is fairly controlled (Serre calls this case peu ramifie) if and only if $\nu_p(q) \equiv 0 \mod p$. So we will make the following ad hoc definition.

Definition 26.3. Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p)$ be a representation. We will say that ρ is peu ramifie at p if the extension at p is given by $\mathbb{Q}_p(\zeta, u^{1/p})$ where ζ is a p^{th}-root of unity and $u \in \mathbb{Z}_p$ is a p-adic unit.

We now apply this analysis to the Frey curve and deduce:

Proposition 26.4. *Let $a^p + b^p + c^p = 0$ be a solution to Fermat's last theorem such that a, b, c are pairwise coprime and $abc \neq 0$ and let E be the Frey elliptic curve associated to it. Then the representation*

$$\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p) \tag{26.5}$$

is unramified for all prime $\ell \neq 2, p$ and for $\ell = p$ the representation is peu ramifie.

Proof. The is immediate from the fact that $\Delta = (abc)^p$ up to a power of 2 and $\nu_\ell(q) = \nu_\ell(\Delta_E)$ and hence the result. □

27. Fermat's Last Theorem

Theorem 27.1 (Serre)**.** *Serre's conjecture implies Fermat's Last Theorem.*

Proof. Serre's conjecture 23.3 implies that the representation $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{F}_p)$ associated to the corresponding Frey elliptic curve is modular of some level $N(\rho)$, weight $k(\rho)$ and nebentype χ. If we follow Serre's recipe (see Remark 23.5) for calculating these invariants we see that $N(\rho) = 2$ because $N(\rho)$ is divisible by only those primes $\ell \neq p$ for which ρ is ramified. Thus by Proposition 26.2 we see that $N(\rho) = 2$. The fact that ρ is peu ramifie at p gives that $k(\rho) = 2$ and recipe for nebentype gives and $\chi = 1$.

Thus the representation associated to the Frey curve arises from a newform of weight 2, level 2 and nebentype 1. But it is known (and fairly elementary to prove) that there are no such forms (see for instance [36]). Thus we arrive at contradiction. This proves the theorem. □

References

[1] B. Birch and H. P. F. Swinnerton Dyer. Notes on elliptic curves I. *J. Reine Angew. Math.*, 212:7–25, 1963.

[2] B. Birch and H. P. F. Swinnerton Dyer. Notes on elliptic curves II. *J. reine angew. Math.*, 218:79–108, 1965.

[3] C. Breuil, B. Conrad, F. Diamond, and R. Taylor. Modular elliptic curves over **Q**: Wild 3-adic exercises. *Preprint*, 2000.

[4] H. Carayol. Sur les représentations ℓ-adiques associées aux formes modulaires de Hilbert. *Ann. Sci. École Norm. Sup. (4)*, 19(3):409–468, 1986.

[5] H. Carayol. Sur les représentations galoisiennes modulo ℓ attachées aux formes modulaires. *Duke Math. J.*, 59(3):785–801, 1989.

[6] N. Chebotarev. Die bestimmung der dichtigkeit einer menge von primzahlen, welche zu einer gegebenen substitutionsklasse gehören. *Math. Ann.*, 95:151–228, 1925.

[7] B. Conrad. Ramified deformation problems. *Duke Math. J.*, 97(3):439–513, 1999.

[8] B. Conrad, F. Diamond, and R. Taylor. Modularity of certain potentially Barsotti-Tate Galois representations. *J. Amer. Math. Soc.*, 12(2):521–567, 1999.

[9] G. Cornell, J. Silverman, and G. Stevens, editors. *Modular forms and Fermat's last theorem*. Springer-Verlag, 1995.

[10] C. Curtis and I. Reiner. *Representation theory of finite groups and associated algebras.* Interscience, New-York-London, 1962.

[11] P. Deligne. Formes modulaires et représentations ℓ-adiques. In *Sem. Bourbaki*, volume 179 of *Lecture Notes in Mathematics*, Berlin, February 1969. Springer-Verlag.

[12] P. Deligne. La conjecture de Weil I. *Publ. Math. I.H.E.S*, 43:5–77, 1974.

[13] F. Diamond. The refined conjecture of Serre. In *Elliptic curves, modular forms and Fermat's Last Theorem*, volume 1 of *Number Theory*, pages 22–37, Cambridge, MA., 1993. International Press.

[14] F. Diamond. On deformation rings and Hecke rings. *Ann. of Math. (2)*, 144(1):137–166, 1996.

[15] F. Diamond. The Taylor-Wiles construction and multiplicity one. *Invent. Math.*, 128(2):379–391, 1997.

[16] F. Diamond, H. Darmon, and R. Taylor. *Fermat's last theorem*, pages 1–154. Current developments in mathematics. International Press, Cambridge, MA, 1997.

[17] F. Diamond and K. Kramer. Modularity of a family of elliptic curves. *Math. Res. Let.*, 2(3):299–304, 1995.

[18] B. Dwork. On the rationality of the zeta function of an algebraic variety. *Amer. Jour. of Math.*, 82:631–648, 1960.

[19] B. Edixhoven. The weight in serre's conjectures on modular forms. *Invent. Math.*, 109(3):563–594, 1992.

[20] G. Frey. Rationale Punkte auf Fermatkurven und getwisteten Modulkurven. *Journal reine angew. Math.*, 331:185–191, 1982.

[21] G. Frey. Links between stable elliptic curves and certain Diophantine equations. *Ann. Univ. Saraviensis, Ser. Math.*, 1:1–40, 1986.

[22] G. Frey. Links between elliptic curves and solutions of $A - B = C$. *J. Indian Math. Soc. N.S.*, 51:117–145, 1987.

[23] H. Hasse. *Mathematische Abhandlungen*, volume 1-3. Walter de Gruyter, Berlin-New-York, 1975.

[24] H. Hida. Iwasawa modules attached to congruences of cusp forms. *Ann. Scient. de l'E. N.S.*, 19:231–273, 1986.

[25] H. Hida. *Elementary theory of L-functions and Eisenstein series*, volume 26 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1993.

[26] K. Ireland and M. Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New-York, second Edition edition, 1990.

[27] N. Katz. An overview of Deligne's proof. In *Proc. Symp. Pure Math.*, volume 28. AMS, 1976.

[28] N. Koblitz. *A course in number theory and cryptography*, volume 114 of *Graduate Texts in Mathematics*. Springer-Verlag, New-York, second Edition edition, 1994.

[29] S. Lang. *Introduction to Modular forms*, volume 222 of *Grundlehren de Mathematishen Wissenschaften*. Springer-Verlag, Berlin, corrected and reprint 1995 edition, 1976.

[30] M. Laska. An algorithm for finding a minimal Weierstrass equation for an elliptic curve. *Math. Comp.*, 38:257–260, 1982.

[31] S. Lubkin. A p-adic proof of Weil's conjectures. *Ann. of Math. (2)*, 87:195–255, 1968.

[32] B. Mazur. Modular curves and the Eisenstein ideal. *Publ. Math. IHES*, 47:33–186, 1977.

[33] B. Mazur and A. Wiles. On p-adic analytic families of modular forms. *Comp. Math.*, 59:231–264, 1986.

[34] Barry Mazur. Deforming Galois Representations. In Y. Ihara, K. Ribet, and J.-P. Serre, editors, *Galois groups over* $\mathbf{Q}$, number 16 in Mathematical Sciences Research Institute Publications, pages 385–438, Berlin, 1989. Springer-Verlag.

[35] S. Ramanujan. *Collected Mathematical Papers*.

[36] R. Rankin. *Modular forms and functions*. Cambridge University Press, Cambridge, 1977.

[37] K. Ribet. On modular representations of $Gal(\bar{\mathbf{Q}}/\mathbf{Q})$ arising from mdoular forms. *Invent. Math.*, 100(2):431–476, 1990.

[38] J.-P. Serre. *Abelian ℓ-adic representations*. Benjamin, New York.

[39] J.-P. Serre. Congruences et formes modulaires (d'aprés h. p. f. Swinnerton-Dyer). In *Sém. Bourbaki*, number 416, 1971/72.

[40] J.-P. Serre. *Local fields*, volume 67 of *Graduate Texts in Mathematics*. Springer, Berlin, Springer-Verlag, 1979.

[41] J.-P. Serre. Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$. *Duke Math. Journal*, 54(1):179–230, 1987.

[42] G. Shimura. *Introduction to arithmetic theory of automorphic forms*. Princeton University Press, 1971.

[43] Joseph Silverman. *The arithmetic of elliptic curves*, volume 106 of *Graduate Text in Mathematics*. Springer-Verlag, Berlin, 1985.

[44] J. Tate. The arithmetic of elliptic curves. *Invent. Math.*, 23:179–206, 1974.

[45] J. Tate. Algorithm for determining the type of a singular fibre in an elliptic pencil. In *Modular functions of one variable IV*, volume 476 of *Lecture Notes in Mathematics*, pages 33–52, Berlin, 1975. Springer-Verlag.

[46] R. Taylor and A. Wiles. Ring theoretic properties of certain Hecke algebras. *Ann. of Math. (2)*, 141:553–572, 1995.

[47] André Weil. Sur un théorème de Mordell. *Bull. Sci. Math.*, 54:182–191, 1930.

[48] André Weil. Number of solutions of equations in finite fields. *Bull. AMS*, 55:497–508, 1949.

[49] A. Wiles. Modular elliptic curves and Fermat's last theorem. *Ann. of Math. (2)*, 141:443–551, 1995.

Kirti Joshi, Department of Mathematics, University of Arizona, Tucson, AZ 85721

E-mail address: `kirti@math.arizona.edu`

Notes on Ribet's Converse to Herbrand

Chandrashekhar Khare

1. Statement of the theorem

Let p be an odd prime. It is called irregular if and only if p divides the class number of $\mathbf{Q}(\mu_p)$. By Kummer's criterion this happens if and only if p divides the numerator of the kth Bernoulli number B_k for an even k between 2 and $p-3$ (note that the denominator is prime to p because of the von-Staudt-Clausen theorem). Recall that the Bernoulli numbers B_n are defined by:

$$\frac{t}{e^t-1}+\frac{t}{2}-1=\Sigma_{n\geq 2}\frac{B_n}{n!}t^n.$$

The first few Bernoulli numbers are

$$B_4=\frac{-1}{30},B_6=\frac{1}{42},B_8=\frac{-1}{30},B_{10}=\frac{5}{66},B_{12}=\frac{-691}{2730}.$$

Let A be the ideal class group of $\mathbf{Q}(\mu_p)$ and let C be the $\mathbf{F}_p$-vector space A/A^p. This has an action of the Galois group $\Delta := Gal(\mathbf{Q}(\mu_p)/\mathbf{Q}) \equiv (\mathbf{Z}/p\mathbf{Z})^*$. We define the mod p cyclotomic character $\chi : \Delta \to (\mathbf{Z}/p\mathbf{Z})^*$ by $g.\zeta = \zeta^{\chi(g)}$. Note that χ generates the character group of Δ with the characters taking values in $\overline{\mathbf{F}}_p^*$.

As Δ has order prime to p we have a canonical decomposition of C as

$$C=\oplus_{i \pmod{p-1}}C(\chi^i)$$

where $C(\chi^i)$ is the χ^i-isotypical component of C as a Δ-module. Note that $C(\chi^i)=e_{\chi^i}C$ where

$$e_{\chi^i}=\frac{1}{p-1}\Sigma_{g\in\Delta}\chi^{-i}(g)g.$$

The main theorem proven by Ribet in [R] is:

Theorem 1 *Let k be an even integer, $2 \leq k \leq p-3$. Then p divides the numerator of B_k if and only if $C(\chi^{1-k}) \neq 0$.*

It was known classically by Herbrand, refining Stickelberger's theorem, that if $C(\chi^{1-k}) \neq 0$, then p divides (numerator of) B_k (Section 3 of Chapter 1 of [L]). The converse was also well-known assuming Vandiver's conjecture that $\mathbf{Q}(\mu_p)^+$ has class number prime to p. The theorem is also a consequence of the Main Conjecture of Iwasawa theory which was proved for abelian number fields by Mazur-Wiles ([MW]). The proof of Ribet, and its reinterpretation and extension in [W], was a significant clue for the work of Mazur-Wiles. Now its possible to give technically simpler proofs of this result using the important technique of Euler systems developed by Kolyvagin (see Rubin's appendix in [L]). But the proof of Ribet is still valuable as it *explicitly* constructs abelian, unramified extensions of exponent p of $\mathbf{Q}(\mu_p)$ with controlled behaviour.

Note that by class field theory we have an isomorphism via the Artin symbol, that we denote by Art, of C with the maximal unramified abelian p-extension E of $\mathbf{Q}(\mu_p)$ (note that the p-torsion of A can be identified with C as a Δ-module). The abelian Galois group $H := Gal(E/\mathbf{Q}(\mu_p))$ has an action of the group Δ by conjugation, that also acts on C as seen above. The Artin reciprocity map is equivariant for the action of Δ. Namely we have

$$Art : g.\mathsf{c} \in C \to gArt(\mathsf{c})g^{-1} \in H.$$

Thus under the hypothesis that p divides B_k $(2 \leq k \leq p-3)$ to construct a non-trivial element in $C(\chi^{1-k})$ it is enough to construct an unramified abelian p-extension $E/\mathbf{Q}(\mu_p)$ (so $Gal(E/\mathbf{Q}(\mu_p))$ is a $\mathbf{Z}/p\mathbf{Z}$ vector space) such that Δ acts on it via χ^{1-k}.

We easily see that then $Gal(E/\mathbf{Q})$ is the semi-direct product of $(\mathbf{Z}/p\mathbf{Z})^r$ (for some positive integer r) by $\mathbf{Z}/p\mathbf{Z}^*$, with the action given by $g.a = \chi^{1-k}(g)a$, $g \in \Delta, a \in Gal(E/\mathbf{Q}(\mu_p))$.

We claim that Theorem 1 follows from:

Theorem 2 *Suppose $p|B_k$. Then there is a representation*

$$\rho : G_{\mathbf{Q}} \to GL_2(\mathbf{F}),$$

where $\mathbf{F}$ is a finite extension of $\mathbf{F}_p$ with the properties:

(i) ρ is unramified at all primes different from p.

(ii) ρ is a reducible non-semisimple representation of the form

$$\begin{pmatrix} 1 & * \\ 0 & \chi^{k-1} \end{pmatrix}$$

with the $*$ non-trivial: another way to say this is the order of the image of ρ is divisible by p.

(iii) Let D be a decomposition group of p in $Gal(\overline{\mathbf{Q}}/\mathbf{Q})$. Then the order of $\rho(D)$ is prime to p: namely the representation ρ when restricted to D is semisimple.

We justify the claim:

If we define E' to be the fixed field of the kernel of ρ, then the fixed field $\mathbf{Q}(\mu_p)^{\otimes(k-1)}$ of the subgroup consisting of matrices of the form

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

is a subfield of $\mathbf{Q}(\mu_p)$ (of order $\frac{(p-1)}{(p-1,k-1)}$), and $Gal(\mathbf{Q}(\mu_p)^{\otimes(k-1)}/\mathbf{Q})$, which is the quotient through which χ^{1-k} factors, acts on $H' := Gal(E/\mathbf{Q}(\mu_p)^{\otimes(k-1)})$ by χ^{1-k}. Because of (ii), H' is a group of $(p, \cdots, p)$ type. The extension $E/\mathbf{Q}(\mu_p)^{\otimes(k-1)}$ is unramified as (i) implies that it is unramified outside p, while (iii) implies that the primes above p split in the extension $E/\mathbf{Q}(\mu_p)^{\otimes(k-1)}$ as this is an extension of $(p, \cdots, p)$ type. Now if we define E to be the compositum of E' and $\mathbf{Q}(\mu_p)$, then as $\mathbf{Q}(\mu_p)$ and E' are linearly disjoint extensions of $\mathbf{Q}(\mu_p)^{\otimes(k-1)}$, the extension $E/\mathbf{Q}(\mu_p)$ has the desired properties.

2. Strategy of proof

It remains only to prove Theorem 2! The existence of the representation ρ is subtle as one wants a two dimensional mod p representation of the Galois group of $\mathbf{Q}$ that is *not* semisimple while its restriction to D is semisimple.

Ribet uses the 2 dimensional mod p representations that arise from the reduction mod p of the p-adic representation attached to cusp forms f that are eigenvectors for Hecke operators. These generally tend to be irreducible unless f is *congruent* to an Eisenstein series mod p.

In particular, consider the Eisenstein series

$$E_k = -B_k/2k + \Sigma_n \sigma_{k-1}(n) q^n,$$

where $q = e^{2\pi iz}$. As $p|B_k$, E_k mod p "looks like" a cusp form, as mod p it vanishes at infinity. In fact there is a cuspidal eigenform $f \in S_k(SL_2(\mathbf{Z}))$ that is "congruent" to E_k mod p. Using the form f, Ribet constructs the representation ρ. The properties (i) and (ii) are not hard to prove, but (iii) requires very delicate results from algebraic geometry ([Ra]).

In fact when Ribet worked out his results enough was not known about mod p representations coming from cuspforms of weight bigger than 2, and he was forced to work at weight 2 using (by now) well-known principles that "mod p everything is weight 2". But now because of recent results proven by Faltings, Jordan ([FJ]), and the theory of Fontaine-Laffaille ([FL]) it seems possible to work directly in higher weights k ($k \leq p-3$). We will give indications of how (i) and (ii) are proved below, and hand wave our way through (iii)!

3. The proof

3.1 Galois representations attached to cuspforms

Let $N \geq 1$, $k \geq 2$ be integers, $\epsilon : (\mathbf{Z}/N\mathbf{Z})^* \to \mathbf{C}^*$ a character. Then we consider the space of $S_k(\Gamma_0(N), \epsilon) \subset S_k(\Gamma_1(N))$ of cuspforms of weight k for the congruence subgroup

$$\Gamma_0(N) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z}) \mid c \equiv 0 \ (N),$$

with character ϵ. These are holomorphic functions f on the upper half-plane $\mathcal{H} = \{z \in \mathbf{C} | im(z) > 0\}$ such that

$$f(\frac{az+b}{cz+d}) = \epsilon(d)(cz+d)^k f(z)$$

for

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N),$$

with the condition that f vanishes at all cusps. This latter condition simply means that $f(\frac{az+b}{cz+d})$ tends to 0 whenever $im(z) \to \infty$ for any matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z}).$$

There is another useful algebraic way of looking at cuspforms f of weight 2. Consider the open Riemann surface $\mathcal{H}/\Gamma_1(N)$. This can be viewed as an affine curve and one can compactify it to get a projective curve $X_1(N)$. We can view this sitting in $\mathbf{P}^n$ for some n, and it is a theorem that the equations which define this curve can be chosen to be stable under the action of the Galois group of $\mathbf{Q}$. A remark for the experts: we will implicitly work with Shimura's canonical model for these modular curves (and denote them by $X_1(N)$) over $\mathbf{Q}$. As a Riemann surface we have the uniformisation

$$\pi : \mathcal{H} \cup \mathbf{P}^1(\mathbf{Q}) \to X_1(N).$$

If we consider a holomorphic differential ω on the curve $X_1(N)$ its pull-back $\pi^*(\omega)$ can be written as $f(z)dz$ (on the curve $X_1(N)$, ω looks like that only locally), and as $\alpha^*(dz) = d(\alpha(z)) = (cz+d)^{-2}dz$ with $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbf{Z})$, we see that f has the "right symmetries". As dz" $=$ "dq/q we see that as ω is a homolomorphic differential on $X_1(N)$, f is forced to be a cuspform. Thus we have an interpretation of the space of cuspforms of weight 2 as the space of holomorphic differentials etc, an interpretation that will be useful later. For higher weights there is also a similar interpretation except that we have to use differentials with values in non-trivial coefficient systems.

As the element

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_0(N)$$

we can develop f in a Fourier series

$$f = \Sigma_n a_n(f) q^n.$$

We have an action of Hecke operators T_n on the space $S_k(\Gamma_0(N), \epsilon)$: explicitly for most primes r the action of T_r is given by

$$f|T_r = \Sigma_n a_{nr}(f) q^n + \epsilon(r) r^{k-1} \Sigma_n a_n(f) q^{nr}.$$

The Hecke operators generate a commutative algebra that we shall denote by $h_1(N)$. Inside the space of cusp forms we also have the lattice of cusp forms with Fourier coefficients in $\mathbf{Z}[\epsilon]$: these are preserved by the Hecke operators, and the Hecke algebra acting on them is a free $\mathbf{Z}$-module of finite rank.

Consider $f \in S_k(\Gamma_0(N), \epsilon)$ which is an eigenform for almost all T_r's where r is a prime. One can then show that the Fourier expansion of f has coefficients which lie in a number field. Further for any automorphism $\sigma \in G_{\mathbf{Q}}$, $f^\sigma := \Sigma_n a_n(f)^\sigma q^n$ is in $S_k(\Gamma_0(N), \epsilon^\sigma)$. This is not evident, but follows from a cohomological interpretation of cusp forms. Be that as it may, the big result here is that associated to f, and any prime ℓ there is a representation

$$\rho_f : G_{\mathbf{Q}} \to GL_2(E),$$

where E is a finite extension of $\mathbf{Q}_\ell$, that is characterised upto semisimplification by:

1. ρ_f is unramified at almost all primes r
2. For almost all primes r the characteristic polynomial of $\rho_f(Frob_r)$ is $x^2 - a_r(f)x + \epsilon(r)r^{k-1}$

Note that amongst the excluded primes in the phrase "for almost all primes" is the prime ℓ. This result is due to Eichler, Shimura and Deligne.

It is not clear if the representation ρ_f is semisimple. But in fact we have:

Theorem 3 *The representation ρ_f is absolutely irreducible.*

Proof. We will prove this only for even weights k, though it is true even for odd weights. This follows from the Ramanujan bounds, proven by Deligne, that for an eigencuspform

$$|a_r(f)| \leq 2r^{(k-1)/2},$$

for almost all primes r. Now if the representation were reducible we would have that the semisimplification would be the sum of two ℓ-adic characters $\chi^r\epsilon_1$ and $\chi^s\epsilon_2$ for integers r, s (by results about Hecke characters in [S]), with χ the ℓ-adic cyclotomic character of $G_{\mathbf{Q}}$ giving the action on roots of unity whose order is some power of ℓ, and ϵ_i finite order characters of $G_{\mathbf{Q}}$. The cyclotomic character χ has the property that for all primes $t \neq \ell$, $\chi(Frob_t) = t$. Comparing determinant characters for ρ_f we deduce that $r + s = k - 1$ and r and s are unequal as k is even: this contradicts the Ramanujan bounds.

The representation ρ_f is continuous with respect to the profinite topology on $G_{\mathbf{Q}}$ (the open subgroups are the subgroups of finite index: the group is totally disconnected and compact), and the ℓ-adic topology on $GL_2(E)$. As the group $GL_2(\mathcal{O}_E)$, with $\mathcal{O}_E$ the ring of integers of E, is open, the inverse image of it under ρ_f is a subgroup H of finite index of $G_{\mathbf{Q}}$: thus H stabilises a lattice L' in E^2 under the action of ρ_f. If we let L be the sum of the translates of L' under the coset representaives of H in $G_{\mathbf{Q}}$, we see that $G_{\mathbf{Q}}$ stabilises L. With respect to a basis of $L = \mathcal{O}_E e_1 + \mathcal{O}_E e_2$, the representation takes values in $GL_2(\mathcal{O}_E)$. We can reduce this integral model of the representation modulo the maximal ideal of $\mathcal{O}_E$, to get a representation $\overline{\rho}_f : G_{\mathbf{Q}} \to GL_2(\mathbf{F})$, where $\mathbf{F}$ is a finite field of characteristic ℓ.

Note that there are many choices of lattices L which $G_{\mathbf{Q}}$ stabilises, and thus $\overline{\rho}_f$ depends on the choice of L. But we have the theorem of Brauer-Nesbitt:

Theorem 4 *The semisimplification $\overline{\rho}_f^{ss}$ of the reduction mod l of ρ_f is well-defined, i.e., does not depend on the choice of lattice.*

Now choose ℓ to be the prime p that we are interested in. Assume now that ρ_f is such that its reduction w.r.t. some lattice L is reducible. Then Theorem 4 implies that its reduction w.r.t. any lattice is reducible. The following proposition is crucial to Ribet's work:

Proposition 1 *Assume ρ_f is such that its reduction with respect to some lattice $L \subset E^2$ (and hence all lattices) stabilised by $G_{\mathbf{Q}}$ is reducible with semisimplification isomorphic to $\phi_1 \oplus \phi_2$, for ϕ_i characters of $G_{\mathbf{Q}} \to \overline{\mathbf{F}}_p^*$. Then there is a lattice L' such that the reduction of ρ_f with respect to L' is not semisimple and of the form $\begin{pmatrix} \phi_1 & * \\ 0 & \phi_2 \end{pmatrix}$, for a specific choice of ϕ_1 and ϕ_2.*

Proof. For the proof of this the crucial ingredient is Theorem 3: ρ_f is irreducible.

Let π be a uniformiser of $\mathcal{O}_E$. Note the conjugation formula

$$P \begin{pmatrix} a & \pi b \\ c & d \end{pmatrix} P^{-1} = \begin{pmatrix} a & b \\ \pi c & d \end{pmatrix},$$

where $P := \begin{pmatrix} 1 & 0 \\ 0 & \pi \end{pmatrix}$. Because of this we may assume that the reduction of a chosen integral model of the representation is of the form $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$.

We first claim that we can choose a $\rho_f(G_{\mathbf{Q}})$-lattice L so that the mod p representation is of the form $\begin{pmatrix} \phi_1 & * \\ 0 & \phi_2 \end{pmatrix}$, rather than $\begin{pmatrix} \phi_2 & * \\ 0 & \phi_1 \end{pmatrix}$. This follows from the conjugation formula

$$Q_k \begin{pmatrix} a & b \\ c & d \end{pmatrix} Q_k^{-1} = \begin{pmatrix} d & -c/\pi^k \\ -\pi^k b & a \end{pmatrix},$$

where $Q_k := \begin{pmatrix} 0 & 1 \\ -\pi^k & 0 \end{pmatrix}$. Choose k to be the lowest power of π which divides all the lower left-corner entries of $\rho_f(g)$ for all $g \in G_{\mathbf{Q}}$ in its matricial representation with respect to some lattice. As ρ_f is irreducible k is a non-negative integer, and because of the above assumption, k is positive. From this the claim follows. Now we *fix* a lattice L so that the reduction of ρ_f with respect to it is of the form $\begin{pmatrix} \phi_1 & * \\ 0 & \phi_2 \end{pmatrix}$.

For the sake of contradiction let us assume that the reduction of ρ_f with respect to all $G_{\mathbf{Q}}$-stable lattices is semisimple, i.e., is of the form $\begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$, where the $*$ are 1-dimensional characters. Choosing the integral model of ρ_f given by L, if we conjugate $\rho_f(G_{\mathbf{Q}}) \subset GL_2(\mathcal{O}_E)$ by a matrix M such that $M\rho_f(G_{\mathbf{Q}})M^{-1} \subset GL_2(\mathcal{O}_E)$ then its mod p reduction is again reducible and semisimple.

To get a contradiction we inductively define a converging sequence of matrices $M_i = \begin{pmatrix} 1 & t_i \\ 0 & 1 \end{pmatrix}$, such that that $M_i\rho_f(G_{\mathbf{Q}})M_i^{-1}$ consists of elements of $GL_2(\mathcal{O}_E)$ whose lower left entries are divisible by π and upper right entries are divisible by π^i. Then the limit $M = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ of the M_i's (with $lim_i t_i = t$) conjugates $\rho_f(G_{\mathbf{Q}})$ into the lower triangular subgroup of $GL_2(E)$.

The inductive hypothesis may be rephrased as: $P^iM_i\rho_f(G_{\mathbf{Q}})M_i^{-1}P^{-i}$ consists of integral matrices whose lower left corner is divisible by π^{i+1}. The reduction of this mod π is upper-triangular. As all reductions are assumed to be semisimple, and they can always be assumed to be upper triangular, there is a unipotent matrix $U = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}$ that diagonalises the reduction mod π of $P^iM_i\rho_f(G_{\mathbf{Q}})M_i^{-1}P^{-i}$. Thus $UP^iM_i\rho_f(G_{\mathbf{Q}})M_i^{-1}P^{-i}U^{-1}$ consists of matrices whose upper right corner is divisible by π while its lower left corner is still divisible by π^{i+1}. Thus $(P^{-i}UP^iM_i)\rho_f(G_{\mathbf{Q}})(P^{-i}UP^iM_i)^{-1}$ consists of integral matrices whose lower left corner is divisible by π and upper right corner entries are divisible by π^{i+1}. We can continue the induction by setting

$$M_{i+1} = P^{-i}UP^iM_i = \begin{pmatrix} 1 & t_i + \pi^i u \\ 0 & 1 \end{pmatrix},$$

and we are done after observing that conjugating by M_i does not change the order with which the characters ϕ_i appear on the diagonal.

3.2 Congruences between cuspforms and Eisenstein series

Consider the Eisenstein series

$$E_k = -B_k/2k + \Sigma_n \sigma_{k-1}(n)q^n.$$

This is a modular form for the group $SL_2(\mathbf{Z})$. We would like to prove that there is a cuspform $f \in S_k(SL_2(\mathbf{Z}))$ such that the Fourier expansion of f

has algebraic integers as coefficients and if we fix a place $\wp$ above p we have the congruence

$$a_r(f) \equiv \sigma_{k-1}(r)(\wp)$$

for almost all primes r. We may try to do this by trying to find a modular form E of weight k for $SL_2(\mathbf{Z})$ with integral Fourier coefficients such that its constant term is a unit at $\wp$, and then by considering $E_k - uB_k/2kE$, for a $\wp$-unit u, we can get a cuspform f, as $\mathcal{H}/SL_2(\mathbf{Z})$ has only one cusp. This is the procedure Ribet follows. But as Kirti Joshi has observed there is a simpler argument (a similar argument occurs in Section 2.2 of [S1]) as follows:

We consider a polynomial $f = \Sigma_{i=1}^{n} c_i \Delta^i$ for some n in the Δ function

$$\Delta := q\Pi(1-q^n)^{24} = \Sigma_{n\geq 1}\tau(n)q^n,$$

with no constant term, with coefficients $c_i = a_i E_4^{c_i} E_6^{d_i}$, with $a_i \in \mathbf{Z}$ and c_i, d_i non-negative integers, and such that $E_k - f \equiv 0 \; mod(\wp)$. Observe that the semigroup generated by $4c_i + 6d_i$, with c_i, d_i non-negative integers consists of all even integers greater than 4. We can find such a polynomial as the constant terms $-B_4/8$ and $-B_6/6$ of E_4 and E_6 are rational numbers with numerator 1, and the fact that, if a modular form of weight k on $SL_2(\mathbf{Z})$ has a sufficient number (roughly $k/12$) of its terms in its Fourier expansion divisible by p, then all its terms are divisible by p. This kind of argument also proves the Ramanujan congruence $\tau(n) \equiv \sigma_{11}(n) \pmod{691}$ (see also [S1]).

It is also possible, as is mentioned in passing in [R], to give a more pure-thought proof of this result along the following lines. Consider the reduction mod $\wp$ of $\overline{E}_k$. Now we can consider this as a differential form on the curve X (that is the projective line, and the compactification of $\mathcal{H}/SL_2(\mathbf{Z})$ by a point at infinity as a Riemann surface) but now with mod p coefficients. Further the "differential form" $\overline{E_k(dq/q)^{\otimes\frac{k}{2}}}$ has no poles and thus gives a holomorphic $k/2$-form on X with mod p coefficients. By general results of Mazur (the q-expansion principle) it follows that $\overline{E}_k$ can be regarded as a global section of $\Omega_{/\mathbf{F}}^{\otimes(k/2)}$ where $\Omega_{/\mathbf{F}}$ is the canonical sheaf with $\mathbf{F}$ coefficients. Now consider the map

$$H^0(X, \Omega_{/\mathcal{O}_E}^{\otimes(k/2)}) \to H^0(X, \Omega_{/\mathbf{F}}^{\otimes(k/2)}).$$

It is then well-known (Section 2.1.2 of [C]), that this map is surjective, assuming that $p \geq 5$.

Theorem 5 *If $p|B_k$ (for $2 < k \leq p-3$), then there is a cuspform $f \in S_k(SL_2(\mathbf{Z}))$ whose Fourier coefficients are algebraic and $f \equiv E_k \pmod{\wp}$.*

But this is not good enough for us as Galois representations are attached to eigencuspforms rather than just cuspforms. Thus we have to prove:

Corollary 1 *If $p|B_k$ (for $2 < k \leq p-3$), then there is an eigencuspform $f \in S_k(SL_2(\mathbf{Z}))$ whose Fourier coefficients are algebraic and such that $f \equiv E_k$ $(\mathrm{mod}\ \wp)$.*

Proof. This is a lemma due to Deligne-Serre. We consider the space $S_k(SL_2(\mathbf{Z}), \mathcal{O}_E) = S_k(SL_2(\mathbf{Z}), \mathbf{Z}) \otimes \mathcal{O}_E$ and its mod p reduction that we denote by $S_k(SL_2(\mathbf{Z}), \mathbf{F})$. We consider the $\mathcal{O}_E$ algebra h generated by the Hecke operators T_r for r prime to p: this is commutative and has no nilpotent elements. We have to show that any mod p eigenform of the Hecke algebra in the latter space lifts to an eigenform of the characteristic 0 (for E sufficiently large). A mod p eigenform gives a homomorphism $h \to \mathbf{F}$ and that corresponds to a maximal ideal m. Consider the set of minimal prime ideals contained in m: at least one of them does not contain p, as h is reduced, and this provides the characteristic 0 lift that we are after. Alternatively as h is reduced we can apply the going-up theorem to conclude.

Corollary 2 *If $p|B_k$ (for $2 < k \leq p-3$), then there is an eigencuspform $f \in S_k(SL_2(\mathbf{Z}))$ whose Fourier coefficients are algebraic and such that $a_r(f) \equiv \sigma_{k-1}(r) mod(\wp)$ for almost all primes r.*

3.3 The representation ρ

After the work of the previous section a representation ρ with the properties (i) and (ii) of Theorem 2 follow easily. Namely consider an eigencuspform f as of the previous corollary. Then the semisimplification of the reduction of the mod $\wp$ reduction of ρ_f and the semisimple representation $\tau := 1 \oplus \chi^{k-1}$ of $G_{\mathbf{Q}}$ have the same characteristic polynomials for the Frobenius elements at almost all primes r. This together with the Cebotarev density theorem and a theorem of Brauer-Nesbitt, Theorem 4, implies that the semisimplification of the reduction of ρ_f and τ are isomorphic. This together with Theorems 1 and 3 implies that there exists a ρ with the property (ii) of Theorem 2. The property (i) follows by a general fact as f is a cuspform of level 1 and the curve X has good reduction everywhere. But the property (iii) is tricky. As noted above we have to exclude information at p when we consider the

p-adic representation attached to f. But it is exactly information at p that one needs!

Ribet does this by reducing to weight 2 and uses results of Raynaud ([Ra]) on finite flat group schemes over finite extensions of $\mathbf{Z}_p$ with ramification less than $p-1$, to conclude that ρ restricted to D leaves stable 2 distinct lines, and hence is semisimple.

Working in higher weights results of the type proven by Faltings-Jordan ([FJ]), together with results of Fontaine-Laffaille ([FL]), will allow one to deduce property (iii) in a similar manner. The crucial point is that we have the trivial Galois module as a submodule of ρ while [FL] and [FJ] imply that the trivial module is a quotient.

This completes the brief sketch of Ribet's converse to Herbrand.

REFERENCES

[C] Carayol, H., *Formes modulaires et représentations galoisiennes avec valeurs dans un anneau local complet*, in *p-adic monodromy and the Birch and Swinnerton-Dyer conjecture*, Contemp. Math. 165 (1994), American Math Soc., 213–237.

[FJ] Faltings, Gerd; Jordan, Bruce W. *Crystalline cohomology and* $\mathrm{GL}(2,\mathbf{Q})$, Israel J. Math. 90 (1995), no. 1-3, 1–66.

[FL] Fontaine, Jean-Marc; Laffaille, Guy *Construction de représentations p-adiques*, Ann. Sci. Ecole Norm. Sup. (4) 15 (1982), no. 4, 547–608.

[L] Lang, Serge, *Cyclotomic fields I and II*, Combined second edition. With an appendix by Karl Rubin. Graduate Texts in Mathematics, 121. Springer-Verlag, New York-Berlin, 1990. xviii+433 pp.

[MW] Mazur, B.; Wiles, A., *Class fields of abelian extensions of* $\mathbf{Q}$, Invent. Math. 76 (1984), no. 2, 179–330

[Ra] Raynaud, Michel, *Schémas en groupes de type* $(p,\cdots,p)$, Bull. Soc. Math. France 102 (1974), 241–280.

[R] Ribet, Kenneth A., *A modular construction of unramified p-extensions of* $\mathbf{Q}(\zeta_p)$, Invent. Math. 34 (1976), no. 3, 151–162.

[S] Serre, Jean-Pierre, *Abelian ℓ-adic representations and elliptic curves*, With the collaboration of Willem Kuyk and John Labute. Revised reprint of the 1968 original. Research Notes in Mathematics, 7. A K Peters, Ltd., Wellesley, MA, 1998.

[S1] Serre, Jean-Pierre, *Un interprétation des congruences relatives á la fonction τ de Ramanujan*, Oeuvres III, no. 80, 498–511, Springer Verlag, 1984.

[W] Wiles, Andrew, *Modular curves and the class group of* $\mathbf{Q}(\mu_p)$, Invent. Math. 58 (1980), no. 1, 1–35.

Chandrashekhar Khare
School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India.
e-mail: shekhar@math.tifr.res.in

Vandiver's Conjecture via K-theory

Eknath Ghate

1. Introduction

The purpose of this note is to describe some recent work by Soulé [8] on Vandiver's conjecture[1] which uses K-theory.

Let us start by recalling the conjecture. The letter p will always denote an odd prime in what follows.

Conjecture 1 (Vandiver's Conjecture) *Let h^+ denote the class number of the maximal totally real subfield $\mathbb{Q}(\zeta_p)^+$ of $\mathbb{Q}(\zeta_p)$. Then $p \not| h^+$.*

At the outset, we should perhaps remind the reader that if p is a 'Vandiver prime', that is an odd prime for which Vandiver's conjecture holds, then much of the theory of the p^{th}-cyclotomic field becomes much 'easier'. For instance, for such p, the proof of the main conjecture is routine (see Theorem 10.16 of [10]).[2]

Here is another example, for which we will need some notation. Let

$$\omega : \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}) \xrightarrow{\sim} (\mathbb{Z}/p)^\times$$

denote the Teichmuller character. Recall that ω is the canonical character of $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ given by the formula

$$\zeta_p^\sigma = \zeta_p^{\omega(\sigma)},$$

for $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. Let σ_a denote the pre-image of $[a] \in (\mathbb{Z}/p)^\times$ under ω.

It will be convenient to regard ω as a p-adic object as follows. Note that $(\mathbb{Z}/p)^\times$ is isomorphic to μ_{p-1}, the group of $(p-1)^{\text{st}}$ roots of 1. By Hensel's lemma, $\mu_{p-1} \subset \mathbb{Z}_p^\times$. Thus we may regard ω as a character

$$\omega : \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q}) \to \mathbb{Z}_p^\times \subset \mathbb{Z}_p.$$

[1] Although Conjecture 1 is attributed to Vandiver, it apparently was already stated by Kummer in a letter to Kronecker in the middle of the 19th century (see the Remark on page 158 of [10]).

[2] The main conjecture was established independently of Vandiver's conjecture by Mazur and Wiles [5] by studying the reductions of modular curves. An alternative proof was given by Kolyvagin and Rubin (see [7], or Chapter 15 of [10]), using the more elementary, but ingenious, method of Euler systems.

Let A be the p-Sylow subgroup of the ideal class group of $\mathbb{Q}(\zeta_p)$. By using a system of orthogonal idempotents of $\mathbb{Z}_p[\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})]$, we may decompose A into 'eigenspaces' for the natural action of the Galois group $\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$ on A (see section 6.3 of [10]):

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

with $A_i = \{a \in A \,\big|\, \sigma(a) = \omega^i(\sigma)a, \text{ for all } \sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})\}$.

Let $B_k \in \mathbb{Q}$ denote the k^{th} Bernoulli number, and v_p denote the normalized p-adic valuation of $\mathbb{Q}_p$, with $v_p(p) = 1$. We have[3]:

Theorem 1 *(Herbrand-Ribet) Let i be an odd integer with $1 \leq i \leq p-2$. Then*

$$A_i \neq 0 \iff v_p(B_{p-i}) > 0.$$

For i as in Theorem 1 above, we have (see Corollary 5.15 of [10]):

$$B_{1,\omega^{-i}} \equiv \frac{B_{p-i}}{p-i} \pmod{p},$$

where[4]

$$B_{1,\omega^{-i}} := \frac{1}{p}\sum_{a=1}^{p-1} a\omega^{-i}(\sigma_a) \in \mathbb{Z}_p.$$

Thus the following theorem is a refinement of the Herbrand-Ribet theorem[5]:

Theorem 2 *(Mazur-Wiles) Let i be an odd integer with $1 \leq i \leq p-2$. Then*

$$\mathrm{card}(A_i) = p^{m_i},$$

where $m_i = v_p(B_{1,\,\omega^{-i}})$.

However, even more is conjectured to be true:

Conjecture 2 *(Iwasawa) When i is odd, $A_i \xrightarrow{\sim} \mathbb{Z}/p^{m_i}$ is cyclic.*

[3] See the articles of Katre [2] and Khare [3] in these proceedings, as well as [10], for various proofs of Theorem 1.

[4] The fact that $B_{1,\omega^{-i}}$ lies in $\mathbb{Z}_p$ and not just $\mathbb{Q}_p$ is forced on us by Theorem 2.

[5] Theorem 2 is a consequence of the main conjecture.

As it turns out (see Corollary 10.15 of [10]), Iwasawa's conjecture is true when p is a Vandiver prime.

The above discussion shows that it is more than simply a matter of curiosity to investigate the validity of Vandiver's conjecture. Numerically, it has been checked that all $p \leq 4 \times 10^6$ are Vandiver primes. However, apparently, this is not sufficient evidence for one to believe that Vandiver's conjecture holds for *all* p. Indeed, a heuristic argument of Washington (see the Remark on page 158 of [10]) shows that the exceptions to Vandiver's conjecture are very rare: the number of exceptions one expects in the range $3 \leq p \leq 4 \times 10^6$ is only 1.36....!

Let us now rephrase Conjecture 1 in a form that will render it more manageable. It is an exercise to check that it is equivalent to the following:

Conjecture 3 (Vandiver's Conjecture) *Let p be an odd prime. Then $A_i = 0$, for all even integers i with $0 \leq i \leq p-3$.*

To place things in context, let us recall that it is well known that $A_0 = A_1 = 0$ and that, moreover, when i is odd, $A_i = 0 \iff p \nmid B_{p-i}$ (Herbrand-Ribet theorem). Thus Vandiver's Conjecture says that, on the other hand, when i is even, A_i always vanishes!

As mentioned already, in this note we would like to describe recent work by Soulé on Vandiver's conjecture which uses K-theory.

The story starts with a pretty result of Kurihara [4], who proved that the 'top' even eigenspace always vanishes:

Theorem 3 *(Kurihara) $A_{p-3} = 0$.*

The idea of Kurihara's proof is to note that there is a surjective map

$$K_4(\mathbb{Z}) \otimes \mathbb{Z}/p \to A_{p-3}, \tag{1}$$

and that $K_4(\mathbb{Z})$ is not too big.[6] Last year Soulé [8] was able to extend Kurihara's result. He showed that if n is small (and odd) compared to p then $A_{p-n} = 0$. More precisely, he showed:

Theorem 4 *(Soulé) Assume $n > 1$ is odd. If $\log p > n^{224n^4}$, then*

$$A_{p-n} = 0.$$

[6] At the time that Kurihara wrote [4] it was known that $K_4(\mathbb{Z})$ is a finite abelian group whose p primary components were 0, for $p \neq 2, 3$. This was enough to deduce Theorem 3. However recently Rognes has shown that in fact $K_4(\mathbb{Z}) = 0$.

The basic idea of Soulé's proof is very similar to Kurihara's. He notes that the 'Chern map' (of which (1) above is a special case):

$$K_{2n-2}(\mathbb{Z}) \otimes \mathbb{Z}/p \to A_{p-n} \tag{2}$$

is surjective. On the other hand the finite abelian group $K_{2n-2}(\mathbb{Z})$, is (essentially) the $(2n-2)^{\text{th}}$ homology group of $SL_N(\mathbb{Z})$ for N large. Classical Voronoi 'reduction theory' gives an explicit cell decomposition of the compactification of the locally symmetric space attached to $SL_N(\mathbb{Z})$. With this in hand, Soulé now implements the following simple remark of Gabber: one may bound the torsion in the homology of a finite CW-complex X purely in terms of data associated with the cellular chain complex $C.(X)$ of X, such as the number of cells of a fixed degree and the number of faces of each cell. This yields an explicit upper bound for the primes p dividing the order of $K_{2n-2}(\mathbb{Z})$: this is the bound that appears in the statement of Theorem 4 above.

Note that because of the inherent surjectivity of the map (2), the Soulé-Kurihara method has natural limitations: one can only expect it to yield rough results such as Theorem 4 above. On the other hand, as far as we are aware, Theorems 3 and 4 are really the first results towards Vandiver's conjecture of a general nature.

I would like to thank Dinesh Thakur for encouraging me to write up these notes, and V. Srinivas for his comments on a first draft.

2. A quick introduction to K-theory

Let R be a commutative ring with 1. In this section we will introduce the K-groups $K_i(R)$ $(i \geq 0)$ attached to R, and describe some of their properties when R is the ring of integers of a number field. References for some of the material described here are Srinivas' book [9] (especially Chapters 1 and 2), and Rosenberg's book [6].

2.1 $K_0(R)$

Here we simply recall the definition of $K_0(R)$. Let $\mathcal{F}$ denote the free abelian group on isomorphism classes of projective R-modules, and let $\mathcal{R}$ denote the subgroup generated by the elements

$$[P \oplus Q] - [P] - [Q],$$

where P and Q are projective R-modules, and [ ] denotes an isomorphism class. Then we set

$$K_0(R) = \mathcal{F}/\mathcal{R}.$$

2.2 Classifying spaces

To introduce the higher K-groups we will need the notion of a classifying space of a discrete group G, which we introduce now.

It is a fact that if G is a group, regarded as a discrete topological group, then there exists a contractible CW-complex X on which G acts freely and cellularly (so properly discontinuously), so that the quotient X/G is a CW-complex (see Theorem 5.1.15 of [6] for an explicit construction of X). We now make the:

Definition 1 *The* classifying space of G *is the quotient space* $BG := X/G$.

It is a fact that BG is well defined up to homotopy equivalence (Theorem 5.1.5. of [6]). Also BG is a $K(G,1)$-space (see Corollary 5.1.25 of [6]). That is, it is a connected space with

$$\pi_1(BG,x) = G \text{ and } \pi_m(BG,x) = 0, \text{ for } m > 1.$$

Here x is a base point, which we will drop from the subsequent notation.

2.3 The plus construction

Let $\mathrm{GL}_n(R)$ denote the ring of invertible $n \times n$ matrices with entries in R. Then $\mathrm{GL}_n(R) \subset \mathrm{GL}_{n+1}(R)$ via the embedding

$$A \mapsto \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}.$$

Let $\mathrm{GL}(R) = \lim_{n\to\infty} \mathrm{GL}_n(R)$, where the limit is taken with respect to these embeddings.

Regard $\mathrm{GL}(R)$ as a topological group with the discrete topology, and let $\mathrm{BGL}(R)$ denote the classifying space of $\mathrm{GL}(R)$. Recall that $\mathrm{BGL}(R)$ is a $K(\mathrm{GL}(R),1)$ space: i.e. $\mathrm{BGL}(R)$ is a connected space with $\pi_1(\mathrm{BGL}(R)) = \mathrm{GL}(R)$, and $\pi_m(\mathrm{BGL}(R)) = 0$, for $m > 0$.

Now one constructs another space $\mathrm{BGL}(R)^+$ from $\mathrm{BGL}(R)$ by attaching two and three cells. This process is called the plus construction, and is described on page 19 of [9]. Here we will be content in describing some of the properties of $\mathrm{BGL}(R)^+$, that we summarize in the following theorem:

Theorem 5 *1. Let* $\mathrm{E}(R)$ *denote the subgroup of* $\mathrm{GL}(R)$ *generated by the elementary matrices (at finite level, these matrices are just* $n \times n$ *matrices with diagonal entries equal to 1 and at most one non-zero off-diagonal entry). Then* $\mathrm{E}(R)$ *is the commutator subgroup of* $\mathrm{GL}(R)$, *and is a perfect normal subgroup of* $\mathrm{GL}(R)$. *Moreover,*

$$\pi_1(\mathrm{BGL}(R)^+) = \mathrm{GL}(R)/\mathrm{E}(R).$$

2. For each $m \geq 0$ we have

$$\mathrm{H}_m(\mathrm{BGL}(R)^+, \mathbb{Z}) = \mathrm{H}_m(\mathrm{BGL}(R), \mathbb{Z}) = \mathrm{H}_m(\mathrm{GL}(R), \mathbb{Z}).$$

2.4 Higher K-groups

We may now give (one of) Quillen's definition's of the higher K-groups of R:

Definition 2 *For each $m \geq 1$, set $K_m(R) := \pi_m(\mathrm{BGL}(R)^+)$.*

We note that in particular $K_m(R)$ is an abelian group for $m \geq 0$.

2.5 K-theory of rings of integers

Let F be a number field, and let $\mathcal{O}_F$ denote the ring of integers of F. The next theorem shows that one might expect that the higher K-groups of $\mathcal{O}_F$ should contain much interesting information about F:

Theorem 6 *1. $K_0(\mathcal{O}_F) = \mathbb{Z} \oplus Cl(F)$, where $Cl(F)$ denote the class group of F.*

2. $K_1(\mathcal{O}_F) = \mathcal{O}_F^\times$, the group of units of $\mathcal{O}_F$.

Quillen had shown that, in general, the abelian groups $K_m(\mathcal{O}_F)$ are finitely generated. Their ranks were subsequently computed by Borel [1]:

Theorem 7 *(Borel) Let r_1 (respectively r_2) denote the number of embeddings of K into $\mathbb{R}$ (respectively $\mathbb{C}$). Then the ranks of $K_m(\mathcal{O}_F)$ are as follows:*

$$\mathrm{rk}(K_0(\mathcal{O}_F)) = 1, \quad \mathrm{rk}(K_1(\mathcal{O}_F)) = r_1 + r_2 - 1 \quad \text{and}$$

$$\mathrm{rk}(K_m(\mathcal{O}_F)) = \begin{cases} r_1 + r_2 & \text{if } m = 4i+1 > 1, \\ r_2 & \text{if } m = 4i+3 > 1, \\ 0 & \text{if } m = 2i. \end{cases}$$

On the other hand, almost nothing is known about the torsion subgroups of $K_m(\mathcal{O}_F)$. The following theorems summarizes our current state of ignorance when $F = \mathbb{Q}$.

Theorem 8 *The K-theory of $\mathbb{Z}$ computed to date is: $K_0(\mathbb{Z}) = \mathbb{Z}$, $K_1(\mathbb{Z}) = \mathbb{Z}/2$, $K_2(\mathbb{Z}) = \mathbb{Z}/2$, $K_3(\mathbb{Z}) = \mathbb{Z}/48$, and $K_4(\mathbb{Z}) = 0$.*

2.6 The Hurewicz map

For computational purposes, we will need the Hurewicz maps ($m \geq 1$)

$$\text{Hurewicz} : \pi_m(X) \to \mathrm{H}_m(X, \mathbb{Z}),$$

which are homomorphisms from the homotopy groups of a CW-complex X, to the homology groups of X. Roughly, they are defined as follows (see Appendix A of [9] for further details). A typical element $[f]$ of $\pi_m(X)$ is a homotopy class of a continuous map $f : S_m \to X$, where S_m is the m-dimensional sphere. We have an induced map

$$\mathrm{H}_m(f) : \mathrm{H}_m(S_m, \mathbb{Z}) \to \mathrm{H}_m(X, \mathbb{Z}),$$

and we set Hurewicz$([f]) = \mathrm{H}_m(f)(\omega)$ where ω is the standard generator (corresponding to a choice of orientation) of $\mathrm{H}_m(S_m, \mathbb{Z}) = \mathbb{Z}$.

It is not true in general that the Hurewicz maps are isomorphisms, though this does hold for $m \geq 2$ when X is $(m-1)$-connected, that is, when $\pi_j(X) = 0$ for $j \leq m-1$. When $m = 1$, and X is 0-connected, that is when X is connected, the kernel of the Hurewicz map is just the commutator subgroup of $\pi_1(X)$, and in this case the Hurewicz map gives an explicit isomorphism $\pi_1(X)^{\text{ab}} = \mathrm{H}_1(X, \mathbb{Z})$.

In our situation the Hurewicz map is a homomorphism (cf. Theorem 5):

$$\text{Hurewicz} : K_m(\mathcal{O}_F) = \pi_m(\mathrm{BGL}(\mathcal{O}_F)^+) \to \mathrm{H}_m(\mathrm{GL}(\mathcal{O}_F), \mathbb{Z}).$$

This map is not injective[7], but when $F = \mathbb{Q}$ we have the following (see the remarks in Section 2.5 of [8] and the references there):

Proposition 1 *The kernel of*

$$\text{Hurewicz} : K_m(\mathbb{Z}) \to \mathrm{H}_m(\mathrm{GL}(\mathbb{Z}), \mathbb{Z})$$

is a finite abelian group, with non-zero p-primary components only for p smaller than the integral part of $(m+1)/2$.

[7] Srinivas has remarked that the kernel of the Hurewicz map for $\mathrm{BGL}(R)^+$ is always a torsion group. In fact it is a theorem of Milnor and Moore that the kernel of Hurewicz is torsion when X is an H-space. We refer the reader to Appendix A of [9] for the definition and properties of H-spaces.

3. The Chern map

We now show how the map (2) is constructed. Unfortunately, we will have to be somewhat brief since we ourselves do not understand some of the details.

Let X be a scheme over $\mathbb{Z}[\frac{1}{p}]$, and let

$$\mathrm{H}^k_{\mathrm{et}}(X, \mathbb{Z}_p(n)) \tag{3}$$

denote the étale cohomology groups of X with coefficients in the n^{th} Tate twist of the group of p-adic integers.

Let us explain what we mean by this in the situation that matters to us, namely when $X = \mathrm{Spec}(\mathbb{Z}[1/p])$. We need some notation. Let $\mathbb{Q}^{p,\infty}$ denote the maximal extension of $\mathbb{Q}$ unramified outside p and ∞. Let ϵ denote the cyclotomic character

$$\epsilon : \mathrm{Gal}(\mathbb{Q}^{p,\infty}/\mathbb{Q}) \twoheadrightarrow \mathrm{Gal}(\mathbb{Q}(\mu_{p^\infty})/\mathbb{Q}) \xrightarrow{\sim} \mathbb{Z}_p^\times \subset \mathbb{Z}_p.$$

Then $\mathbb{Z}_p(n)$ is the $\mathrm{Gal}(\mathbb{Q}^{p,\infty}/\mathbb{Q})$-module $\mathbb{Z}_p$ with action:

$$g \cdot a = \epsilon(g)^n a,$$

where $g \in \mathrm{Gal}(\mathbb{Q}^{p,\infty}/\mathbb{Q})$ and $a \in \mathbb{Z}_p$. Then, when $X = \mathbb{Z}[1/p]$, the group (3) above is nothing but the continuous Galois cohomology group

$$\mathrm{H}^k(\mathrm{Gal}(\mathbb{Q}^{p,\infty}/\mathbb{Q}), \mathbb{Z}_p(n)).$$

One may also speak of the K-theory of the scheme X. The exact definition[8] does not concern us here, since when $X = \mathrm{Spec}(R)$ is affine (as a scheme over $\mathrm{Spec}(\mathbb{Z}[1/p])$, then $K_m(X) = K_m(R)$. There is also the notion of the étale K-groups, $K^{\mathrm{et}}_m(X)$, of X, whose definition I don't know. However, I do know that these are $\mathbb{Z}_p$-modules which come equipped with maps

$$K_m(X) \otimes \mathbb{Z}_p \to K^{\mathrm{et}}_m(X). \tag{4}$$

Finally Dwyer and Friedlander have shown that these gadgets are connected by an Atiyah-Hirzebruch type[9] spectral sequence:

$$E_2^{rs} = \mathrm{H}^r_{\mathrm{et}}(X, \mathbb{Z}_p(-s/2)) \Longrightarrow K^{\mathrm{et}}_{-r-s}(X),$$

[8] See Chapters 3 and 4 of [9] if you are interested in the definition!

[9] The name is because it is the exact analog of Atiyah-Hirzebruch spectral sequence connecting the singular homology of a space X with the topological K-theory of X.

or re-indexing (set $r = k$, and $s = -2n$)

$$E_2^{kn} = \mathrm{H}^k_{\mathrm{et}}(X, \mathbb{Z}_p(n)) \Longrightarrow K^{\mathrm{et}}_{2n-k}(X).$$

We now have the following

Theorem 9 *Say $X = \mathbb{Z}[1/p]$. Assume that $n > 0$ and p is odd. Then*

$$\mathrm{H}^k_{\mathrm{et}}(X, \mathbb{Z}_p(n)) = 0,$$

unless $k = 1$ or 2.

Now set $m = 2n-k$. Then Theorem 9 shows that when $X = \mathrm{Spec}(\mathbb{Z}[1/p])$, the spectral sequence above degenerates, and so there are surjective maps:

$$K^{\mathrm{et}}_m(X) \twoheadrightarrow \mathrm{H}^k_{\mathrm{et}}(X, \mathbb{Z}_p(n)), \tag{5}$$

for $m = 2n-1$ or $2n-2$.

We are interested in the case when $k = 2$, and $n > 1$ is odd. In this case $m = 2n-2$, and in particular m is even. By Theorem 7, we see that $K_m(\mathbb{Z})$ is a finite abelian group. Also we have the following (see Section 1 of [4]):

Proposition 2 *Let $X = \mathrm{Spec}(\mathbb{Z}[1/p])$. Suppose $n > 1$ is odd. Then*

$$\mathrm{H}^2_{\mathrm{et}}(X, \mathbb{Z}_p(n)) \otimes \mathbb{Z}/p = A_{p-n}.$$

Combining the natural maps

$$K_{2n-2}(\mathbb{Z}) \to K_{2n-2}(\mathbb{Z}[1/p]) \to K_{2n-2}(\mathbb{Z}[1/p]) \otimes \mathbb{Z}_p$$

with the maps (4), (5), and the above proposition, we get a map

$$K_{2n-2}(\mathbb{Z}) \otimes \mathbb{Z}/p \to A_{p-n}.$$

It is (apparently) a fact (due to Soulé and Dwyer-Friedlander) that this map is surjective, and this is the map (2) that we have called the Chern map in the Introduction.

4. Voronoi's reduction theory Fix $N \geq 2$. Let V_N denote the space of $N \times N$ real symmetric matrices.

Recall that a symmetric matrix A is called *positive semi-definite* if $vAv^t \geq 0$, for all v, and is called *positive definite* if in addition $vAv^t = 0 \iff v = 0$.

Let P_N denote the subset of V_N of all positive definite symmetric matrices. Note that $\mathbb{R}_+^\times$ acts on V_N by scalar multiplication. Set $X_N = P_N/\mathbb{R}_+^\times$. Then $X_N = SL_N(\mathbb{R})/SO_N(\mathbb{R})$ is the symmetric space for $SL_N(\mathbb{R})$.

Let P_N^* denote the subset of V_N of all symmetric positive semi-definite matrices, with rational null-space (that is $\ker(A)$ is spanned by vectors in $\mathbb{Q}^N$). Set $X_N^* = P_N^*/\mathbb{R}_+^\times$. We have the following commutative diagram of spaces:

$$\begin{tikzcd}
P_N \arrow[r, phantom, "\subset"] \arrow[d] & P_N^* \arrow[d, "\pi"] \\
X_N \arrow[r, phantom, "\subset"] & X_N^*,
\end{tikzcd}$$

where π denotes the projection map.

Now $SL_N(\mathbb{Z})$ acts on P_N^* as follows:

$$g \cdot A = gAg^t,$$

where $g \in SL_N(\mathbb{Z})$ and $A \in P_N^*$. P_N is clearly preserved under this action. Set $Y_N = X_N/SL_N(\mathbb{Z})$ and $Y_N^* = X_N^*/SL_N(\mathbb{Z})$.

Definition 3 *Let $A \in P_N$. Set*

$$\begin{aligned}
\mu(A) &:= \min\{vAv^t \mid v \in \mathbb{Z}^N \subset \mathbb{R}^N\}, \\
m(A) &:= \{v \in \mathbb{Z}^N \setminus 0 \mid vAv^t = \mu(A)\}.
\end{aligned}$$

Definition 4 *Let $A \in P_N$. Then say A is* perfect *if $\mu(A) = 1$, and if whenever $B \in P_N$ with $\mu(B) = 1$ and $m(A) = m(B)$, then $B = A$.*

Note that each element $v \in \mathbb{Z}^N \setminus 0$ determines an element $\hat{v} = v^t v \in P_N^*$.

Definition 5 *Given any finite subset $B \subset \mathbb{Z}^N \setminus 0$, the* convex hull of B *is the set $\pi\left(\{\sum_j \lambda_j \hat{v}_j \mid v_j \in B, \lambda_j \geq 0\}\right)$.*

When A is perfect, let $\sigma(A)$ denote the convex hull of $m(A)$. We may now state the main theorem of Voronoi reduction theory:

Theorem 10 *(Voronoi)*

1. *Up to conjugation by $SL_N(\mathbb{Z})$, there are only finitely many perfect forms.*

2. *The cells $\sigma(A)$ and their intersections, as A varies through the set of perfect forms, gives a cell decomposition of X^*_N, invariant under $SL_N(\mathbb{Z})$.*

The above theorem says that the space $Y^*_N = X^*_N/\mathrm{SL}_N(\mathbb{Z})$ is a finite CW-complex. Soulé has computed explicit upper bounds for the number of cells of a fixed dimension, and the number of faces of such cells:

Proposition 3 *There exist explicit constants $c(k, N)$ and $f(k, N)$ such that*

1. *The number of $SL_N(\mathbb{Z})$-conjugacy classes of k-dimensional cells in the Voronoi cell decomposition of X^*_N is bounded by $c(k, N)$, and,*
2. *Any k-dimensional cell has at most $f(k, N)$ faces.*

Proof: The proof is easy: we refer the reader to Propositions 1 and 2 of [8] for details.

5. A key Lemma

The following simple lemma (it is a good exercise to try and prove it for yourself) is really at the heart of the whole proof:

Lemma 1 *Let $\phi : \mathbb{Z}^a \to \mathbb{Z}^b$ be a $\mathbb{Z}$-linear map. Let $Q = \mathrm{coker}(\phi)$. Let $\{e_i \,\big|\, 1 \le i \le a\}$ denote the standard basis of $\mathbb{Z}^a$, and let $I \subset \{1, \dots, a\}$ be such that $\{\phi(e_i) \,\big|\, i \in I\}$ is a basis for* $\mathrm{image}(\phi) \otimes \mathbb{R}$. *Then*

$$\mathrm{card}(Q_{\mathrm{tors}}) \le \prod_{i \in I} ||\phi(e_i)||.$$

Now let X be a finite CW-complex. Let $(C_{.}(X), \partial_{.})$ denote its cellular chain complex. Recall that $C_k(X)$ is a free $\mathbb{Z}$-module of finite rank, with basis, say, Σ_k, and that there are boundary maps

$$\partial_{k+1} : C_{k+1}(X) \to C_k(X). \tag{6}$$

Suppose that

$$\partial_{k+1}(\sigma) = \sum_{\sigma' \in \Sigma_k} n_{\sigma\sigma'}\, \sigma',$$

for $\sigma \in \Sigma_{k+1}$. Set

$$\begin{aligned} a(k) &= \min\left(\mathrm{card}(\Sigma_{k+1}), \mathrm{card}(\Sigma_k)\right), \\ b(k) &= \max\left(1, \max_{\sigma \in \Sigma_{k+1}} \left(\sum_{\sigma' \in \Sigma_k} n^2_{\sigma\sigma'}\right)^{1/2}\right). \end{aligned}$$

The following corollary follows immediately from Lemma 1:

Corollary 1 *(Gabber) We have*

$$\operatorname{card}(\mathrm{H}_k(C.,\partial.)_{\mathrm{tors}}) \leq b(k)^{a(k)}.$$

Proof: Note that

$$\mathrm{H}_k(C.,\partial.) = \frac{\ker(\partial_k)}{\operatorname{image}(\partial_{k+1})} \subset Q = \operatorname{coker}(\partial_{k+1}).$$

Also $||\partial_{k+1}(\sigma)|| \leq b(k)$. Thus

$$\operatorname{card}(\mathrm{H}_k(C.,\partial.)_{\mathrm{tors}}) \leq \operatorname{card}(Q_{\mathrm{tors}}) \leq b(k)^{a(k)}.$$

Let us apply this corollary in our situation: namely when $X = Y_N^*$. We choose our basis set Σ_k to be a set of representatives of the conjugacy classes of k-dimensional cells in the Voronoi cell decomposition. Note that if $\sigma \in \Sigma_{k+1}$, the absolute values $|n_{\sigma\sigma'}|$ of the integers appearing in (6) above are at most the number of faces τ of σ which are conjugate to σ'. So

$$\left(\sum_{\sigma' \in \Sigma_k} n_{\sigma\sigma'}^2\right)^{1/2} \leq \sum_{\sigma' \in \Sigma_k} |n_{\sigma\sigma'}| \leq f(k+1,N),$$

by Proposition 3. The same proposition shows that

$$\operatorname{card}(\Sigma_k) \leq c(k,N).$$

Consequently, in our situation, we may rephrase Corollary 1 as:

Corollary 2 *We have*

$$\operatorname{card}(\mathrm{H}_k(Y_N^*,\mathbb{Z})_{\mathrm{tors}}) \leq f(k+1,N)^{c(k+1,N)}.$$

6. Proof of Theorem 4

In this section, we tie together the previous sections and give a sketch of the proof of Theorem 4.

Let p be an odd prime, and let $n > 1$ be odd. Recall that the map (2)

$$K_{2n-2}(\mathbb{Z}) \otimes \mathbb{Z}/p \to A_{p-n}$$

is surjective, and that, moreover, the abelian group $K_{2n-2}(\mathbb{Z})$ is finite. We would therefore like to get a bound for the torsion in $K_{2n-2}(\mathbb{Z})$.

The Hurewicz map (see Section 2.6)

$$\text{Hurewicz} : K_m(\mathbb{Z}) \to \mathrm{H}_m(\mathrm{GL}(\mathbb{Z}), \mathbb{Z}).$$

converts our task into a question of computing homology groups (easy), rather than computing homotopy groups (more difficult). Indeed modulo 'small primes' (cf. Proposition 1 above), which we can ignore, the Hurewicz map is in fact injective.

Moreover, it is a fact that the homology groups of $\mathrm{GL}(\mathbb{Z})$ are 'stable' (see the references in Section 2.5 of [8]). Thus we have:

$$\mathrm{H}_m(\mathrm{GL}(\mathbb{Z}), \mathbb{Z}) = \mathrm{H}_m(\mathrm{GL}_N(\mathbb{Z}), \mathbb{Z}),$$

for N large, in fact for $N \geq 2m + 1$.

On the other hand there is an exact sequence

$$1 \longrightarrow \mathrm{SL}_N(\mathbb{Z}) \longrightarrow \mathrm{GL}_N(\mathbb{Z}) \xrightarrow{\det} \{\pm 1\} \to 1,$$

and so a simple application of the Hochschild-Serre spectral sequence shows that up to a power of 2 (which again we can ignore) we have

$$\mathrm{card}(\mathrm{H}_m(\mathrm{GL}_N(\mathbb{Z}), \mathbb{Z})) = \mathrm{card}(\mathrm{H}_m(\mathrm{SL}_N(\mathbb{Z}), \mathbb{Z})).$$

Now, modulo some more small primes, we have

$$\mathrm{card}(\mathrm{H}_m(\mathrm{SL}_N(\mathbb{Z}), \mathbb{Z})) = \mathrm{card}(\mathrm{H}_m(Y_N, \mathbb{Z})). \tag{7}$$

This would have been an exact equality, except for the fact that $\mathrm{SL}_N(\mathbb{Z})$ has some elements of finite order. By passing to a torsion-free normal subgroup Γ of $\mathrm{SL}_N(\mathbb{Z})$ of finite index (divisible by exactly the primes which divide the cardinalities of the elements of finite order in $\mathrm{SL}_N(\mathbb{Z})$), and noting that the analog of (7) holds for Γ, we may deduce (7) itself, by 'taking invariants'.

But, by Corollary 2 we have

$$\mathrm{card}(\mathrm{H}_m(Y_N^*, \mathbb{Z})_{\mathrm{tors}}) \leq f(m+1, N)^{c(m+1,N)}.$$

Now a technical argument (see [8], proof of Theorem 1) allows us to deduce that the cardinality of the homology of the non-compact space $Y_N \subset Y_N^*$ may also be bounded explicitly. Thus there is a constant $A(m, N)$, related to $f(m+1, N)^{c(m+1,N)}$, such that

$$\mathrm{card}(\mathrm{H}_m(Y_N, \mathbb{Z})) \leq A(m, N).$$

An explicit computation of $A(m,N)$ for $N = 2m+1$ and $m = 2n-2$ (see [8], Lemma 2)) shows that if p is large compared to n, namely $p > \exp(n^{224n^4})$, then $p \not| \operatorname{card}(\mathrm{H}_m(Y_N, \mathbb{Z}))$. By the remarks above, and the surjectivity of the map (2), we have:

$$A_{p-n} = 0,$$

for such p. This 'finishes' the proof of Theorem 4.

REFERENCES

1. A. Borel, *Stable real cohomology of arithmetic groups*, Ann. Scient. Ec. Norm. Sup., **7**, 235-272, 1974.
2. S. A. Katre, *Gauss-Jacobi Sums and Stickelberger's Theorem*, This volume.
3. C. Khare, *Notes on Ribet's converse to Herbrand*, This volume.
4. M. Kurihara, *Some remarks on conjectures about cyclotomic fields and K-groups of Z*, Compositio Math., **81**, 223-236, 1992.
5. B. Mazur and A. Wiles, *Class fields of abelian extensions of Q*, Invent. Math, **76**, 179-330, 1984.
6. J. Rosenberg, *Algebraic K-Theory and Its Applications*, GTM,bf 147, Springer-Verlag, Berlin-New York, 1994.
7. K. Rubin, *Appendix to Cyclotomic fields I and II, by Serge Lang*, GTM, **121**, Springer-Verlag, Berlin-New York, 1990.
8. C. Soulé, *Perfect forms and Vandiver's conjecture*, Preprint, http://www.math.uiuc.edu/K-theory, 1998.
9. V. Srinivas, *Algebraic K-theory*, Second Edition, Progress in Mathematics, **90**, Birkhäuser, 1993.
10. L. Washington, *Introduction to cyclotomic fields*, Second edition, Springer-Verlag, Berlin-New York, 1996.

Eknath Ghate
School of Mathematics, Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India.
e-mail: eghate@math.tifr.res.in

A ‘universal’ Torsor for a Finite Group

Nitin Nitsure

Abstract

Let n be a positive integer, let $\mathbf{G}_{m,\mathbb{Z}[1/n]} = \operatorname{Spec} \mathbb{Z}[1/n][t, t^{-1}]$ be the multiplicative group scheme over $\mathbb{Z}[1/n]$, and let $(\,)^n : \mathbf{G}_{m,\mathbb{Z}[1/n]} \to \mathbf{G}_{m,\mathbb{Z}[1/n]}$ be the n th power morphism. The Hilbert theorem 90 implies that this morphism $(\,)^n : \mathbf{G}_{m,\mathbb{Z}[1/n]} \to \mathbf{G}_{m,\mathbb{Z}[1/n]}$ has the following property: Given any field K such that $char(K)$ does not divide n and K contains a primitive n th root of unity, any field extension L/K which is Galois with Galois group cyclic of order n can be obtained as a pull-back of the n th power morphism
$(\,)^n : \mathbf{G}_{m,\mathbb{Z}[1/n]} \to \mathbf{G}_{m,\mathbb{Z}[1/n]}$ via a morphism $u : \operatorname{Spec}(K) \to \mathbf{G}_{m,\mathbb{Z}[1/n]}$.

There is nothing special about the cyclic group; in fact, the following much more general result exists. For each finite group Γ, there exists a certain étale locally trivial Γ-torsor $U_{\Gamma,\mathbb{Z}} \to U_{\Gamma,\mathbb{Z}}/\Gamma$ such that for every field K, every étale locally trivial Γ-torsor over K is obtainable as a pull-back. Note that there is no restriction on the base field K, and moreover the total space of the Γ-torsor on K need not be connected (that is, L need not be a field).

When Γ is abelian, this result is a crucial step in the Lang-Rosenlicht theorem that any abelian extension of the function field of a curve is the pull-back of a covering of a generalized Jacobian of the curve (where the curve is geometrically irreducible, reduced, smooth, projective over a finite field).

What follows is an expository account of the construction and the universal property of the Γ-torsor $U_{\Gamma,\mathbb{Z}} \to U_{\Gamma,\mathbb{Z}}/\Gamma$.

1. Quotient by the free action of a finite group

The reader is assumed to be familiar with the elements of the language of schemes. In particular, we shall use the concepts of valued points, and morphisms which are of finite type, separated, finite, proper, flat, faithfully flat, unramified (separable), and étale. All of the above is nicely explained in Mumford’s introductory textbook ‘The Red Book of Varieties and Schemes’ (an inexpensive Indian edition, published by Narosa, exists).

For a morphism $\pi : X \to S$ of schemes, $Aut_S(X)$ will denote the group of all automorphisms $\varphi : X \to X$ of the scheme X which satisfy $\pi \circ \varphi = \pi$. Let Γ be a finite group. A **right-action** of Γ on X/S is by definition a group homomorphism $\rho : \Gamma^{op} \to Aut_S(X)$ of the opposite group of Γ into $Aut_S(X)$. Unless otherwise indicated, all actions on schemes (respectively, on rings) will be right-actions (respectively, left-actions). A right-action

$\rho : \Gamma^{op} \to Aut(X)$ of Γ on an affine scheme $X = \mathrm{Spec}(A)$ corresponds to a left-action of Γ on the ring A defined as follows. Note that any $f \in A$ can be regarded as a morphism $f : X \to \mathbf{A}^1_{\mathbb{Z}}$ from X to the affine line $\mathbf{A}^1_{\mathbb{Z}}$. We put $\gamma f = f \circ \rho(\gamma) : X \to \mathbf{A}^1_{\mathbb{Z}}$, which can be seen to define a left-action of Γ on A by ring automorphisms.

By definition, if a group Γ acts on a scheme X, and if $x \in X$, then the **decomposition group** at x is the subgroup $D_x \subset \Gamma$ consisting of all γ with $\gamma(x) = x$, that is, D_x is the set-theoretic isotropy at x for the action of γ on the underlying set of the scheme X. Any element $\gamma \in D_x$ induces an automorphism $\gamma^* : \kappa(x) \to \kappa(x)$ of the residue field at x. The **inertia group** at x is the subgroup $I_x \subset D_x$ consisting of all γ such that the automorphism $\gamma^* : \kappa(x) \to \kappa(x)$ is identity. We say that Γ **acts freely** on X if for each $x \in X$ the inertia group I_x is trivial. Equivalently, Γ acts freely on X if for every field K, the action of Γ on the set $X(K)$ of K-valued points of X is free (in the usual set-theoretic sense).

Let there be given a scheme X over a base S, and let a finite group Γ act (on the right) on X over S. Let $X \times \Gamma = \coprod_\Gamma X$ be the disjoint union of copies X_γ of X, indexed by $\gamma \in \Gamma$. We define a morphism

$$\alpha : X \times \Gamma \to X \times_S X$$

as follows. If $x : T \to X_\gamma$ is a T-valued point of X_γ, we define $\alpha(x)$ to be the T-valued point $(x, x\gamma)$ of $X \times_S X$. This uniquely determines α.

We say that a given action makes $X \to S$ an **étale locally trivial Γ-torsor over the base** S (or a **Γ-torsor over** S **in the étale topology**) if $X \to S$ is finite, étale, surjective, and moreover the above morphism $\alpha : X \times \Gamma \to X \times_S X$ is an isomorphism.

Note that if $X \to S$ is an étale locally trivial Γ-torsor, then for any morphism $T \to S$ of schemes, the base change $X \times_S T \to T$ has a natural structure of an étale locally trivial Γ-torsor (called as the **pull back** of $X \to S$ under $T \to S$).

We say that the Γ-torsor $X \to S$ is **trivial** if there exists a Γ-equivariant isomorphism $X \to S \times \Gamma$ over the base S, where the action of Γ on $S \times \Gamma$ is by right translation (an element γ maps S_δ identically to $S_{\delta\gamma}$).

The following exercise connects the notion of a torsor with basic field theory.

Exercise. Let L/K be a field extension, and Γ a finite group acting on $\mathrm{Spec}(L)$ over $\mathrm{Spec}(K)$. Show that under this action $\mathrm{Spec}(L)$ is an étale Γ-torsor over $\mathrm{Spec}(K)$ if and only if L/K is a finite Galois extension, with Galois group $Gal(L/K)$ isomorphic to Γ via the given action.

Theorem 1. *Let k be a noetherian ring, let A be a finite type k-algebra, and let Γ be a finite group of k-automorphisms of A such that Γ acts freely* *on the scheme* Spec(A). *Let* $A^\Gamma \subset A$ *be the invariant subring. Then the induced morphism* $\pi : \mathrm{Spec}(A) \to \mathrm{Spec}(A^\Gamma)$ *of schemes is finite étale, and together with the* Γ*-action,* π *is a* Γ*-torsor in the étale topology.*

Proof. Step 1 : *A* is finite over A^Γ, and A^Γ is of finite type over *k*. Let $a_1, \ldots, a_n$ be algebra generators for A over k. Then for each i, the element a_i satisfies the monic polynomial

$$f_i(t) = \prod_{\gamma \in \Gamma} (t - \gamma(a_i)) \in A^\Gamma[t]$$

Let $C \subset A$ be the k-subalgebra generated by the coefficients of all the f_i. Then C is of finite type over k, so C is noetherian as k is noetherian. As A is finite type and integral over C, it follows that A is finite over C. Note that $C \subset A^\Gamma \subset A$, so A^Γ is finite over C as C is noetherian. In particular, A^Γ is finite type over k. As A is finite over C, it follows that A is finite over A^Γ. We have thus proved that $\pi : \mathrm{Spec}(A) \to \mathrm{Spec}(A^\Gamma)$ is a finite morphism of affine schemes of finite type over k.

Step 2 : Flat base change and invariants. Let B be a flat A^Γ-algebra. Consider the action of Γ on $A \otimes_{A^\Gamma} B$ given by $\gamma(a \otimes b) = \gamma(a) \otimes b$. Consider the left-exact sequence of A^Γ-modules

$$0 \to A^\Gamma \to A \to \oplus_{\gamma \in \Gamma} A$$

where the last map sends a to the n-tuple $(a - \gamma_1(a), \ldots, a - \gamma_n(a))$. As B is A^Γ-flat, tensoring gives the left-exact sequence

$$0 \to B \to A \otimes_{A^\Gamma} B \to \oplus_{\gamma \in \Gamma} A \otimes_{A^\Gamma} B$$

which shows the equality $B = (A \otimes_{A^\Gamma} B)^\Gamma$.

Step 3 : Base change preserves freeness of action. Let Γ act on X/S, and let $f : S' \to S$ be any morphism. Let $X' = X \times_S S'$, and let $f' : X' \to X$ be the projection. Consider the induced action of Γ on X'/S'. For any $x' \in X'$, let $x = f'(x')$. Then it is clear that the decomposition group $D_{x'}$ is contained in D_x. Next, let $\gamma \in D_{x'}$. Note that $f' : X' \to X$ induces an inclusion $\kappa(x) \hookrightarrow \kappa(x')$, and the homomorphism $\gamma : \kappa(x') \to \kappa(x')$ restricts under this inclusion to the homomorphism $\gamma : \kappa(x) \to \kappa(x)$. Hence we get an inclusion of inertia groups $I_{x'} \subset I_x$. It follows that if the action of Γ on X is free, then the action of Γ on X' is also free.

Step 4 : Reduction to the case where A^Γ is local. For any prime $\mathbf{p} \subset A^\Gamma$, the local ring $(A^\Gamma)_\mathbf{p}$ is flat over A^Γ. We denote $A \otimes_{A^\Gamma} (A^\Gamma)_\mathbf{p}$ simply by $A_\mathbf{p}$ as usual (this will be a semi-local ring, as it is finite over the local ring $(A^\Gamma)_{\mathbf{p}}$). To prove A is étale over A^Γ, we just have to prove that $A_{\mathbf{p}}$ is étale over $(A^\Gamma)_{\mathbf{p}}$ for every prime $\mathbf{p} \subset A^\Gamma$. By step 2, $(A^\Gamma)_{\mathbf{p}} = (A_{\mathbf{p}})^\Gamma$. By step 3, freeness of action holds for Γ acting on $A_{\mathbf{p}}$. Hence we are reduced to the case where A^Γ is local.

Step 5 : Reduction to the case where A^Γ is strictly henselian. Now, let B be a strict henselization of the local ring A^Γ. Note that as B is faithfully flat over A^Γ, A is étale over A^Γ if $A \otimes_{A^\Gamma} B$ is étale over B. By steps 2 and 3, we are then reduced to the case where A^Γ is strictly henselian.

Step 6 : Proof when A^Γ is strictly henselian. As already proved, A is finite over A^Γ. Hence by henselian property of A^Γ, the A^Γ-algebra A is a direct product of A^Γ-algebras $A_1 \times \ldots \times A_r$, where each A_i is a henselian local ring, finite over A^Γ. Let $\mathbf{m}_i$ denote the maximal ideal of A_i. Then

$$\mathbf{n}_i = A_1 \times \ldots A_{i-1} \times \mathbf{m}_i \times A_{i+1} \times \ldots \times A_r$$

are all the maximal ideals of A, and for each $\gamma \in \Gamma$ and each $\mathbf{n}_i$, we have $\gamma(\mathbf{n}_i) = \mathbf{n}_j$ for some j. Suppose $\gamma(\mathbf{n}_i) = \mathbf{n}_i$, in other words, γ lies in the decomposition group $D_{\mathbf{n}_i}$. Then γ induces an automorphism of the residue field $A_i/\mathbf{n}_i$ over the base $A^\Gamma/\mathbf{m}$, where $\mathbf{m}$ is the maximal ideal of A^Γ. By assumption of strict henselianness, $A^\Gamma/\mathbf{m}$ is separably closed, so the finite extension $A_i/\mathbf{n}_i$ is purely inseparable, so $\gamma \in D_{\mathbf{n}_i}$ induces identity on $A_i/\mathbf{n}_i$, hence $D_{\mathbf{n}_i}$ is the same as the inertia group $I_{\mathbf{n}_i}$. By assumption of free action, it follows that each $D_{\mathbf{n}_i}$ is trivial.

As $\mathrm{Spec}(A_i)$ are the connected components of $\mathrm{Spec}(A)$, any automorphism γ maps each A_i isomorphically onto some A_j, with $\gamma(\mathbf{n}_i) = \mathbf{n}_j$. We claim that Γ acts transitively on the set of all $\mathbf{n}_i$. Otherwise, consider the element $a = (a_1, \ldots, a_r) \in A$ where $a_i = 1 \in A_i$ if $\mathbf{n}_i$ is in the orbit of $\mathbf{n}_1$, and $a_i = 0$ otherwise. Then clearly $a \in A^\Gamma$. Now suppose $\mathbf{n}_i$ is not in the orbit of $\mathbf{n}_1$. The map $A^\Gamma/\mathbf{m} \to A_i/\mathbf{n}_i$ sends $a \mapsto 0$, so $a \in \mathbf{m}$. But then under $A^\Gamma/\mathbf{m} \to A_1/\mathbf{n}_1$ we would have $a \mapsto 0$, which is a contradiction as $a_1 = 1$. This proves transitivity.

Each $D_{\mathbf{n}_i}$ is trivial, hence for each i, j there exists a unique $\gamma \in \Gamma$ with $\gamma(\mathbf{n}_i) = \mathbf{n}_j$. In particular, for each i there is a unique γ_i with $\gamma_i(\mathbf{n}_1) = \mathbf{n}_i$. This gives an isomorphism $\varphi_i : A_1 \to A_i$. If we write $A_1 = R$ say, then it follows that A is the product $R \times \ldots \times R$ of r copies of R, and Γ acts by permuting factors in a transitive way. The invariant subring A^Γ is therefore the diagonal embedding of R into $R \times \ldots \times R$. Hence A is étale over A^Γ.

Note The hypothesis that Γ acts freely on $\mathrm{Spec}(A)$ was very important in the above. Without this hypothesis, $\pi : \mathrm{Spec}(A) \to \mathrm{Spec}(A^\Gamma)$ **may not even be flat**. For example, with base $k = \mathbb{C}$, take the action of $\Gamma = \mathbb{Z}/(2)$ on the polynomial ring $\mathbb{C}[x, y]$ sending $x \mapsto$

$-x$ and $y \mapsto -y$. Then the module $\mathbb{C}[x,y]$ is not flat over the ring $\mathbb{C}[x,y]^\Gamma = \mathbb{C}[x^2,xy,y^2]$, as its generic rank is 2, but the fiber over the point $\mathbf{m} = (x^2,xy,y^2)$ has rank 3.

Remark 2. If $\pi : Y \to X$ is a Γ-torsor in the étale topology for a finite group Γ, then in general Γ is only a subgroup of the deck transformation group $Aut_X(Y)$ of the finite étale covering $Y \to X$. For example, if X is connected, and Y is the trivial torsor $X \times \Gamma$, then the deck transformation group $Aut_X(Y)$ is the permutation group S_n, where n is the order of Γ. However, if Y is connected, then we have $\Gamma = Aut_X(Y)$, as in that case any deck transformation is determined by its effect on a single closed point.

2. The structure of Γ-torsors over a field K

Summary The first part of this section is just an exercise in so called 'Galois descent' (see for example chapter 2 of Milne's book 'Étale Cohomology' for the basics of Galois descent and its relation with Galois cohomology). Let E/K be a finite Galois field extension, let Γ be a finite group, let $\varphi : Gal(E/K) \to \Gamma$ be a group homomorphism, and let A/K be the Γ-torsor obtained by 'extension of structure group' from the $Gal(E/K)$-torsor E/K via φ. Every étale Γ-torsor over a field K arises this way, where moreover we can choose E so that φ is injective. When φ is chosen to be injective, A is isomorphic to the direct product E^r as a K-algebra, where r is the index of $image(\varphi)$ in Γ, and we explicitly write the Γ-action on $A = E^r$ which makes it a Γ-torsor.

The final result of this section (Theorem 6 below) gives the existence of an element $c \in A$ such that the determinant $\det(\gamma_i\gamma_j^{-1}(c))$ is a unit in A.

Let Γ be a finite group, K be a field, E/K a finite Galois extension field, and $\varphi : Gal(E/K) \to \Gamma$ a homomorphism of groups. The set $Maps(\Gamma, E)$ of all set maps $c : \Gamma \to E$ becomes a commutative E-algebra under pointwise operations. We denote $c(\gamma)$ by c_γ. Let $Maps_\varphi(\Gamma, E) \subset Maps(\Gamma, E)$ be the subring of φ-equivariant maps, that is, those c which satisfy

$$g(c_\gamma) = c_{\varphi(g)\gamma} \text{ for all } g \in Gal(E/K),\ \gamma \in \Gamma$$

The ring $Maps_\varphi(\Gamma, E)$ is a K-algebra under pointwise operations. We define a left-action of Γ on $Maps(\Gamma, E)$ by

$$(\alpha c)_\beta = c_{\alpha\beta} \text{ for all } \alpha, \beta \in \Gamma$$

This makes $Maps(\Gamma, E)$ a trivial Γ-torsor over E.

Note that the K-subalgebra $Maps_\varphi(\Gamma, E) \subset Maps(\Gamma, E)$ is invariant under the left-action of Γ.

We have the following structure theorem for Γ-torsors on K.

Lemma 3. *If Γ is a finite group, K is a field, E/K a finite Galois field extension, and $\varphi : Gal(E/K) \to \Gamma$ a group homomorphism, then the K-algebra $Maps_\varphi(\Gamma, E)$ with the left-action of Γ as defined above is an étale* *locally trivial Γ-torsor over K. Its pull-back to E is the trivial Γ-torsor $Maps(\Gamma, E)$ over E.*

Conversely, if L is a K-algebra together with a left Γ-action which makes it an étale locally trivial Γ-torsor over K, and if E/K is a finite Galois field extension such that it pulls back to a trivial Γ-torsor over E, then there exists a group homomorphism $\varphi : Gal(E/K) \to \Gamma$ such that the given Γ-torsor is isomorphic to the Γ-torsor $Maps_\varphi(\Gamma, E)$ over K constructed above.

Proof. It is clear from its construction that $Maps_\varphi(\Gamma, E)$ indeed has the desired properties. Conversely, if a Γ-torsor over K pulls back to a trivial Γ-torsor on E where E/K is a finite Galois field extension, then it corresponds to a 1-cocycle φ in $H^1(Gal(E/K), \Gamma)$.
As Γ is a constant group, we have $H^1(Gal(E/K), \Gamma) = Hom(Gal(E/K), \Gamma)$, so we regard φ as a group homomorphism $Gal(E/K) \to \Gamma$. Now by 'descending' the trivial Γ-torsor $Maps(\Gamma, E)$ over E by the cocycle φ, we get the torsor $Maps_\varphi(\Gamma, E)$ over K constructed above.

Remark 4. With the above notation, if $N = ker(\varphi) \subset Gal(E/K)$, then we can replace E with the invariant subfield E^N, and $\varphi : Gal(E/K) \to \Gamma$ by the induced homomorphism $Gal(E/K)/N \to \Gamma$ which is injective. Hence we can always assume that $\varphi : Gal(E/K) \to \Gamma$ is injective.

We now more directly describe the structure of the Γ-torsor $Maps_\varphi(\Gamma, E)$ over K, where by the above remark, we can assume that $\varphi : Gal(E/K) \to \Gamma$ is injective without any loss of generality. Let $image(\varphi) = D \subset \Gamma$. Let $D\sigma_1, \ldots, D\sigma_r$ be the distinct right cosets of D in Γ, where $r = (\Gamma : D)$ is the index of D in Γ. Consider the r-fold direct product E^r, which is a ring under componentwise operations, and is an E-algebra (hence also a K-algebra) via the diagonal map $\Delta : E \to E^r$. For $1 \leq k \leq r$, let $p_k : E^r \to E$ be the projections, which are E-algebra homomorphisms, and let $f_k : E \to E^r$ be the inclusions (with $p_k f_k = \mathrm{id}_E$ and $p_i f_j = 0$ for $i \neq j$) which are merely E-linear maps of vector spaces. We define an E-algebra homomorphism $Maps_\varphi(\Gamma, E) \to E^r$ by sending $c \mapsto x$ where $p_k(x) = c_{\sigma_k}$, which is an isomorphism, with inverse $E^r \to Maps_\varphi(\Gamma, E)$ defined by sending $x \mapsto c$ where $c_{g\sigma_k} = g(c_{\sigma_k}) = g(p_k(x))$ where $g \in D$. These homomorphisms are inverses of each other, so define an isomorphism $Maps_\varphi(\Gamma, E) \to E^r$. Via this isomorphism, the left-action of Γ on $Maps_\varphi(\Gamma, E)$ gives the following left-action of Γ on E^r. Any element $x \in E^r$ is the sum of elements of the form $f_k(a) \in E^r$, where $a \in E$. Given any $\gamma \in \Gamma$, there exists a unique i with $1 \leq i \leq r$ and a unique $g \in D$ such that

$$\gamma = \sigma_i^{-1} g \sigma_k$$

Then we define $\gamma f_k(a) \in E^r$ by putting

$$p_j \sigma_i^{-1} g \sigma_k f_k(a) = \begin{cases} g(a) & \text{if } j = i \\ 0 & \text{otherwise.} \end{cases}$$

Trace map on Γ-torsors

For any finite K-algebra A we define the trace map

$$Trace_{A/K} : A \to K$$

as follows. For $x \in A$, consider the K linear map $A \to A$ which sends $y \mapsto xy$. Then $Trace_{A/K}(x)$ is the trace of this linear map. If E/K is a finite Galois field extension, then it is known (elementary fact) that for all $x \in E$,

$$Trace_{E/K}(x) = \sum_{g \in Gal(E/K)} g(x)$$

We now calculate trace for the Γ-torsor E^r above. For $f_k(a) \in E^r$ we have

$$\begin{aligned} p_j \left(\sum_{\gamma \in \Gamma} \gamma f_k(a) \right) &= \sum_{\gamma \in \Gamma} p_j \gamma f_k(a) \\ &= \sum_{1 \leq i \leq r} \sum_{g \in D} p_j \sigma_i^{-1} g \sigma_k f_k(a) \\ &= \sum_{g \in D} g(a) \\ &= Trace_{E/K}(a) \end{aligned}$$

As the above holds for all j, we get

$$\sum_{\gamma \in \Gamma} \gamma f_k(a) = \Delta(Trace_{E/K}(a))$$

For any $x \in E^r$, we have $x = \sum_k f_k p_k(x)$, hence summing the above gives the equality

$$\sum_{\gamma \in \Gamma} \gamma x = \sum_{1 \leq k \leq r} \Delta(Trace_{E/K}(p_k(x))) = \Delta(Trace_{E^r/K}(x))$$

Hence we have proved the following.

Proposition 5. *Let Γ be a finite group, let K be a field, and let A be an étale Γ-torsor over K. Then we have equality of K-linear maps*

$$\sum_{\gamma\in\Gamma}\gamma \;=\; Trace_{A/K} : A \to K$$

Next, suppose that A is a finite separable K-algebra. We define a symmetric K-bilinear map

$$T_{A/K} : A \times A \to K : (x, y) \mapsto Trace_{A/K}(xy)$$

If $A = L$ is a field, then by linear independence of characters, the map $Trace_{L/K} : L \to K$ is non-zero, so it follows that $T_{L/K} : L \times L \to K$ is non-degenerate. In general, $A = L_1 \times \dots L_r$ is a direct product of such field extensions, and $(A, T_{A/K})$ is the orthogonal direct sum of the $(L_i, T_{L_i/K})$, so again $T_{A/K} : A \times A \to K$ is non degenerate.

Normal basis theorem for Γ-torsors

By the normal basis theorem for finite Galois field extensions E/K, there exists an element $a \in E$ such that $g_1(a), \dots, g_m(a)$ is a K-linear basis for E, where $\{g_1, \dots, g_m\} = Gal(E/K)$. Now with the previous notation, consider the Γ-torsor E^r over K. It is clear from the description of the Γ-action on E^r that the element $c = f_1(a)$ of E^r has the property that the elements $\gamma(c)$, for $\gamma \in \Gamma$, form a K-linear basis of E^r.

Existence of $c \in A$ such that $\det(\gamma_i\gamma_j^{-1}(c))$ is a unit

We are at last ready to prove the main result that we want.

Theorem 6. *Let $\Gamma = \{\gamma_1, \dots, \gamma_n\}$ be a finite group, let K be a field, and let A be an étale Γ-torsor over K. For any $c \in A$, consider the $n \times n$ matrix $(\gamma_i\gamma_j^{-1}(c))$ over A. There exists an element $c \in A$ such that this matrix is invertible, that is,*

$$\det(\gamma_i\gamma_j^{-1}(c)) \quad \textit{is a unit in } A.$$

Proof. Let $e_1, \dots, e_n$ be any K-linear basis for A. Consider the $n \times n$ matrix $M = (\gamma_i(e_j))$ over A. As proved above, the trace map is given by $\sum_\Gamma \gamma$. Hence we have

$$({}^tMM)_{j,k} = \sum_i \gamma_i(e_j)\gamma_i(e_k) = \sum_i \gamma_i(e_je_k) = Trace_{A/K}(e_je_k)$$

Hence ${}^tMM$ is the matrix of the bilinear form $T_{A/K} : A \times A \to K$. As this K-bilinear form is non-degenerate, $\det({}^tMM)$ is a unit in K, hence a unit in A. Hence $\det(M)$ is also a unit in A, as $(\det(M))^2 = \det({}^tMM)$.

As shown earlier (the normal basis theorem for Γ-torsors), there exists an element $c \in A$ such that $e_j = \gamma_j^{-1}(c)$ is a K-linear basis for A. For this basis, $M = (\gamma_i\gamma_j^{-1}(c))$, and so the element $\det(\gamma_i\gamma_j^{-1}(c))$ is a unit in A.

Units in a group ring

Now that adequate preparation – in particular, a careful formulation using the concept of a torsor – has been made, the rest of these notes essentially follow Serre's 1959 book 'Groupes algébriques et corps de classes'. It seems that the concept of an étale locally trivial torsor was not available then – it was formalized (by Serre himself) somewhat later.

Let Γ be a finite group. For any commutative ring k, let $k[\Gamma]$ denote the group ring of Γ with coefficients k. We define a functor

$$U_\Gamma : \mathit{Rings} \longrightarrow \mathit{Groups} : k \mapsto U_\Gamma(k) = \text{Units in } k[\Gamma]$$

Lemma 7. *The above functor U_Γ is represented by an affine group scheme $U_{\Gamma,\mathbb{Z}}$ of finite type over $\mathbb{Z}$. The underlying scheme of the group scheme $U_{\Gamma,\mathbb{Z}}$ can be embedded as an open subscheme of the affine n-space $\mathbf{A}^n_{\mathbb{Z}}$, where n is the order of Γ. In particular, $U_{\Gamma,\mathbb{Z}}$ is irreducible and smooth over* Spec($\mathbb{Z}$) *of relative dimension n. The constant group scheme $\Gamma_{\mathbb{Z}}$ is a closed subgroup scheme of $U_{\Gamma,\mathbb{Z}}$.*

Proof. Let $\Gamma = \{\gamma_1, \dots, \gamma_n\}$. It can be seen that an element $x = \sum x_{\gamma_i}\gamma_i \in k[\Gamma]$ is invertible if and only if the $n \times n$-matrix $M(x)$ over k with entries

$$M(x)_{i,j} = x_{\gamma_i\gamma_j^{-1}}$$

is invertible (lies in $GL_n(k)$). From this it follows that the scheme $U_{\Gamma,\mathbb{Z}}$ can be constructed as the open subscheme of the affine space $\mathbf{A}^n_{\mathbb{Z}}$, which is the inverse image of $GL_{n,\mathbb{Z}}$ under the morphism $M : \mathbf{A}^n_{\mathbb{Z}} \to \mathbf{A}^{n\times n}_{\mathbb{Z}}$ which maps the k-valued point x to the k-valued point $M(x)$. As the morphism M is affine, and as $GL_{n,\mathbb{Z}}$ is affine open in $\mathbf{A}^{n\times n}_{\mathbb{Z}}$, its inverse image $U_{\Gamma,\mathbb{Z}}$ is affine open in $\mathbf{A}^n_{\mathbb{Z}}$.

The group Γ acts on the scheme $U_{\Gamma,\mathbb{Z}}$ by right translation, and let

$$\pi : U_{\Gamma,\mathbb{Z}} \to U_{\Gamma,\mathbb{Z}}/\Gamma$$

be the quotient. The theorem 1 and Remark 2, together with the following lemma 8, imply that the morphism $\pi : U_{\Gamma,\mathbb{Z}} \to U_{\Gamma,\mathbb{Z}}/\Gamma$ is a Γ-torsor in the étale topology (in particular π is finite étale), with deck transformation group Γ.

Lemma 8. *The right translation action of* Γ *on* $U_{\Gamma,\mathbb{Z}}$ *is free.*

Proof. Let K be a field, and consider a K-valued point $x \in U_{\Gamma,\mathbb{Z}}(K)$. This corresponds to an invertible element $x \in K[\Gamma]$. Hence for any $\gamma \neq 1$, we have $x\gamma \neq x$, so Γ acts freely on the set $U_{\Gamma,\mathbb{Z}}(K)$.

‘Universal’ property of $U_{\Gamma,\mathbb{Z}} \to U_{\Gamma,\mathbb{Z}}/\Gamma$

The reason for our interest in the étale Γ-torsor $U_{\Gamma,\mathbb{Z}} \to U_{\Gamma,\mathbb{Z}}/\Gamma$ is that it has the following property. I have put ‘universal’ in quotes, because there is no uniqueness for the morphism $u : \mathrm{Spec}(K) \to U_{\Gamma,\mathbb{Z}}/\Gamma$.

Theorem 9. *Let* K *be any field and let* L *be a* Γ*-torsor over* K *in the étale topology. Then there exists a morphism of schemes* $u : \mathrm{Spec}(K) \to U_{\Gamma,\mathbb{Z}}/\Gamma$ *and a* Γ*-equivariant morphism of schemes* $v : \mathrm{Spec}(L) \to U_{\Gamma,\mathbb{Z}}$ *such that the following diagram is cartesian.*

$$\begin{tikzcd}
\mathrm{Spec}(L) \arrow[r, "v"] \arrow[d] & U_{\Gamma,\mathbb{Z}} \arrow[d] \\
\mathrm{Spec}(K) \arrow[r, "u"] & U_{\Gamma,\mathbb{Z}}/\Gamma
\end{tikzcd}$$

In other words, for any field K*, every étale* Γ*-torsor over* $\mathrm{Spec}(K)$ *is a pull-back of* $U_{\Gamma,\mathbb{Z}} \to U_{\Gamma,\mathbb{Z}}/\Gamma$.

Proof. Let $x_{\gamma_1}, \ldots, x_{\gamma_n}$ be indeterminates, and let $\mathbf{A}^n_{\mathbb{Z}}$ be the affine space $\mathrm{Spec}\ \mathbb{Z}[x_{\gamma_1}, \ldots, x_{\gamma_n}]$. By definition, $U_{\Gamma,\mathbb{Z}} \subset \mathbf{A}^n_{\mathbb{Z}}$ is the open subscheme which is the complement of the divisor D defined by $\det(x_{\gamma_i\gamma_j^{-1}}) = 0$, in other words, the coordinate ring of $U_{\Gamma,\mathbb{Z}}$ is

$$A = \mathbb{Z}[x_{\gamma_1}, \ldots, x_{\gamma_n}, 1/\det(x_{\gamma_i\gamma_j^{-1}})]$$

The action of the group Γ by right-translation on $U_{\Gamma,\mathbb{Z}}$ corresponds to its left-action on the polynomial ring $\mathbb{Z}[x_{\gamma_1}, \ldots, x_{\gamma_n}]$ by permuting variables, defined by

$$\alpha(x_\beta) = x_{\beta\alpha^{-1}}$$

where $\alpha, \beta \in \Gamma$. For any ring L, an L-valued point of $U_{\Gamma,\mathbb{Z}}$ is given by a ring homomorphism

$$v : \mathbb{Z}[x_{\gamma_1}, \ldots, x_{\gamma_n}, 1/\det(x_{\gamma_i\gamma_j^{-1}})] \to L$$

Let $x_\gamma \mapsto v_\gamma$ under the above homomorphism. Then note that we can choose the n elements $v_{\gamma_i} \in L$ arbitrarily, subject to the only requirement that

$$\det(v_{\gamma_i\gamma_j^{-1}}) \text{ is a unit in the ring } L.$$

If we are also given a left-action of Γ on L then the homomorphism v is Γ-equivariant if and only if the following is satisfied for all $\alpha, \beta \in \Gamma$

$$v(\alpha(x_\beta)) = \alpha(v(x_\beta))$$

Substituting $\alpha(x_\beta) = x_{\beta\alpha^{-1}}$ from above, we get $v_{\beta\alpha^{-1}} = \alpha(v_\beta)$. In particular,

$$v_\gamma = \gamma^{-1}(v_e)$$

where $e \in \Gamma$ is the identity. Hence an equivariant homomorphism $v : \mathbb{Z}[x_{\gamma_1}, \ldots, x_{\gamma_n}, 1/\det(x_{\gamma_i\gamma_j^{-1}})] \to L$ is determined by an arbitrary element $c = v_e \in L$, provided the element c satisfies the only condition that

$$\det(\gamma_j\gamma_i^{-1}(c)) \ \text{ is a unit in the ring } L.$$

Now assume that there exists such an element c, and let $v : H^0(U_{\Gamma,\mathbb{Z}}, \mathcal{O}_{U_{\Gamma,\mathbb{Z}}}) \to L$ be defined by $x_\gamma \mapsto \gamma^{-1}(c)$. As this ring homomorphism is Γ-equivariant, it induces a ring homomorphism

$$H^0(U_{\Gamma,\mathbb{Z}}/\Gamma, \mathcal{O}_{U_{\Gamma,\mathbb{Z}}/\Gamma}) = (H^0(U_{\Gamma,\mathbb{Z}}, \mathcal{O}_{U_{\Gamma,\mathbb{Z}}}))^\Gamma \to L^\Gamma = K$$

of the invariant subrings. Hence we have a commutative diagram

$$\begin{tikzcd}
\mathrm{Spec}(L) \arrow[r, "v"] \arrow[d] & U_{\Gamma,\mathbb{Z}} \arrow[d] \\
\mathrm{Spec}(K) \arrow[r, "u"] & U_{\Gamma,\mathbb{Z}}/\Gamma
\end{tikzcd}$$

As both columns are Γ-torsors and the top map is Γ-equivariant, the above rectangle is cartesian.

If L is a field, then by normal basis theorem there exists such an element c. When L is a more general étale Γ torsor over the field K, the existence of such a c is given by the theorem 6. This completes the proof of the theorem 9.

REFERENCES

1. Milne : *Étale Cohomology*, Princeton Univ Press, 1980.
2. Mumford : *The Red Book of Varieties and Schemes*, Springer-Narosa, 1995.
3. Serre : *Groupes algébriques et corps de classes*, Hermann, 1959.

Nitin Nitsure
School of Mathematics
Tata Institute of Fundamental Research
Homi Bhabha Road, Mumbai 400 005, India.
e-mail: nitsure@math.tifr.res.in

On the Coefficients of Cyclotomic Polynomials

R. Thangadurai

1. Properties of Cyclotomic Polynomials

Cyclotomy is the process of dividing a circle into equal parts, which is precisely the effect obtained by plotting the n-th roots of unity in the complex plane.

For integers $n \geq 1$, we know that $X^n - 1 = \prod_{m=0}^{n-1}(X - e^{\frac{2\pi i m}{n}})$ over $\mathbb{C}$. The n-th *cyclotomic polynomial* can be defined as

$$\Phi_n(X) = \prod_{m=1,(m,n)=1}^{n} (X - e^{\frac{2\pi i m}{n}})$$

where $e^{\frac{2\pi i}{n}}$ is a primitive n-root of unity. Clearly, the degree of $\Phi_n(X)$ is $\phi(n)$ where ϕ is the Euler totient function. We have $X^n - 1 = \prod_{d|n} \Phi_d(X)$.

Lemma 1.1 *The cyclotomic polynomial $\Phi_n(X)$ is a monic polynomial over integers.*

Proof. We use induction to prove this result. We have $\Phi_1(X) = X - 1$. We assume that the result is true for all $d < n$ and we prove the result for n. By the induction hypothesis, we have $F(X) \overset{\text{def}}{=} \prod_{d<n,d|n} \Phi_d(X) \in \mathbb{Z}[X]$ and its leading coefficient is 1. As $F(X)$ is monic, by division algorithm, $\exists\ h(X), r(X) \in \mathbb{Z}[X]$ such that $h(X)$ is monic and $X^n - 1 = F(X)h(X) + r(X)$, where $r(X) = 0$ or $\deg r(X) < \deg F(X)$.

But, $X^n - 1 = F(X)\Phi_n(X)$. Therefore, by uniqueness of quotient and remainder in $\mathbb{C}[X]$, we must have $h(X) = \Phi_n(X)$. Also it is clear that $\Phi_n(X)$ has leading coefficient 1. □

The *Möbius function*, $\mu(n)$, is defined by

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^k & \text{if } n = p_1 p_2 \cdots p_k \text{ for distinct primes } p_i\ , \\ 0 & \text{otherwise.} \end{cases}$$

Note that it can be easily seen that $\mu(mn) = \mu(m)\mu(n)$ whenever $(m, n) = 1$. Also, $\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise} \end{cases}$

Lemma 1.2 If $\mu(n)$ denotes the Möbius function, then,

$$\Phi_n(X) = \prod_{d|n}(X^{n/d} - 1)^{\mu(d)} = \prod_{d|n}(X^d - 1)^{\mu(n/d)}.$$

Proof. We shall prove that if $f(n) = \prod_{d|n} g(d)$, then $g(n) = \prod_{d|n} f(n/d)^{\mu(d)}$.

$$\begin{aligned}\text{We have, } \quad \prod_{d|n} f(n/d)^{\mu(d)} &= \prod_{d|n}\left(\prod_{m|(n/d)} g(m)\right)^{\mu(d)}\\ &= \prod_{m|n}\left(\textstyle\prod_{d|(n/m)} g(m)^{\mu(d)}\right)\\ &= \prod_{m|n} g(m)^{\sum_{d|(n/m)}\mu(d)} \;=\; g(n).\end{aligned}$$

Since $X^n - 1 = \prod_{d|n} \Phi_d(X)$, we are done. □

Lemma 1.3
(i) If $n = p_1^{a_1} p_2^{a_2} \cdots p_\ell^{a_\ell}$, $a_i > 0$, *and* $N = p_1 p_2 \cdots p_\ell$, *then* $\Phi_n(X) = \Phi_N(X^{n/N})$.
(ii) If $n > 1$ *and* $(2, n) = 1$, *then* $\Phi_{2n}(X) = \Phi_n(-X)$.
(iii) For all positive integers $n > 1$, *we have* $X^{\phi(n)}\Phi_n(1/X) = \Phi_n(X)$.

Proof. (i) Since $\mu(m) = 0$ for all integers m which are not square free, we have,

$$\begin{aligned}\Phi_n(X) &= \prod_{d|n}(X^{n/d} - 1)^{\mu(d)} = \prod_{d|n, d|N}(X^{n/d} - 1)^{\mu(d)}\\ &= \prod_{d|N}((X^{n/N})^{N/d} - 1)^{\mu(d)} = \Phi_N(X^{n/N}).\end{aligned}$$

This proves part (i).
(ii) Consider

$$\begin{aligned}\Phi_{2n}(X) &= \prod_{d|(2n)} (X^d - 1)^{\mu(2n/d)}\\ &= \prod_{2|d}(X^d - 1)^{\mu((2n)/d)} \prod_{d|n}(X^d - 1)^{\mu((2n)/d)}\\ &= \prod_{d|n}\left[(X^d - 1)^{\mu(2n/d)}(X^{2d} - 1)^{\mu(n/d)}\right]\\ &= \prod_{d|n}(X^d + 1)^{\mu(n/d)}, \text{ as } \mu(2m) = -\mu(m) \; for\ odd\ m\\ &= \prod_{d|n}(-X^d - 1)^{\mu(n/d)} = \Phi_n(-X).\end{aligned}$$

(iii) Now, consider

$$\Phi_n(1/X) = \prod_{d|n}(1/X^d - 1)^{\mu(n/d)} = \prod_{d|n}(1 - X^d)^{\mu(n/d)} \prod_{d|n}(1/X^d)^{\mu(n/d)}.$$

Therefore we get,

$$\begin{aligned} X^{\sum_{d|n} d\mu(n/d)}\Phi_n(1/X) &= \prod_{d|n}(-1)^{\mu(n/d)}(X^d-1)^{\mu(n/d)} \\ &= (-1)^{\sum_{d|n}\mu(n/d)}\prod_{d|n}(X^d-1)^{\mu(n/d)} = \Phi_n(X). \end{aligned}$$

Since $\sum_{d|n} d\mu(n/d) = \phi(n)$, we get the result. □

2. The Coefficients of Cyclotomic Polynomials

Since $\Phi_n(X)$ is a polynomial with degree $\phi(n)$, we can write

$$\Phi_n(X) = \sum_{i=0}^{\phi(n)} a_n(i)X^i$$

where $a_n(i)$ denotes the i-th coefficient.

Lemma 2.1
(i) $a_n(i) \in \mathbb{Z}$ for all $i,\ 0 \le i \le \phi(n),\ n \in \mathbb{N}$.
(ii) $a_n(i) = a_n(\phi(n)-i)$ for all $i,\ 0 \le i \le \phi(n),\ n\ (>1) \in \mathbb{N}$. That is, the coefficients of cyclotomic polynomials are symmetric.

Proof. (i) follows from Lemma 1.1. Also (ii) follows from Lemma 1.3(iii) immediately. □

Remark 2.2
(1) Lemma 1.3(i) says that

$$a_n(i) = \begin{cases} a_N(iN/n) & \text{if } \frac{n}{N}|i \\ 0 & \text{otherwise.} \end{cases}$$

(2) From Lemma 1.3(ii) we get for odd $n > 1$, $a_{2n}(i) = (-1)^i a_n(i)$.
(3) When $n = p$ a prime number, from Lemma 1.2 we have

$$\Phi_p(X) = (X^p-1)/(X-1) = X^{p-1} + X^{p-2} + \cdots + X + 1.$$

Hence $a_p(i) = 1$ for all $i = 0, 1, \cdots, p-1$.

Thus in any investigation about the coefficients of cyclotomic polynomials we can reduce our enquiry to the case when n is odd, square-free and composite.

When $n = p$ a prime number, as we had seen earlier,

$$a_p(i) = \begin{cases} 1 & \text{if } i = 0, 1, \cdots, p-1 \\ 0 & \text{for all integers } i > p-1. \end{cases}$$

We shall now pass on to the next interesting case when $n = pq$ where p and q are two distinct odd prime numbers. Here are two explicit examples:

$$\Phi_{15}(X) = X^8 - X^7 + X^5 - X^4 + X^3 - X + 1$$

$$\text{and } \Phi_{21}(X) = X^{12} - X^{11} + X^9 - X^8 + X^6 - X^4 + X^3 - X + 1.$$

In 1883, Migotti [27] showed that all $a_{pq}(i) \in \{0, \pm 1\}$. Marion Beiter [5] and [8] gave a criterion on i for $a_{pq}(i)$ to be $0, 1$ or -1. Also Carlitz [11] computed the number of non-zero $a_{pq}(i)$'s. Here, we shall give a simpler proof of the following theorem due to Lam and Leung [19].

Theorem 2.3 *Let r and s be non-negative integers such that $(p-1)(q-1) = rp + sq$ written uniquely. Then we have*

$$\Phi_{pq}(X) = \left(\sum_{i=0}^{r} X^{ip}\right)\left(\sum_{j=0}^{s} X^{jq}\right) - \left(\sum_{i=r+1}^{q-1} X^{ip}\right)\left(\sum_{j=s+1}^{p-1} X^{jq}\right) X^{-pq}.$$

Also, for $0 \leq k \leq (p-1)(q-1)$, we have
(1) $a_{pq}(k) = 1$ if and only if $k = ip + jq$ for some $i \in [0, r]$ and $j \in [0, s]$;
(2) $a_{pq}(k) = -1$ if and only if $k + pq = ip + jq$ for some $i \in [r+1, q-1]$ and $j \in [s+1, p-1]$; and
(3) $a_{pq}(k) = 0$ otherwise.

The number of terms of the former two kinds are, respectively, $(r+1)(s+1)$ and $(p-s-1)(q-r-1)$, with difference 1.

Proof. We know that $\phi(pq) = (p-1)(q-1)$ can be expressed uniquely in the form $rp + sq$ where r, s are non-negative integers (see for instance [23], Page 22, Ex. 4). Since $(p-1)(q-1) = rp + sq$, it is clear that $r \leq q-2$ and $s \leq p-2$.

Now, we shall prove that

$$\Phi_{pq}(X) = \left(\sum_{i=0}^{r} X^{ip}\right)\left(\sum_{j=0}^{s} X^{jq}\right) - \left(\sum_{i=r+1}^{q-1} X^{ip}\right)\left(\sum_{j=s+1}^{p-1} X^{jq}\right) X^{-pq}.$$

Let $\zeta = e^{2i\pi/(pq)}$ be a primitive pq-th root of unity. Then since $\zeta^p = e^{2i\pi/q}$ and $\zeta^q = e^{2i\pi/p}$, we have $\Phi_p(\zeta^q) = \Phi_q(\zeta^p) = 0$. That is, we have

$$\sum_{i=0}^{q-1} (\zeta^p)^i = 0 = \sum_{j=0}^{p-1} (\zeta^q)^j.$$

Therefore,

$$\sum_{i=0}^{r} (\zeta^p)^i = -\sum_{i=r+1}^{q-1} (\zeta^p)^i \text{ and } \sum_{j=0}^{s} (\zeta^q)^j = -\sum_{i=s+1}^{p-1} (\zeta^q)^j.$$

Hence multiplying these two, we get the identity

$$\left(\sum_{i=0}^{r}(\zeta^p)^i\right)\left(\sum_{j=0}^{s}(\zeta^q)^j\right)-\left(\sum_{i=r+1}^{q-1}(\zeta^p)^i\right)\left(\sum_{i=s+1}^{p-1}(\zeta^q)^j\right)=0.$$

Thus ζ is a zero of the polynomial

$$f(X) := \left(\sum_{i=0}^{r}X^{pi}\right)\left(\sum_{j=0}^{s}X^{qj}\right)-\left(\sum_{i=r+1}^{q-1}X^{pi}\right)\left(\sum_{i=s+1}^{p-1}X^{qj}\right)X^{-pq}. \qquad (1)$$

Since $rp+sq=(p-1)(q-1)$, the first product in (1) is a monic polynomial of degree $(p-1)(q-1)$. In the second product, the lowest term has degree $(r+1)p+(s+1)q-pq=rp+sq+p+q-pq=1$ and its highest term has degree $(q-1)p+(p-1)q-pq=(p-1)(q-1)-1$. Hence the second product is also a monic polynomial of degree $(p-1)(q-1)-1$. Therefore $f(X)\in\mathbb{Z}[X]$ is a monic polynomial of degree $(p-1)(q-1)=\phi(pq)$. Moreover, we know that $f(\zeta)=0$. If ζ' is any other primitive pq-th root of unity, then also we have $f(\zeta')=0$. Since $f(X)$ is monic polynomial of degree $\phi(pq)$ with $f(e^{2i\pi m/(pq)})=0$ for all integers m such that $(m,pq)=1$, we must have $f(X)=\Phi_{pq}(X)$.

Now note that if $i,i'\in[0,q-1]$, $j,j'\in[0,p-1]$, and $ip+jq$ is equal to $i'p+j'q$ or $i'p+j'q-pq$, then $q|(i-i')$ and $p|(j-j')$. This implies that $i=i'$ and $j=j'$.

If we expand the products in equation (1), then using the above note, the rest of the assertions follow immediately. □

Remark 2.4 Theorem 2.3 together with our earlier observations proves that the coefficients of the first 104 cyclotomic polynomials are all $\pm1,0$.

Corollary 2.5 *Assume that $q>p$, and let $\ell=(p-1)(q-1)/2$. Then the middle coefficient $a_{pq}(\ell)$ of $\Phi_{pq}(X)$ is $(-1)^r$.*

Proof. By Remark 2.2(2), we can assume that $p>2$. Since $(p-1)(q-1)=rp+sq$, r and s have the same parity. If r is even, then $\ell=(r/2)p+(s/2)q$. Therefore, by Theorem 2.3, we have $a_{pq}(\ell)=1$. If r is odd, then so is s, and we can write,

$$\ell+pq=\left(\frac{r+q}{2}\right)p+\left(\frac{s+p}{2}\right)q.$$

Since $r\le q-2$ and $s\le p-2$, we have $(r+q)/2\in[r+1,q-1]$ and $(s+p)/2\in[s+1,p-1]$. Therefore by Theorem 2.3, we have $a_{pq}(\ell)=-1$. Note that when $p=2$, by Remark 2.2(2), we have $a_{2q}(\ell)=(-1)^{\ell}a_q(\ell)=(-1)^{(q-1)/2}=(-1)^r$. (since $2r+sq=q-1\Longrightarrow r=(q-1)/2$). □

Thus, Theorem 2.3 finishes the problem of finding the values of the coefficients of cyclotomic polynomials explicitly in the case when $n = pq$ where p and q are two distinct odd primes.

If n is a product of more than two distinct primes, then the explicit values of the coefficients are not known in general. But in the case when $n = pqr$, some good amount of progress has been made. Let us discuss this case briefly.

In 1895, Bang [2] proved that the upper bound for the magnitude of the coefficients of $\Phi_{pqr}(X)$ where p, q, r are odd primes such that $p < q < r$ is $p - 1$. Then, in 1968, Marion Beiter [6] and Bloom [9] simultaneously established $(p+1)/2$ as the upper bound in the special case where q and/or r is congruent to ±1 modulo p. In 1971, Beiter [7] gave the following better general bound.

Theorem 2.6 [7] *The magnitude of the largest coefficient of $\Phi_{pqr}(X)$ where p, q, r are odd primes such that $p < q < r$ is less than or equal to $p - k$ or $p - (k+1)$ for $p = 4k+1$ or $4k+3$ respectively.*

We shall skip the proof of this theorem.

Remark 2.7 Note that when $p = 3$, Theorem 2.6 says that $|a_{3qr}(i)| \leq 2$ for all i. Remark 2.5 together with this, we see that the first cyclotomic polynomial $\Phi_{105}(X)$ where we can look for a non-zero coefficient whose magnitude is not just one; but two. In fact this is the case. Indeed, it was shown by Migotti [27] in 1883 that the coefficient of X^7 in the 105-th cyclotomic polynomial is equal to -2. In fact, the 105th cyclotomic polynomial is as follows:

$$\begin{aligned}\Phi_{105}(X) \quad = \quad & X^{48} + X^{47} + X^{46} - X^{43} - X^{42} - 2X^{41} - X^{40} - X^{39} + \\ & X^{36} + X^{35} + X^{34} + X^{32} + X^{31} - X^{28} - X^{26} - X^{24} - \\ & X^{22} - X^{20} + X^{17} + X^{16} + X^{15} + X^{14} + X^{13} + X^{12} - \\ & X^{9} - X^{8} - 2X^{7} - X^{6} - X^{5} + X^{2} + X + 1\end{aligned}$$

Later, P. Erdös [13] showed that $a_n(i) = 0, \pm 1$ for all i and for all $n < 105$ and that $a_{105}(7) = 2$. Also M. Endo [12] proved that $(k, n) = (7, 105)$ is the smallest pair for which $|a_n(k)| > 1$ by a different method.

Conjecture (Beiter, 1971) *$a_{pqr}(i) \leq (p+1)/2$ for all i and for any $p < q < r$ and this upper bound is the best possible.*

Indeed, Beiter remarks in the same paper [7] that the above conjecture is true for $p = 3, 5$ and for any $q < r$. In support of the above conjecture, Möller [29] proved the following theorem.

Theorem 2.8 *Let* $3 < p < q < r$ *be prime numbers satisfying* $q \equiv 2 \pmod{p}$ *and* $r = \frac{1}{2}(mpq - 1)$ *for some integer* m. *Then,*

$$a_{pqr}((p-1)(qr+1)/2) = \frac{1}{2}(p+1).$$

Recently, W. Bosma [10] has written an expository article on the various methods which are helpful in computing cyclotomic polynomials and its coefficients.

Thus from the above theorems and remarks, it appears that the growth of the magnitude of the coefficients of cyclotomic polynomials is very slow. However, it is not very clear at this stage whether the coefficients can take arbitrarily large values.

Schur [33] was the first one to show that there are cyclotomic polynomials whose coefficients are arbitrarily large. If we let $A(n) = \max_m |a_n(m)|$, then Schur showed that $\limsup_{n \longrightarrow \infty} A(n) = \infty$.

We shall give a trivial upper bound for $|a_n(m)|$ for all m in terms of n alone as follows.

Lemma 2.9 *We have* $\log(A(n)) \ll \sqrt{n}$.

Proof. Since using Lemma 1.2, it can be seen that the coefficient of z^m in $\prod_{d \geq 1}(1 - z^d)^{-1}$ is greater than or equal to $|a_n(m)|$ and the former is nothing but $p(m)$ where $p(m)$ is the number of partitions of m, we get $|a_n(m)| \leq p(m)$ for all m.

This inequality together with the Hardy-Ramanujan [18] asymptotic formula for $p(m)$ in the form

$$\log|p(m)| \sim \pi\sqrt{2/3}\,\sqrt{m} \quad \text{as} \quad m \to \infty$$

implies the estimate $\log(a_n(m)) \ll \sqrt{m}$. Since $a_n(m) = 0$ for all $m > n$, the above estimate yields the bound that $\log(A(n)) \ll \sqrt{n}$. □

Since $A(p) = 1$, the only lower bound for $A(n)$ which is valid for all n is the trivial bound $A(n) \geq 1$.

Erdös [14] and [15] has shown that occasionally the coefficients can get very large indeed. More precisely, he showed that $\exists\ c > 0$ such that

$$\log A(n) \gg \exp\left(\frac{c \log n}{\log\log n}\right).$$

Using the refinement of the above argument in Lemma 2.9, Bateman improved the bound in Lemma 2.9 which gives the following bound for the maximum of the absolute values of the coefficients of cyclotomic polynomials:

$$\log A(n) < \left\{\exp\left((\log 2 + o(1))\frac{\log n}{\log\log n}\right)\right\}.$$

The constant $\log 2$, here, is the best possible. This was first asked by P. Erdös [14] and then shown by Vaughan [31].

In 1981, Bateman, Pomerence and Vaughan [4] have refined these results by giving estimates for $A(n)$ in terms of prime factors of n. More recently, Maier [26] showed that for any function $\chi(n)$ tending to infinity, the inequality $A(n) \leq n^{\chi(n)}$ for almost all n.

On the other hand, Maier [24] had earlier proved that, for any function $\epsilon(n)$ defined for all positive integers such that $\epsilon(n)$ tends to zero as n tends to infinity, the inequality $A(n) \geq n^{\epsilon(n)}$ holds except perhaps for a set of positive integers of zero natural density. This settled a long-standing conjecture ($A(n) \to \infty$ for almost all n) of Erdös. Later, he [25] proved that for any $N > 0$, there exists a positive constant $C(N)$ depending on N such that the lower density of the set of n's for which the inequality $A(n) \geq n^N$ is at least $C(N)$. Therefore, Maier's upper bound for $A(n)$ is the best possible one.

In the proof of Lemma 2.9, we first gave an upper bound for $|a_n(m)|$ which is independent of n. More precisely, we proved that $|a_n(m)| \leq p(m)$ where $p(m)$ is the number of partitions of m. Indeed, Möller [28] showed that $|a_n(m)| \leq p(m) - p(m-2)$.

Now we define a dual function (which was first considered by Erdös and Vaughan [16])

$$B(m) = \max_n |a_n(m)|.$$

Note that in the definition of $B(m)$, we can replace maximum by limit supremum. This is because $a_n(m) = a_{npq}(m)$ for all primes p and q with $(n,p) = 1 = (m,q)$ and they are greater than m. Hence from the arguments given in the proof of Lemma 2.9, we can conclude that

$$\log B(m) \ll \sqrt{m}.$$

The first non-trivial result in this direction is due to Erdös and Vaughan [16] who showed that

$$|a_n(m)| < \exp\left\{\left(\tau^{1/2} + o(1)\right) m^{1/2}\right\}$$

uniformly in n as m tends to infinity, where

$$\tau = \prod_p \left(1 - \frac{2}{p(p+1)}\right).$$

They also further showed that for every large m

$$\log B(m) \ll \sqrt{\frac{m}{\log m}}$$

and conjectured that $\log(B(m)) = o(m^{1/2})$.

Vaughan [31] has obtained a sharper bound for infinitely many m; viz.

$$\limsup_{n\to\infty} \left(m^{-1/2}(\log m)^{1/4}\log(B(m))\right) > 0.$$

Montgomery and Vaughan [30] proved the conjecture of Erdös *et al.* in this connection, by proving that $B(m)$ is of exact order $m^{1/2}(\log m)^{-1/4}$.

Recently, Bachman [1] improved the work of Montgomery and Vaughan. He derived the asymptotic formula

$$\log B(m) = C_o \frac{\sqrt{m}}{(\log m)^{1/4}}\left(1 + O\left(\frac{\log\log m}{\sqrt{\log m}}\right)\right).$$

Though some coefficients of cyclotomic polynomials can grow arbitrarily large, it is not still apparent that the collection of all of $a_n(m)$ for all n and m can cover the whole set of integers. This was proved by Jiro Suzuki [34] in 1987.

Theorem 2.10 [34]

$$\mathbb{Z} = \{a_n(k)| \quad k, n \in \mathbb{N}\}.$$

Proof. Let us first prove the following claim. The claim says that if t is any integer greater than 2, then there exist t distinct primes $p_1 < p_2 < \cdots < p_t$ such that $p_1 + p_2 > p_t$.

Assume the contrary, that is, there exists an integer $t > 2$ for which the claim is false. For this t, given any t distinct primes $p_1 < p_2 \cdots < p_t$, we have $p_1 + p_2 \le p_t$. This implies $2p_1 < p_t$. Therefore, for any given integer k, the number of primes between 2^{k-1} and 2^k is always less than t. This is because if we have t distinct primes between 2^{k-1} and 2^k, then we have $p_1 > 2^{k-1} \Longrightarrow 2p_1 > 2^k > p_t$ which is not true by our assumption. Hence the number of primes less than 2^k is $\pi(2^k) < kt$ which is false by prime number theorem, since $\pi(x) > x/\log x$ for all $x \ge 17$. Thus the claim is true.

Now we shall prove the theorem. Let t be any odd positive integer greater than 2. From the above claim, we can find t distinct primes $p_1 < p_2 < \cdots < p_t$ such that $p_1 + p_2 > p_t$.

Let $p = p_t$ and $n = p_1 p_2 \cdots p_t$. Now consider $\Phi_n(X)$. We have, $\Phi_n(X) = \prod_{d|n}(X^d - 1)^{\mu(n/d)}$. We go modulo X^{p+1} and since n is square-free integer, because of the conditions on these set of primes, whenever $d \ne p_i, 1$ for all

$i = 1, 2, \cdots, t$ we have

$$\begin{aligned}\Phi_n(X) &= \prod_{d|n}(X^d-1)^{\mu(n/d)} \equiv \prod_{i=1}^{t}\frac{(X^{p_i}-1)}{(X-1)} \pmod{X^{p+1}}.\\ &\equiv \frac{(1-X^p)}{(1-X)}(1-X^{p_1})\cdots(1-X^{p_{t-1}}) \pmod{X^{p+1}}.\\ &\equiv (1+X+\cdots+X^{p-1})(1-X^{p_1}-\cdots-X^{p_{t-1}}) \pmod{X^{p+1}}.\end{aligned}$$

This yields that $a_n(p) = -t+1$ and $a_n(p-2) = -t+2$. Hence if we let

$$\mathbf{S} := \{a_n(m) \mid \forall\ n, m \in \mathbb{N}\},$$

then, **S** contains $\{\ell \in \mathbb{Z} \mid \ell \leq -1\}$ as t varies over all the odd integer greater than or equal to 3. By Theorem 2.3, already we know $\{0, \pm 1\} \subset \mathbf{S}$. In order to prove that **S** contains all positive integers greater than or equal to 2, consider $\Phi_{2n}(X)$ where $n = p_1p_2\cdots p_t$. By Lemma 1.3(ii), we have $a_{2n}(p) = (-1)^p a_n(p) = t-1$ and $a_{2n}(p-2) = (-1)^{p-2}a_n(p-2) = t-2$. Hence by varying t over all the odd integers ≥ 3, we see that **S** contains all the positive integers greater than or equal to 1. □

Theorem 2.10 says that given any integer k, then there exist natural numbers n and m for which $a_n(m) = k$. In 1991, Grytczuk and Tropak [17] considered the following problem:

Given integer k such that $|k| \geq 2$, find the minimal m for which there exists a natural number n such that $a_n(m) = k$.

If m is one such, then for all n, we must have $a_n(r) \neq k$ for all $r < m$.

For example, if $k = -2$, then we know that $m = 7$ is the minimal integer for which $a_{105}(7) = -2$.

From Lemma 1.2, we know that

$$\Phi_n(X) = \prod_{d|n}(1-X^d)^{\mu(n/d)} = \prod_{d=1}^{\infty}(1-X^d)^{\mu(n/d)}$$

by setting $\mu(n/d) = 0$ whenever n/d is not an integer.

From this identity, it follows that, for a square-free integer n, the value $a_n(m)$ depends only on the values of $\mu(n), \mu(n/d)$ and on the primes less than $m+1$ which happent to divide n.

Using this identity, we can derive a formula for $a_n(m)$ for a fixed m as follows.

$$a_n(1) = -\mu(n), a_n(2) = 1/2\mu(n)(\mu(n)-1) - \mu(n/2)$$

$$a_n(3) = 1/2\mu(n)^2 - 1/2\mu(n) + \mu(n/2)\mu(n) - \mu(n/3), \cdots$$

This method has been used by D. H. Lehmer [20] and H. Möller [28].

A. Grytczuk and B. Tropak [17] derived a recurrence relation for the coefficients of n-cyclotomic polynomial as follows.

$$a_n(m) = -\frac{1}{m}\sum_{\ell=0}^{m-1} a_n(\ell)T_{m-\ell}$$

where $T_{m-\ell} = \mu(n)\mu((n, m-\ell))\phi((n, m-\ell))$ with $a_n(0) = 1$.

Using this recurrence relation, they found for $k = \pm 2, \pm 3, \cdots, \pm 9$ and 10, the minimal values of m for which there exist n such that $a_n(m) = k$.

Acknowledgement: I would like to thank Professor S. A. Katre for carefully going through the manuscript and pointing out several corrections.

References

[1] G. Bachman, *On the coefficients of cyclotomic polynomial*, Mem. Amer. Math. Soc., 106, no. 510 (1993).

[2] A. S. Bang, *Om ligningen* $\Phi_n(X) = 0$, Tidsskrift for Math., 6-12, (1985).

[3] P. T. Bateman, *Note on the coefficients of cyclotomic polynomial*, Bull. Amer. Math. Soc., 55, 1180-1181 (1949).

[4] P. T. Bateman, C. Pomerance and R. C. Vaughan, *On the coefficients of cyclotomic polynomial*, Coll. Math. Soc. J. Bolyai, 34, Topics in classical number theory, 171-202, Budapest, (1981).

[5] M. Beiter, *The midterm coefficient of the cyclotomic polynomial* $\Phi_{pq}(X)$, Amer. Math. Monthly, 71, 769-770 (1964).

[6] M. Beiter, *Magnitude of the coefficients of the cyclotomic polynomials* $\Phi_{pqr}(X)$, Amer. Math. Monthly, 75, 370-372 (1968).

[7] M. Beiter, *Magnitude of the coefficients of the cyclotomic polynomials* $\Phi_{pqr}(X)$ *II*, Duke Math. Jour., 38(3), 591-594 (1971).

[8] M. Beiter, *Coefficients in the Cyclotomic polynomial for numbers with at most three distinct odd primes in their factorization*, The catholic university of America Press, Washington, (1960).

[9] D. M. Bloom, *On the coefficients of the cyclotomic polynomials,* Amer. Math. Monthly, 75, 372-377 (1968).

[10] W. Bosma, *Computation of cyclotomic polynomials with Magma*, in: W. Bosma et al (eds.) Computational algebra and Number Theory, Netherlands: Kluwer Academic Publishers, 216-225 (1995).

[11] L. Carlitz, *The number of terms in the cyclotomic polynomial* $\Phi_{pq}(X)$, Amer. Math. Monthly, 73, 979-981 (1966).

[12] M. Endo, *On the coefficients of the cyclotomic polynomials,* Comment. Math. Univ. St. Pauli, 23, 121-126, (1974/75).

[13] P. Erdös, *On the coefficients of the cyclotomic polynomials,* Bull. Amer. Math. Soc., 52, 179-181, (1946).

[14] P. Erdös, *On the coefficients of the cyclotomic polynomials,* Portugal. Math. 8, 63-71 (1949).

[15] P. Erdös, *On the growth of the cyclotomic polynomials in the interval* $(0, 1)$*,* Proc. Glasgow Math. Assoc. 3, 102-104 (1957).

[16] P. Erdös and R. C. Vaughan, *Bounds for* r*-th coefficients of cyclotomic polynomials*, J. London Math. Soc.(2) 8, 393-400 (1974).

[17] A. Grytczuk and B. Tropak, *A numerical method for the determination of the cyclotomic polynomial coefficients*, in: A. Pethö et al (eds.), Computational Number Theory, Berlin: de Gruyter, 15-19 (1991).

[18] G. H. Hardy and S. Ramanujan, *Asymptotic formulae in combinatorial analysis*, Proc. London Math. Soc. 17, 75-115 (1918).

[19] T. Y. Lam and K. H. Leung, *On the cyclotomic polynomial* $\Phi_{pq}(X)$, Amer. Math. Monthly, 562-564 (1996).

[20] D. H. Lehmer, *Some properties of cyclotomic polynomials*, J. Math. Anal. Appl., 15, 105-117 (1966).

[21] E. Lehmer, *On the magnitude of coefficients of the cyclotomic polynomials*, Bull. Amer. Math. Soc., 42 (1936).

[22] H. W. Lenstra, Jr, *Vanishing sums of roots of unity*, Proc. Bicentennial Congress Wiskundig Genootshapp. Part II, Math. Centre Tracts, 101, Math. Centrum, Amsterdam, 249-268 (1979).

[23] W. LeVeque, *Topics in number theory*, Vol 1, Addison-Wesley, Reading, Mass., (1956).

[24] H. Maier, *The coefficients of cyclotomic polynomials*, Analytic number theory, Proc. of a Conf. in Honor of P. T. Bateman, Prog. Math. 85, 349-366 (1990).

[25] H. Maier, *Cyclotomic polynomials with large coefficients*, Acta Arith., 64, 227-235 (1993).

[26] H. Maier, *The size of the coefficients of cyclotomic polynomials*, Analytic number theory, Proc. of a Conf. in Honor of P. T. Bateman, Prog. Math. , 633-639 (1995).

[27] A. Migotti, *Aur Theorie der Kreisteilungsgleichung*, Z. B. der Math.-Naturwiss, Classe der Kaiserlichen Akademie der Wissenschaften, Wien, 87, 7-14 (1883).

[28] H. Möller, *Über die i-ten Koeffizienten der Kreisteilungspolynome*, Math. Ann., 188, 26-38 (1970).

[29] H. Möller, *Über die Koeffizienten des n-ten Kreisteilungspolynome*, Math. Z., 119, 34-40 (1971).

[30] H. L. Montgomery and R. C. Vaughan, *The order of the m-th coefficients of cyclotomic polynomials*, Glasgow Math. J., 143-159 (1985).

[31] R. C. Vaughan, *Bounds for the coefficients of cyclotomic polynomials*, Michigan Math. J. 21, 289-295 (1975).

[32] R. C. Vaughan, *Adventures in arithmetick, or: How to make good use of a Fourier transform*, The Math. Intelligencer, 9(2), 53-60 (1987).

[33] I. Schur, *Letter to Landau* (1935). (see [21]).

[34] J. Suzuki, *On coefficients of cyclotomic polynomials*, Proc. Japan Acad., 63, Ser. A, 279-280 (1987).

R. Thangadurai
Mehta Research Institute
Chhatnag Road, Jhusi
Allahabad 211 019, India.

Current Address:
Institute of Mathematical Sciences
Tharamani
Chennai 600 113
e-mail: thanga@imsc.ernet.in

Overview and Interconnections

Dinesh S. Thakur

In this workshop, we went through several topics, probably rapidly for those who saw them for the first time. So now, we will just take an overview of what was done, see how different topics were connected with each other, what the main tools were, and mention briefly a few more simple applications and techniques which we could not cover earlier for the lack of time.

Starting with our three motivating problems mentioned in the first lecture, we were led to the study of number fields, and in particular to quadratic, cyclotomic and Kummer extensions: Quadratic reciprocity found natural proofs coming from comparison of factorization laws in quadratic and cyclotomic extensions, regular polygons could be constructed by studying the relevant cyclotomic extensions as successive quadratic ones and Fermat equation could be handled (for regular primes) via study of cyclotomic and Kummer extensions.

Basic structure theorems of number theory (unique factorization of ideals, finiteness of class group, structure theorems for the unit group) were provided, in the lectures on Dedekind domains. We saw usefulness of the basic tools of localization and completions. We saw Kronecker's theorem on how the prime decomposition in an extension can (essentially) be decided by factoring a polynomial modulo that prime. We studied structures of $\mathbf{Q}_p$, $\mathbf{Q}_p^*$, $\mathbf{Z}_p$ etc., classified unramified and totally ramified extensions, studied the concepts of decomposition and inertia groups, Frobenius, discriminant, ramification, prime splitting, its reformulation in terms of Galois groups, using Frobenius conjugacy classes (elements in case of abelian extensions). The existence of nice decomposition laws for quadratic, cyclotomic and Kummer extensions was explained by unifying feature of having abelian Galois group in the class field theory lectures.

We know that $\mathbf{C}$ has no finite non-trivial extension, while $\mathbf{R}$ has $\mathbf{C}$ as its only non-trivial finite extension and $\mathbf{Q}_p$ has many finite extensions, but all with solvable Galois groups (we also saw in problem sessions that for odd p, there is no extension of $\mathbf{Q}_p$ with the Galois group $(\mathbf{Z}/p\mathbf{Z})^3$, as a corollary to local class field theory and structure theorem of $\mathbf{Q}_p^*$ (see also the proof of Lemma 14.8 in Washington using Kummer theory)) and only finitely many of a given degree, in contrast to the global case of $\mathbf{Q}$. As a simple corollary to the determination of the Galois group of $\mathbf{Q}(\zeta_n)$ and structure theorems for abelian groups and for $(\mathbf{Z}/n\mathbf{Z})^*$, we saw that all finite abelian groups occur as Galois groups of cyclotomic extensions over $\mathbf{Q}$. The Kronecker-Weber theorem proved the converse. It is conjectured that every finite group occurs as a Galois group of an extension of $\mathbf{Q}$.

We saw that we have nice tools such as Hensel's lemma (finding a p-adic root by successive Newton approximation method) to solve equations in local fields and then looked at local-global principle example of Hasse-Minkowski theorem. The obstruction to local-global principle (the class group in the case of units and the Tate-Shafarevich group in the case of elliptic curves) is an important object of study.

We looked at zeta and L-functions encoding unique factorization in their simple sum and product descriptions. The special values and leading terms of these simply defined functions encode very interesting arithmetic information: Bernoulli numbers, class number formula that residue at $s = 1$ of $\zeta_K(s)$ is $2^{r_1}(2\pi)^{r_2}hR/w_K\sqrt{|d|}$. The idea of the proof was that the defining sum over ideals for the zeta decomposes into h equal ideal classes contributions, each computed by looking at the limiting sums as integrals and computing the resulting volumes, as in the Dedekind domains lectures.

The regulator R, which is a certain determinant of logarithms of the units is usually hard to compute. But for $\mathbf{Q}$ or an imaginary quadratic field, the rank zero unit group leads to a trivial regulator. Since the Riemann zeta function has a simple pole with residue 1 at $s = 1$, the factorization into L-functions simplifies the left side of the class number formula, e.g., in the case of quadratic fields it becomes $L(1, \chi)$, where χ corresponds to the field.

As an application, the familiar calculation

$$\frac{\pi}{4} = \tan^{-1}(1) = \int_0^1 \frac{dx}{1+x^2} = \int_0^1 (1 - x^2 + x^4 + \cdots)\, dx = 1 - \frac{1}{3} + \frac{1}{5} - \cdots$$

which calculates the L-function for the quadratic character corresponding to $\mathbf{Q}(i)$, implies that its class number is 1 by the class number formula. For $\mathbf{Q}(\sqrt{-5})$, the L-value is $1+1/3+1/7+1/9-1/11-\cdots$ which is approximately 1.4, whereas the right side of the formula is $h\pi/\sqrt{20}$, which is approximately $.7h$ giving $h = 2$, (since h has to be an integer), as we had verified by ideal manipulations using Minkowski bounds, in the problem sessions. Such approximate calculations also allow to calculate exactly ($2h$-th power of) the fundamental unit ϵ of the real quadratic field: More precisely the class number formula gives $\epsilon^{2h} = e^{\sqrt{d}L(1,\chi)}$ which can be calculated approximately, leading to the exact calculation of (trace) integer $A = \epsilon^{2h} + \epsilon^{-2h}$, so that ϵ^{2h} is a root of $x^2 - Ax + 1 = 0$.

It should be mentioned that using character sum manipulations, class number formula simplifies further.

Thus, for the imaginary quadratic fields of discriminant $d = -p$, where p is a prime, we have $h = R_p - N_p$ or $(R_p - N_p)/3$ according as $p \equiv 7$ or $p \equiv 3$ modulo 8, where R_p (respectively N_p) is the number of quadratic residues (respectively non-residues) modulo p within 1 to $(p-1)/2$. No simpler proof of even $R_p > N_p$ is known.

In the case of cyclotomic fields, comparing the class number formulas for K and K^+, we cancel out regulators and get a nice useful formula for the ratio h^-.

The class field theory, the powerful theorems of which we just stated and illustrated, sets up a useful correspondence between the class fields (i.e., the finite abelian extensions) and ideal groups. The simplest and yet powerful example is the Hilbert class field, that is the class field corresponding to the principal ideal group. By the basic properties of the correspondence, we see that it is the maximal abelian everywhere unramified extension, its Galois group is isomorphic to the class group (via the Frobenius map, and hence) exactly the principal primes split in it. Another important theorem about it is the principal ideal theorem, which says that every ideal in the ground field becomes principal in it (but of course it may have non-principal ideals). So we may try to get a tower of Hilbert class fields, and if it stops (i.e., if at some stage the Hilbert class field has class number one), then we get rid of the class number problem, as all the relevant ideals are now principal, at the expense of going to an extension. This sometimes works, but quite often the Hilbert class field tower is infinite. (We do not know a good 'if and only if' criterion).

If we are just interested in making every ideal principal in an extension, there are easier ways: Take I_i be ideal representatives of the ideal classes giving a basis for the class group of a number field F and let o_i be the order of the class corresponding to I_i. Let $I_i^{o_i} = (\alpha_i)$ and let β_i be an o_i-th root of α_i. Then $F(\{\beta_i\})$ is a degree h_F extension in which all ideals of F become principal. But in general this is not the Hilbert class field and we loose other interesting properties that the Hilbert class field has.

We have seen, as an application of Minkowski's discriminant bounds, that the maximal unramified extension of $\mathbf{Q}$ is $\mathbf{Q}$ itself. Again it is not known when exactly the maximal unramified extension is finite or infinite. Infinite class field tower examples lead to infinitely many such examples. For example, any $\mathbf{Q}(\sqrt{d})$, with d being square free and product of 8 or more distinct primes, has infinite class field tower. As an application of Odlyzko discriminant bounds, we can see for example that the maximal unramified extension (and the Hilbert class field) of $\mathbf{Q}(\sqrt{-5})$ is $\mathbf{Q}(\sqrt{-5}, \sqrt{5})$. (Exercise: Verify this by calculating basis of algebraic integers, discriminants and using Odlyzko discriminant bounds. Without Odlyzko bounds or class field theory, verify also that every ideal becomes principal (only one ideal needs to be checked) and with a little more work, verify that it is of class number one, in fact. Verify also that the recipe above need not give the Hilbert class field in this case. (See Washington 11.4 and exercise 11.2). In this case, we could verify existence (and determination) of the Hilbert class field without class field theory ideas. In general, it seems difficult.

Kummer congruences on zeta values at negative integers lead to p-adic interpolation of the zeta function. The Kummer congruences can be thought of as reflection of Fermat little theorem congruences on $n^{p^{k-1}(p-1)} \equiv 1$ modulo p^k, if we think that even after analytic continuation the zeta values retain formal formula $\zeta(-k) = \sum n^k$. The integration approach (for other approaches see Washington or the article 'Modular forms and related objects' by Harold Stark in CMS Conf. Proc., 7) to the interpolation that we saw justifies this intuition. In fact, before the concept of analytic continuation was clear, Euler used this idea to give heuristic calculation of $\zeta(1-k)$, for $k > 0$:

$$\zeta(1-k) = \sum n^{k-1} = (d/dt)^{k-1} \sum e^{nt}|_{t=0} = -B_k/k$$

with the last equality implied by the fact that the geometric sum $1/(1-e^t)-1$ is basically the generating function for the Bernoulli numbers (or rather $B_k/k!$ depending on how you normalize).

Iwasawa theory lectures explained, how analogies (see 7.4 and 13.6 of Washington or Iwasawa's original paper for more) between number fields and function fields motivate the Iwasawa theory, its main conjecture, conjecture about $\mu = 0$ in the class number growth formula. In this original analogy, since the constant field extension tower of a function field is obtained by adjoining roots of unity, we also look at such a tower over a number field. There are different kind of analogies and ideas back and forth between Iwasawa theory and Carlitz-Hayes-Drinfeld explicit cyclotomic theory over function fields leading to a quite active interesting area of research. (We refer to two surveys on these aspects: J. Algebra 81 (1983), 107-149 by David Goss and Contemporary Math. 174 (1994), 157-165, Ed. Jones J. and Childress N., by me).

It is usually hard to get information on class numbers. But, by his simple structure theorem Iwasawa was able to get precise growth estimate $e_n = \lambda n + \mu p^n + \nu$ for n large, for the largest power p^{e_n} dividing the class number at the n-th layer in any $\mathbf{Z}_p$-extension. Now $e_n = 0$ for the $\mathbf{Z}_p$-extension of $\mathbf{Q}$ (and for any base F which has only one prime above p and class number not divisible by p). But in fact, in the traditional function field analogy, the constant field p^n- extension tower is the analogue of the cyclotomic tower, as it is also obtained by adjoining roots of unity. It is easy to calculate class number growth there (Washington 7.4), because of Weil's results on Frobenius eigenvalues. This led Iwasawa to conjecture by analogy that $\mu = 0$ for the cyclotomic $\mathbf{Z}_p$-extension for any number field base. Ferrero and Washington (and later Sinnott) proved this for abelian extensions of $\mathbf{Q}$. Another conjecture, which comes through this analogy, is that λ remains bounded, if we vary p (for cyclotomic $\mathbf{Z}_p$-extensions) over a fixed number field and known only for the base $\mathbf{Q}$, as λ is identically 0 then. Greenberg conjectures more generally that λ is identically 0 for cyclotomic $\mathbf{Z}_p$-extension over totally real number fields.

We encountered two other important open problems: The *Leopoldt conjecture* on independence of units in the p-adic context says that p-adic regulator is non-zero or equivalently that there are exactly $r_2 + 1$ (unconditionally, the number is between $r_2 + 1$ and $r_1 + 2r_2$) independent $\mathbf{Z}_p$ extensions for any number field. The conjecture is known (Washington 5.32) for an abelian extension of $\mathbf{Q}$ (or of an imaginary quadratic field), as an application of p-adic transcendence theory. The *Kummer-Vandiver conjecture* states that p does not divide h^+, the class number of $\mathbf{Q}(\zeta_p)^+$.

It has many important *consequences*: Since it implies that all even components of A, the p-sylow subgroup of the class group of $\mathbf{Q}(\zeta_p)$ are zero, by the reflection theorem we proved, we get that the odd components are cyclic. Combining the annihilation information in Herbrand theorem with the p-adic class number formula, we then see that the i-th component is then isomorphic to $\mathbf{Z}_p/B_{1,w^{-i}}\mathbf{Z}_p$ for odd i between 3 and $p-2$. That the size of both the sides is the same is a hard (unconditional) theorem of Mazur-Wiles as a consequence of their proof of main conjecture of Iwasawa theory. (Another proof is due to ideas of Thaine, Kolyavagin and Rubin). The cyclicity is still unknown unconditionally. Note that even this size comparison proves the Ribet's converse to Herbrand that we saw. In fact, the Ribet's theorem was one of the starting points to Mazur-Wiles. Similarly, we see as implication of Vandiver that the odd part A^- of A is isomorphic to R^-/I^- as R-modules (again the index formula we saw, follows by comparing the sizes of two isomorphic objects), so that the Stickelberger ideal gives all the relations in A^-. Vandiver also implies this for the whole cyclotomic tower, it gives the full main conjecture as easy implication. Some important known quasi-isomorphisms in the theory become isomorphisms under Vandiver hypothesis. We refer to Washington 10.3 for more details.

Note that given even i, since p does not divide B_i for large enough p, Herbrand's theorem implies that $(p-i)$-th component of A vanishes, for large enough p. A similar consequence was explained using some ideas from K-theory, for odd i.

We saw how useful the Gauss and Jacobi sums are: Study of their factorization leads to Stickelberger theorem on ideal class annihilators for the cyclotomic fields. They can be used for power reciprocity proofs.

We saw reciprocity laws in various general contexts. Here is a nice application to the Fermat equation. We first state the *Artin-Hasse explicit reciprocity* law for p-th power residue symbol $(\alpha/\beta) \in \mu_p$:

With $\zeta = \zeta_p$, $\lambda = 1 - \zeta$ as usual, for $\alpha, \beta \in \mathbf{Z}[\zeta]$, $(\alpha, \beta) = 1$ and $\alpha \equiv 1$ modulo λ and $\beta \equiv 1$ modulo p, we have $(\beta/\alpha)(\alpha/\beta)^{-1} = \zeta^{tr(\eta)}$ where $\eta = (\beta - 1)(\alpha - 1)/(p\lambda)$. Here tr denotes the trace from $\mathbf{Q}(\zeta)$ to $\mathbf{Q}$. Note that for $p = 2$, we have $\lambda = 2$ and this reduces to the usual statement of quadratic reciprocity.

Application: If we have a first case solution to the p-th Fermat equation, i.e., $x^p + y^p = z^p$, $(x, y, z) = 1$ and xyz is not divisible by p, then $q^{p-1} \equiv 1$ modulo p^2 for any prime q dividing xyz. (The special case $q = 2$ is called *Wieferich criterion*. Only primes $p < 3 \times 10^9$ which satisfy $2^{p-1} \equiv 1$ modulo p^2 are 1093 and 3511. So the first case follows for the rest).

Proof: By assumption, $x + \zeta^i y = I_i^p$, for ideal I_i. So $\alpha := 1 - y\lambda/(x+y) \equiv 1\ (\lambda)$ is a p-th power of a fractional ideal I and hence $(\beta/\alpha) = (\beta/I)^p = 1$, for any β prime to α. In particular, put $\beta := q^{p-1} \equiv 1\ (p)$. Without loss of generality, $q|y$ and so $\alpha \equiv 1\ (q)$. Hence $(\alpha/q) = 1$, so that $(\alpha/\beta) = 1$ and by the reciprocity law, we get $(q^{p-1} - 1)tr((\alpha - 1)/\lambda)/p \equiv 0\ (p)$. But $tr((\alpha - 1)/\lambda) = -(p-1)y/(x+y) \not\equiv 0\ (p)$, proving the claim.

Urge for more and more general and refined reciprocity laws led to so-called *non-abelian class field theory and Langlands program.* We saw that in the abelian case, irreducible representations of the Galois groups lead to characters. In the next stage, we look at Galois group representations with values in $Gl_2(\mathbf{C})$ rather than $Gl_1(\mathbf{C}) = \mathbf{C}^*$ i.e., move from the commutative domains of numbers to the non-commutative domain of 2 by 2 matrices. It turns out that there are modular forms of weight one with q-expansion $\sum a_n q^n$ with a_p being the trace of Frobenius at p viewed as the corresponding matrix. So, for example, p splits in the corresponding extension, if $p = 2$. So the generalized congruence conditions for splitting in the class field theory get replaced by such conditions governed by modular forms. (The analogy is clearer in the adelic setting: The automorphic forms (closely related to modular forms) are then Gl_2 analogs of Gl_1 (Hecke) characters). The modular forms are still manageable interesting objects: They form a finite dimensional space.

As for the eventual proof of the Fermat's last theorem by Wiles, the techniques went way beyond the cyclotomic theory, but nonetheless there is *historical continuation of motivation and techniques* from the cyclotomic theory:

In Kummer's approach, as we saw, hypothetical non-trivial solution to the Fermat equation for the exponent p gave a Kummer extension of degree p of $\mathbf{Q}(\zeta_p)$ which was unramified everywhere and hence could not exist for the regular primes. In (Hellagouarch, Frey, Serre, Ribet and Taylor-) Wiles approach, the hypothetical non-trivial solution gives rise to an elliptic curve (with discriminant essentially a p-th power) whose p-torsion points give a field extension with Galois group inside $Gl_2(\mathbf{Z}/p\mathbf{Z})$, unramified outside 2 and p and only 'mildly' ramified at p etc.. (In fact, as we saw using the Tate curve, this ramification analysis boiled down to that for a Kummer extension of a cyclotomic one). We used class field theory to rule out the extension in the Kummer case, for regular primes. Now we use non-abelian class field theory (and modular forms mentioned above) to rule this extension, for all primes. The problem is that non-abelian class field theory is not well-developed yet (just as Kummer did not have class field theory at his disposal and had to take detours) to have an easy classification doing the job, so Wiles had to invent and use a lot of techniques to do the job. (In fact, the crucial non-abelian part of Langlands program that got used (due to Langlands and Tunnel) is still in 'solvable' domain, done using class field theory, but the point is that the correct framework of ideas it provides helps).

To continue listing inputs from cyclotomic theory ideas to the techniques and motivations, we first note that Eichler, Shimura, Deligne, Serre's works connecting Galois representations and modular forms, Ribet's proof of converse to Herbrand theorem are some of the starting points of the circle of ideas used eventually. Next, Hida (and Mazur) developed Iwasawa theory in this non-abelian context: Just as Iwasawa theory takes the advantage of structural simplicity once we pass to inverse limit over the tower, they studied the liftings of the representations in rings of matrices with entries in large p-adic rings where we can deform them and use geometric techniques. We looked at Kolyavagin's Euler system argument to control class groups, class number formulas, main conjecture etc. These have their counterparts in elliptic curves theory, which got used. (See Washington's article 'Number fields and elliptic curves' and 'Modular forms and Fermat's last theorem' volume for a good exposition of related ideas).

Dinesh Thakur
Department of Mathematics
University of Arizona
Tucson, AZ 85721
USA
e-mail: thakur@math.arizona.edu

Bibliography

1. S. D. Adhikari, *An Introduction to Commutative Algebra and Number Theory*, Narosa Publishing House, 1999.

2. Tom M. Apostol, *Modular Functions and Dirichlet series in Number Theory*, GTM **41**, Springer-Verlag, Berlin - Heidelberg - New York, 1990.

3. A. O. L. Atkin and J. Lehner, *Hecke operators on* $\Gamma_0(m)$, Math. Ann. **185** (1970), 134–160.

4. Z. Borevich and I. Shafarevich, *Number Theory*, Academic Press, 1966.

5. J. W. S. Cassels and A. Fröhlich (Eds.), *Algebraic Number Theory*, Academic Press, 1967.

6. G. Cornell, J. H. Silverman and G. Stevens (Eds.), *Modular Forms and Fermat's Last Theorem*, Springer, 2000.

7. David A. Cox, *Primes of the form* x^2+ny^2, John Wiley & Sons, (1989).

8. F. Diamond, H. Darmon, and R. Taylor. *Fermat's last theorem*, pages 1–154, Current developments in mathematics. International Press, Cambridge, MA, 1997.

9. Jody Esmonde and M. Ram Murty, *Problems in Algebraic Number Theory*, Springer-Verlag, 1999.

10. Gerd Faltings, Bruce W. Jordan, *Crystalline cohomology and* GL(2, **Q**), Israel J. Math. 90 (1995), no. 1-3, 1–66.

11. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Fourth Ed., Cambridge Uni. Press, 1975.

12. H. Hida, *Elementary theory of L-functions and Eisenstein series*, volume 26 of London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 1993.

13. K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, GTM 2nd edition, Springer-Verlag, New York Inc., 1990.

14. H. Iwaniec, *Topics in classical automorphic forms*, Graduate Studies in Mathematics **17**, American Mathematical Society, Providence, RI, 1997.

15. K. Iwasawa, *Lectures on p-adic L-functions*, Annals of Math. Studies **74**, Princeton University Press, Princeton, 1972.

16. Gerald J. Janusz, *Algebraic Number Fields*, Graduate Studies in Mathematics, Vol. **7**, Second Ed., American Mathematical Society, (1996).

17. N. I. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Graduates Texts in Mathematics **97**, Springer-Verlag, Berlin - Heidelberg - New York, 1984.

18. N. I. Koblitz, *p-adic numbers, p-adic analysis and zeta-functions* GTM **58**, Second Ed., Springer Verlag. 1984.

19. Serge Lang, *Introduction to Modular Forms*, Grundl. Math. Wiss. **222**, Springer - Verlag, Berlin - Heidelberg - New York, 1976.

20. Serge Lang, *Cyclotomic Fields I and II* GTM **121**, Springer-Verlag, 1990.

21. Serge Lang, *Algebra*, 3^{rd} Ed. (Addison - Wesley), 1994.

22. Serge Lang, *Algebraic Number Theory*, GTM 110, Springer-Verlag, Second Edition, 1994.

23. W. -W. Li, *Newforms and functional equations*, Math. Ann. **212** (1975), 285 315.

24. Daniel A. Marcus, *Number Fields*, Springer-Verlag, 1977.

25. B. Mazur and A. Wiles, *Class fields of abelian extensions of* **Q**. Invent. Math. **76** (1984), no. 2, 179–330

26. Milne, *Etale Cohomology*, Princeton Univ Press, 1980.

27. T. Miyake, *Modular Forms*, Springer–Verlag, 1989.

28. Mumford, *The Red Book of Varieties and Schemes*, Springer-Narosa, 1995.

29. Raghavan Narasimhan, S. Raghavan, S.S. Rangachari and Sundar Lal, *Algebraic Number Theory*, TIFR pamphlet, 1966.

30. J. Neukirch, Class Field Theory, Springer-Verlag, 1986.

31. Andrew P. Ogg, *Survey of modular functions of one variable.* Notes by F. van Oystaeyen. Modular functions of one variable, I, pp. 1–35. Lecture Notes in Math. Vol. **320**, Springer, Berlin, 1973. Corrections: Modular functions of one variable, IV, p. 145. Lecture Notes in Math., Vol. **476**, Springer, Berlin, 1975.

32. R. Rankin, *Modular forms and functions*, Cambridge University Press, Cambridge, 1977.

33. Paulo Ribenboim, *13 Lectures on Fermat's Last Theorem,* Springer-Verlag New York, 1979.

34. Kenneth A. Ribet, *A modular construction of unramified p-extensions of* $\mathbf{Q}(\zeta_p)$, Invent. Math. 34 (1976), no. 3, 151–162.

35. J. Rosenberg, *Algebraic K-Theory and Its Applications*, GTM,**147**, Springer-Verlag, Berlin-New York, 1994.

36. K. Rubin, *Appendix to Cyclotomic fields I and II, by Serge Lang*, GTM, **121**, Springer-Verlag, Berlin-New York, 1990.

37. J. -P. Serre, *Groupes algébriques et corps de classes*, Hermann, 1959.

38. J. -P. Serre, *A Course in Arithmetic*, GTM **7**, Springer-Verlag, 1973.

39. J. -P. Serre, *Local fields*, GTM **67**, Springer-Verlag, Berlin-New York, 1979.

40. J. -P. Serre, *Abelian ℓ-adic representations and elliptic curves*, With the collaboration of Willem Kuyk and John Labute. Revised reprint of the 1968 original. Research Notes in Mathematics, 7. A K Peters, Ltd., Wellesley, MA, 1998.

41. Parvati Shastri (Ed.) *Introduction to Class Field Theory*, Lecture Notes of the Instructional School on Algebraic Number Theory, held in the Department of Mathematics, University of Mumbai, December 1994-January 1995.

42. G. Shimura, *Introduction to the arithmetic theory of automorphic functions*, Princeton Univ. Press, Princeton, 1971.

43. G. Shimura, *Abelian varieties with complex multiplication and modular functions*, Princeton Mathematical Series, **46**, Princeton Univ. Press, Princeton, 1999.

44. G. Shimura and Y. Taniyama, *Complex multiplication of abelian varieties and its applications to number theory*, Publications of the Math. Society of Japan, **6**, The Math. Society of Japan, Tokyo, 1961.

45. C. Soulé, *Perfect forms and Vandiver's conjecture*, Preprint, http://www.math.uiuc.edu/K-theory, 1998.

46. V. Srinivas, *Algebraic K-theory*, Second Edition, Progress in Mathematics, **90**, Birkhäuser, 1993.

47. H. M. Stark, *Modular forms and related objects*, Number Theory (Montreal, Que.,1985), 421-455, CMS Conf. Proc., 7, Amer. Math. Soc., Providence, R.I., 1987.

48. H. Stichtenoth, *Algebraic Number Fields and Codes*, Springer-Verlag.

49. J. Tate, Problem 9: The General Reciprocity Law, Proceedings of Symp. in Pure Math. Vol. **28**, 1976, pp. 311-322.

50. R. Taylor and A. Wiles, *Ring theoretic properties of certain Hecke algebras* Ann. of Math. (2), **141** 553–572, 1995.

51. J. Tate, *Les conjectures de Stark sur les fonctions L d'Artin en $s = 0$*, Birkhäuser, 1984.

52. L. C. Washington, *Number fields and elliptic curves*, Number Theory and Applications, 245-278 NATO Adv. Sci. Inst. Ser. C : Math. Sci., **265**, Kluwer Acad. Publ., 1989.

53. L. C. Washington, *Introduction to Cyclotomic Fields*, Second Ed., GTM **83**, Springer-Verlag New York Inc., 1997.

54. W. Wei, *Moduli fields of CM-motives applied to Hilbert's 12th problem*, Preprint, Available on the world wide web at http://www.mathematik.uni-bielefeld.de/sfb343/preprints/pr94070.ps.gz, 1994.

55. Andrew Wiles, *Modular curves and the class group of $\mathbf{Q}(\mu_p)$*, Invent. Math. **58**, no. 1, 1–35, (1980).

56. Andrew Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. (2), **141** no. 443–551, (1995).

The present volume is a collection of refereed articles based on lectures on cyclotomic fields and related topics by several mathematicians working in leading institutions in India and abroad. These lectures were given in a Summer School in Pune organized jointly by Department of Mathematics, University of Pune and Bhaskaracharya Pratishthana, Pune, in June 1999.

The aim of this volume is to provide an accessible reference to an important area of Number Theory for students and teachers of Mathematics. The volume deals with topics in algebraic number theory such as: Reciprocity laws, Fermat's last theorem, Kronecker-Weber theorem, *L*-functions, Stickelberger's theorem, Vandiver's conjecture and Iwasawa theory.

Contributors include:

S.D. Adhikari
Eknath Ghate
Kirti Joshi
S.V. Kanetkar
S.A. Katre
C. Khare
M.J. Narlikar
Nitin Nitsure
Ravi Raghunathan
C.S. Rajan
B. Ramakrishnan
Parvati Shastri
R. Sridharan
R. Sujatha
B. Sury
Dinesh Thakur
R. Thangadurai
C.S. Yogananda