FIELDS OF DIVISION POINTS OF ELLIPTIC CURVES

RELATED TO COATES-WILES

by

Rajiv Gupta

B.Math. University of Waterloo
(1979)

Submitted to the Department of
Mathematics
in Partial Fulfillment of the
Requirements of the
Degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1983

Signature of Author: Signature Redacted
Department of Mathematics, May 23, 1983

Certified by: Signature Redacted
Harold M. Stark, Thesis Supervisor

Accepted by: Signature Redacted
Nesmith C. Ankeny, Chairman, Departmental Committee on Graduate Students

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to my supervisor, Harold Stark. This thesis would not have been possible without his guidance and many valuable suggestions. I would also like to thank my parents for their love and support,and for encouraging me to do what I enjoy. Finally, thanks for financial support to the Natural Sciences and Engineering Research Council of Canada, Imperial Oil Ltd., and M.I.T.

TABLE OF CONTENTS

Acknowledgments ii

Abstract iv

I. INTRODUCTION 1

II. PRELIMINARIES 6

2.1 Elliptic Curves 6
2.2 Complex Multiplication 10
2.3 Eisenstein and Almost-Eisenstein Criteria 15
2.4 Some Polynomial Discriminants 18

III. FIELDS OF DIVISION POINTS 20

3.1 The Field of π^n-Division Points 20
3.2 The Construction of K_n 26
3.3 Character Relations 33

IV. THE CONDUCTOR OF K_n/k_n 39

4.1 A Conductor-Discriminant Relation 39
4.2 Evaluation of the Conductor 41
4.3 Ramification at Some Location 46

V. THE COATES-WILES THEOREM 52

5.1 Elliptic Units 52
5.2 Units and Class Fields 53
5.3 Conclusion 55

References 57

FIELDS OF DIVISION POINTS OF ELLIPTIC CURVES

RELATED TO COATES-WILES

by

RAJIV GUPTA

Submitted to the Department of Mathematics on
May 23, 1983 in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy in
Mathematics

ABSTRACT

The conjectures of Birch and Swinnerton-Dyer predict the behavior of the Hasse-Weil zeta function of an elliptic curve defined over the rational numbers at one. One of the major advances in the study of these conjectures is the result of Coates and Wiles. For elliptic curves with complex multiplication by the ring of integers of a complex quadratic field k of class number one, they show that the zeta function vanishes at one if there is a rational point of infinite order. A version of this result due to Harold Stark is an attempt to understand it from an analytic viewpoint. Stark's proof assumes that for infinitely many primes π in k, a certain abelian extension of the field of π-division points of the elliptic curve is ramified. This assumption is analogous to Wieferich's criterion for the first case of Fermat's last theorem.

We remove this assumption by extending the argument up the tower of fields of division points considered by Coates and Wiles. The analogous abelian extension of the field of π^n-division points is ramified for some n. We determine the least such n and for this n find that the conductor of the corresponding extension is $\mathfrak{P}^2$, where $\mathfrak{P}$ is the unique prime ideal above (π). This follows from a discriminant bound obtained using Kronecker's limit formulas and a discriminant-conductor relation which is a consequence of relations between induced characters on relative Galois groups. The $\mathfrak{P}^2$ conductor result and the structure of certain units (Robert's elliptic units) combine to give the theorem of Coates and Wiles without most of the machinery they use.

Our method works for first and second degree primes and avoids the complications of anomalous primes which Coates and Wiles need to consider. We are also able to calculate conductors at arbitrary locations in the tower.

Thesis Supervisor: Dr. Harold Stark

Title: Professor of Mathematics

CHAPTER ONE

INTRODUCTION

The theory of the Hasse-Weil zeta function of an elliptic curve is replete with conjectures but relatively lacking in general results. Foremost in the state of uncertainty about these zeta functions is the question of where they exist. They are defined by Euler products which converge in a half plane and are not known in general to be continuable to the whole complex plane. Weil has conjectured that the zeta function of an elliptic curve over $\mathbb{Q}$ is in fact a familiar object, namely the Mellin transform of a modular form.

By drawing parallels with zeta functions of number fields, we may expect that the zeta function $L_E(s)$ of an elliptic curve E stores arithmetical information about E in its behavior at special points. In this spirit, and based on extensive numerical evidence, Birch and Swinnerton-Dyer ([5]) in 1965 conjectured that, for E over $\mathbb{Q}$, $L_E(s)$ has a zero at $s = 1$ of order equal to the rank of E. This conjecture is especially remarkable in light of the fact that the half plane where $L_E(s)$ is initially defined is $\{s \mid \mathrm{Re}(s) > 3/2\}$.

For elliptic curves with complex multiplication, many of these conjectures can be replaced by theorems. Building on the classical theory as developed by Kronecker, Weber, and others, Deuring made major advances in the theory of complex multiplication in the 1950's ([7], [8], [9]). In particular, he showed that if E has complex multiplication then $L_E(s)$ has an analytic continuation to the entire complex plane and satisfies a functional equation.

In 1977 Coates and Wiles ([6]) proved

THEOREM. *Suppose* E *has complex multiplication by the ring of integers of an imaginary quadratic field* k *of class number one and that* E *is defined over* k *or* $\mathbb{Q}$. *Then if* E *has positive rank over* k *or* $\mathbb{Q}$, *its zeta function over* k *or* $\mathbb{Q}$ *respectively vanishes at* $s = 1$.

Considering the full statement of the Birch and Swinnerton-Dyer conjecture, this is a relatively weak result but it is the strongest evidence to date in support of the Birch and Swinnerton-Dyer conjecture.

In ([16]), Stark attempts to understand this important result from a more analytic viewpoint. The central idea is the same, namely to show that $L_E(1) = 0$ by showing that some associated number $\widetilde{L_E(1)}$ is divisible by infinitely many primes. Here, $\widetilde{L_E(1)}$ is in k and is zero if and only if $L_E(1)$ is zero. In both versions, $\widetilde{L_E(1)}$ is brought into play by the structure of certain elliptic units, but Stark replaces the formal groups arguments in ([6]) by a simple congruence. The major simplification involves the explicit observation that if, for π a prime in k, the field of π-division points of E has an abelian extension of conductor $\mathfrak{P}^2$ (where $\mathfrak{P}$ is the unique prime above (π)), then π divides $\widetilde{L_E(1)}$. This extension is constructed by dividing a rational point of infinite order by π. Assuming that this extension is ramified, Stark is able to show it has conductor $\mathfrak{p}^2$.

Our goal in this thesis is to remove this major assumption by investigating more carefully the structure of the tower of fields of division points considered by Coates and Wiles. We prove, as do Coates and Wiles, that for some n the extension K_n of the field k_n of π^n-division points is ramified, where K_n is obtained by dividing a rational point of infinite order by π^n. Moreover, we find the conductor of the abelian extensions K_n/k_n for all n and determine the first n for which ramification occurs. For this n, we prove a "conductor $\mathfrak{p}^2$" result which implies that π divides $\widetilde{L_E(1)}$. This happens for all but finitely many π and hence $L_E(1) = 0$.

The methods we use are considerably different from those used by Coates and Wiles. We are able to find our conductors using "conductor-discriminant relations" and certain key polynomial discriminants calculated by Stark using Kronecker's limit formula. Eisenstein and "almost-Eisenstein" criteria play central roles in our arguments. We avoid the machinery of formal groups, Lubin-Tate theory, and local class field theory used by Coates and Wiles. Concerning class field theory, we use only the most elementary facts about ray class fields.

While Coates and Wiles consider only first degree primes, our method allows us to include second degree primes as well. It certainly suffices to consider only first degree primes for the purposes of showing $L_E(1) = 0$, but we believe that any treatment of curves without complex multiplication using these methods will be modelled on the second degree prime case. Most of the time, we are able to treat first and second degree primes simultaneously. We are also able to include the (most likely infinite) set of anomalous primes which Coates and Wiles need to eliminate from consideration.

Following Birch and Swinnerton-Dyer, we deal only with elliptic curves over $\mathbb{Q}$. The reader interested in the full generality of the Coates-Wiles theorem will note that our arguments extend easily to elliptic curves over k. We believe that our methods should extend to elliptic curves (not defined over $\mathbb{Q}$) with complex multiplication by an order of class number greater than one. Nicole Arthaud ([1]) has generalized the Coates-Wiles result slightly, but still with the class number one assumption, and promises to remove the class number one assumption under certain milder assumptions in a subsequent paper. We also generalize the Coates-Wiles theorem somewhat by proving it for elliptic curves over $\mathbb{Q}$ with complex multiplication by a non-maximal order.

Chapter 2 consists mainly of background material needed later. In chapter 3 we construct the fields K_n and k_n and prove some facts we need about them. We also prove a character relation which leads to the crucial conductor-discriminant relation. In chapter 4 we present our main result, the calculation of the conductor of K_n/k_n. Finally, in chapter 5 we prove the Coates-Wiles theorem using the results of chapter 4.

CHAPTER TWO

PRELIMINARIES

In this chapter we present some background material on elliptic curves as well as some facts which will prove useful in the next chapters. We draw special attention to the "almost Eisenstein" criteria of section 2.3 and the calculation of certain polynomial discriminants in section 2.4. The first two sections deal with well known facts about elliptic curves and complex multiplication; for our purposes the important result here is Proposition 2.2.4 which tells us something about the form of multiplication by π.

2.1 Elliptic Curves

We gather some basic facts about elliptic curves. See Tate's excellent survey article ([18]) or Lang ([11]) for details. We confine ourselves to elliptic curves over $\mathbb{Q}$; the general form for such a curve is

$$(2.1.1) \qquad Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$$

where the coefficients are rational. For most purposes it suffices to deal with the simpler Weierstrass normal form

$$(2.1.2) \qquad y^2 = 4x^3 - g_2x - g_3$$

obtained by letting

$$X = x - \frac{a_1^2 + 4a_2}{12} \quad , \quad Y = \frac{1}{2}(y - a_1X - a_3) \quad .$$

The discriminant of (2.1.2) is $\Delta = g_2^3 - 27g_3^2$ which is non-zero.

The equation (2.1.1) for E is unique up to a coordinate change

$$X = u^2X' + r \quad , \quad Y = u^3Y' + su^2X' + t$$

where $u,r,s,t \in \mathbb{Q}$ with $u \neq 0$. Under some coordinate change, it is always possible to find a model (2.1.1) for E such that all the a_i are integers and $|\Delta|$ is minimal. The associated Weierstrass normal form (2.1.2) is unique subject to these restraints and we call it the minimal Weierstrass equation for E; even though g_2 and g_3 may have denominators dividing 6^3, Δ is integral and we call it the discriminant $\Delta(E)$ of E. The j-invariant of E is

$$j(E) = \frac{1728\, g_2^3}{\Delta(E)} \quad .$$

This quantity is indeed invariant under coordinate change and in fact parametrizes isomorphism classes of elliptic curves over an algebraically closed field.

From now on, unless otherwise stated, we take E to be given by a minimal Weierstrass equation. Associated to E is a lattice $\Omega \subset \mathbb{C}$ such that if $\wp(z)$ is the Weierstrass $\wp$-function for this lattice, then the map

$$\lambda : \mathbb{C}/\Omega \longrightarrow E(\mathbb{C}) , \qquad z \longmapsto (p(z), p'(z))$$

is an isomorphism between the torus $\mathbb{C}/\Omega$ and the complex points of E (including the point $\lambda(0) = (\infty , \infty)$). This map allows us to define an addition of points on E. If $(x_1, y_1) = \lambda(z_1)$ and $(x_2, y_2) = \lambda(z_2)$ we let

$$(x_1, y_1) \oplus (x_2, y_2) = (x_1 \oplus x_2, \ y_1 \oplus y_2) = \lambda(z_1 + z_2) \ ;$$

the origin for this group law is (∞ , ∞). The following theorem makes this addition more concrete.

ADDITION THEOREM.

$$p(z_1 + z_2) = -p(z_1) - p(z_2) + \frac{1}{4}\left(\frac{p'(z_1) - p'(z_2)}{p(z_1) - p(z_2)}\right)^2 \ .$$

We thus have $x_1 \oplus x_2$ as a rational function of x_1, x_2, y_1, and y_2. Differentiating the above formula yields $y_1 \oplus y_2$ as a rational function in these quantities as well.

The Addition Theorem shows that the set $E(\mathbb{Q})$ of rational points of E is a subgroup of $E(\mathbb{C})$. In 1925 Mordell proved that $E(\mathbb{Q})$ is finitely generated. This result was generalized by Weil and is known as the

MORDELL-WEIL THEOREM. *If* E *is defined over a number field* K , *the group* $E(K)$ *of* K-*rational points of* E *is finitely generated*.

In fact, Weil showed that this result holds for abelian varieties.

In light of the Mordell-Weil Theorem it makes sense to make

DEFINITION 2.1.3. *The rank* $r(E)$ *of* E *over* $\mathbb{Q}$ *is the rank of* $E(\mathbb{Q})$ *mod torsion*, *i.e.* $E(\mathbb{Q}) \cong \mathbb{Z}^{r(E)} \oplus$ torsion.

We now define the Hasse-Weil zeta function of E.

DEFINITION 2.1.4.

$$L_E(s) = \prod_{p \nmid \Delta(E)} \left(1 - \frac{p+1-N_p}{p^s} + \frac{1}{p^{2s-1}}\right)^{-1} \prod_{p \mid \Delta(E)} \left(1 - \frac{t_p}{p^s}\right)^{-1} .$$

Here, s is a complex variable, the products are over primes, N_p is the number of points on E mod p , and t_p is ± 1 or 0. N_p is one more than (because of the point (∞, ∞)) the number of solutions in $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$ to a minimal equation (2.1.1); for $p \neq 2, 3$ we may use the minimal Weierstrass equation to find N_p. Thanks to Hasse, who proved $|p+1-N_p| \leq 2\sqrt{p}$, it is easy to see that $L_E(s)$ is defined for $\mathrm{Re}\ s > 3/2$. In fact much more is believed.

CONJECTURE 2.1.5. $L_E(s)$ *has an analytic continuation to the whole complex plane and satisfies a functional equation*.

Even though we are not justified in doing so, let us put $s = 1$ in Definition 2.1.4. Ignoring the $\prod_{p \mid \Delta(E)}$ part we get

$$L_E(1) \approx \prod_{p \nmid \Delta(E)} \left(\frac{N_p}{p}\right)^{-1} = \prod \frac{p}{N_p} \quad .$$

Now if $r(E)$ is large, we expect that N_p is large so that $L_E(1)$ is small. Based on similar heuristics and extensive numerical calculations, Birch and Swinnerton-Dyer made in ([5]).

CONJECTURE 2.1.6. $L_E(s)$ _has a zero at_ $s = 1$ _of order_ $r(E)$.

Here we are of course assuming that Conjecture 2.1.5 is true.

An expression for the lead coefficient in the Taylor series about $s = 1$ is formulated in a more general version of this conjecture. This expression involves the order of the Tate-Shaferevitch group, which is conjectured to be finite. This leads Tate to remark ([18]) "This remarkable conjecture relates the behavior of a function at a point where it is not at present known to be defined to the order of a group which is not known to be finite !"

2.2 Complex Multiplication

We now discuss an important special class of elliptic curves which we deal with in this thesis, those with complex multiplication. For more on the rich theory of complex multiplication, see for example Weber ([19]) or Lang ([11]).

DEFINITION 2.2.1. *E has multiplication by* $\beta \in \mathbb{C}$ *if* $\beta\Omega \subset \Omega$. *E has complex multiplication if it has multiplication by some* $\beta \notin \mathbb{Z}$.

Deuring ([8]) proved

THEOREM 2.2.2. *If E has complex multiplication, Conjecture* 2.1.5 *is true.*

Deuring proved this by showing $L_E(s)$ is a Hecke L-function with Grössencharacter. Thus for elliptic curves with complex multiplication it at least makes sense to formulate Conjecture 2.1.6.

We note that E clearly has multiplication by rational integers. If E has complex multiplication (CM) then

$$\Gamma = \{\beta \mid E \text{ has multiplication by } \beta\}$$

is an order in the ring of integers $\mathfrak{O}_k$ of an imaginary quadratic field k and $\Omega = \omega_o \Gamma$ for some $\omega_o \in \Omega$. For some $c \in \mathbb{Z}$, called the conductor of Γ, $\Gamma = \mathbb{Z} + c\delta\mathbb{Z}$ where $\{1, \delta\}$ is an integral basis for $\mathfrak{O}_k$. The j-invariant j(E) depends only on Γ and is in $\mathbb{Q}$ exactly when the class number of Γ is one. There are precisely 13 orders in imaginary quadratic fields with class number one. Besides the 9 maximal orders $\mathfrak{O}_k$ as k ranges over the imaginary quadratic fields of class number one, the orders with conductors 2 and 3 in $\mathbb{Q}(\sqrt{-3})$

and the orders with conductor 2 in $\mathbb{Q}(i)$ and $\mathbb{Q}(\sqrt{-7})$ have class number one.

This means that over the algebraic closure of $\mathbb{Q}$ there are exactly 13 isomorphism classes of elliptic curves definable over $\mathbb{Q}$ with complex multiplication. To decide whether a given elliptic curve has CM, we only need to compute its j-invariant and check whether it is one of the corresponding 13 values. These values, which are in fact not only rational but integral, may be found in Weber ([19]). For example, the curves $y^2 = 4x^3 - 4Dx$ with j-invariant 1728 have CM by the Gaussian integers.

Not surprisingly, given the terminology, if E has multiplication by β we can define a multiplication by β on the points $E(\mathbb{C})$. Namely, we define

$$\beta \odot (x,y) = (\beta \odot x, \beta \odot y) = (\wp(\beta z), \wp'(\beta z))$$

where $(x,y) = (\wp(z), \wp'(z))$. The inclusion $\Omega\beta \subset \Omega$ implies that this multiplication is well-defined; $\wp(\beta z)$ and $\wp'(\beta z)$ are elliptic functions for the lattice Ω.

The theory of the Weierstrass $\wp$-function tells us that $\wp(\beta z)$ is a rational function in $\wp(z)$. Investigation of the poles of $\wp(\beta z)$ in fact yields

LEMMA 2.2.3. Suppose E has multiplication by β. Then

$$\wp(\beta z) = \frac{f(\wp(z))}{g(\wp(z))}$$

where f *and* g *are polynomials of degree* $|\beta|^2$ *and* $|\beta|^2 - 1$ *respectively.*

We rewrite this more suggestively as $\beta \odot x = f(x)/g(x)$. For $|\beta|^2$ odd, the polynomial $g(x)$ has double roots at the points $\wp\left(\frac{\omega}{\beta}\right)$, $\omega \in \Omega \sim \beta\Omega$, so $g(x)$ is a square. We normalize f and g by setting $g(x) = \psi_\beta(x)^2$ where

$$\psi_\beta(x) = \beta \prod_{\omega} \left(x - \wp\left(\frac{\omega}{\beta}\right)\right) .$$

Here, the product ranges over $\omega \in \Omega \sim \beta\Omega$ mod $\beta\Omega$, mod ± 1, where by "mod ± 1" we mean that if ω_1, ω_2 occur in the product then $\omega_1 \not\equiv \pm\omega_2$ mod $\beta\Omega$. The fact that $\wp(z)$ is even means that $\psi_\beta(x)$ is well-defined. With this normalization we have $\beta \odot x = \varphi_\beta(x)/\psi_\beta(x)^2$ where φ_β is monic and φ_β, ψ_β have coefficients in $\mathbb{Q}$ or k according as $\beta \in \mathbb{Z}$ or $\mathcal{O}_k$. For $|\beta|^2$ even it is still possible to define ψ_β so that ψ_β^2 is a polynomial over $\mathbb{Q}$ or k and $\beta \odot x$ is as above.

The following result tells us much more about the polynomials φ_β and ψ_β in the case that β is a prime in $\mathcal{O}_k$. Our future investigations rely heavily upon it.

PROPOSITION 2.2.4. *Suppose* E *has complex multiplication by an order* $\Gamma \subset \mathcal{O}_k$ *and let* $\pi \in \Gamma$ *be a prime of norm* q *relatively prime to* $6\,\Delta(E)$. *Then*

$$\varphi_{\pi}(x) \equiv x^{q} \bmod(\pi)$$

and

$$\psi_{\pi}(x) \equiv \text{constant} \not\equiv 0 \bmod(\pi)$$

where these congruences hold coefficient-wise.

This result is most efficiently proved by invoking the notion of Frobenius automorphisms. With E and π as above we may reduce the minimal Weierstrass equation for $E \bmod(\pi)$; we then have a curve $\overline{E}$ over $\mathbb{F}_q$. The Frobenius automorphism of $\overline{E}$ is

$$\sigma_{(\pi)} : (\bar{x}, \bar{y}) \longmapsto (\bar{x}^{q}, \bar{y}^{q}) \quad .$$

In this setting Deuring ([8]) (see also Lang ([11], p. 138)) proved

THEOREM 2.2.5. *Using a bar to denote reduction* $\bmod(\pi)$, *we have* $\overline{\pi' \odot (x,y)} = \sigma_{(\pi)}(\bar{x}, \bar{y})$ *for some generator* π' *of* (π) *in* Γ.

Remarks. (1) If π is second degree and has norm p^2, then $\pi' = -p$.

(2) Since $\wp(z)$ is even and $\wp'(z)$ is odd, $\overline{(-\pi') \odot (x,y)} = (\bar{x}^{q}, -\bar{y}^{q})$.

Using these remarks we see that $\overline{\pi \odot x} = \bar{x}^{q}$ except possibly when Γ contains units other than ± 1; even then we have

FACT 2.2.6. *If* $\Gamma = \mathfrak{O}_{\mathbb{Q}(i)}$ *then* $\overline{\pi \odot x} = \bar{\rho}\bar{x}^{q}$ *where* $\rho^{2} = 1$ *and if* $\Gamma = \mathfrak{O}_{\mathbb{Q}(\sqrt{-3})}$ *then* $\overline{\pi \odot x} = \bar{\rho}\bar{x}^{q}$ *where* $\rho^{3} = 1$.

This fact may be found in Stark ([15]) where an elementary proof of Theorem 2.2.5 (at least for π first degree) is given.

Proposition 2.2.4 now follows immediately from

$$\frac{\overline{\varphi_\pi(x)}}{\overline{\psi_\pi(x)}^2} = \frac{\bar{x}^q + \cdots}{(\bar{\pi}\bar{x}^{(q-1)/2} + \cdots)^2} = \bar{\rho}\bar{x}^q \quad .$$

This last equality is an equality of rational functions so indeed $\overline{\varphi_\pi(x)} = \bar{x}^q$ and $\overline{\psi_\pi(x)}^2 = \bar{\rho}^{-1}$ (= $\bar{1}$ most of the time).

2.3 Eisenstein and Almost-Eisenstein Criteria

In this section we state for later reference the well-known Eisenstein criterion and prove some elementary facts about polynomials which are "almost-Eisenstein." To inspire our proofs of these almost-Eisenstein criteria we briefly sketch the proof of the Eisenstein criterion. Here, $\|$ denotes "exactly divides" and by the discriminant disc (f) of a polynomial $f(x)$ we mean the discriminant of the associated monic polynomial, which we recall is the square of a van der Monde determinant.

THEOREM 2.3.1 (Eisenstein irreducibility criterion). *Let* k *be a number field,* $\mathfrak{P}$ *a prime ideal in* $\mathfrak{O}_k$, *and* $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_o \in \mathfrak{O}_k[x]$. *Suppose that* $\mathfrak{P} \nmid a_n$, $\mathfrak{P} \mid a_{n-1}, \ldots,$ $\mathfrak{P} \mid a_1$ *and* $\mathfrak{P} \| a_o$ *and let* θ *be a root of* $f(x)$. *Then*

(1) $f(x)$ *is irreducible over* k.

(2) $\mathfrak{P}$ *ramifies totally in* $K = k(\theta)$ - *say* $\mathfrak{P} = \mathfrak{p}^n$.

(3) $\wp \,\|\, \theta$.

(4) *The power of* $\mathfrak{P}$ *in disc* (f) *is the power of* $\mathfrak{P}$ *in the field discriminant* D(K/k).

Proof. We reduce to the case $h(k) = 1$ by localizing at $\mathfrak{P}$; see Lang ([10], p. 65) for justification. Then (1) follows easily and (2) and (3) follow by setting $\wp = (\mathfrak{P}, a_n \theta)$. For (4), the assumption $h(k) = 1$ means there is an integral basis $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ for $\mathcal{O}_K/\mathcal{O}_k$, so there exists an $n \times n$ matrix M over $\mathcal{O}_k$ such that

$$(2,3,2) \qquad (1, a_n\theta, \dots, (a_n\theta)^{n-1}) = (\alpha_1, \alpha_2, \dots, \alpha_n) M .$$

If $g(x)$ has root $a_n\theta$, this leads to

$$(2,3,3) \qquad \operatorname{disc}(g) = D(K/k)(\det M)^2 .$$

The same power of $\mathfrak{P}$ occurs in disc(g) and disc(f). One checks by induction that $\Sigma_{j=0}^{n-1} c_j(a_n\theta)^j \equiv 0 \bmod \mathfrak{P}$ for $c_j \in \mathcal{O}_k$ forces $c_j \equiv 0 \bmod \mathfrak{P}$ for all j, so $\mathfrak{P} \nmid \det M$ which gives (4).

Stark ([17]) noticed the following partial converse to the Eisenstein criterion.

THEOREM 2.3.4. *Let* $k, \mathfrak{P}, f(x), \theta, K$ *be as in the statement of Theorem* 2.3.1 *but now suppose that* $\mathfrak{P} \nmid a_n$, $\mathfrak{P} \mid a_{n-1}, \dots, \mathfrak{P} \mid a_1$ *and* $\mathfrak{P}^2 \mid a_o$ *and that* f(x) *is irreducible. Then the power of* $\mathfrak{P}$ *in* disc(f) *is greater than the power of* $\mathfrak{P}$ *in* D(K/k).

Proof. We consider two cases.

Case 1. $\mathfrak{P}$ does not ramify totally in K. Then all ramification indices of primes above $\mathfrak{P}$ are $\leq n-1$. From

$$\theta^n = -\frac{1}{a_n}(a_{n-1}\theta^{n-1} + \cdots + a_o)$$

we see that $\mathfrak{P} \mid (a_n\theta)^n$, whence $\mathfrak{P} \mid (a_n\theta)^m$ for some $m < n$.

Case 2. $\mathfrak{P} = \mathfrak{p}^n$ in K. Since the norm $N_{K/k}((a_n\theta)) = (a_o a_n^{n-1})$ is divisible by $\mathfrak{P}^2$, we have $\mathfrak{p}^2 \mid a_n\theta$. As above, $\mathfrak{P} \mid (a_n\theta)^m$ for some $m < n$.

Now, as in the proof of Theorem 2.3.1,we may assume by localizing that $\{\alpha_1, \ldots, \alpha_n\}$ is an integral basis for $\mathcal{O}_K/\mathcal{O}_k$ and for some matrix $M \in M_n(\mathcal{O}_k)$ (2.3.2) holds. The m^{th} column of M has entries all divisible by $\mathfrak{P}$ since $\mathfrak{P} \mid (a_n\theta)^m$, so we have $\mathfrak{P} \mid \det M$. Equation (2.3.3) completes the proof.

It is classical that the powers of $\mathfrak{P}$ in disc(f) and D(K/k) differ by an even number. If this difference is 2 we can say more.

THEOREM 2.3.5. *In addition to the assumptions in Theorem 2.3.4, assume that the powers of* $\mathfrak{P}$ *in* disc(f) *and* D(K/k) *differ by 2 and* $n \geq 4$. *Then* $\mathfrak{P}$ *factors as* $\mathfrak{P} = \mathfrak{p}_1^{n-1}\mathfrak{p}_2$ *in* K. *Moreover, if* $\mathfrak{P}^e \parallel a_o$ *then* $\mathfrak{p}_1 \parallel \theta$ *and* $\mathfrak{p}_2^{e-1} \parallel \theta$.

Proof. From the proof of Theorem 2.3.4 we know that $\mathfrak{P} \mid (a_n \theta)^m$ for some $m < n$. But then $\mathfrak{P} \mid (a_n \theta)^{m+1}, \ldots, \mathfrak{P} \mid (a_n \theta)^{n-1}$ and arguing as before we see that $\mathfrak{P}^{n-m} \mid \det M$, whence the powers of $\mathfrak{P}$ in disc(f) and D(K/k) differ by at least $2(n-m)$. This shows that $\mathfrak{P} \mid (a_n \theta)^{n-1}$ but $\mathfrak{P} \nmid (a_n \theta)^{n-2}$, which means some prime above $\mathfrak{P}$ has ramification index at least $n-1$. $\mathfrak{P} = \wp^n$ is ruled out since then $\wp^2 \mid a_n \theta$ implies $\mathfrak{P} \mid (a_n \theta)^m$ where m is the least integer $\geq n/2$ and $m < n-1$. Thus $\mathfrak{P} = \wp_1^{n-1} \wp_2$ where $\wp_1$ and $\wp_2$ have norm $\mathfrak{P}$. Finally, if $\wp_1^{e_1} \| \theta$ and $\wp_2^{e_2} \| \theta$ then e_1, e_2 are positive, $e_1 + e_2 = e$, and $e_1 \geq 2$ contradicts $\mathfrak{P} \nmid (a_n \theta)^{n-2}$. This completes the proof.

2.4 Some Polynomial Discriminants

We state here, with a slight misprint corrected, the following important result due to Stark ([16]):

THEOREM 2.4.1. Suppose M is an odd integer and ν_1, ν_2 range over $M^{-1}\Omega \bmod \Omega$. Then

$$\prod_{\nu_1 \neq \nu_2} (\wp(z+\nu_1) - \wp(z+\nu_2)) = \pm M^{M^2} \wp'(Mz)^{M^2-1} \Delta(E)^{(2M^2-3)(M^2-1)/12}.$$

Also, if E has complex multiplication by a first degree prime π of odd norm p and ν_1, ν_2 run through $\pi^{-1}\Omega \bmod \Omega$, then

$$\prod_{\nu_1 \neq \nu_2} (\wp(z+\nu_1) - \wp(z+\nu_2)) = \pm \pi^p \wp'(\pi z)^{p-1} \Delta(E)^{(2p-3)(p-1)/12} .$$

In the sequel, these formulas will enable us to calculate certain key polynomial and field discriminants. See Stark ([16]) for the pretty analytic proof of this theorem, which is based upon Kronecker's limit formula.

CHAPTER THREE

FIELDS OF DIVISION POINTS

Continuing with the notation of chapter 2, we take E to be an elliptic curve over $\mathbb{Q}$ with complex multiplication. We suppose E is given by the minimal Weierstrass equation $y^2 = 4x^3 - g_2x - g_3$ parametrized by $x = \wp(z)$, $y = \wp'(z)$. Our goal in this chapter is to define certain fields of division points and prove some basic facts about them. In section 1 we introduce k_n, the field of π^n-division points, and calculate the (π)-part of its discriminant. In section 2 we define for positive rank curves an extension K_n of k_n and show it is "big as possible." The last section deals with a character relation which will be crucial for the calculation of the conductor of K_n/k_n in the next chapter.

3.1 <u>The Field of π^n-Division Points</u>

We assume E has complex multiplication by the order Γ in $\mathfrak{O}_k$. Let $\pi \in \Gamma$ be a first or second degree prime of norm $q = p$ or p^2 and assume π does not divide $6\Delta(E)$. If π is second degree, we always take $\pi = p$.

DEFINITION 3.1.1. *The set of π^n-division points of E is*

$$E_{\pi^n} = \{(\wp(\nu), \wp'(\nu)) \mid \nu \in \pi^{-n}\Omega \pmod{\Omega}\} \quad .$$

It is easily checked that there are q^n π^n-division points. For $\nu \in \pi^{-n}\Omega$ we will call $p(\nu)$ an x-π^n-division value and denote a π^n-division point by (x_n, y_n). We have $\pi^n \odot (x_n, y_n) = (\infty, \infty)$ and call (x_n, y_n) primitive if $\pi^{n-1} \odot (x_n, y_n) \neq (\infty, \infty)$ (recall (∞, ∞) is the origin for the group law). The coordinates of finite division points are algebraic; the finite x-π^n-division values are just the roots of the $[(q^n-1)/2]^{th}$ degree polynomial $\psi_{\pi^n}(x)$ introduced in section 2.2.

DEFINITION 3.1.2. *The field of π^n-division points of E, denoted k_n, is k with all coordinates of finite π^n-division points adjoined.*

The field k_n is a normal extension of k whose structure is well known. We say something about the size of k_n in the next lemma - it essentially tells us that k_n is not as big as we might initially suspect.

LEMMA 3.1.3. *Let (x_n, y_n) be a primitive π^n-division point. Then $k_n = k(x_n, y_n)$.*

Proof. First note that x_n gives us all x-π^n-division values. Indeed, if $\alpha_1, \ldots, \alpha_{(q^n-1)/2}$ are representatives for $\Gamma/\pi^n \setminus \{0\}$ mod ± 1, then the numbers $\alpha_i \odot x_n$ are the $(q^n-1)/2$ distinct finite x-π^n-division values, and are in $k(x_n)$ by the facts stated in section

2.2. Next, let $(x_n', y_n') \neq \pm 1 \odot (x_n, y_n)$ be another π^n-division point. By the addition theorem,

$$x_n' \oplus x_n = -x_n' - x_n + \frac{1}{4}\left(\frac{y_n' - y_n}{x_n' - x_n}\right)^2 .$$

Using $y_n^2 = 4x_n^3 - g_2 x_n - g_3$ and similarly for $y_n'^2$, we solve this linearly for y_n' in terms of y_n, x_n, x_n', and $x_n' \oplus x_n$. By the first observation above x_n' and $x_n' \oplus x_n$ are in $k(x_n)$, so $y_n' \in k(x_n, y_n)$. This proves the lemma.

Since $y_n^2 \in k(x_n)$, we get $[k_n : k(x_n)] \leq 2$ and $[k_n : k] \leq q^n - 1$. In fact, the latter degree is even smaller.

THEOREM 3.1.4. k_n <u>is a degree</u> $q^{n-1}(q-1)$ <u>extension of</u> k <u>in which</u> (π) <u>ramifies totally. The relative field discriminant</u> $D(k_n/k)$ <u>has</u> (π)<u>-part</u> $(\pi)^{nq^n - (n+1)q^{n-1}}$.

The first statement of this theorem is well known; Coates-Wiles ([6]) deduce it from Lubin-Tate theory. We present an alternate proof using Eisenstein polynomials. Using Theorem 2.4.1 we are also able to calculate the crucial discriminant. Given an extension M/k we let $D_{(\pi)}(M/k)$ denote the (π)-part of $D(M/k)$. Fix a period $\omega_o \in \Omega \smallsetminus \pi\Omega$ and let $(x_n, y_n) = (\wp(\omega_o/\pi^n), \wp'(\omega_o/\pi^n))$, a primitive π^n-division point. The theorem will follow from

LEMMA 3.1.5. *We have for all* n

(1) $[k(x_n): k] = q^{n-1}(q-1)/2$.

(2) $(\pi) = \mathfrak{B}_n^{q^{n-1}(q-1)/2}$ *ramifies totally in* $k(x_n)$.

(3) $\mathfrak{B}_n^{-1} \parallel x_n$.

(4) $D_{(\pi)}(k(x_n)/k) = (\pi)^{\frac{1}{2}(nq^n - (n+1)q^{n-1} - 1)}$.

(5) $[k_n : k(x_n)] = 2$, $\mathfrak{B}_n = \mathfrak{P}_n^2$ *ramifies in this extension, and* $\mathfrak{P}_n^{-3} \parallel y_n$.

Proof. We proceed by induction. The case $n = 1$ is done in Stark ([16]) but we outline his argument to inspire our proof. x_1 is a root of the polynomial $\psi_\pi(x) = \pi x^{(q-1)/2} + b_1 x^{(q-3)/2} + \cdots + b_{(q-1)/2}$ where $b_i \in k$. We know by Proposition 2.2.4 that

$$\pi \mid b_1, \pi \mid b_2, \ldots, \pi \mid b_{(q-3)/2} \text{ but } \pi \nmid b_{(q-1)/2} .$$

$\psi_\pi(x)$ is *reverse* Eisenstein with respect to (π); x_1^{-1} is the root of an Eisenstein polynomial. By the Eisenstein criterion (Theorem 2.3.1), (1), (2), and (3) hold for $n = 1$. Now $\mathfrak{B}_1^{-1} \parallel x_1$ implies that

$$\mathfrak{B}_1^{-3} \parallel y_1^2 = 4x_1^3 - g_2 x_1 - g_3$$

so y_1 cannot be in $k(x_1)$. Thus $k_1/k(x_1)$ is quadratic and $\mathfrak{B}_1 = \mathfrak{P}_1^2$ in k_1 so that $\mathfrak{P}_1^{-6} \parallel y_1^2$ and $\mathfrak{P}_1^{-3} \parallel y_1$, giving (5). Since $p \nmid [k(x_1): k] = (q-1)/2$ and (π) ramifies totally, we have $D_{(\pi)}(k(x_1)/k) = (\pi)^{(q-3)/2}$, which gives (4) for $n = 1$.

We now take $n > 1$. Then x_n is a solution of

$$\pi \odot x = \frac{\varphi_\pi(x)}{\psi_\pi(x)^2} = x_{n-1}$$

so the polynomial $\varphi_\pi(x) - x_{n-1}\psi_\pi(x)^2$, a q^{th} degree monic polynomial over $k(x_{n-1})$, has root x_n. It will be more convenient to work with $f(x) = x_{n-1}^{-1}\varphi_\pi(x) - \psi_\pi(x)^2$. Using (3) of the induction hypothesis, $\mathfrak{B}_{n-1} \| x_{n-1}^{-1}$, and recalling the information Proposition 2. 2. 4 tells us about $\varphi_\pi(x)$ and $\psi_\pi(x)$, we see that $f(x)$ is reverse Eisenstein with respect to $\mathfrak{B}_{n-1}$. This means $[k(x_n) : k(x_{n-1})] = q$, $\mathfrak{B}_{n-1} = \mathfrak{B}_n^q$ in $k(x_n)$, and $\mathfrak{B}_n^{-1} \| x_n$ so (1), (2), and (3) hold. Arguing as in the case $n = 1$, we see that (5) holds as well. Finally, by (4) of Theorem 2. 3. 1 we know that the $\mathfrak{B}_{n-1}$-part of the discriminant of the polynomial for x_n^{-1} over $k(x_{n-1})$ is equal to the $\mathfrak{B}_{n-1}$-part of $D(k(x_n)/k(x_{n-1}))$, which we denote $D_{\mathfrak{B}_{n-1}}(k(x_n)/k(x_{n-1}))$. The polynomial discriminant is

$$\pm \prod_{\substack{\nu_1, \nu_2 \in \pi^{-1}\Omega \,(\mathrm{mod}\ \Omega) \\ \nu_1 \neq \nu_2}} \left[\frac{1}{\wp\left(\frac{\omega_o}{\pi^n} + \nu_1\right)} - \frac{1}{\wp\left(\frac{\omega_o}{\pi^n} + \nu_2\right)} \right]$$

$$= \pm \frac{\prod_{\nu_1 \neq \nu_2} \left[\wp\left(\frac{\omega_o}{\pi^n} + \nu_2\right) - \wp\left(\frac{\omega_o}{\pi^n} + \nu_1\right) \right]}{\prod_{\nu_1 \neq \nu_2} \wp\left(\frac{\omega_o}{\pi^n} + \nu_1\right) \wp\left(\frac{\omega_o}{\pi^n} + \nu_2\right)} \quad .$$

By Theorem 2.4.1, the numerator of this expression is

$$\pm \pi^q p'\left(\frac{\omega_o}{\pi^{n-1}}\right)^{q-1} \Delta(E)^{(2q-3)(q-1)/12} .$$

As $\mathfrak{P}_{n-1}^{-3} \parallel y_{n-1} = p'(\omega_o/\pi^{n-1})$, we have $(\pi)^q \mathfrak{P}_{n-1}^{-3(q-1)}$ in the numerator. Also, each $p(\omega_o/\pi^n + \nu_i)$ is divisible exactly by $\mathfrak{B}_n^{-1}$ so the product over $q(q-1)$ terms in the denominator has $\mathfrak{B}_n^{-2q(q-1)}$. We thus have

$$\begin{aligned} D_{\mathfrak{B}_{n-1}}(k(x_n)/k(x_{n-1})) &= (\pi)^q \mathfrak{P}_{n-1}^{-3(q-1)} \mathfrak{B}_n^{2q(q-1)} \\ &= (\pi)^q \mathfrak{B}_{n-1}^{-\frac{3}{2}(q-1)} \mathfrak{B}_{n-1}^{2(q-1)} \\ &= (\pi)^q \mathfrak{B}_{n-1}^{(q-1)/2} . \end{aligned}$$

It will be convenient not to simplify further by replacing (π) by $\mathfrak{B}_{n-1}^{q^{n-2}(q-1)/2}$. Using

$$D(k(x_n)/k) = [N_{k(x_{n-1})/k}(D(k(x_n)/k(x_{n-1})))] \cdot D(k(x_{n-1})/k)^q ,$$

we have

$$\begin{aligned} D_{(\pi)}(k(x_n)/k) &= [N_{k(x_{n-1})/k}((\pi)^q \mathfrak{B}_{n-1}^{(q-1)/2})]\left[(\pi)^{\frac{1}{2}[(n-1)q^{n-1}-nq^{n-2}-1]}\right]^q \\ &= (\pi)^{q^{n-1}(q-1)/2} (\pi)^{(q-1)/2} (\pi)^{\frac{1}{2}[(n-1)q^n - nq^{n-1} - q]} \\ &= (\pi)^{\frac{1}{2}(nq^n-(n+1)q^{n-1} - 1)} . \end{aligned}$$

This gives (4) and completes the proof of the lemma.

Proof of Theorem 3.1.4. All that is left to do is to compute the discriminant. Statement (5) of Lemma 3.1.5 shows that $D_{\mathfrak{B}_n}(k_n/k(x_n)) = \mathfrak{B}_n$ so

$$D_{(\pi)}(k_n/k) = [N_{k(x_n)/k}(\mathfrak{B}_n)]\left[(\pi)^{\frac{1}{2}(nq^n-(n+1)q^{n-1}-1)}\right]^2$$

$$= (\pi)^{nq^n-(n+1)q^{n-1}} .$$

This completes the proof.

Remark. If Γ is maximal and $k \neq \mathbb{Q}(i)$ or $\mathbb{Q}(\sqrt{-3})$ it is known that $k(x_n)$ is the ray class field of $k \bmod (\pi)^n$. This knowledge could be used to give an alternate calculation of $D_{(\pi)}(k_n/k)$.

The Galois group $G(k_n/k)$ is just $(\Gamma/(\pi)^n)^{\times} \cong (\mathcal{O}_k/(\pi)^n)^{\times}$, a group of order $q^{n-1}(q-1)$. The action is given by

$$x_n \circ \alpha = \alpha \odot x_n , \quad y_n \circ \alpha = \alpha \odot y_n .$$

3.2 The Construction of K_n

We now suppose that E with CM by $\Gamma = \mathbb{Z} + c\delta\mathbb{Z}$ (where $\{1, \delta\}$ is an integral basis for $\mathcal{O}_k/\mathbb{Z}$) has rank $r > 0$. Take π to be as in section 3.1 with the added restraint $\pi \nmid D(k/\mathbb{Q})$ - in fact this already follows from $\pi \nmid 6\Delta(E)$. Let $P_1, \ldots, P_r \in E(\mathbb{Q})$

be a basis for $E(\mathbb{Q})$ modulo torsion.

LEMMA 3.2.1. <u>There does not exist a point</u> $P \in E(k)$ <u>such that</u> $\pi \odot P = P_1$.

<u>Proof.</u> Here, we use a bar to denote complex conjugation and $\equiv$ to indicate equality mod torsion. We first claim that $P_1, \ldots, P_r$ are linearly independent over Γ, for suppose

$$\alpha_1 \odot P_1 + \cdots + \alpha_r \odot P_r \equiv (\infty, \infty)$$

where $\alpha_i \in \Gamma$. Then, since $P_i \in E(\mathbb{Q})$,

$$\overline{\alpha}_1 \odot P_1 + \cdots + \overline{\alpha}_r \odot P_r \equiv (\infty, \infty) .$$

Adding these equations and using the independence of the P_i over $\mathbb{Z}$ yields $\operatorname{Re} \alpha_i = 0$ for all i and $\operatorname{Im} \alpha_i = 0$ follows immediately.

Next, we claim that if $P \in E(k)$, there exist $\alpha_1, \ldots, \alpha_r \in \Gamma$ such that

$$c(\delta - \overline{\delta}) \odot P \equiv \sum_{i=1}^{r} \alpha_i \odot P_i \quad .$$

This is because $P \oplus \overline{P}$ and $(c\delta \odot P) \oplus \overline{(c\delta \odot P)} = (c\delta \odot P) \oplus (c\overline{\delta} \odot \overline{P})$ are in $E(\mathbb{Q})$ so there exist $n_i, m_i \in \mathbb{Z}$ such that

$$P \oplus \overline{P} \equiv \sum_{i=1}^{r} n_i P_i$$

and

$$(c\delta \odot P) \oplus (c\bar{\delta} \odot \bar{P}) \equiv \sum_{i=1}^{r} m_i P_i \quad .$$

Multiplying $P \oplus \bar{P}$ by $c\bar{\delta}$ and subtracting $(c\delta \odot P) \oplus (c\bar{\delta} \odot \bar{P})$ proves our claim.

Now suppose there is a $P \in E(k)$ with $P_1 = \pi \odot P$. Then

$$\begin{aligned} c(\delta - \bar{\delta}) \odot P_1 &= \pi \odot (c(\delta - \bar{\delta}) \odot P) \\ &\equiv \pi \odot \sum_{i=1}^{r} \alpha_i \odot P_i = \sum_{i=1}^{r} \pi\alpha_i \odot P_i \, , \end{aligned}$$

so $c(\delta - \bar{\delta}) = \pi\alpha_1$, but our restrictions on π mean $\pi \nmid c(\delta - \bar{\delta})$ (recall c is 1, 2, or 3), thus proving the lemma.

Remark. This also shows E has rank at least $2r$ over k.

We now fix a rational point $Q_o = (X_o, Y_o) = (\wp(w), \wp'(w))$ which is part of an integral basis for $E(\mathbb{Q})$ mod torsion. Let $X_n = \wp(w/\pi^n)$, $Y_n = \wp(w/\pi^n)$, and $Q_n = (X_n, Y_n)$ so that $\pi^n \odot Q_n = Q_o$. Also, let $K_n = k_n(X_n, Y_n)$; we have $K_{n-1} \subset K_n$.

LEMMA 3.2.2. *The field* K_n *is* $k_n(X_n)$ *and is a normal extension of* k.

Proof. Let $g(x) = \varphi_{\pi^n}(x)/\psi_{\pi^n}(x)^2$. We have

$$\pi^n \odot x = \wp(\pi^n z) = g(x)$$

where $x = \wp(z)$, and differentiating this with respect to z gives

$$\pi^n \wp'(\pi^n z) = g'(x)\wp'(z) = g'(x)y \quad .$$

Putting $z = w/\pi^n$, we solve for Y_n and get $Y_n = \pi^n Y_o/g'(X_n)$ which shows that $Y_n \in k_n(X_n)$. The proof of the lemma is complete after we note that the conjugates of Q_n are merely translations of Q_n by π^n-division points and the coordinates of these translations are in K_n by the addition theorem.

We record for future reference the formula for multiplication by π^n:

LEMMA 3.2.3. $\pi^n \odot (x,y) = (g(x), g'(x)y/\pi^n)$ _where_ $g(x) = \varphi_{\pi^n}(x)/\psi_{\pi^n}(x)^2$.

The Galois group $G_n = G(K_n/k)$ is a subgroup of

$$H_n = \left\{\begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix} \in GL_2(\Gamma/(\pi)^n)\right\} \quad .$$

If (x_n, y_n) is a fixed primitive π^n-division point and $\sigma \in G_n$, then $(x_n, y_n)^\sigma = \beta \odot (x_n, y_n)$ and $X_n^\sigma = X_n \oplus (\alpha \odot x_n)$ for some $\alpha \in \Gamma/(\pi)^n$, $\beta \in (\Gamma/(\pi)^n)^\times$; the embedding of G_n in H_n is $\sigma \longmapsto \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}$.

It is an interesting general question to determine when Galois groups of fields of division points, such as the ones we are considering, are "big as possible." Using cohomological methods, one can usually say that they are big as possible except finitely often (see for example Bashmakov ([3], [4]), Ribet ([12]), Serre ([14])). In our case we in fact can show, using elementary methods

THEOREM 3.2.4. Assuming Q_o is part of a basis of $E(\mathbb{Q})$ mod torsion, we have $G_n = H_n$.

We need a couple of preliminary lemmas. Note that elements of $\Gamma/(\pi)^n$ naturally project to elements of $\Gamma/(\pi)$.

LEMMA 3.2.5. Suppose we have a map $f\colon (\Gamma/(\pi)^n)^\times \to \Gamma/(\pi)$ satisfying $f(\beta_1\beta_2) = f(\beta_2) + \beta_2 f(\beta_1)$. Then if π is first degree or $n=1$, there is a $\gamma \in \Gamma/(\pi)$ such that $f(\beta) = \gamma(\beta-1)$. If π is second degree, we still have $f(\beta) = 0$ whenever $\beta \equiv 1 \bmod \pi$.

Proof. If π is first degree or $n=1$, the multiplicative group $(\Gamma/(\pi)^n)^\times$ is cyclic and we pick a generator λ. We have

$$f(\lambda^2) = f(\lambda) + \lambda f(\lambda)$$

and more generally, by induction,

$$\begin{aligned} f(\lambda^m) &= f(\lambda)(1 + \lambda + \cdots + \lambda^{m-1}) \\ &= \frac{f(\lambda)}{\lambda - 1}(\lambda^m - 1) . \end{aligned}$$

The first statement of the lemma follows by setting $\gamma = f(\lambda)/(\lambda-1)$.

Suppose now $\pi = p$ is second degree, $n > 1$, and we have a

$\beta \in (\Gamma/(p)^n)^\times$ with $\beta \equiv 1 \bmod p$. It is readily checked that for some $\lambda \in (\Gamma/(p)^n)^\times$ not congruent to 1 mod p, β equals λ^{p-1}. For example, in the case $n=2$, if $\beta = 1 + \alpha p$ ($\alpha \in \Gamma/(p)$), then we can find $\eta \in \Gamma/(p)$ satisfying $(2+\eta p)^{p-1} \equiv 1 + \alpha p \bmod p^2$ and $\lambda = 2 + \eta p$ works. As above, we get

$$f(\beta) = f(\lambda^{p-1}) = [f(\lambda)/(\lambda - 1)](\lambda^{p-1} - 1) = [f(\lambda)/(\lambda - 1)](\beta - 1)$$

and we have the second statement of the lemma.

LEMMA 3.2.6. $\mathfrak{U} = \{\alpha \in \Gamma/(\pi)^n \mid \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \in G_n\}$ _is an ideal in_ $\Gamma/(\pi)^n$.

Proof. Clearly, $\mathfrak{U}$ is a subgroup of $\Gamma/(\pi)^n$. To see that $\mathfrak{U}$ is an ideal, note that since $G(k_n/k)$ is all of $(\Gamma/(\pi)^n)^\times$, we have for any $\beta \in (\Gamma/(\pi)^n)^\times$ some $\alpha_1, \alpha_2 \in \Gamma/(\pi)^n$ with

$$\begin{pmatrix} 1 & \alpha_1 \\ 0 & \beta \end{pmatrix}, \begin{pmatrix} 1 & \alpha_2 \\ 0 & \beta^{-1} \end{pmatrix} \in G_n .$$

Multiplying these on the right by all $\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \in G_n$ shows that we in fact have

$$\begin{pmatrix} 1 & \alpha_1 + \alpha \\ 0 & \beta \end{pmatrix}, \begin{pmatrix} 1 & \alpha_2 + \alpha \\ 0 & \beta^{-1} \end{pmatrix} \in G_n$$

for all $\alpha \in \mathfrak{U}$. Now

$$\begin{pmatrix} 1 & \alpha_1 \\ 0 & \beta \end{pmatrix}\begin{pmatrix} 1 & \alpha_2 \\ 0 & \beta^{-1} \end{pmatrix} = \begin{pmatrix} 1 & \alpha_2 + \alpha_1 \beta^{-1} \\ 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & \alpha_2 \\ 0 & \beta^{-1} \end{pmatrix} \begin{pmatrix} 1 & \alpha_1 \\ 0 & \beta \end{pmatrix} = \begin{pmatrix} 1 & \alpha_1 + \alpha_2\beta \\ 0 & 1 \end{pmatrix}$$

so $\alpha_2 + \alpha_1\beta^{-1}$ and $\alpha_1 + \alpha_2\beta$ are in $\mathfrak{U}$. In particular, $\alpha_2 = -\alpha_1\beta^{-1} + \alpha_o$ for some $\alpha_o \in \mathfrak{U}$. Now fix α_1 and let α be any element of $\mathfrak{U}$; then $-\alpha_1\beta^{-1} + \alpha$ is a legitimate choice for α_2 above. Taking $\alpha_2 = -\alpha_1\beta^{-1} + \alpha$ we have

$$\alpha_1 + \alpha_2\beta = \alpha_1 + (-\alpha_1\beta^{-1} + \alpha)\beta = \alpha\beta \in \mathfrak{U}$$

and this holds for all $\alpha \in \mathfrak{U}$ and $\beta \in (\Gamma/(\pi)^n)^\times$. Since any element of $\Gamma/(\pi)^n$ is the sum of two units, $\mathfrak{U}$ is an ideal.

Proof of Theorem 3.2.4. $\mathfrak{U}$ is an ideal in the (local) ring $\Gamma/(\pi)^n$ so $\mathfrak{U} = (\pi^m)$ for some $m \geq 0$. We want to show $m = 0$ so suppose to the contrary that $m \geq 1$. Taking $\beta \in (\Gamma/(\pi)^n)^\times$ and $\begin{pmatrix} 1 & \alpha_1 \\ 0 & \beta \end{pmatrix}, \begin{pmatrix} 1 & \alpha_2 \\ 0 & \beta^{-1} \end{pmatrix} \in G_n$ we have, as in the proof of Lemma 3.2.6, $\alpha_1 + \alpha_2\beta \in \mathfrak{U}$ so

$$\alpha_1 \equiv -\alpha_2\beta \bmod \pi.$$

Sending $\beta \mapsto \alpha_1$ thus defines a map f from $(\Gamma/(\pi)^n)^\times \to \Gamma/(\pi)$;

$$\begin{pmatrix} 1 & \alpha_1 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1 & \alpha_1' \\ 0 & \beta' \end{pmatrix} = \begin{pmatrix} 1 & \alpha_1' + \beta'\alpha_1 \\ 0 & \beta\beta' \end{pmatrix}$$

shows that f satisfies the hypothesis of Lemma 3.2.5.

In the case $n = 1$, Lemma 3.2.5 gives

$$G_1 = \left\{ \begin{pmatrix} 1 & \gamma(\beta - 1) \\ 0 & \beta \end{pmatrix} \middle| \beta \in (\Gamma/(\pi))^{\times} \right\} \quad .$$

If $\gamma = 0$, then $X_1 \in k$ and Lemma 3.2.3 shows $Y_1 \in k$ as well, but this contradicts Lemma 3.2.1. If $\gamma \neq 0$, $\gamma(\beta - 1)$ ranges over $q-1$ residue classes mod π and X_1 has $q-1$ conjugates over k. If α is a representative for the q^{th} residue class, then $X_1' = X_1 \oplus (\alpha \odot x_1)$ and $Y_1' = Y_1 \oplus (\alpha \odot y_1)$ lie in k and $\pi \odot (X_1', Y_1') = (X_o, Y_o)$, contradicting Lemma 3.2.1 again. Thus $G_1 = H_1$.

If $n > 1$, then Lemma 3.2.5 shows that any $\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \in G_n$ with $\beta \equiv 1 \bmod \pi$ has $\alpha \equiv 0 \bmod \pi$. But G_1 is a quotient of G_n, namely

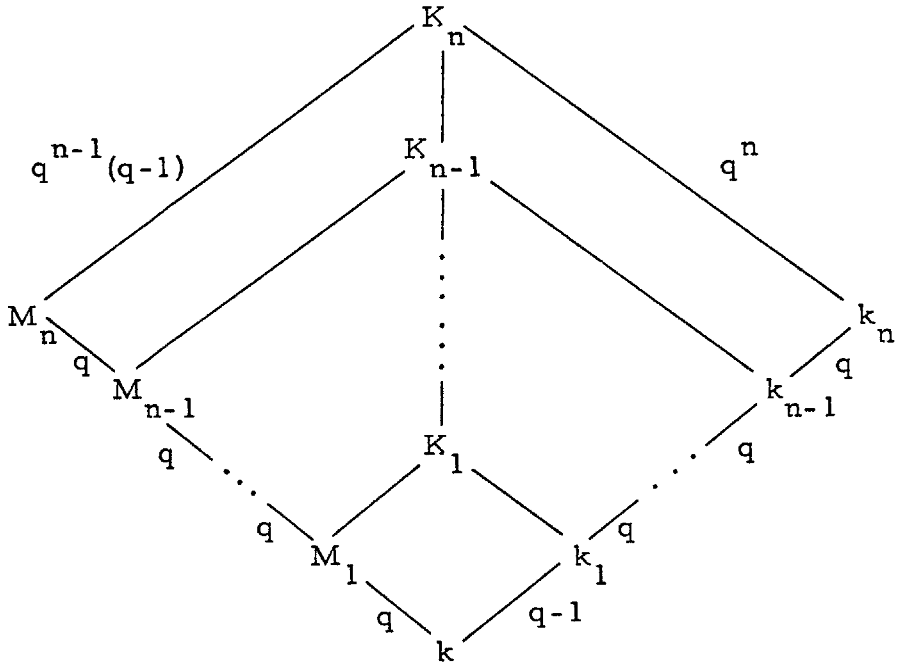
$$G_1 = G_n \Big/ \left\{ \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \equiv \text{Id mod } \pi \right\}$$

and the maximality of G_1 contradicts what we have deduced above. This completes the proof of the theorem.

3.3 Character Relations

We start this section by drawing a picture. Let $M_n = k(X_n)$ and recall the definitions of k_n and K_n from sections 3.1 and 3.2. It will be helpful to keep this picture in mind as we proceed.

(3.3.1)



We know the degrees are as indicated by our work in sections 3.1 and 3.2; we also know $G_n = G(K_n/k) = \left\{\begin{pmatrix} 1 & * \\ 0 & * \end{pmatrix} \in GL_2(\Gamma/(\pi)^n)\right\}$. Clearly,

$$G(K_n/k_n) = \left\{\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}\right\} \quad ,$$

$$G(K_n/M_n) = \left\{\begin{pmatrix} 1 & 0 \\ 0 & * \end{pmatrix}\right\} \quad ,$$

and a little thought shows

$$G(K_n/M_{n-1}) = \left\{\begin{pmatrix} 1 & \alpha \\ 0 & * \end{pmatrix} \middle| \ \alpha \equiv 0 \bmod \pi^{n-1}\right\} \quad .$$

Note that $G(K_n/k_n)$ is the abelian group $\Gamma/(\pi)^n$ which is $\mathbb{Z}/p^n\mathbb{Z}$ for π first degree and $\mathbb{Z}/p^n\mathbb{Z} \times \mathbb{Z}/p^n\mathbb{Z}$ for π second degree.

We are interested in finding the conductor of K_n/k_n. The following proposition will aid us greatly in our endeavors.

PROPOSITION 3.3.2. *Let* χ *be any first degree character of* $G(K_n/k_n)$ *satisfying* $\chi^{p^{n-1}} \neq 1$. *Then*

$$1^*(K_n/M_n) = 1^*(K_n/M_{n-1}) + \chi^*$$

where $1(K_n/-)$ *denotes the trivial character on the associated Galois group and* $*$ *denotes inducing to* G_n.

Proof. We first compute the values of $1^*(K_n/M_n)$ and $1^*(K_n/M_{n-1})$. Let $\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \in G_n$. By the definition of induced characters, we have

$$1^*(K_n/M_n)\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} =$$

$$\frac{1}{q^{n-1}(q-1)}\left[\#\left\{\begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} \in G_n \,\middle|\, \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix}^{-1}\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}\begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} \in G(K_n/M_n)\right\}\right].$$

We compute

$$\begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix}^{-1}\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}\begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} = \begin{pmatrix} 1 & u+\alpha v-\beta u \\ 0 & \beta \end{pmatrix}$$

so we need to find the number of solutions $(u, v) \in \Gamma/(\pi)^n \times (\Gamma/(\pi)^n)^\times$ to

$$(3.3.3) \qquad (1-\beta)u + \alpha v \equiv 0 \bmod \pi^n \quad .$$

If $\beta = 1$, then (3.3.3) has no solutions unless $\alpha = 0$ and then of course all (u, v) are solutions. We have

$$1^*(K_n/M_n)\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} = \begin{cases} \dfrac{|G_n|}{|G(K_n/M_n)|} = q^n & \text{if } \alpha = 0 \\ 0 & \text{otherwise} \end{cases} .$$

If $\beta \neq 1$, for some m we have $\pi^m \| (1-\beta)$ where $0 \le m \le n-1$. The congruence (3.3.3) has no solutions unless $\pi^m | \alpha$ and in this case (3.3.3) becomes

$$\frac{(1-\beta)}{\pi^m} u + \frac{\alpha v}{\pi^m} \equiv 0 \bmod \pi^{n-m} \quad .$$

Given each of the $q^{n-1}(q-1)$ v's, we solve this for $u \bmod \pi^{n-m}$, giving q^m solutions for u (mod π^n). We have for $\beta \neq 1$, $\pi^m \| (1-\beta)$,

$$1^*(K_n/M_n)\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} = \begin{cases} q^m & \text{if } \pi^m | \alpha \\ 0 & \text{otherwise} . \end{cases}$$

For $1^*(K_n/M_{n-1})$ we need to solve (3.3.3) but now mod π^{n-1}; $1^*(K_n/M_{n-1})$ is $q^{-n}(q-1)^{-1}$ times the number of solutions. Exactly the same argument with n replaced by $n-1$ shows that for $\beta \equiv 1$ mod π^{n-1},

$$1^*(K_n/M_{n-1})\begin{pmatrix}1 & \alpha\\ 0 & \beta\end{pmatrix} = \begin{cases} q^{n-1} & \text{if } \pi^{n-1} \mid \alpha \\ 0 & \text{otherwise}\end{cases}$$

and for $\beta \not\equiv 1 \bmod \pi^{n-1}$, $\pi^m \parallel (1-\beta)$ $(0 \le m \le n-2)$,

$$1^*(K_n/M_{n-1})\begin{pmatrix}1 & \alpha\\ 0 & \beta\end{pmatrix} = \begin{cases} q^{m} & \text{if } \pi^{m} \mid \alpha \\ 0 & \text{otherwise} \end{cases} .$$

Looking at these values, we observe that $1^*(K_n/M_n)$ and $1^*(K_n/M_{n-1})$ differ only when $\beta = 1$. Letting $\psi = 1^*(K/M_n) - 1^*(K/M_{n-1})$, we have

$$\psi\begin{pmatrix}1 & \alpha\\ 0 & \beta\end{pmatrix} = \begin{cases} q^n - q^{n-1} & \text{if } \beta = 1,\ \alpha = 0 \\ -q^{n-1} & \text{if } \beta = 1,\ \pi^{n-1} \parallel \alpha \\ 0 & \text{otherwise} . \end{cases}$$

Note that ψ assumes the value $-q^{n-1}$ exactly $q-1$ times so

$$\langle \psi, \psi \rangle = \frac{1}{q^{2n-1}(q-1)} \left[(q^n - q^{n-1})^2 + (q-1)q^{2(n-1)} \right]$$
$$= 1 ,$$

and ψ is irreducible. By Frobenius reciprocity,

$$\langle \psi, \chi^* \rangle_{G_n} = \left\langle \psi \,\middle|\, G(K_n/k_n) ,\ \chi \right\rangle_{G(K_n/k_n)}$$
$$= \frac{1}{q^n} \left[(q^n - q^{n-1}) - q^{n-1} \sum_{\pi^{n-1} \parallel \alpha} \chi \begin{pmatrix}1 & \alpha\\ 0 & 1\end{pmatrix} \right] .$$

To evaluate the sum, re-write it as

$$\sum_{\beta \in (\Gamma/(\pi))\sim\{0\}} \chi\begin{pmatrix}1 & \pi^{n-1}\beta\\ 0 & 1\end{pmatrix} = \sum_{\beta \in (\Gamma/(\pi))\sim\{0\}} \chi^{p^{n-1}}\begin{pmatrix}1 & \beta\\ 0 & 1\end{pmatrix} .$$

This equality is clear if π is second degree and is seen easily for π first degree by replacing β by $\bar{\pi}^{n-1}\beta$. But $\chi^{p^{n-1}}$ is a non-trivial first degree character of $\Gamma/(\pi)$ ($= \mathbb{Z}/p\mathbb{Z}$ for π first degree, $(\mathbb{Z}/p\mathbb{Z})^2$ for π second degree) so the sum is -1. This gives $\langle \psi, \chi^* \rangle = 1$ whence $\psi = \chi^*$ since both characters have the same degree and ψ is irreducible. Recalling what ψ is, the proposition follows.

Using ψ and characters from the quotient groups G_m, $m < n$, it is quite easy to find the character table of G_n. We will not need this so we leave it as an exercise for the reader (if he cares).

Next we see how Proposition 3.3.2 leads to a conductor-discriminant relation which allows us to find the conductor of K_n/k_n.

CHAPTER FOUR

THE CONDUCTOR OF K_n/k_n

We are now ready for our main result, the evaluation of the conductor of K_n/k_n. Throughout, we assume that $\pi \nmid 6\Delta(E)D(k/\mathbb{Q})$ and that $Q_o = (X_o, Y_o)$ is part of a basis for $E(\mathbb{Q})$ mod torsion (in particular we are assuming $r(E) > 0$) so that the results of chapter 3 hold. We will see in the next chapter how the Coates-Wiles theorem follows from the results of this chapter.

4.1 A Conductor-Discriminant Relation

We recall that class field theory associates to any first degree character χ of an abelian Galois group a conductor $\mathfrak{F}(\chi)$ which is an ideal in the base field.

PROPOSITION 4.1.1. *Let* χ *be any first degree character of* $G(K_n/k_n)$ *such that* $\chi^{p^{n-1}} \neq 1$. *Then*

$$D(M_n/k) = D(M_{n-1}/k)\,D(k_n/k)\,N_{k_n/k}(\mathfrak{F}(\chi)) \;.$$

As we shall see shortly, this result follows easily from Proposition 3.3.2 and results in Artin ([2]). In this paper, Artin defines "conductors" of characters on arbitrary (i.e. not necessarily abelian) Galois groups. We use $\mathfrak{F}$ to denote these conductors as well; this is justified since they agree with the conductors from class field theory in the abelian case. The following theorem summarizes the results in ([2]) we need. As before, * denotes inducing from a subgroup to the whole group.

THEOREM 4.1.2. _Let_ K/k _be normal and suppose_ $K \supset M \supset k$. _If_ χ_1, χ_2 _are characters of_ $G(K/k)$ _and_ ψ _is a character of_ $G(K/M)$ _then_

(1) $\mathfrak{F}(\chi_1 + \chi_2) = \mathfrak{F}(\chi_1)\,\mathfrak{F}(\chi_2)$.

(2) $\mathfrak{F}(\psi^*) = D(M/k)^{\psi(1)}\, N_{M/k}(\mathfrak{F}(\psi))$.

(3) $\mathfrak{F}(1) = (1)$.

Proof of Proposition 4.1.1. Take conductors of both sides of

$$1^*(K_n/M_n) = 1^*(K_n/M_{n-1}) + \chi^*,$$

which we know by Proposition 3.3.2,and apply the properties (1), (2), (3) above.

Remark. We can give an alternate proof of Proposition 4.1.1 using the theory of L-functions, for the character relation above implies

$$\zeta_{M_n}(s) = \zeta_{M_{n-1}}(s)\, L(s, \chi) .$$

Proposition 4.1.1 follows, up to absolute values, by comparing fudge factors in the functional equations of these series. A little more work gets rid of the absolute values.

4.2 Evaluation of the Conductor

Recall that (X_o, Y_o) is a rational point of infinite order and $\pi^n \odot (X_n, Y_n) = (X_o, Y_o)$.

LEMMA 4.2.1. *Suppose π does not divide the numerators or denominators of X_o or Y_o. Then the numerators and denominators of X_n, Y_n are also relatively prime to (π).*

Proof. By Lemma 3.2.3,

$$(X_{n-1}, Y_{n-1}) = \pi \odot (X_n, Y_n) = \left(g(X_n),\ g'(X_n)\frac{Y_n}{\pi}\right)$$

where $g(x) = \varphi_\pi(x)/\psi_\pi(x)^2$. Using our knowledge of $\varphi_\pi(x)$ and $\psi_\pi(x)$ (Proposition 2.2.4), we see immediately from $X_{n-1} = g(X_n)$ that if X_{n-1} is relatively prime to (π), then so is X_n. By simply differentiating $g(x)$ we also see that $g'(X_n)$ is integral at primes above (π) and $\pi | g'(X_n)$. Using $Y_{n-1} = (g'(X_n)/\pi) Y_n$ we observe that if Y_{n-1} is relatively prime to (π) then the numerator of Y_n is relatively prime to (π). Since

$$Y_n^2 = 4X_n^3 - g_2 X_n - g_3$$

implies that any prime above (π) dividing the denominator of Y_n must also divide the denominator of X_n, we are done.

Recall that $\mathfrak{P}_n$ is the unique prime above (π) in k_n. We will need the following two results, the second of which implies that the conductor of K_n/k_n is a power of $\mathfrak{P}_n$.

PROPOSITION 4.2.2. *For some* n, *the extension* K_n/k_n *is ramified.*

PROPOSITION 4.2.3. *The extension* K_n/k_n *is unramified outside of* $\mathfrak{P}_n$.

These results may be found in Coates-Wiles ([6]) but we give another proof of Proposition 4.2.2 in the next section using our methods. Proposition 4.2.3 follows from the fact that E has "good reduction everywhere" over k_1 — see [6]. We remark that Theorem 2.4.1 simplifies the argument in ([6]) to eliminate primes above $\overline{\pi}$.

We assume from now on that π does not divide the numerators or denominators of X_o or Y_o. As before, we let $D_{(\pi)}(M/k)$ denote the (π)-part of $D(M/k)$. Our main theorem is a generalization of Theorem 1 in Stark ([16]). For convenience we set M_o, K_o, $k_o = k$.

THEOREM 4.2.4. *Let* $n \geq 0$ *and suppose* K_i/k_i *is unramified for all* $i < n$. *If* K_n/k_n *is ramified, then it has conductor* $\mathfrak{P}_n^2$. *If* K_n/k_n *is unramified, then*

$$D_{(\pi)}(M_n/k) = (\pi)^{nq^n - 2q^{n-1} - \cdots - 2q - 2}$$

where for $n = 0$ *we make the convention* $nq^n - 2q^{n-1} - \cdots - 2q - 2 = 0$.

Proof. We proceed by induction. The case $n = 0$ is a triviality, so take $n > 0$. The extension M_n/M_{n-1} is obtained by dividing X_{n-1} by π, i.e. adding a solution of $\varphi_\pi(x)/\psi_\pi(x)^2 = X_{n-1}$. We are adding a root of the monic q^{th} degree polynomial

$$f_n(x) = \varphi_\pi(x) - X_{n-1}\,\psi_\pi(x)^2$$

which we know is irreducible by Theorem 3.2.4. The discriminant of $f_n(x)$ has (π)-part $(\pi)^q$ by Theorem 2.4.1 and Lemma 4.2.1 (take $z = w/\pi^n$ in Theorem 2.4.1); here (π)-part means the part of disc (f_n) above (π). Moreover, $f_n(x)$ has coefficients which are integral at (π) so if $(\pi) = \mathfrak{P}_1^{e_1}\,\mathfrak{P}_2^{e_2}\cdots\mathfrak{P}_g^{e_g}$ in M_{n-1}, the $\mathfrak{P}_i$-part $D_{\mathfrak{P}_i}(M_n/M_{n-1})$ of $D(M_n/M_{n-1})$ divides $\mathfrak{P}_i^{qe_i}$. Since polynomial and field discriminants differ by squares (at least locally), the quotient is an even power of $\mathfrak{P}_i$. Using

$$D(M_n/k) = N_{M_{n-1}/k}\,(D(M_n/M_{n-1}))\,D(M_{n-1}/k)^q$$

and the induction hypothesis we see that

$$D_{(\pi)}(M_n/k)\;\Big|\;(\pi)^{q^n}\left[(\pi)^{(n-1)q^{n-1}-2q^{n-2}-\cdots-2}\right]^q .$$

Simplify to get

$$D_{(\pi)}(M_n/k)\;\Big|\;(\pi)^{nq^n-2q^{n-1}-\cdots-2q}$$

where for $n = 1$ we interpret the exponent as being q. We know that the quotient of the above power of (π) by $D_{(\pi)}(M_n/k)$ is an even power of (π).

On the other hand, by Proposition 4.1.1 we have

$$D_{(\pi)}(M_n/k) = D_{(\pi)}(M_{n-1}/k)\, D_{(\pi)}(k_n/k)\, N_{k_n/k}(\mathfrak{F}(\chi))$$

for χ any first degree character of $G(K_n/k_n)$ with $\chi^{p^{n-1}} \neq 1$ (Proposition 4.2.3 tells us $\mathfrak{F}(\chi)$ is a power of $\mathfrak{P}_n$). By Theorem 3.1.4 and the induction hypothesis, we have

$$D_{(\pi)}(M_{n-1}/k)\, D_{(\pi)}(k_n/k) = (\pi)^{(n-1)q^{n-1}-2q^{n-2}-\cdots-2}(\pi)^{nq^n-(n+1)q^{n-1}}$$

$$= (\pi)^{nq^n-2q^{n-1}-2q^{n-2}-\cdots-2} .$$

Now a miracle has happened, namely the bound we have for the power of (π) in $D(M_n/k)$ and the exponent immediately above differ exactly by 2. This shows that either $\mathfrak{F}(\chi) = \mathfrak{P}_n^2$ or $\mathfrak{F}(\chi) = (1)$ and $D_{(\pi)}(M_n/k) = (\pi)^{nq^n-2q^{n-1}-\cdots-2q-2}$. The conductor of K_n/k_n being the ℓ.c.m. of the conductors of the characters of $G(K_n/k_n)$, we have our result.

Combining Theorem 4.2.4 and Proposition 4.2.2 gives

THEOREM 4.2.5. *For some* n, K_n/k_n *has conductor* $\mathfrak{P}_n^2$.

For the purposes of proving the Coates-Wiles theorem, the above result is the important one. It turns out that Theorem 4.2.5

holds for exactly *one* n ; in the next section we find an explicit description of this n.

We say that a prime $\wp$ in M_n above (π) is *relative first degree* if $N_{M_n/k}(\wp) = (\pi)$. The proof of Theorem 4.2.4 immediately yields

COROLLARY 4.2.6. *Suppose* $K_1/k_1, \ldots, K_{n-1}/k_{n-1}$ *are unramified and* $\wp$ *is a prime above* (π) *in* M_{n-1}. *If the powers of* $\wp$ *in* $D(M_n/M_{n-1})$ *and* $\mathrm{disc}(f_n)$ *differ, then this difference is* 2 *and* $\wp$ *is relative first degree. Moreover, this happens for exactly one prime above* (π) *if* K_n/k_n *is unramified and otherwise does not occur.*

Here, f_n is as introduced in the proof of Theorem 4.2.4. In the next section we will say more about the factorization of (π) in M_n.

We close this section with

PROPOSITION 4.2.7. *Suppose* $\pi = p$ *is second degree,* $K_1/k_1, \ldots, K_{n-1}/k_{n-1}$ *are unramified, and* K_n/k_n *has conductor* $\mathfrak{P}_n^2$. *Then for any field* L *with* $k_n \subset L \subset K_n$ *and* $[K_n : L] = p$, *the conductor of* L/k_n *is* $\mathfrak{P}_n^2$ *as well.*

Proof. Let $H \subset G(K_n/k_n)$ be the subgroup corresponding to L. The characters of $G(L/k_n) = G(K_n/k_n)/H$ correspond to the characters of $G(K_n/k_n)$ trivial on H. Since $G(K_n/k_n) \cong \mathbb{Z}/p^n\mathbb{Z} \times \mathbb{Z}/p^n\mathbb{Z}$ and $|H| = p$, for some such character χ we have $\chi^{p^{n-1}} \neq 1$. The proof of Theorem 4.2.4 shows that χ has conductor $\mathfrak{P}_n^2$ and hence so does L/k_n.

4.3 Ramification at Some Location

In this section we will prove Proposition 4.2.2. Our method of proof enables us to describe the explicit factorization of (π) in M_n and calculate the conductor of K_n/k_n for all n. Almost-Eisenstein criteria (section 2.3) form the heart of our arguments. In this section we suppose π has been chosen (as guaranteed by Theorem 2.2.5) so that $\pi \odot x \equiv x^q \bmod (\pi)$.

Recall that we have monic irreducible q^{th} degree polynomials

$$f_n(x) = \varphi_\pi(x) - X_{n-1}\psi_\pi(x)^2 \in M_{n-1}[x]$$

generating M_n/M_{n-1}, i.e. with root X_n. It will be convenient to introduce the translations of $f_n(x)$

$$g_n(x) = f_n(x + X_o) \quad ;$$

$g_n(x)$ has $X_n - X_o$ as a root. Here, as before, we will assume π is relatively prime to the numerators and denominators of X_o and Y_o. The coefficients of $x^{q-1}, \ldots, x$ in $f_n(x)$ are divisible by π but the constant coefficient is not. The coefficients of $x^{q-1}, \ldots, x$

in $g_n(x)$ are also divisible by π. Letting a_1 be the constant coefficient of $g_1(x)$ we have

$$a_1 = \varphi_\pi(X_o) - X_o \psi_\pi(X_o)^2$$

$$\equiv X_o^q - X_o \equiv 0 \mod (\pi)$$

because $\varphi_\pi(x) \equiv x^q$ and $\psi_\pi(x)^2 \equiv 1 \mod (\pi)$ by our choice of π. Since $g_1(x)$ is irreducible, $a_1 \neq 0$ and we may find e so that $\pi^e \| a_1$. We use $\wp_{n,i}$ to denote a prime above (π) in M_n ; recall that we say $\wp_{n,i}$ is relative first degree if $N_{M_n/k}(\wp_{n,i}) = (\pi)$.

THEOREM 4.3.1. _The extensions_ $K_1/k_1, \ldots, K_{e-1}/k_{e-1}$ _are unramified. For_ $n \geq e$, K_n/k_n _is ramified and has conductor_ $\mathfrak{P}_n^{2q^{n-e}}$. _Also_, (π) _is a product of relative first degree primes in_ M_n. _For_ $n \leq e-1$

$$(1) \qquad (\pi) = \wp_{n,1}^{q^{n-1}(q-1)} \wp_{n,2}^{q^{n-2}(q-1)} \cdots \wp_{n,n}^{q-1} \wp_{n,n+1}$$

and for $n \geq e$

$$(2) \qquad (\pi) = \wp_{n,1}^{q^{n-1}(q-1)} \wp_{n,2}^{q^{n-2}(q-1)} \cdots \wp_{n,e-1}^{q^{n-e+1}(q-1)} \wp_{n,e}^{q^{n-e+1}} .$$

In the case $n = 0$ _we interpret_ (1) _as_ $(\pi) = \wp_{n,n+1}$ _and if_ $e = 1$ _we interpret_ (2) _as_ $(\pi) = \wp_{n,1}^{q^n}$.

Proof. The key to the proof is an investigation of the constant coefficient a_n of $g_n(x)$; we know $g_n(x)$ is monic and the other coefficients are all divisible by π. We have

$$a_n = \varphi_\pi(X_o) - X_{n-1}\,\psi_\pi(X_o)^2$$

$$= a_1 - (X_{n-1} - X_o)\,\psi_\pi(X_o)^2 \quad . \qquad (4.3.2)$$

Suppose $e = 1$. Then $g_1(x)$ is Eisenstein with respect to (π) and by the Eisenstein criterion (π) ramifies totally to M_1, say $(\pi) = \wp_{1,1}^{q}$, $\wp_{1,1} \parallel (X_1 - X_o)$, and $\operatorname{disc}_{(\pi)}(g_1) = D_{(\pi)}(M_1/k)$ where as usual the subscript indicates "(π)-part of."

Now suppose $e > 1$. By the almost-Eisenstein criterion (Theorem 2.3.4), $\operatorname{disc}_{(\pi)}(g_1) \neq D_{(\pi)}(M_1/k)$ so by Corollary 4.2.6, K_1/k_1 is unramified and

$$\operatorname{disc}_{(\pi)}(g_1) = (\pi)^2\, D_{(\pi)}(M_1/k) \, .$$

By Theorem 2.3.5, $(\pi) = \wp_{1,1}^{q-1}\,\wp_{1,2}$ in M_1 with $\wp_{1,1} \parallel (X_1 - X_o)$ and $\wp_{1,2}^{e-1} \parallel (X_1 - X_o)$. Equation (4.3.2) shows that $\wp_{1,1} \parallel a_2$ and since $\wp_{1,2}^{e} \mid a_1$, we also get $\wp_{1,2}^{e-1} \parallel a_2$. The polynomial $g_2(x)$ is Eisenstein with respect to $\wp_{1,1}$ so $\wp_{1,1} = \wp_{2,1}^{q}$ in M_2. If $e = 2$ then $g_2(x)$ is Eisenstein with respect to $\wp_{1,2}$ as well; if $e > 2$ then $g_2(x)$ is almost-Eisenstein with respect to $\wp_{1,2}$. In the latter case, the almost-Eisenstein criterion, Corollary 4.2.6, and Theorem 2.3.5 now applied to $(g_2(x), \wp_{1,2})$ instead of $(g_1(x), (\pi))$ show that K_2/k_2 is unramified and $\wp_{1,2} = \wp_{2,2}^{q-1}\,\wp_{2,3}$ in M_2 with

$\mathfrak{P}_{2,2} \parallel (X_2 - X_o)$ and $\mathfrak{P}_{2,3}^{e-2} \parallel (X_2 - X_o)$. We also know $\mathfrak{P}_{2,1} \parallel (X_2 - X_o)$. Using (4.3.2) as before we see that $\mathfrak{P}_{2,1} \parallel a_3$, $\mathfrak{P}_{2,2} \parallel a_3$, and $\mathfrak{P}_{2,3}^{e-2} \parallel a_3$; $g_3(x)$ is Eisenstein with respect to $\mathfrak{P}_{2,1}$ and $\mathfrak{P}_{2,2}$ and also with respect to $\mathfrak{P}_{2,3}$ if $e = 3$, otherwise it is almost-Eisenstein with respect to $\mathfrak{P}_{2,3}$.

Repeated applications of the above argument give the first statement of Theorem 4.3.1 and the factorization (1) for $n < e$. In M_{e-1} we have

$$(\pi) = \mathfrak{P}_{e-1,1}^{q^{e-2}(q-1)} \mathfrak{P}_{e-1,2}^{q^{e-3}(q-1)} \cdots \mathfrak{P}_{e-1,e-1}^{q-1} \mathfrak{P}_{e-1,e}$$

where now $g_e(x)$ is Eisenstein with respect to <u>all</u> $\mathfrak{P}_{e-1,i}$. We have $\mathfrak{P}_{e-1,i} = \mathfrak{P}_{e,i}^q$ and (4.3.2) shows $g_{e+1}(x)$ is Eisenstein with respect to all $\mathfrak{P}_{e,i}$ - repeating this argument gives (2) for $n \geq e$.

To finish the proof of Theorem 4.3.1, note that the bound $(\pi)^q$ for $D_{(\pi)}(M_n/M_{n-1})$ we had in the proof of Theorem 4.2.4 now becomes exact for $n \geq e$ because $g_n(x)$ is Eisenstein with respect to all primes above (π). For $n \geq e$ we thus have

$$D_{(\pi)}(M_n/k) = N_{M_{n-1}/k}((\pi)^q)\, D_{(\pi)}(M_{n-1}/k)^q$$

$$= (\pi)^{q^n} D_{(\pi)}(M_{n-1}/k)^q \quad .$$

Using

$$D_{(\pi)}(M_{e-1}/k) = (\pi)^{(e-1)q^{e-1}-2q^{e-2}-\cdots-2}$$

(from Theorem 4.2.4) we get

$$D_{(\pi)}(M_n/k) = (\pi)^{nq^n-2q^{n-1}-\cdots-2q^{n-e+1}}.$$

The conductor-discriminant relation

$$D(M_n/k) = D(M_{n-1}/k)\,D(k_n/k)\,N_{k_n/k}(\mathfrak{F}(\chi))$$

and Theorem 3.1.4 then yield $\mathfrak{F}(\chi) = \mathfrak{P}_n^{2q^{n-e}}$ for any character χ with $\chi^{p^{n-1}} \neq 1$, so K_n/k_n has conductor $\mathfrak{P}_n^{2q^{n-e}}$ as claimed. This completes the proof of the theorem.

Theorem 4.3.1 certainly encompasses Proposition 4.2.2 so we have accomplished the goal of this section. Although we used Proposition 4.2.3 to conclude $\mathfrak{F}(\chi) = \mathfrak{P}_n^{2q^{n-e}}$ above, a closer inspection shows that we in fact have a proof of Proposition 4.2.2 independent of Proposition 4.2.3.

It is reasonable to expect that $e = 1$ most of the time so that ramification occurs at the first level of the tower. However, this is reminiscent of the famous

OPEN QUESTION. *Are there infinitely many primes* p *such that* $p \parallel (2^{p-1} - 1)$?

Wieferich showed that $p \,\|\, (2^{p-1} - 1)$ implies the truth of the first case of Fermat's Last Theorem for exponent p; it is not known whether the first case of Fermat's Last Theorem holds infintely often. Thus, we expect it to be very difficult to decide if K_1/k_1 is ramified infinitely often, but as far as the Coates-Wiles result is concerned, ramification at some level is all we need.

CHAPTER FIVE

THE COATES-WILES THEOREM

We will now use the results of chapter 4 to prove the Coates-Wiles theorem. Crucial to the argument is the structure of certain units living in k_n. As before, π is a first or second degree prime in k not dividing $6\,\Delta(E)\,D(k/\mathbb{Q})$ or the numerators or denominators of X_o or Y_o.

5.1 Elliptic Units

Robert ([13]) constructed certain "elliptic units" in fields of π^n-division points. These units are constructed by evaluating certain elliptic functions at division points of the lattice Ω. The important fact we need is

THEOREM 5.1.1. *Outside of finitely many* π's *there is a unit* $u_n \in k_n$ *such that*

$$u_n \equiv 1 + \widetilde{L_E(1)}\,\frac{y_n}{2x_n^2} \bmod \mathfrak{P}_n^2$$

where $\widetilde{L_E(1)} \in k$ *is integral at* π, *does not depend on* π, *and is zero if and only if* $L_E(1) = 0$.

This result follows from results in Stark ([16]); u_n is just the elliptic function $h(z)$ in ([16]) evaluated at some primitive

π^n-division point of the lattice Ω. Lemma 3 in ([16]) gives the structure of u_n above. $\widetilde{L_E(1)}$ is essentially, i.e. up to some transcendental factor, $L_E(1)$ with a finite number of Euler factors missing.

In section 3.1 we saw that $\mathfrak{P}_n^{-3} \parallel y_n$ and $\mathfrak{P}_n^{-2} \parallel x_n$ so $\mathfrak{P}_n \parallel (y_n/2x_n^2)$ and $u_n \equiv 1 \bmod \mathfrak{P}_n$. We have

COROLLARY 5.1.2. $u_n \equiv 1 \bmod \mathfrak{P}_n^2$ *if and only if* $\pi \mid \widetilde{L_E(1)}$.

Proof. Clearly, $u_n \equiv 1 \bmod \mathfrak{P}_n^2$ if and only if $\mathfrak{P}_n \mid \widetilde{L_E(1)}$ and since $\widetilde{L_E(1)} \in k$, this happens if and only if $\pi \mid \widetilde{L_E(1)}$.

5.2 Units and Class Fields

We suppose $r(E) > 0$ and let n be the unique location in the tower (3.3.1) such that K_n/k_n has conductor $\mathfrak{P}_n^2$ - we know n exists by Theorem 4.2.5 and is unique by Theorem 4.3.1.

PROPOSITION 5.2.1. *Any unit in* k_n *which is* $\equiv 1 \bmod \mathfrak{P}_n$ *is* $\equiv 1 \bmod \mathfrak{P}_n^2$.

Proof. Let $G(\mathfrak{U})$, $K(\mathfrak{U})$ denote the ray class group, respectively field, of k_n mod $\mathfrak{U}$. We have

$$|G(\mathfrak{U})| = [K(\mathfrak{U}): k_n] = \frac{h(k_n)\,\phi(\mathfrak{U})}{\epsilon(\mathfrak{U})} ,$$

where $h(k_n)$ is the class number of k_n, ϕ is the Euler ϕ function, and $\varepsilon(\mathfrak{U})$ is the number of incongruent units of k_n mod $\mathfrak{U}$. The group $G(\mathfrak{P}_n)$ is a quotient of $G(\mathfrak{P}_n^2)$ and

$$[K(\mathfrak{P}_n^2): K(\mathfrak{P}_n)] = \frac{|G(\mathfrak{P}_n^2)|}{|G(\mathfrak{P}_n)|} = \frac{\phi(\mathfrak{P}_n^2)}{\phi(\mathfrak{P}_n)} \cdot \frac{\varepsilon(\mathfrak{P}_n)}{\varepsilon(\mathfrak{P}_n^2)}$$

$$= \frac{q}{\varepsilon(\mathfrak{P}_n^2)/\varepsilon(\mathfrak{P}_n)} .$$

We now consider two cases.

<u>Case 1</u>. π is first degree. Then

$$\frac{|G(\mathfrak{P}_n^2)|}{|G(\mathfrak{P}_n)|} = \frac{p}{\varepsilon(\mathfrak{P}_n^2)/\varepsilon(\mathfrak{P}_n)} = 1 \text{ or } p,$$

and since K_n/k_n has conductor $\mathfrak{P}_n^2$, the above quotient must be p. This gives $\varepsilon(\mathfrak{P}_n^2) = \varepsilon(\mathfrak{P}_n)$, which easily implies our proposition.

<u>Case 2</u>. $\pi = p$ is second degree. Then

$$\frac{|G(\mathfrak{P}_n^2)|}{|G(\mathfrak{P}_n)|} = \frac{p^2}{\varepsilon(\mathfrak{P}_n^2)/\varepsilon(\mathfrak{P}_n)} = 1,\ p,\ \text{or } p^2$$

and the conductor of K_n/k_n being $\mathfrak{P}_n^2$ rules out $|G(\mathfrak{P}_n^2)|/|G(\mathfrak{P}_n)| = 1$. We need to rule out $|G(\mathfrak{P}_n^2)|/|G(\mathfrak{P}_n)| = p$. Note that since K_{n-1}/k_{n-1} is unramified, $K' = k_n K_{n-1}$ is an unramified extension of k_n. Also, K_n/K' is a (p,p) extension and we have the following picture:

$$\begin{tikzcd}
& K(\mathfrak{P}_n^2) \arrow[dl, dash] \arrow[dr, dash] & \\
K(\mathfrak{P}_n) \arrow[dr, dash] & & K_n \arrow[dl, dash, "p^2"] \\
& K' &
\end{tikzcd}$$

We know $[K(\mathfrak{P}_n^2) : K(\mathfrak{P}_n)]$ is p or p^2. Let $G = G(K(\mathfrak{P}_n^2)/K')$, $H_1 = G(K(\mathfrak{P}_n^2)/K_n)$, and $H_2 = G(K(\mathfrak{P}_n^2)/K(\mathfrak{P}_n))$. By Proposition 4.2.7, $K(\mathfrak{P}_n) \cap K_n = K'$, so $H_1 H_2 = G$. If $H_1 \cap H_2 = \{1\}$ as well, then $H_2 \cong G/H_1$ which has order p^2, so $[K(\mathfrak{P}_n^2) : K(\mathfrak{P}_n)] = p^2$. If $H_1 \cap H_2 \neq \{1\}$ our picture looks like

$$\begin{tikzcd}
& K(\mathfrak{P}_n^2) \arrow[d, dash] & \\
& N \arrow[dl, dash] \arrow[dr, dash] & \\
K(\mathfrak{P}_n) \arrow[dr, dash] & & K_n \arrow[dl, dash] \\
& K' &
\end{tikzcd}$$

where $K(\mathfrak{P}_n^2)/N$ and $N/K(\mathfrak{P}_n)$ are non-trivial,so $[K(\mathfrak{P}_n^2) : K(\mathfrak{P}_n)] = p^2$ again. The equality $\epsilon(\mathfrak{P}_n^2) = \epsilon(\mathfrak{P}_n)$ then finishes the proof of Case 2.

5.3 Conclusion

Proposition 5.2.1 and Corollary 5.1.2 combine to show that if $r(E) > 0$ then $\pi \mid \widetilde{L_E(1)}$ for all π outside a finite bad list. There are infinitely many π (even for non-maximal orders), so

$\widetilde{L_E(1)} = L_E(1) = 0$ and we have proved the Coates-Wiles theorem.

We may very well wonder whether our methods extend to curves without complex multiplication. The fields in question become much bigger for non-CM curves. In fact Serre ([14]) shows that for non-CM curves the field of p-division points is an extension of $\mathbb{Q}$ with Galois group $GL_2(\mathbb{Z}/p\mathbb{Z})$ except for finitely many p. Also, the ramification index of p is $p-1$, $p(p-1)$, or p^2-1.

The last instance is that of "good reduction of height 2" ; from our point of view it is the case when the Frobenius automorphism associated to p has trace zero and Proposition 2. 2. 4 holds for multiplication by p. According to the Sato-Tate conjecture, we should be in this case infinitely often. Stark ([16]) indicates how in this situation we can get a "conductor $\mathfrak{p}^2$" result at the first level of the tower, if ramification occurs at this level. We expect that it should be possible to extend this result up the tower as we have done here for CM curves. Our investigations on this question are rather preliminary and we hope to report on them in the future. Another question that invites investigation is whether there are units for non-CM curves that have a structure which yields information about $L_E(1)$.

References

[1] Arthaud, N.: On Birch and Swinnerton-Dyer's conjecture for elliptic curves with complex multiplication I. Compositio Math. 37, Fasc. 2, 209-232 (1978).

[2] Artin, E.: Die gruppentheoretische Struktur der Diskriminanten algebraischer Zahlkörper. J. Reine Angew. Math. 164, 1-11 (1931).

[3] Bashmakov, M.: Un théorème de finitude sur la cohomology des courbes elliptiques. C. R. Acad. Sci. Paris Sér. A-B 270, A999-A1000 (1970).

[4] Bashmakov, M.: The cohomology of abelian varieties over a number field. Russian Math. Surveys 27 (6), 25-70 (1972).

[5] Birch, B., Swinnerton-Dyer, P.: Notes on elliptic curves II. J. Reine Angew. Math. 218, 79-108 (1965).

[6] Coates, J., Wiles, A.: On the conjecture of Birch and Swinnerton-Dyer. Inventiones Math. 39, 223-251 (1977).

[7] Deuring, M.: Die Typen der Multiplikatorenringe elliptischer Funktionenkörper. Abh. Math. Sem. Univ. Hamb. 16, 32-47 (1949).

[8] Deuring, M.: Die Zetafunktion einer algebraischen Kurve vom Geschlechte Eins, I, II, III, IV. Nachr. Akad. Wiss. Gott., 1953, 1955, 1956, 1957.

[9] Deuring, M.: Die Klassenkörper der komplexen Multiplikation. Enz. Math. Wiss., Band I-2, Heft 10, Teil II, Teubner, Stuttgart, 1958.

[10] Lang, S.: Algebraic number theory, Addison-Wesley, Reading, 1970.

[11] Lang, S.: Elliptic functions, Addison-Wesley, Reading, 1973.

[12] Ribet, K.: Dividing rational points on abelian varieties of CM-type. Compositio Math. 33, Fasc. 1, 69-74 (1976).

[13] Robert, G.: Unités elliptiques. Bull. Soc. Math. France, Mémoire 36, 1973.

[14] Serre, J-P.: Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. Inventiones Math. 15, 259-331 (1972).

[15] Stark, H.: M.I.T. course notes on elliptic curves and complex multiplication, 1981.

[16] Stark, H.: The Coates-Wiles theorem revisited. Number theory related to Fermat's last theorem, Progress in Math. vol. 26, Birkhäuser, Boston, 1982.

[17] Stark, H.: The analytic theory of algebraic numbers, Springer-Verlag, to appear.

[18] Tate, J.: The arithmetic of elliptic curves. Inventiones Math. 23, 179-206 (1974).

[19] Weber, H.: Lehrbuch der Algebra. III. 3 Aufl., Braunschweig, 1908.