THETA FUNCTIONS AND DIVISION POINTS ON ABELIAN VARIETIES OF DIMENSION TWO

by

David R. Grant
A. B., Princeton University
(1981)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS OF THE
DEGREE OF

DOCTOR OF PHILOSOPHY
IN MATHEMATICS

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 1985

Signature redacted

Signature of Author _____
Department of Mathematics
May 20, 1985

Signature redacted

Certified by _____
Harold M. Stark
Thesis Supervisor

Signature redacted

Accepted by _____
Nesmith C. Ankeny
Chairman, Departmental Committee on Graduate Students

Acknowledgment

I want to express my gratitude to my teacher, adviser, and friend, Harold Stark, who shared not only his time and insight, but also his home.

Dedication

This thesis is dedicated to an earlier teacher, adviser, and friend -- my father.

THETA FUNCTIONS AND DIVISION POINTS ON ABELIAN VARIETIES OF DIMENSION TWO

by

DAVID R. GRANT

Submitted to the Department of Mathematics
on May 20, 1985 in partial fulfillment of the
requirements for the Degree of Doctor of Philosophy in
Mathematics

ABSTRACT

The fields generated by torsion points on abelian varieties defined over a number field have long been an object of interest. For elliptic curves, these fields are generated by the Weierstrass $\wp$-function and its derivative restricted to p-division points and play a crucial role in the Coates-Wiles theorem. Stark and Gupta produced a version of the theorem which utilizes rational primes and implicitly uses the modular properties of the $\wp$-function.

For Jacobians of curves of genus two we introduce a generalization of the $\wp$-function and express it in terms of theta functions, extending to genus two a classical elliptic formula. We discuss its modular properties when restricted to p-division points and relate it to the discriminant of the curve.

For abelian varieties of dimension two with complex multiplication by certain number fields, we generalize some of Gupta's results in the elliptic case, calculating discriminants in the tower of fields of p^n-division values. We also derive a character equation which governs the relations between these discriminants and those in a tower generated by a p^n-division value of a point of infinite order in the Mordell-Weil group.

Thesis Supervisor: Dr. Harold M. Stark

Title: Professor of Mathematics

Introduction

> A destiny that leads the English to the Dutch is strange enough; but one that leads from Epsom into Pennsylvania, and thence to the hills that shut in Altamont over the proud coral cry of the cock, and the soft stone smile of an angel, is touched by that dark miracle of chance which makes new magic in a dusty world.
>
> - Thomas Wolfe, Look Homeward, Angel [W]

It may appear a strange destiny that combines the mathematics of two centuries into one thesis: From an earlier era, we have the theory of theta functions, and from more recent times, the algebraic theory of complex multiplication. Our work, though not complete, is consistent. For our driving interest has been in the fields generated by torsion points -- and points whose multiples are a point of infinite order -- on an abelian variety of dimension two defined over a number field, and for this quest both theories can play a role.

The study of such fields is quite classical, and arises, for instance, in the proof of the Mordell-Weil theorem. For elliptic curves with complex multiplication, Coates and Wiles [CW] have exploited these fields to make a dent in the conjectures of Birch and Swinnerton-Dyer. Subsequently, Stark [St] and Gupta [G] gave another version of their proof. For abelian varieties of dimension two with complex multiplication by certain number fields, we extend some of Stark and Gupta's results, principally in the calculation of

discriminants in towers of fields generated by torsion points -- and the derivation of a character equation which governs the relations between this tower and a tower of fields generated by points dividing a point of infinite order in the Mordell-Weil group. This is done in chapter two where more details are provided in a separate introduction.

Stark and Gupta's version utilizes rational primes which remain inert in the field of complex multiplication. Their results make use of the fact that the towers of fields in question are basically generated by the Weierstrass p-function evaluated at p-division points -- which is an elliptic modular form of level p. If an elliptic curve E is analytically isomorphic to $\mathbb{C}/\Lambda$, Λ a lattice generated by 1 and τ, then Stark found [St]:

$$\prod_{\substack{u \neq \pm v \in \frac{1}{p}\Lambda \\ u, v \neq 0}} (p(u) - p(v)) = \pm p^{-2(p^2-3)} \Delta(E)^{(p^2-1)(p^2-3)/6}$$

where $\Delta(E)$ is the discriminant of the curve in its Weierstrass form. This equation between modular forms can be viewed as an analytic explanation of the role of primes of bad reduction of the elliptic curve in the fields generated by p-division points.

Besides elliptic curves, those of genus 2 are best understood. They are all hyperelliptic, and their moduli space (over $\mathbb{C}$) is dense in the fundamental domain of the Siegel upperhalf plane modulo the

action of the symplectic group. This often allows one to extend functions on the moduli space to Siegel modular functions. In particular, the discriminant of the curve is a modular form of weight 10.

This enables us to pose many questions by analogy with the elliptic case. Are there functions defined on the Jacobians of curves of genus 2 which:

1) take on algebraic values at torsion points when the curve is defined over a number field;

2) are modular forms of a certain level when restricted to p-division points; and

3) have the discriminant of the curve arising in the symmetrized product of the difference of the function evaluated at all p-division points?

There are two functions which we study in an attempt to meet these criteria. We denote them by $p(u)$ and $p(u, v)$ (§6). The former is to our tastes the nicer generalization of the Weierstrass p-function. It meets criterion 1, and it behaves like a meromorphic modular form on the moduli space of curves of genus 2, but we do not yet have a meromorphic continuation of it to the whole Siegel upper-half plane. The latter, $p(u, v)$ -- closely allied with $p(u)$ -- is the determinant of the difference of functions and is not a difference of functions itself. Yet, we have a meromorphic continuation of it to the whole Siegel upper-half plane, and we are able to isolate the role of the

discriminant in the product of $p(u, v)$ over pairs of p-division points.

In the first section of chapter 1 we review the classical theory of hyperelliptic curves and their Jacobians as they relate to curves of genus 2. In section 2 we gather certain facts about theta functions and their role in building up functions on the Jacobian. Mumford [M1] is a general reference for these sections. The function theory on the Jacobian is discussed more carefully in section 3, while in section 4 we cull together some analytic expressions for functions on the Jacobian which seem to put genus 2 practically on equal footing with genus 1. The chief reference is a wonderful old text by Baker* [Bak]. In particular, we isolate a function, $p(u)$, which is the "unique" even function with a pole of precisely order three along the theta divisor.

Some standard facts about modular forms are culled together in section 5. Exploiting the role of a change in symplectic basis on our curve, in section 6 we determine the "modular" properties of $p(u)$ and its allied function $p(u, v)$. In sections 7 and 9 we provide $p(u, v)$ with an expression in terms of theta functions which is analogous to the one which holds for the difference of the Weierstrass p-function evaluated at two points. We take a break in section 8 to

---

*I wish to thank several members of the Harvard English Department, all named Rosenberg, for getting this material into my hands.

see how one of these formulae relates to the addition law on the Jacobian. Finally, in section 10 we write $p(u, v)$ as the ratio of two modular forms, $N'(u, v)$, $D'(u, v)$, and discuss how the discriminant function appears in both when a product is taken over pairs of p-division values.

It seems a tradition among writers on theta functions to note that although the multitude of identities may seem at first chaotic, with a little time and patience they assume a logical pattern. I too believe this, but make no pretense that our calculations in sections 7 and 9 are anything less than unwieldy. The symmetry and coherence of the resulting formulae, however, seem to bear some testimony to the worth of the undertaking. Whether these will have any arithmetic application remains to be seen. They certainly have the applications enunciated by Landau [Land, p. 33].

TABLE OF CONTENTS

Chapter 1

§1: Curves of Genus Two and Their Jacobians

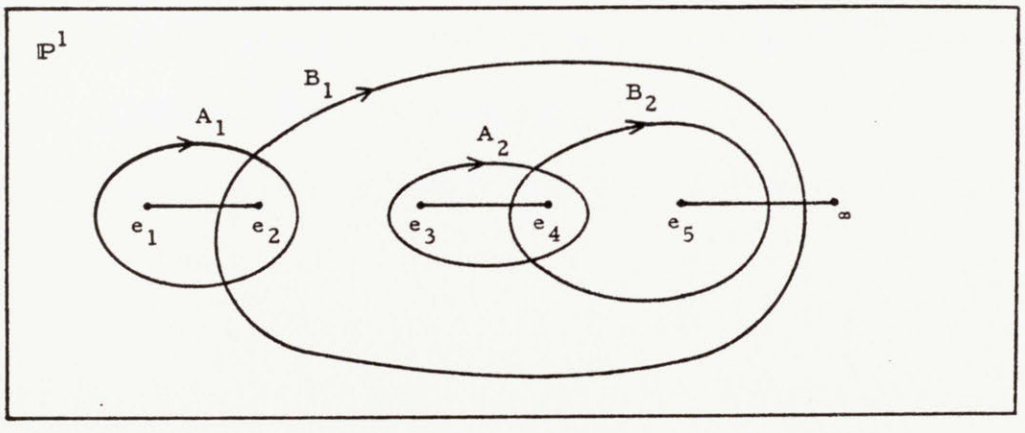
Let C be a curve of genus 2 defined over $\mathbb{C}$. Topologically, C is a 2-1 covering of $\mathbb{P}^1$, branched at six points. The isomorphism class of C is determined only up to projective equivalence of the six branch points; henceforth we will assume one branch point is at the point at infinity, ∞, and the other five at points e_i, $i = 1, \ldots, 5$.

In that case, C has a model as an affine curve

$$(1.1.0) \qquad y^2 = f(x) = \prod_{i=1}^{5} (x - a_i) = x^5 + b_1 x^4 + b_2 x^3 + b_3 x^2 + b_4 x + b_5$$

where x and y are functions on C, with poles only at ∞ of orders 2 and 5, respectively, and $a_i = x(e_i)$. We think of points P on C as solutions (x, y) to (1.1.0). C is endowed with an automorphism of order 2, the hyperelliptic "flip," mapping $P = (x, y)$ to $\overline{P} = (x, -y)$.

For hyperelliptic curves there is an explicit way to choose a symplectic basis for the first homology group, $H^1(C, \mathbb{Z})$. We "cut" the surface between e_1 and e_2, e_3 and e_4, and e_5 and ∞. Consider the following paths on $\mathbb{P}^1$:



We lift these paths to form a basis for $H^1(C, \mathbb{Z})$. The B paths follow different sheets over $\mathbb{P}^1$ as they traverse the cuts. If we let $\cdot$ denote intersection multiplicity, then by construction

$$A_i \cdot A_j = A_i \cdot B_j = B_i \cdot B_j = 0 \quad \text{for} \quad i \neq j \ ,$$

and if we defined intersection with the correct orientation, $A_i \cdot B_i = 1$.

If we now "cut" C along the basis, we are left with a simply-connected domain. A basis for the differentials of the first kind on C is given by:

$$\mu_1 = \frac{dx}{y} \ , \quad \mu_2 = \frac{xdx}{y}$$

and we define

$$\omega_{ij} = \int_{A_j} \mu_i \ , \quad \omega'_{ij} = \int_{B_j} \mu_i$$

Standard calculations show:

i) $\det \omega \neq 0$,

ii) $\tau = \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{12} & \tau_{22} \end{pmatrix} = \omega^{-1} \omega'$ is in $\mathfrak{H}^{(2)}$,

the Siegel upper-half plane of degree two. That is, τ is a symmetric matrix with positive definite imaginary part.

We will often want to normalize the differentials of the first kind by defining

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \omega^{-1} \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

then

$$\int_{A_j} v_i = \delta_{ij}, \quad \text{and} \quad \int_{B_j} v_i = \tau_{ij} \ .$$

Algebraically, we define the Jacobian of C, denoted $J(C)$, as the group of divisors of degree zero modulo linear equivalence. Since C is of genus 2, every divisor of degree zero is linearly equivalent to one of the form

$$P_1 + P_2 - 2 \cdot \infty \tag{1.1.1}$$

The expression (1.1.1) is almost unique. In fact, letting $\sim$ denote linear equivalence,

$$P_1 + P_2 - 2 \cdot \infty \sim P_3 + P_4 - 2 \cdot \infty$$

if and only if $\{P_1, P_2\} = \{P_3, P_4\}$, or $P_2 = \overline{P}_1$ and $P_4 = \overline{P}_3$; $P + \overline{P} \sim 2 \cdot \infty$ for all P. So the Jacobian can be considered $C \times C$ modulo the action switching P_1 and P_2 in (1.1.1), with the locus $(P, \overline{P})$ "blown down" to a single point at the origin of $J(C)$ [M2].

Analytically, we think of $J(C)$ as $\mathbb{C}^2$ modulo the lattice Λ generated by the period columns

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \ \begin{pmatrix} \tau_{11} \\ \tau_{12} \end{pmatrix}, \ \begin{pmatrix} \tau_{12} \\ \tau_{22} \end{pmatrix} \ .$$

The map from divisor classes to $\mathbb{C}^2/\Lambda$ is given by

$$\Phi : P_1 + P_2 - 2 \cdot \infty \longmapsto \int_\infty^{P_1} + \int_\infty^{P_2} (v_1, v_2) \text{ modulo } \Lambda$$

We embed C into $J(C)$ by

$$P - \infty \longmapsto \int_\infty^{P} (v_1, v_2) \text{ modulo } \Lambda$$

We will now explicitly give $J(C) - \Phi(C)$ (here "$-$" denotes set-theoretic difference) the structure of an affine variety.

The idea is to define a set of auxiliary polynomials for the divisor classes which correspond to $J(C) - \Phi(C)$. We then coordinatize the polynomials to endow $J(C) - \Phi(C)$ with the structure of an affine variety. The idea goes back to Jacobi, and we will follow the exposition of Mumford [M1]. The divisors which correspond to $J(C) - \Phi(C)$ are $\text{Div}^+ = \{P_1 + P_2 - 2 \cdot \infty \mid P_i \neq \infty, P_2 \neq \overline{P}_1\}$. If $D = P_1 + P_2 - 2 \cdot \infty$ is such a divisor, $P_i = (x_i, y_i)$, we define:

$$u_D(t) = (t - x_1)(t - x_2) \tag{1.1.2}$$

$$v_D(t) = \begin{cases} y_1 \left(\dfrac{t - x_2}{x_1 - x_2} \right) + y_2 \left(\dfrac{t - x_1}{x_2 - x_1} \right) & \text{for } x_1 \neq x_2 \\ \left(\dfrac{f'(x_1)}{2y_1} \right) t - \dfrac{x_1 f'(x_1)}{2y_1} + y_1 & \text{for } x_1 = x_2 \end{cases}$$

$\left(\text{so that } v_D(x_1) = y_1 \text{ , } \frac{d}{dt} v_D(t)\Big|_{t=x_1} = \frac{d}{dt}\sqrt{f(t)}\Big|_{t=x_1}\right)$. By construction: $f(t) - v_D^2(t) = u_D(t)\, w_D(t)$ for some polynomial $w_D(t)$. Let

(1.1.3)
$$u(t) = t^2 + u_1 t + u_2$$
$$v(t) = v_1 t + v_2$$

where the u_1, u_2, v_1, v_2 are undetermined coefficients. Then performing the division:

$$f(t) - v(t)^2 = u(t)\,[\text{cubic polynomial}] + R_1(u_i, v_j)\,t + R_2(u_i, v_j) \text{ .}$$

We see that the equations $R_1(u_i, v_j) = R_2(u_i, v_j) = 0$ turn the set of polynomials (u, v), u monic quadratic, v linear, such that $f - v^2 = uw$, into an affine variety. We've shown (u_D, v_D) lie on this variety, and in fact, there is a bijection between Div^+ and such pairs (u, v) [M1, II, p. 3.19]. This is how we endow Div^+ with the structure of an affine variety.

We note that:

(1.1.4)
$$u_1 = -(x_1 + x_2),\ u_2 = x_1 x_2,\ v_1 = \frac{y_1 - y_2}{x_1 - x_2},\ v_2 = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}$$

Therefore we find that the ring of regular functions on $J(C) - \Phi(C)$ is:

$$(1.1.5) \qquad A(J(C) - \Phi(C)) = \mathbb{C}[u_1, u_2, v_1, v_2]$$

$$= \mathbb{C}\left[x_1 + x_2,\ x_1 x_2,\ \frac{y_1 - y_2}{x_1 - x_2},\ \frac{x_2 y_1 - x_1 y_2}{x_1 - x_2}\right].$$

Since a finite number of translates of $J(C) - \Phi(C)$ cover all of $J(C)$, we can form an atlas that gives $J(C)$ the structure of an algebraic variety. The details are in [M1, II.]. We note that this implies the field of functions on $J(C)$ is given by the field of fractions of (1.1.5). We return to this field in §3. First we want to come up with another characterization of $\Phi(C)$.

§2: Theta Functions

Let $\tau \in \mathfrak{H}^{(2)}$, $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ in $\mathbb{C}_1^2$ and a, b column vectors in $\mathbb{Q}^2$. We define the (2-dimensional) theta function with characteristic $\begin{bmatrix} a \\ b \end{bmatrix}$ by

$$(1.2.0)\quad \theta\begin{bmatrix} a \\ b \end{bmatrix}(z,\tau) = \sum_{n \in \mathbb{Z}^2} e^{\pi i\, {}^t(n+a)\tau(n+a) + 2\pi i\, {}^t(n+a)(z+b)}$$

where n is written as a column vector. The function converges absolutely and uniformly on sets of bounded $\operatorname{im} z_i$ and $\operatorname{im} \tau > c\, I_{2\times 2}$ for some $c > 0$. As such, it is analytic on $\mathbb{C}^2 \times \mathfrak{H}^{(2)}$.

Translational Formulae (1.2.1). Let p, q be column vectors in $\mathbb{Z}^2$.

i) $\theta\begin{bmatrix} a \\ b \end{bmatrix}(z+\tau p+q) = e^{-\pi i\, {}^t p\tau p - 2\pi i\, {}^t p(z+b) + 2\pi i\, {}^t a q}\, \theta\begin{bmatrix} a \\ b \end{bmatrix}(z,\tau)$.

ii) $\theta\begin{bmatrix} a+p \\ b+q \end{bmatrix}(z,\tau) = e^{2\pi i\, {}^t a q}\, \theta\begin{bmatrix} a \\ b \end{bmatrix}(z,\tau)$

iii) $$\frac{\theta\begin{bmatrix} a \\ b \end{bmatrix}(z+\tau p+q)}{\theta\begin{bmatrix} a' \\ b' \end{bmatrix}(z+\tau p+q)} = \frac{\theta\begin{bmatrix} a+p \\ b+q \end{bmatrix}(z)}{\theta\begin{bmatrix} a'+p \\ b'+q \end{bmatrix}(z)}\, e^{-2\pi i\, {}^t p(b-b')}$$

We will use these freely throughout chapter 1. As already should be apparent, we frequently drop τ from the notation.

When $2\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{Z}^4$, we call $\begin{bmatrix} a \\ b \end{bmatrix}$ a theta characteristic (when considered as such, we usually only worry about a, b modulo 1). For such an $\begin{bmatrix} a \\ b \end{bmatrix}$,

$$\theta\begin{bmatrix} a \\ b \end{bmatrix}(-z, \tau) = e^{4\pi i\, {}^t ab}\, \theta\begin{bmatrix} a \\ b \end{bmatrix}(z, \tau)$$

and $\begin{bmatrix} a \\ b \end{bmatrix}$ is called <u>odd</u> or <u>even</u> depending on whether $e^{4\pi i\, {}^t ab}$ is minus or plus one (i.e., whether $\theta\begin{bmatrix} a \\ b \end{bmatrix}(z)$ is an odd or even function). Of course, $\theta\begin{bmatrix} a \\ b \end{bmatrix}(0, \tau) = 0$ for $\begin{bmatrix} a \\ b \end{bmatrix}$ odd. We will see later that $\theta\begin{bmatrix} a \\ b \end{bmatrix}(0, \tau) \neq 0$ for $\begin{bmatrix} a \\ b \end{bmatrix}$ even and τ a period matrix (this is peculiar to genera 1 and 2). There are 6 odd and 10 even theta characteristics. We note that by (1.2.1), for $\begin{bmatrix} a \\ b \end{bmatrix}$ a theta characteristic, $\theta\begin{bmatrix} a \\ b \end{bmatrix}^2(z)$ depends only on $\begin{bmatrix} a \\ b \end{bmatrix}$ modulo 1. Theta functions can be used in several ways to build up functions on $J(C)$, when τ is a period matrix of C. To wit:

(1.2.2) For $a_i, b_i, a_i', b_i' \in \mathbb{Z}^2$, $\sum a_i = \sum a_i'$, $\sum b_i = \sum b_i'$ modulo 1

$$\frac{\prod_i \theta\begin{bmatrix} a_i \\ b_i \end{bmatrix}(z, \tau)}{\prod_i \theta\begin{bmatrix} a_i' \\ b_i' \end{bmatrix}(z, \tau)}$$

is a function on $\mathbb{C}^2/\Lambda$ by (1.2.1), as is

(1.2.3) $$\frac{\partial^2}{\partial z_i \partial z_j} \log \theta\begin{bmatrix} a \\ b \end{bmatrix}(z, \tau)$$

where $i, j = 1$ or 2. In fact, all functions on a complex abelian variety can be built from the ratio of theta functions. We want to

express the regular functions on $J(C) - \Phi(C)$ which we found in §1 in terms of theta functions. Perhaps the key ingredient is the Riemann vanishing theorem, which states that there is an odd theta characteristic $\delta = \begin{bmatrix} \delta' \\ \delta'' \end{bmatrix}$, such that $\theta[\delta](z, \tau)$ vanishes to the first order precisely along $\Phi(C)$. Which odd theta characteristic plays the role of δ depends on the choice of basis for $H^1(C, \mathbb{Z})$ employed to define the Jacobian. One of the reasons we picked our basis so carefully is that we can identify δ for our embedding. The calculation is tedious, laborious, and classical [Bak; M1, II]. The result is $\delta = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$. The other five odd theta characteristics also play a central role in the function theory. There are 16 points T of order two in $J(C)$; the origin, $a_i - \infty$, $i = 1, \ldots, 5$; and $a_i + a_j - 2 \cdot \infty$, $i, j = 1, \ldots, 5$, $i \neq j$. Applying Φ, we get

$$\Phi(T) = \tau \eta'_T + \eta''_T \mod (\Lambda), \quad \text{for some } \eta'_T, \eta''_T \in \frac{1}{2}\mathbb{Z}^2 .$$

This defines an isomorphism from points of order two on $J(C)$ to theta characteristics. By a similarly painstaking calculation [M1, II] one can calculate $\begin{bmatrix} \eta'_T \\ \eta''_T \end{bmatrix}$ for $T = a_i - \infty$, which we denote by $\begin{bmatrix} \eta'_i \\ \eta''_i \end{bmatrix} = \eta_i$.

$$(1.2.4)\qquad \begin{bmatrix} \eta'_1 \\ \eta''_1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \eta'_2 \\ \eta''_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 0 \end{bmatrix}, \begin{bmatrix} \eta'_3 \\ \eta''_3 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \eta'_4 \\ \eta''_4 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \eta'_5 \\ \eta''_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

When T is the origin, we denote $\eta_T = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = [0]$ by η_0. It is easy to show that $\delta + \eta_i$, $i = 1, \ldots, 5$, are precisely the other five odd theta characteristics. For $\begin{bmatrix} a \\ b \end{bmatrix}$ a theta characteristic:

$$\theta \begin{bmatrix} \delta + a \\ b \end{bmatrix} (0, \tau) = 0 \iff \theta[\delta](\tau a + b, \tau) = 0$$

$$\iff \tau a + b \text{ is on } \Phi(C)$$

$$\iff \begin{bmatrix} a \\ b \end{bmatrix} = \eta_i \text{ for some } i .$$

It's important to note that two of these η_i are distinguished by being odd theta characteristics themselves, η_2 and η_4. However, the terminology of ordering is so strong that a_2, a_4, ∞ will be called <u>even</u> branch points, and a_1, a_3, a_5 <u>odd</u> branch points.

Although the choice of δ depends on our homology basis, we will call its zeroes, that is, $\Phi(C)$, <u>the</u> theta divisor and denote it by θ. Using (1.2.2) and restricting denominators to $\theta[\delta](z, \tau)$, it's easy to imagine how we can start building up regular functions on $J(C) - \theta$. The coordinate functions of $u_D(t)$ $(1, u_1, u_2)$ can be found by evaluating $u_D(t)$ at three distinct points. We will take as our choice the three odd branch points a_1, a_3, a_5.

We pull out a theorem from [M1, II, p. 3.113]. Let $U = \{1,3,5\}$, V a set of 3 indices in $\{1,\dots,5\}$, and let $U \circ V$ denote the symmetric difference $U \circ V = U \cup V - (U \cap V)$. Then for $k \in V$, $z = \Phi(D)$,

$$(1.2.5) \qquad u_D(a_k) = (-1)^{4\,{}^t\delta'\eta_k'' + 4\,{}^t\eta_{U\circ V}''\eta_k'} \cdot \prod_{\substack{i \in V \\ i \neq k}} (a_k - a_i) \frac{\theta[\eta_{U\circ V} + \eta_k]^2(0)\,\theta[\delta + \eta_k]^2(z)}{\theta[\eta_{U\circ V}]^2(0)\,\theta[\delta]^2(z)}$$

It is worth a moment to comment about the sign. We first observe that from (1.2.4), letting $s(k,i) = (-1)^{\eta_k''\eta_i'}$

$$(a_k - a_i) = s(k,i)\langle a_k - a_i\rangle$$

where

$$\langle a_k - a_i\rangle = \begin{cases} a_k - a_i & \text{if } k < i \\ a_i - a_k & \text{if } i < k \end{cases}$$

Indeed, $s(k,i)\,s(i,k) = -1$. Note also that modulo 1, $\eta_1 + \eta_3 + \eta_5 = \delta = \eta_2 + \eta_4$, so $\eta_{U\circ V} = \delta + \eta_V$. Now let's apply (1.2.5) with $V = \{k,i,j\}$, $z = \tau(\eta_\ell' + \eta_m') + \eta_\ell'' + \eta_m''$:

$$(a_k - a_\ell)(a_k - a_m) = (-1)^{4\,{}^t\delta'\eta_k'' + 4\,{}^t(\delta'' + \eta_V'')\eta_k'} (a_k - a_i)(a_k - a_j) \cdot \frac{\theta[\delta + \eta_i + \eta_j]^2(0)\,\theta[\delta + \eta_k + \eta_\ell + \eta_m]^2(0)}{\theta[\delta + \eta_k + \eta_i + \eta_j]^2(0)\,\theta[\delta + \eta_\ell + \eta_m]^2(0)}\, s(k,\ell)\, s(k,m)$$

Since by (1.2.1):

$$\frac{\theta[\delta+\eta_k]^2(z)}{\theta[\delta]^2(z)} = \frac{\theta[\delta+\eta_k+\eta_\ell+\eta_m]^2(0)}{\theta[\delta+\eta_\ell+\eta_m]^2(0)} e^{4\pi i \eta''_k(\eta'_\ell+\eta'_m)}$$

Now

$$(-1)^{4^t\delta'\eta''_k + 4^t(\delta''+\eta''_V)\eta'_k} = (-1)^{4^t(\delta'+\eta'_k)(\delta''+\eta''_k) - 4^t\delta''\delta'} (-1)^{4^t(\eta''_i+\eta''_j)\eta'_k}$$

$$= s(i,k)\, s(j,k)$$

since both δ and $\delta+\eta_k$ are odd theta characteristics. Therefore:

$$(1.2.6)\quad \frac{\langle a_k-a_\ell\rangle\langle a_k-a_m\rangle}{\langle a_k-a_i\rangle\langle a_k-a_j\rangle} = \frac{\theta[\delta+\eta_i+\eta_j]^2(0)\,\theta[\delta+\eta_k+\eta_\ell+\eta_m]^2(0)}{\theta[\delta+\eta_i+\eta_j+\eta_k]^2(0)\,\theta[\delta+\eta_\ell+\eta_m]^2(0)}$$

This holds independent of coincidence in our choice of i,j,ℓ,m. We will return to this later. For now it will suffice to consider the special case of (1.2.5), where $V=\{1,3,5\}$, (so $U\circ V=\phi$), $k\in U$. We have:

$$(1.2.7)\quad u_D(a_k) = (-1)^{4^t\delta'\eta''_k} \prod_{\substack{i\in U\\ i\neq k}} (a_k-a_i) \frac{\theta[\eta_k]^2(0)\,\theta[\delta+\eta_k]^2(z)}{\theta[0]^2(0)\,\theta[\delta]^2(z)}$$

where $z=\Phi(D)$.

To pull out u_1, u_2, we resort to a partial fractions decomposition

(1.2.8)
$$\frac{u_D(a_k)}{\prod_{k\in U}(t-a_k)} = \sum_{k\in U}\frac{x_k}{t-a_k}$$

where

$$x_k = \frac{u_D(a_k)}{\prod_{\substack{i\in U\\ i\neq k}}(a_k-a_i)}$$

and so (1.2.8) is equal to (since η_k is an even characteristic)

$$\prod_{\substack{i\in U\\ i\neq k}} s(k,i)\ \frac{\theta[\eta_k]^2(0)\ \theta[\delta+\eta_k]^2(z)}{\theta[0]^2(0)\ \theta[\delta]^2(z)} \qquad \text{and ,}$$

$$t^2+u_1t+u_2 = \sum_{k\in U} x_k \prod_{\substack{i\in U\\ i\neq k}}(t-a_i)$$

$$= \left(\sum_k x_k\right)t^2 + \sum_k\left(\sum_{i\neq k} - a_i\right)x_kt + \sum_k \sum_{i,j\neq k} a_ia_jx_k$$

So $\Sigma_k x_k = 1$, and therefore:

(1.2.9)
$$u_1 = \sum_k a_kx_k - \left(\sum_k x_k\right)(a_1+a_3+a_5) = \sum_k a_kx_k - \sum_k a_k$$

and
$$u_2 = \sum_k a_ia_jx_k \ , \quad i,j\neq k\ .$$

As for v_1, v_2, they'll turn out to be derivatives of u_1 and u_2 by a certain differential operator on $J(C)$. We'll pinpoint the operator later, but first we should look more closely at the function theory of $J(C)$.

§3: Functions on the Jacobian

The functions on $J(C)$ are precisely the functions on $C \times C$ which are invariant under the action of the symmetric group S_2. The functions of C are generated by x, y, so let x_i, y_i be the corresponding functions on two copies of C, C_i, $i = 1, 2$. Then the function field of $J(C)$ consists of those elements of $\mathbb{C}(x_1, x_2, y_1, y_2)$ invariant under the transposition of the subscripts 1 and 2. The transposition is an automorphism of order 2, so $K(J)$, the function field of $J(C)$, is a subfield of index 2. Clearly, $L = \mathbb{C}(x_1x_2, x_1 + x_2, y_1 + y_2, y_1y_2)$ is contained in $K(J)$, and $L(x_1)/L$ is at worst a quadratic extension. Also, it's easy to see that $x_2, y_1^2, y_2^2 \in L(x_1)$. Therefore $y_1^2 - y_2^2 = (y_1 - y_2)(y_1 + y_2) \in L(x_1)$, hence also $y_1 - y_2$, so y_1, y_2, too. Therefore $K(J) = \mathbb{C}(x_1x_2, x_1 + x_2, y_1 + y_2, y_1y_2)$. The discerning reader will note that the inclusion of y_1y_2 as a generator is redundant, but we keep it in tow because we wish to find the subfield $E(J)$ of even functions on $J(C)$, i.e. those invariant under $(x_1, y_1) + (x_2, y_2) - 2 \cdot \infty \to (x_1, -y_1) + (x_2, -y_2) - 2 \cdot \infty$.

We see just as quickly that $E(J) = \mathbb{C}(x_1x_2, x_1 + x_2, y_1y_2)$, since $K(J)/E(J)$ is quadratic, and $y_1 + y_2$ is quadratic over $\mathbb{C}(x_1x_2, x_1 + x_2, y_1y_2)$. ($(y_1 + y_2)^2 = y_1^2 + y_2^2 + 2y_1y_2$, and $y_1^2 + y_2^2 = f(x_1) + f(x_2)$ which is a symmetric polynomial in x_1 and x_2). Finally we note $(y_1y_2)^2 \in \mathbb{C}(x_1 + x_2, x_1x_2)$ and $x_1 + x_2, x_1x_2$ are algebraically independent functions on $J(C)$. Therefore we have the tower of fields:

$$\mathbb{C}(x_1, x_2, y_1, y_2)$$

$$| \ 2$$

$$K(J) = \mathbb{C}(x_1 + x_2, x_1 x_2, y_1 + y_2)$$

$$| \ 2$$

$$E(J) = \mathbb{C}(x_1 + x_2, x_1 x_2, y_1 y_2)$$

$$| \ 2$$

$$\mathbb{C}(x_1 + x_2, x_1 x_2)$$

Given a divisor D on $J(C)$, let $\ell(D)$ be the dimension of $L(D)$, the vector space of functions on $J(C)$ with poles at worst those of D, i.e., those f such that $(f) \geq -D$. The theta-divisor θ is ample, so we can apply the Riemann-Roch theorem for abelian varieties, which states [Lang 1; p. 99] for varieties of dimension 2, and positive integers r,

$$\ell(r\theta) = r^2$$

(The θ-divisor of a Jacobian is a principal polarization and hence has a trivial pfaffian.) So $L(\theta)$ consists only of the constants, and $L(2\theta)$ has a surplus of three dimensions over $L(\theta)$, two of which are made up by the functions u_1, u_2. The expressions in (1.2.8) and (1.2.9) guarantee their presence in $L(2\theta)$, and they are algebraically, and therefore linearly, independent. To find a third new dimension, we resort to (1.2.3): we can apply $\dfrac{\partial}{\partial z_1^2}$, $\dfrac{\partial^2}{\partial z_1 \partial z_2}$, and $\dfrac{\partial^2}{\partial z_2^2}$ to

$\log \theta(z, \tau)$ to garner three linearly independent functions in $L(2\theta)$. We will actually be slightly more careful in picking our differential operators. First we will introduce the two-dimensional σ-function.

§4: Two-dimensional σ and η functions

This section will cull together some facts from Baker [Bak] with some generalizations of his definitions.

We define the following two differentials of the second kind on C (with poles only at ∞, of orders 4 and 2, respectively).

$$\zeta_1 = \frac{3x^3 + 2b_1x^2 + b_2x}{4y}\,dx$$

$$\zeta_2 = \frac{x^2\,dx}{4y}$$

Corresponding to the symplectic basis (A_i, B_i), (although we won't show this dependence in our notation), we define the matrices η and η' by

$$\eta_{ij} = \int_{A_j} \zeta_i \ , \quad \eta'_{ij} = \int_{B_j} \zeta_i$$

Baker [Bak, pp. 14, 15; our definitions of η and ω differ by a sign or a factor of 2 here and there] shows:

(1.4.0) $$\eta\,{}^t\eta' = \eta'\,{}^t\eta \ , \quad {}^t\eta\omega = {}^t\omega\eta \ ,$$

and the Legendre-esque

$${}^t\omega\eta' - {}^t\eta\omega' = 2\pi i \cdot I_{2\times 2} \ ,$$

a proof of which requires close encounters with differentials of the third kind. We define for $z \in \mathbb{C}^2$

$$(1.4.1)\quad \sigma\begin{bmatrix} a \\ b \end{bmatrix}(z, \omega, \omega') = e^{-\frac{1}{2}\, {}^t z\, \eta \omega^{-1} z}\ \theta\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1} z,\ \omega^{-1}\omega') \ .$$

When $\begin{bmatrix} a \\ b \end{bmatrix} = \delta$, we simply denote this by $\sigma(z, \omega, \omega')$. What we've built into this notation is that σ is a function of z, the curve, and two choices: that of a symplectic basis for $H^1(C, \mathbb{Z})$, and that of a basis of differentials of the first kind. The latter choice will be removed shortly; the effect of the former will be discussed in the next section.

Let L be the lattice generated by the columns of ω and ω'. For a given z not in L, there is a unique divisor on C, $P_1 + P_2 - 2 \cdot \infty$, $P_i = (x_i, y_i)$, such that

$$z \equiv \int_{\infty}^{P_1} + \int_{\infty}^{P_2} (\mu_1, \mu_2) \bmod L \ .$$

Letting

$$p_{ij} = \frac{-\partial^2 \log \sigma(z, \omega, \omega')}{\partial z_i\ \partial z_j} \ ,$$

Baker shows [Bak, p. 38]

$$(1.4.2)\quad p_{22} = \frac{1}{4}(x_1 + x_2),\ p_{12} = -\frac{1}{4} x_1 x_2 \ ,\ p_{11} = \frac{F(x_1, x_2) - 2y_1 y_2}{4(x_1 - x_2)^2}$$

where

$$F(x_1,x_2) = (x_1+x_2)(x_1x_2)^2 + 2b_1(x_1x_2)^2$$

$$+ b_2(x_1+x_2)(x_1x_2) + 2b_3(x_1x_2) + b_4(x_1+x_2) + 2b_5$$

Remark: Baker considers a "Weierstrass" form $y^2 = 4x^5 + \cdots$ to avoid the $1/4$'s in (1.4.2). p_{22} is the genus 2 version of what Mumford calls "the hyperelliptic p-function." Compare [M1, II, §10] where our $\partial/\partial z_2$ is his D_∞. For the record, $D_\infty u_i = v_i$, for $i = 1,2$.

We can now see that $L(2\theta)$ is generated by

$$1, x_1 + x_2, x_1x_2 \quad \text{and} \quad \frac{F(x_1,x_2) - 2y_1y_2}{(x_1 - x_2)^2} ,$$

the latter of which is not redundant since $y_1y_2 \notin \mathbb{C}(x_1x_2, x_1+x_2)$. $F(x_1,x_2)$ has a pole of order 6 on θ, and $(y_1y_2)^2 \in L(10\theta)$, so $y_1y_2 \in L(5\theta)$. Hence the numerator of p_{11} has a pole of order 6 or θ. Therefore $(x_1-x_2)^2$ has a pole of order 4 on θ.

Let H stand for the Hessian operator

$$\begin{bmatrix} \dfrac{\partial^2}{\partial z_1^2} & \dfrac{\partial^2}{\partial z_1 \partial z_2} \\ \dfrac{\partial^2}{\partial z_1 \partial z_2} & \dfrac{\partial^2}{\partial z_2^2} \end{bmatrix} .$$

A function we would like to study more carefully is

$$p = p_{11}p_{22} - p_{12}^2 = \det H \log \sigma(z, \omega, \omega')$$

$$= \frac{1}{4}\left[\frac{(x_1+x_2)[(x_1x_2)^2(x_1+x_2) + 2b_1(x_1x_2)^2 + \cdots - 2y_1y_2] - (x_1x_2)^2(x_1-x_2)^2}{(x_1-x_2)^2}\right]$$

$$= \frac{1}{4}\left[\frac{4(x_1x_2)^3 + 2b_1(x_1x_2)^2(x_1+x_2) + \cdots - 2y_1y_2(x_1+x_2)}{(x_1-x_2)^2}\right]$$

which has a pole of order 3 on θ since the highest terms of $p_{11}p_{22}$ and p_{12}^2 cancel in the difference.

Before we commence our study we must extend some definitions. We note that any basis for the differentials of the first kind on C can be written as $m\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}$ for $m \in GL_2(\mathbb{C})$. For a given fixed symplectic basis, we write

$$\begin{pmatrix}\zeta_1(m)\\ \zeta_2(m)\end{pmatrix} = {}^tm^{-1}\begin{pmatrix}\zeta_1\\ \zeta_2\end{pmatrix}$$

and correspondingly, we get the integrals

$$\eta(m\omega) = {}^tm^{-1}\eta$$

$$\eta(m\omega') = {}^tm^{-1}\eta'$$

which we will extend by linearity to the lattice generated by the columns of $m\omega$ and $m\omega'$:

$$\eta(m\omega A + m\omega' B) = {}^tm^{-1}(\eta(\omega)A + \eta(\omega')B), \quad A, B \text{ integral matrices},$$

where $\eta(\omega) = \eta$, $\eta(\omega') = \eta'$. We can now rewrite (1.4.0) as

$$\eta(\omega)\,{}^t\eta(\omega') = \eta(\omega')\,{}^t\eta(\omega) \tag{1.4.3}$$

$${}^t\omega\,\eta(\omega') - {}^t\eta(\omega)\,\omega' = 2\pi i\, I_{2\times 2}$$

and note that these relations still hold when ω, ω' are replaced by $m\omega$, $m\omega'$.

Finally we can verify the homogeneity of the σ-function:

$$\sigma\begin{bmatrix} a \\ b \end{bmatrix}(mz, m\omega, m\omega')$$

$$= e^{-\frac{1}{2}\,{}^t(mz)\eta(m\omega)(m\omega)^{-1}(mz)}\ \theta\begin{bmatrix} a \\ b \end{bmatrix}((m\omega)^{-1}mz, (m\omega)^{-1}(m\omega'))$$

$$= \sigma\begin{bmatrix} a \\ b \end{bmatrix}(z, \omega, \omega')\ .$$

So the σ-function is independent of the choice of basis of differentials once we normalize z. We write the resulting function as a function of τ:

$$\sigma\begin{bmatrix} a \\ b \end{bmatrix}(z, \omega, \omega') = \sigma\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau).$$

and

$$p(\omega^{-1}z, \tau) = H \log \sigma(\omega^{-1}z, \tau)\ .$$

§5: Modular Forms of Degree Two

We let Γ denote the (degree two) symplectic group $Sp(2, \mathbb{Z})$. That is, matrices $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where A, B, C, D are 2×2 integral matrices satisfying (where I is the 2×2 identity matrix)

$$ {}^t\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} $$

or equivalently, writing out these conditions:

$$ (1.5.0) \qquad {}^tAC = {}^tCA,\ {}^tBD = {}^tDB \text{ and } {}^tAD - {}^tCB = I $$

Generators for Γ are all matrices of the form $\begin{pmatrix} I & B \\ 0 & I \end{pmatrix}$, B integral, and $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}$ [Ma].

We let $\Gamma(p)$ denote the subgroup of those $\gamma \in \Gamma$, such that $\gamma \equiv \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \bmod p$. $\Gamma = \Gamma(1)$. Γ acts on $\mathfrak{H}^{(2)}$ by

$$ \gamma \circ \tau = (A\tau + B)(C\tau + D)^{-1} \quad \text{where } \gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} $$

Let $g(\tau)$ be holomorphic on $\mathfrak{H}^{(2)}$. We say g is a <u>modular form</u> of <u>level p</u> and <u>weight</u> k if:

$$ g(\gamma \circ \tau) = \det(C\tau + D)^k\, g(\tau) \quad \text{for all } \gamma \in \Gamma(p) $$

(If g is meromorphic of weight 0 we call it a <u>modular function</u>.) The ring of modular forms of level 1 was explicitly determined by Igusa [I2] (subsequently by Hammond and Freitag [H, Fr]). It is generated by forms of weight 4, 6, 10, 12 and 35. The form of

weight 10 is related to the discriminant $\Delta(C)$ of our curve and we'll discuss its properties in §10.

So far we've been dealing with marked Riemann surfaces — those with a chosen homology basis. Associated to each basis and curve is our period matrix $\tau \in \mathfrak{H}^{(2)}$. We will show in the next section that a change in basis corresponds to the action of $\tau \to \gamma \circ \tau$ for some $\gamma \in \Gamma$. Conversely, given two period matrices τ, τ' in $\mathfrak{H}^{(2)}$, they correspond to the same curve if and only if there exists a $\gamma \in \Gamma$ such that $\tau' = \gamma \circ \tau$. The moduli space of curves of genus two — the space of isomorphism classes of curves — therefore sits inside the Siegal fundamental domain $F = \mathfrak{H}^{(2)}/\Gamma$. In fact, it is dense. More precisely, $\tau' \in \mathfrak{H}^{(2)}$ is not a period matrix if and only if $\tau' = \gamma \circ \tau$ for some $\gamma \in \Gamma$ where $\tau = \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{12} & \tau_{22} \end{pmatrix}$ satisfies $\tau_{12} = 0$. In such a case, $\mathbb{C}^2/\Lambda$ is isogenous to a product of one-dimensional tori.

Our approach will be to conjure up functions on the moduli space of curves of genus two and then study their extension to all of $\mathfrak{H}^{(2)}$. Our function p has a couple of problems: first it is not defined for τ not in the moduli space, and second, it is not analytic in τ. The latter is not so horrible, for $\theta[\delta]^3(z,\tau)\,p(z,\tau)$ is analytic on the moduli space. (The reason why η is analytic is that our basis can be chosen in a smooth way as τ varies [M1, II, §8]. Then the integrals of differentials along these paths will be holomorphic on the moduli space.) The first problem is more intractable, for although

we will demonstrate in the next section that $p(u)$ transforms like a modular function (of weight 2 and level 2p) on the moduli space when u is a point of order p on $\mathbb{C}^2/\Lambda$, we unfortunately don't yet have a good characterization of how η acts as τ approaches a point in $\mathfrak{H}^{(2)}$ off the moduli space. Therefore we will focus our attention in the later sections to a closely allied function which doesn't depend on η.

Igusa [I1] has shown that θ-functions are the building blocks of modular forms of any level.

Theta functions transform under Γ by:

$$(1.5.1)\quad \theta\begin{bmatrix} a^* \\ b^* \end{bmatrix}(z^*, \tau^*) = \zeta \det(C\tau+D)^{1/2} e^{\pi i\, {}^t z(C\tau+D)^{-1}Cz}\, \theta\begin{bmatrix} a \\ b \end{bmatrix}(z, \tau)$$

where $\zeta = \zeta(\gamma)$ = a complex number of absolute value 1, independent of z and τ

$$z^* = \gamma \circ z = {}^t(C\tau + D)^{-1} z$$

$$\tau^* = \gamma \circ \tau = (A\tau + B)(C\tau + D)^{-1}$$

$$\gamma \circ \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a^* \\ b^* \end{bmatrix} = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix}\begin{bmatrix} a \\ b \end{bmatrix} + \frac{1}{2}\left((C\,{}^t D)_o, (A\,{}^t B)_o\right)$$

and where M_o denotes the row vector consisting of the diagonal entries of M.

The map $\begin{bmatrix} a \\ b \end{bmatrix} \to \begin{bmatrix} a^* \\ b^* \end{bmatrix}$ induces an action of $\Gamma/\Gamma(2)$ on theta characteristics modulo 1. The action has two orbits, one being the 6 odd theta characteristics, the other the 10 even theta characteristics. $|\Gamma/\Gamma(2)| = 720$ and it acts as the symmetric group S_6 on

the odd theta characteristics [I2].

Likewise, $\Gamma/\Gamma(2p)$ acts on $\begin{bmatrix} a \\ b \end{bmatrix}$ modulo 1 where $2pa, 2pb \in \mathbb{Z}^2$. Since $\Gamma(2)/\Gamma(2p) \simeq \Gamma/\Gamma(p)$ for p odd [I1], this action is transitive on all $\begin{bmatrix} a \\ b \end{bmatrix}$ such that $2p\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{Z}^2$ but $2\begin{bmatrix} a \\ b \end{bmatrix} \notin \mathbb{Z}^2$.

§6: Modular Properties of $p(z)$

We will investigate the effect of a change of symplectic basis on σ and p.

Let A_1^*, A_2^*, B_1^*, B_2^* be a new symplectic basis for $H^1(C, \mathbb{Z})$. Then

$$\begin{pmatrix} B_1^* \\ B_2^* \\ A_1^* \\ A_2^* \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ A_1 \\ A_2 \end{pmatrix} \quad \text{for A, B, C, D } 2\times 2 \text{ integral matrices}$$

and the conditions $A_i^* \cdot A_j^* = B_i^* \cdot B_j^* = A_i^* \cdot B_j^* = 0$ for $i \neq j$ and $A_i^* \cdot B_i^* = 1$ are precisely equivalent to (1.5.0); hence

$\gamma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma$.

Let $C = [C_{ij}]$, $D = [D_{ij}]$. Then

$$\omega^* = \left[\int_{A_j^*} \mu_i\right] = \left[\int_{\Sigma_k C_{jk}B_k + \Sigma_k D_{jk}A_k} \mu_i\right] = \left[\sum_k (C_{jk}\omega'_{ik} + D_{jk}\omega_{ik})\right]$$

$$= \omega' {}^tC + \omega {}^tD$$

Likewise: $\omega'^* = \omega' {}^tA + \omega {}^tB$

$$\eta^* = \eta(\omega^*) = \eta(\omega') {}^tC + \eta(\omega) {}^tD$$

$$\eta'^* = \eta(\omega'^*) = \eta(\omega') {}^tA + \eta(\omega) {}^tB$$

And therefore:

$$\tau^* = {}^t\tau^* = {}^t\omega'^* {}^t\omega^{*-1} = (A\,{}^t\omega' + B\,{}^t\omega)(C\,{}^t\omega' + D\,{}^t\omega)^{-1}$$

$$= (A\tau + B)(C\tau + D)^{-1} = \gamma \circ \tau$$

For the convenience of the reader, we'll state a

Hessian Lemma (1.6.0): Let

$$H = \begin{pmatrix} \dfrac{\partial^2}{\partial z_1^2} & \dfrac{\partial^2}{\partial z_1 \partial z_2} \\ \dfrac{\partial^2}{\partial z_1 \partial z_2} & \dfrac{\partial^2}{\partial z_2^2} \end{pmatrix}$$

where $z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2$, A, B 2×2 integral matrices, ${}^tA = A$, g analytic. Then

i) $H({}^t z\, A\, z) = 2A$, and

ii) $H(g(Bz)) = {}^tB((Hg)(Bz))B$.

We now calculate. Recall

$$\sigma\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau) = e^{-\frac{1}{2}\,{}^t z\,\eta(\omega)\omega^{-1} z}\ \theta\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau)$$

By (1.4.0) $\eta(\omega)\omega^{-1}$ is symmetric, so

$$(1.6.1)\quad H\log\left(\sigma\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau)\right) = -\eta(\omega)\omega^{-1} + H\log\left(\theta\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau)\right)$$

Now by (1.5.1)

$$(1.6.2)\quad \theta\begin{bmatrix} a^* \\ b^* \end{bmatrix}({}^t(C\tau+D)^{-1}\omega^{-1}z, \gamma\circ\tau)$$

$$= \zeta \det(C\tau+D)^{1/2} e^{\pi i\, {}^t z\, {}^t\omega^{-1}(C\tau+D)^{-1}C\omega^{-1} z}\ \theta\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau)$$

where $$\begin{bmatrix} a^* \\ b^* \end{bmatrix} = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} + \frac{1}{2}\left((C\,{}^tD)_o, (A\,{}^tB)_o\right)$$

We note that

$$^t(C\tau + D)^{-1}\omega^{-1}z = (\omega(\tau C^t + D^t))^{-1}z = (\omega' C^t + \omega D^t)^{-1}z = \omega^{*-1}z$$

since $\tau = {}^t\tau$. We further note that $(C\tau + D)^{-1}C$ is symmetric, since its inverse, $\tau + C^{-1}D$ is by (1.5.0). So applying H log to (1.6.2) yields:

$$\text{(1.6.3)}\quad H\log\left(\theta\begin{bmatrix} a^* \\ b^* \end{bmatrix}(\omega^{*-1}z, \gamma\circ\tau)\right)$$

$$= -2\pi i\,{}^t\omega^{-1}(C\tau + D)^{-1}C\omega^{-1} + H\log\left(\theta\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau)\right)$$

Combining this with (1.6.1) we get

$$\text{(1.6.4)}\quad H\log\left(\sigma\begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau)\right) - H\log\left(\sigma\begin{bmatrix} a^* \\ b^* \end{bmatrix}(\omega^{*-1}z, \gamma\circ\tau)\right)$$

$$= -\eta(\omega)\omega^{-1} + \eta(\omega^*)\omega^{*-1} - 2\pi i\,{}^t\omega^{-1}(C\tau + D)^{-1}C\omega^{-1}$$

However,

$$-\eta(\omega)\omega^{-1} + \eta(\omega^*)\omega^{*-1} = -\eta(\omega)\omega^{-1} + (\eta(\omega')\,{}^tC + \eta(\omega)\,{}^tD)(\omega'\,{}^tC + \omega\,{}^tD)^{-1}$$

$$= (-\eta(\omega)\omega^{-1}(\omega'\,{}^tC + \omega\,{}^tD) + \eta(\omega')\,{}^tC + \eta(\omega)\,{}^tD)(\omega'\,{}^tC + \omega\,{}^tD)^{-1}$$

$$= (-\eta(\omega)\omega^{-1}\omega' + \eta(\omega'))\,{}^tC(\tau\,{}^tC + {}^tD)^{-1}\omega^{-1}$$

$$= (-\eta(\omega)\,{}^t\omega'\,{}^t\omega^{-1} + \eta(\omega'))(\tau + {}^tD\,{}^tC^{-1})^{-1}\omega^{-1}$$

$$= (-\eta(\omega)\,{}^t\omega' + \eta(\omega')\,{}^t\omega)\,{}^t\omega^{-1}(\tau + C^{-1}D)^{-1}\omega^{-1} \quad \text{by (1.5.0)}$$

$$= 2\pi i\,{}^t\omega^{-1}(C\tau + D)^{-1}C\omega^{-1} \quad \text{by (1.4.3)}$$

Hence both sides of (1.6.4) are zero. Therefore:

$$
\begin{aligned}
{}^{t}\omega^{-1}\left(H \log \sigma \begin{bmatrix} a \\ b \end{bmatrix}\right)(\omega^{-1}z, \tau)\,\omega^{-1} &= H \log \left(\sigma \begin{bmatrix} a \\ b \end{bmatrix}(\omega^{-1}z, \tau)\right) \\
&= H \log \left(\sigma \begin{bmatrix} a^{*} \\ b^{*} \end{bmatrix}(\omega^{*-1}z, \gamma \circ \tau)\right) \\
&= {}^{t}\omega^{*-1}\left(H \log \sigma \begin{bmatrix} a^{*} \\ b^{*} \end{bmatrix}\right)(\omega^{*-1}z, \gamma \circ \tau)\,\omega^{*-1}
\end{aligned}
$$

But

$$
\omega^{-1}\omega^{*} = {}^{t}(C\tau + D) \,, \tag{1.6.5}
$$

so

$$
\begin{aligned}
&\left(H \log \sigma \begin{bmatrix} a \\ b \end{bmatrix}\right)(\omega^{-1}z, \tau) \\
&\quad = (C\tau+D)^{-1}\left(H \log \sigma \begin{bmatrix} a^{*} \\ b^{*} \end{bmatrix}\right)({}^{t}(C\tau+D)^{-1}(\omega^{-1}z), \gamma \circ \tau)\,{}^{t}(C\tau+D)^{-1}
\end{aligned}
$$

Switching variables $(z' = \omega^{-1}z)$ and taking determinants, we have:

$$
\begin{aligned}
&\det\left(H \log \sigma \begin{bmatrix} a \\ b \end{bmatrix}\right)(z', \tau) \\
&\quad = \det\left(H \log \sigma \begin{bmatrix} a^{*} \\ b^{*} \end{bmatrix}\right)({}^{t}(C\tau+D)^{-1}z', \gamma \circ \tau)\,\det(C\tau+D)^{-2}
\end{aligned}
$$

Since we are taking logarithmic derivatives, by (1.2.1), the expression depends only on $\begin{bmatrix} a \\ b \end{bmatrix}$ modulo 1. In particular, when

$$
\begin{bmatrix} a \\ b \end{bmatrix} \equiv \begin{bmatrix} a^{*} \\ b^{*} \end{bmatrix} \equiv \delta \text{ modulo } 1,
$$

$$
p(z', \tau) = p(\gamma \circ z', \gamma \circ \tau)\,\det(C\tau+D)^{-2} \tag{1.6.6}
$$

By reasons similar to those of an upcoming argument, if $pz \in L$, then p transforms like a meromorphic modular form of level $2p$.

However, as we've lamented before, p is not defined on all of $\mathfrak{H}^{(2)}$.

Therefore, we shall shift our attention to:

$$(1.6.7) \qquad p(u,v,\tau) = \det(H \log \sigma(u,\tau) - H \log \sigma(v,\tau))$$

$$= \det(H \log \theta[\delta](u,\tau) - H \log \theta[\delta](v,\tau))$$

$$= \det \begin{vmatrix} p_{11}(u) - p_{11}(v) & p_{12}(u) - p_{12}(v) \\ p_{12}(u) - p_{12}(v) & p_{22}(u) - p_{22}(v) \end{vmatrix}$$

where $u, v \in \mathbb{C}^2$. In taking the difference, we have cancelled out the effect of the η-function and left ourselves within the realm of θ-functions, which are defined for all $\tau \in \mathfrak{H}^{(2)}$.

We must remember that in (1.6.7) the differentiation is taking place with respect to $u\omega$ and $v\omega$. So letting H_u, H_v stand for taking Hessians with respect to u and v, respectively, we have by the Hessian lemma:

$$(\det \omega)^2 p(u,v,\tau) = \det(H_u \log \theta[\delta](u,\tau) - H_v \log \theta[\delta](v,\tau))$$

This provides us with a meromorphic continuation of $(\det \omega)^2 p(u,v,\tau)$ to all of $\mathfrak{H}^{(2)}$. To investigate its modular properties, take $\gamma \in \Gamma$ such that $\gamma \circ \delta = \delta$ (in particular, $\gamma \in \Gamma(2)$) then just as we arrived at (1.6.6), we find:

$$p({}^t(C\tau + D)^{-1} u, {}^t(C\tau + D)^{-1} v, \gamma \circ \tau) = \det(C\tau + D)^2 p(u,v,\tau) .$$

In particular, if u, v are p-division values, i.e.

$$u = \tau\frac{\alpha}{p} + \frac{\beta}{p}, \quad v = \tau\frac{\varepsilon}{p} + \frac{\varphi}{p}, \qquad \alpha, \beta, \varepsilon, \varphi \in \mathbb{Z}^2$$

Then

$$(1.6.8) \qquad {}^t(C\tau+D)^{-1}u = (\gamma \circ \tau)\frac{(D\alpha - C\beta)}{p} + \frac{(-B\alpha + A\beta)}{p}$$

and similarly for v. So if in addition $\gamma \in \Gamma(p)$,

$$p\left(\gamma \circ \tau\frac{\alpha}{p} + \frac{\beta}{p}, \ \gamma \circ \tau\frac{\varepsilon}{p} + \frac{\varphi}{p}, \ \gamma \circ \tau\right)$$

$$= \det(C\tau + D)^2 \ p\left(\tau\frac{\alpha}{p} + \frac{\beta}{p}, \ \tau\frac{\varepsilon}{p} + \frac{\varphi}{p}, \ \tau\right)$$

so $(\det \omega)^2 \, p(u, v)$, $u, v \in \frac{1}{p}\Lambda$, is a meromorphic modular form of level $2p$ and weight 2.

We still have only expressed $(\det \omega)^2 \, p(u, v, \tau)$ in terms of derivatives of theta-functions. We will spend the next three sections finding another expression just in terms of theta-functions — piece by piece.

Remark: ω is a well-defined (analytic) function on the moduli space because we made a definite choice of symplectic basis for $H^1(C, \mathbb{Z})$. Considering the moduli space as sitting inside $\mathfrak{H}^{(2)}/\Gamma$, then ω is a function of τ, but only modulo the action of Γ.

§7: Some Long Calculations

There are two time-honored theta-function identities which we will employ. One is a special case of Riemann's theta formulae, which for dimension two states: [M1, I, p. 214]:

Riemann's Theta Formula (1.7.0): For $a, b, c, d, e, f, g, h \in \mathbb{Q}^2$, $x, y, u, v \in \mathbb{C}^2$,

$$\theta\begin{bmatrix}\frac{a+b+c+d}{2}\\ \frac{e+f+g+h}{2}\end{bmatrix}\left(\frac{x+y+u+v}{2}\right)\theta\begin{bmatrix}\frac{a+b-c-d}{2}\\ \frac{e+f-g-h}{2}\end{bmatrix}\left(\frac{x+y-u-v}{2}\right)\theta\begin{bmatrix}\frac{a-b+c-d}{2}\\ \frac{e-f+g-h}{2}\end{bmatrix}\left(\frac{x-y+u-v}{2}\right)\theta\begin{bmatrix}\frac{a-b-c+d}{2}\\ \frac{e-f-g+h}{2}\end{bmatrix}\left(\frac{x-y-u+v}{2}\right)$$

$$= \frac{1}{4}\sum_{\alpha,\beta \in \frac{1}{2}\mathbb{Z}^2} e^{-2\pi i\, {}^t\beta(a+b+c+d)}\,\theta\begin{bmatrix}a+\alpha\\ e+\beta\end{bmatrix}(x)\,\theta\begin{bmatrix}b+\alpha\\ f+\beta\end{bmatrix}(y)\,\theta\begin{bmatrix}c+\alpha\\ g+\beta\end{bmatrix}(u)\,\theta\begin{bmatrix}d+\alpha\\ h+\beta\end{bmatrix}(v)$$

A formula whose beauty, if not immediately apparent, has yet proved enduring.

The other is due to Thomae, and it relates "thetanulls," $\theta\begin{bmatrix}a\\ b\end{bmatrix}(0, \tau)$ for $\begin{bmatrix}a\\ b\end{bmatrix}$ a theta characteristic, to the branch points a_i on C.

Let S be a set containing an even number of the branch points $\{a_{i_k}\}$ (we will alternately consider S a set of a_{i_k} or merely its indices i_k). Recall that $\eta_S = \Sigma_{i \in S}\, \eta_i$ modulo 1, and $U = \{1, 3, 5\}$.

Thomae's Formula (1.7.1) [Mumford, II, §8]:

$$\theta[\eta_S]^4(0) = \begin{cases} d \prod_{\substack{i<j \\ i,j\in S\circ U}} (a_i - a_j) \prod_{\substack{i<j \\ i,j\notin S\circ U}} (a_i - a_j) & \text{if } \#\, S\circ U = 3 \\ 0 & \text{otherwise} \end{cases}$$

where $d = \pm(\det \omega/2\pi i)^2$. Since we will need this so frequently, we will make a chart of the 10 non-trivial cases.

(1.7.2)

S	S ∘ U	η_S	$i<j$, $i,j \in S\circ U$	$i<j$, $i,j \notin S\circ U$
ϕ	{1,3,5}	$\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$	(13)(15)(35)(24)	
{1,2}	{2,3,5}	$\begin{bmatrix}0\\0\\\frac{1}{2}\\0\end{bmatrix}$	(23)(25)(35)(14)	
{1,4}	{3,4,5}	$\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$	(34)(35)(45)(12)	
{2,3}	{1,2,5}	$\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\\0\end{bmatrix}$	(12)(15)(25)(34)	
{3,4}	{1,4,5}	$\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}$	(14)(15)(45)(23)	
{2,5}	{1,2,3}	$\begin{bmatrix}\frac{1}{2}\\0\\0\\\frac{1}{2}\end{bmatrix}$	(12)(13)(23)(45)	
{4,5}	{1,3,4}	$\begin{bmatrix}0\\\frac{1}{2}\\0\\0\end{bmatrix}$	(13)(14)(34)(25)	
{2,3,4,5}	{1,2,4}	$\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}$	(12)(14)(24)(35)	
{1,2,4,5}	{2,3,4}	$\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}$	(23)(24)(34)(15)	
{1,2,3,4}	{2,4,5}	$\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$	(24)(25)(45)(13)	

We now can begin to calculate the pieces of $p(u, v)$.

Culling together (1.4.2), (1.1.4) and (1.2.9) we have

$$(1.7.3)\quad p_{22}(u) - p_{22}(v) = \frac{1}{4}((x_1+x_2)(u) - (x_1+x_2)(v))$$

$$= -\frac{1}{4}\left(\sum_{k \in U} a_k (x_k(u) - x_k(v))\right)$$

where

$$(1.7.4)\quad x_k(z) = (-1)^{4\,{}^t\delta'\eta_k''} \frac{\theta[\eta_k]^2(0)\ \theta[\delta+\eta_k]^2(z)}{\theta[0]^2(0)\qquad \theta[\delta]^2(z)},\quad k = 1, 3, 5$$

So we need to calculate

$$(1.7.5)\quad x_k(u) - x_k(v)$$

$$= (-1)^{4\,{}^t\delta'\eta_k''} \frac{\theta[\eta_k]^2(0)}{\theta[0]^2(0)} \frac{\theta[\delta+\eta_k]^2(u)\ \theta[\delta]^2(v) - \theta[\delta+\eta_k]^2(v)\ \theta[\delta]^2(u)}{\theta[\delta]^2(u)\ \theta[\delta]^2(v)}$$

We resort to (1.7.0) to tackle the numerator of this last expression. Let $\begin{bmatrix} a \\ e \end{bmatrix} = 2\delta + \eta_k$, $\begin{bmatrix} b \\ f \end{bmatrix} = -\eta_k$, $\begin{bmatrix} c \\ g \end{bmatrix} = \begin{bmatrix} d \\ h \end{bmatrix} = 0$, and plug in $u+v$ for x, $u-v$ for y, and 0 for u and v. Then we have

$$(1.7.6)\quad \theta[\delta]^2(u)\ \theta[\delta+\eta_k]^2(v)$$

$$= \frac{1}{4} \sum_{\alpha, \beta \in \mathbb{Z}^2} e^{-2\pi i\, {}^t\beta(2\delta')}\, \theta\begin{bmatrix} 2\delta+\eta_k+\alpha \\ \beta \end{bmatrix}(u+v)\ \theta\begin{bmatrix} -\eta_k+\alpha \\ \beta \end{bmatrix}(u-v)\ \theta\begin{bmatrix} \alpha \\ \beta \end{bmatrix}^2(0)$$

Now

$$\theta\begin{bmatrix}2\delta+\eta_k+\alpha\\ \beta\end{bmatrix} = e^{2\pi i\, {}^t(2\delta'')(\eta_k'+\alpha)}\,\theta\begin{bmatrix}\eta_k+\alpha\\ \beta\end{bmatrix} \quad \text{by (1.2.1),}$$

likewise

$$\theta\begin{bmatrix}-\eta_k+\alpha\\ \beta\end{bmatrix} = e^{2\pi i\, {}^t(2\eta_k'')\alpha}\,\theta\begin{bmatrix}\eta_k+\alpha\\ \beta\end{bmatrix}$$

so (1.7.6) becomes

(1.7.7) $$\theta[\delta]^2(u)\;\theta[\delta+\eta_k]^2(v)$$

$$= \frac{1}{4}\sum_{\alpha,\beta\,\in\frac{1}{2}\mathbb{Z}^2} \mathrm{sign}(k,\alpha,\beta)\;\theta[\delta+\eta_k](u+v)\;\theta[\delta+\eta_k](u-v)\;\theta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}^2(0)$$

where

$$\mathrm{sign}(k,\alpha,\beta) = e^{-2\pi i\,({}^t\beta(2\delta')+{}^t(\eta_k'+\alpha)(2\delta'')+{}^t(2\eta_k'')\alpha)}$$

In (1.7.7) the six terms with $\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ odd can be dropped since $\theta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}(0) = 0$. We want to take (1.7.7) and subtract from it the same expression with the roles of u and v reversed. The only terms of the right hand side which will survive will be those for which $\theta\begin{bmatrix}\eta_k+\alpha\\ \beta\end{bmatrix}(u-v)$ also changes sign, i.e. those for which $\eta_k+\alpha$ is odd.

(1.7.8) $$\theta[\delta]^2(u)\;\theta[\delta+\eta_k]^2(v) - \theta[\delta]^2(v)\;\theta[\delta+\eta_k]^2(u)$$

$$= \frac{1}{2}\sum_{\substack{\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\text{even}\\ \eta+\begin{bmatrix}\alpha\\ \beta\end{bmatrix}\text{odd}}} \mathrm{sign}(k,\alpha,\beta)\;\theta[\delta+\eta_k](u+v)\;\theta[\delta+\eta_k](u-v)\;\theta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}^2(0)$$

Chart 7.1	$\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ even, $\eta_k + \begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ odd	sign(k, α, β)
k = 1, $\begin{bmatrix}\eta_1'\\ \eta_1''\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ 0\end{bmatrix}$	$\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ 0\end{bmatrix}$	-1
	$\begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}$	-1
	$\begin{bmatrix}0\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}$	1
	$\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}$	1
k = 3, $\begin{bmatrix}\eta_3'\\ \eta_3''\end{bmatrix} = \begin{bmatrix}0\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}$	$\begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\\ 0\\ 0\end{bmatrix}$	-1
	$\begin{bmatrix}0\\ 0\\ 0\\ \frac{1}{2}\end{bmatrix}$	1
	$\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ 0\end{bmatrix}$	1
	$\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}$	-1
k = 5, $\begin{bmatrix}\eta_5'\\ \eta_5''\end{bmatrix} = \begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}$	$\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ \frac{1}{2}\end{bmatrix}$	1
	$\begin{bmatrix}0\\ \frac{1}{2}\\ 0\\ 0\end{bmatrix}$	1
	$\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ 0\end{bmatrix}$	-1
	$\begin{bmatrix}0\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}$	-1

Plugging the signs from chart 7.1 into (1.7.8), and that into (1.7.5), and subsequently that into (1.7.3):

$$(1.7.9)\quad -8(p_{22}(u) - p_{22}(v))\,\theta[0]^2(0)\,\theta[\delta]^2(u)\,\theta[\delta]^2(v)$$

$$= -a_1\,\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\Big[-\theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)\,\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\0\end{bmatrix}^2(0)$$

$$-\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)$$

$$+\theta[\delta+\eta_5](u+v)\,\theta[\delta+\eta_5](u-v)\,\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)$$

$$+\theta[\delta+\eta_3](u+v)\,\theta[\delta+\eta_3](u-v)\,\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\Big]$$

$$+a_3\,\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\Big[-\theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\\0\end{bmatrix}^2(0)$$

$$+\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)\,\theta\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)$$

$$+\theta[\delta+\eta_5](u+v)\,\theta[\delta+\eta_5](u-v)\,\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)$$

$$-\theta[\delta+\eta_1](u+v)\,\theta[\delta+\eta_1](u-v)\,\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\Big]$$

$$-a_5\,\theta\begin{bmatrix}0\\0\\\frac12\\\frac12\end{bmatrix}^2(0)\left[\theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)\,\theta\begin{bmatrix}\frac12\\0\\0\\\frac12\end{bmatrix}^2(0)\right.$$

$$+\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)\,\theta\begin{bmatrix}0\\\frac12\\0\\0\end{bmatrix}^2(0)$$

$$-\theta[\delta+\eta_3](u+v)\,\theta[\delta+\eta_3](u-v)\,\theta\begin{bmatrix}\frac12\\0\\0\\0\end{bmatrix}^2(0)$$

$$\left.-\theta[\delta+\eta_1](u+v)\,\theta[\delta+\eta_1](u-v)\,\theta\begin{bmatrix}0\\\frac12\\\frac12\\0\end{bmatrix}^2(0)\right]$$

$$=\theta[\delta+\eta_1](u+v)\,\theta[\delta+\eta_1](u-v)(a_5-a_3)\,\theta\begin{bmatrix}0\\\frac12\\\frac12\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\0\\\frac12\\\frac12\end{bmatrix}^2(0)$$

$$+\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)$$

$$\cdot\left(a_1\theta\begin{bmatrix}\frac12\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac12\\\frac12\\\frac12\\\frac12\end{bmatrix}^2(0)+a_3\theta\begin{bmatrix}0\\\frac12\\\frac12\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\0\\0\\\frac12\end{bmatrix}^2(0)-a_5\theta\begin{bmatrix}0\\0\\\frac12\\\frac12\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\\frac12\\0\\0\end{bmatrix}^2(0)\right)$$

$$+\theta[\delta+\eta_3](u+v)\,\theta[\delta+\eta_3](u-v)(a_5-a_1)\,\theta\begin{bmatrix}\frac12\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\0\\\frac12\\\frac12\end{bmatrix}^2(0)$$

$$+\theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)$$

$$\cdot\left(a_1\theta\begin{bmatrix}\frac12\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\0\\\frac12\\0\end{bmatrix}^2(0)-a_3\theta\begin{bmatrix}0\\\frac12\\\frac12\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac12\\\frac12\\0\\0\end{bmatrix}^2(0)-a_5\theta\begin{bmatrix}0\\0\\\frac12\\\frac12\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac12\\0\\0\\\frac12\end{bmatrix}^2(0)\right)$$

$$+ \theta[\delta+\eta_5](u+v)\,\theta[\delta+\eta_5](u-v)(a_3-a_1)\,\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)$$

At this juncture, we would like to employ (1.7.2) to plug in values for $\theta\begin{bmatrix}a\\b\end{bmatrix}^2(0)$. Unfortunately, this forces us to choose lots of signs of square roots, which will affect our summations. We do have some information on the consistency of the signs, however.

Let's say we've chosen definite values for $(a_i - a_j)^{1/2}$, and pick the square root of d so that

$$\theta[0]^2(0) = d^{1/2}(a_1-a_3)^{1/2}(a_1-a_5)^{1/2}(a_3-a_5)^{1/2}(a_2-a_4)^{1/2}$$

Then define $e(S)$ by

$$\theta[\eta_S]^2(0) = e(S)\,d^{1/2}\prod_{\substack{i<j\\ i,j\in S\circ U}}(a_i-a_j)^{1/2}\prod_{\substack{i<j\\ i,j\notin S\circ U}}(a_i-a_j)^{1/2}$$

for S satisfying $\#\, S\circ U = 3$. Clearly $e(S)$ depends only on η_S modulo 1. Using induction and (1.2.6) repeatedly, it's not hard to show that

$$(1.7.10)\quad \frac{\prod_{i=1}^{r}\theta[\eta_{S_i}]^2(0)}{\prod_{i=1}^{r}\theta[\eta_{T_i}]^2(0)} \text{ involves no unsquared terms } (a_i-a_j)^{1/2} \iff \sum\eta_{S_i} = \sum\eta_{T_i} \text{ modulo 1.}$$

Since every even theta-characteristic is the sum of two even theta characteristics, and we've normalized $e(0) = 1$, we find $e(S)$ extends to a homomorphism of all theta characteristics into ± 1. In the following sections, we will only evaluate $\theta[\eta_S]^2(0)$ when summing terms with like surds: (1.7.10) and the fact that e is a homomorphism guarantee that we can do this unambiguously. We'll demonstrate with an example.

The coefficient of $\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)$ in the last equality of (1.7.9) is

(1.7.11)

$$\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\left[a_1+a_3\frac{\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)}{\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)}-a_5\frac{\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\\frac{1}{2}\\0\\0\end{bmatrix}^2(0)}{\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)}\right]$$

and we can utilize (1.2.6) twice:

$$\frac{\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)}{\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)}=\frac{\theta\begin{bmatrix}0\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)}{\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)}\cdot\frac{\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)}{\theta\begin{bmatrix}0\\0\\0\\0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)}$$

$$=\frac{(a_1-a_4)(a_1-a_5)}{(a_1-a_2)(a_1-a_4)}\,\frac{(a_1-a_3)(a_2-a_3)}{(a_1-a_3)(a_3-a_5)}=\frac{(a_1-a_5)(a_2-a_3)}{(a_1-a_2)(a_3-a_5)}$$

A similar calculation shows:

$$\frac{\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\;\theta\begin{bmatrix}0\\\frac{1}{2}\\0\\0\end{bmatrix}^2(0)}{\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\;\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)} = \frac{(a_1-a_3)(a_2-a_5)}{(a_1-a_2)(a_3-a_5)}$$

Plugging these back into (1.7.11), and using (1.7.2) to evaluate

$$\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\;\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\quad,$$

we find that the coefficient of $\theta[\delta+\eta_2](u+v)\;\theta[\delta+\eta_2](u-v)$ in (1.7.9) is

$$de(\delta+\eta_2)((a_1-a_4)(a_2-a_4)(a_3-a_4)(a_4-a_5))^{1/2}$$

$$\cdot\,(a_1(a_1-a_2)(a_3-a_5)+a_3(a_1-a_5)(a_2-a_3)-a_5(a_1-a_3)(a_2-a_5))$$

$$= e(\delta+\eta_2)\,d((a_1-a_4)(a_2-a_4)(a_3-a_4)(a_4-a_5))^{1/2}(a_1-a_3)(a_3-a_5)(a_1-a_5)$$

$$= e(\delta+\eta_2)\,d^{1/2}\left(\prod_{i,j\neq 2}\langle a_i-a_j\rangle^{1/2}\right)\theta[0]^2(0)$$

$$= \frac{e(\delta+\eta_2)\,d^{1/2}\,(\Delta(C))^{1/4}\,\theta[0]^2(0)}{\prod_{j\neq 2}\langle a_2-a_j\rangle^{1/2}}$$

using (1.7.2) again. Here $\Delta(C)$ is the discriminant of our curve

$$= \prod_{i<j} (a_i - a_j)^2 \ .$$

The calculations for the coefficients of

$$\theta[\delta + \eta_i](u+v)\ \theta[\delta + \eta_i](u-v)$$

for $i \neq 2$ are no worse. Their verification is left to the masochistic reader. The result is:

Proposition (1.7.12):

$$p_{22}(u) - p_{22}(v) = \frac{1}{8}\,\Delta(C)^{1/4}\, d^{1/2} \sum_{i=1}^{5} \xi_i(u, v)$$

where

$$\xi_i(u, v) = \frac{e(\delta+\eta_i)\,\theta[\delta+\eta_i](u+v)\ \theta[\delta+\eta_i](u-v)(-1)^{i-1}}{\prod_{j\neq i} \langle a_i - a_j \rangle^{1/2}\ \theta[\delta]^2(u)\ \theta[\delta]^2(v)}$$

Now that we've laid the groundwork, the calculation of $p_{12}(u) - p_{12}(v)$ is uniformly but not absolutely horrendous. Combining (1.4.2), (1.1.4) and (1.2.9) we get:

$$p_{12}(u) - p_{12}(v) = -\frac{1}{4}(x_1x_2(u) - x_1x_2(v)) = -\frac{1}{4} \sum_{k \in U} \left(\prod_{\substack{i \in U \\ i \neq k}} a_i \right) (x_k(u) - x_k(v))$$

Using our calculation of $x_k(u) - x_k(v)$ from (1.7.6) to (1.7.9) we obtain: (all we have to do is replace a_i in (1.7.9) by $a_j a_k$, $i \neq j,\ k \in U$):

$$(1.7.13)\quad -8(p_{12}(u) - p_{12}(v))\,\theta[\delta]^2(u)\,\theta[\delta]^2(v)\,\theta[0]^2(0)$$

$$= \theta[\delta+\eta_1](u+v)\,\theta[\delta+\eta_1](u-v)\,a_1(a_3-a_5)\,\theta\begin{bmatrix}0\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}^2(0)$$

$$+ \theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)$$

$$\cdot\left(a_3a_5\,\theta\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ 0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}^2(0) + a_1a_5\,\theta\begin{bmatrix}0\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\ 0\\ 0\\ \frac{1}{2}\end{bmatrix}^2(0) - a_1a_3\,\theta\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\ \frac{1}{2}\\ 0\\ 0\end{bmatrix}^2(0)\right)$$

$$+ \theta[\delta+\eta_3](u+v)\,\theta[\delta+\eta_3](u-v)\,a_3(a_1-a_5)\,\theta\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ 0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}^2(0)$$

$$+ \theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)$$

$$\cdot\left(a_3a_5\,\theta\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ 0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ 0\end{bmatrix}^2(0) - a_1a_5\,\theta\begin{bmatrix}0\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\\ 0\\ 0\end{bmatrix}^2(0) - a_1a_3\,\theta\begin{bmatrix}0\\ 0\\ \frac{1}{2}\\ \frac{1}{2}\end{bmatrix}^2(0)\,\theta\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ \frac{1}{2}\end{bmatrix}^2(0)\right)$$

$$+ \theta[\delta+\eta_5](u+v)\,\theta[\delta+\eta_5](u-v)\,a_5(a_1-a_3)\,\theta\begin{bmatrix}\frac{1}{2}\\ 0\\ 0\\ 0\end{bmatrix}^2(0)\,\theta\begin{bmatrix}0\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}^2(0)$$

Again, the coefficient of $\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)$ in (1.7.13) will lead to a worst case analysis. We find that coefficient is

$$-e(\delta+\eta_2)\,d((a_1-a_4)(a_2-a_4)(a_3-a_4)(a_4-a_5))^{1/2}$$

$$\cdot\,(a_3a_5(a_1-a_2)(a_3-a_5)+a_1a_5(a_1-a_5)(a_2-a_3)-a_1a_3(a_1-a_3)(a_2-a_5))$$

$$=-e(\delta+\eta_2)\,d((a_1-a_4)(a_2-a_4)(a_3-a_4)(a_4-a_5))^{1/2}$$

$$\cdot\,(-a_2(a_1-a_3)(a_1-a_5)(a_3-a_5))$$

$$=-e(\delta+\eta_2)\,d^{1/2}\Delta(C)^{1/4}\,a_2\,\theta[0]^2(0)\prod_{j\neq 2}\langle a_2-a_j\rangle^{-1/2}$$

Again we leave the other coefficients, which are equally pleasant, to the reader. We get

Proposition (1.7.14)

$$p_{12}(u)-p_{12}(v)=\frac{1}{8}\,\Delta(C)^{1/4}\,d^{1/2}\sum_{i=1}^{5}(-a_i)\,\xi_i(u,v)$$

Comparing this with (1.7.12), one might be led to conjecture

$$p_{11}(u)-p_{11}(v)=\frac{1}{8}\,d^{1/2}\,\Delta(C)^{1/4}\sum_{i=1}^{5}a_i^2\,\xi_i(u,v)$$

We caution such a reader that we have no such formulae as (1.1.4) with which to tackle this expression. However, there is a beautiful formula by Baker which will allow us to show that this ain't so very far from wrong. To wit: [Bak, p. 100]

$$(1.7.15)\quad \frac{\sigma(u+v)\,\sigma(u-v)}{\sigma^2(u)\sigma^2(v)}=h\left(p_{22}(u)p_{12}(v)-p_{22}(v)p_{12}(u)+\frac{p_{11}(v)}{4}-\frac{p_{11}(u)}{4}\right)$$

for some constant h. We will calculate h and more fully explain this formula in the next section.

§8: Interlude

If for the moment, p is the Weierstrass p-function, σ the one-dimensional σ-function then (for proof and definitions, see [Lang 2]).

$$(1.8.0) \qquad -\frac{\sigma(u+v)\,\sigma(u-v)}{\sigma^2(u)\sigma^2(v)} = p(u) - p(v)$$

Formulae (1.7.12) and (1.7.14) are attempts to generalize this result by modifying the right-hand side of (1.8.0). Formula (1.7.15), however, is more directly a generalization of (1.8.0) from the left-hand side. Let σ and p revert to their previous definitions.

Define

$$\xi_o(u, v) = \frac{\theta[\delta](u+v)\,\theta[\delta](u-v)}{\theta[\delta]^2(u)\,\theta[\delta]^2(v)}$$

which, as a function of u on $J(C)$, has a pole of order 2 on θ, and zeroes of order one on $\theta \pm v$. Likewise, as a function of v, it has a pole of order 2 on θ, and zeroes of order one on $\theta \pm u$. Therefore, as a function of u, letting $1, p_{11}, p_{12}, p_{22}$ be a basis $b_1, \ldots, b_4$ of $L(2\theta)$,

$$\xi_o(u, v) = \sum A_i(v)\, b_i(u)$$

But since the b_i are linearly independent, we can pick values for u which let us solve for $A_i(v)$, and show that $A_i(v) \in L(2\theta)$. So $A_i(v) = \gamma_{ij} b_j(v)$ and $\xi_o(u, v) = \Sigma\, \gamma_{ij}\, b_i(u) b_j(v)$. Since $\xi_o(v, u) =$

$-\xi_o(u,v)$, $\gamma_{ji} = -\gamma_{ij}$ so $\xi_o(u,v) = \sum_{i<j} \gamma_{ij}(b_i(u)b_j(v) - b_i(v)b_j(u))$. This is how we were first led to (1.7.15).

Baker apparently had several proofs in his day, the one in [Bak] is basically the one we have detailed so far [he figures the constants by finding a differential equation for σ and then expanding. We prefer to plug in 2-division points for u and v, as we will do shortly]. We would like to follow a different tack, one which relates to the addition law on $J(C)$. Since (1.7.15) enjoyed several proofs nearly a century ago, we will seek merely to outline another proof.

A general point of $J(C)$ is of the form $P_1+P_2 - 2\cdot\infty$. The addition law states that given 2 points, $P_1+P_2 - 2\cdot\infty$, $P_3+P_4-2\cdot\infty$, there is a unique point $\overline{P}_5+\overline{P}_6 - 2\cdot\infty$ such that

$$P_1+P_2 - 2\cdot\infty + P_3+P_4 - 2\cdot\infty \sim \overline{P}_5+\overline{P}_6 - 2\cdot\infty$$

or,
$$P_1+P_2+P_3+P_4+P_5+P_6 - 6\cdot\infty \sim 0$$

To determine P_5 and P_6, we can do the following: Given P_1, P_2, P_3, P_4, find a function g on C such that g has a pole of order 6 at ∞, and zeroes on P_i, $i = 1,\ldots,4$. Then the other two zeroes of g will determine P_5 and P_6. Such a g must be a polynomial in x and y, and since these enjoy poles of orders 2 and 5, respectively, at ∞, the general such g is of the form $g = \alpha y - \beta x^3 - \gamma x^2 - \delta x - \epsilon$, or allowing a, b, c, d to be meromorphic on $J(C)$,

(1.8.1)
$$g = y - ax^3 - bx^2 - cx - d .$$

Assume P_i, $i = 1, \ldots, 4$ are not ∞, and are distinct. Then if $P_i = (x_i, y_i)$ we have:

$$y_i = ax_i^3 + bx_i^2 + cx_i + d, \quad i = 1, \ldots, 4$$

so we can solve for the coefficients by Cramer's rule. For example:

$$a = \frac{\begin{vmatrix} y_1 & x_1^2 & x_1 & 1 \\ y_2 & x_2^2 & x_2 & 1 \\ y_3 & x_3^3 & x_3 & 1 \\ y_4 & x_4^2 & x_4 & 1 \end{vmatrix}}{D}, \quad b = \frac{\begin{vmatrix} x_1^3 & y_1 & x_1 & 1 \\ x_2^3 & y_2 & x_2 & 1 \\ x_3^3 & y_3 & x_3 & 1 \\ x_4^3 & y_4 & x_4 & 1 \end{vmatrix}}{D}, \quad \text{where}$$

$$D = \begin{vmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ x_2^3 & x_2^2 & x_2 & 1 \\ x_3^3 & x_3^2 & x_3 & 1 \\ x_4^3 & x_4^2 & x_4 & 1 \end{vmatrix}$$

To find x_5 and x_6, we take solutions to $(1.8.1)$ for a, b, c, d and equate

$$(ax^3 + bx^2 + cx + d)^2 = y^2 = x^5 + b_1x^4 + b_2x^3 + b_3x^2 + b_4x + b_5$$

This results in a sextic

$$a^2x^6 + (2ab - 1)x^5 + \cdots + (d^2 - b_5) = 0 \ .$$

Then

$$(1.8.2) \qquad \sum_{i=1}^{6} x_i = \frac{1 - 2ab}{a^2} \quad , \quad \prod_{i=1}^{6} x_i = \frac{d^2 - b_5}{a^2} \quad , \text{ etc.}$$

This allows us to solve for $x_5 + x_6$ and $x_5 x_6$. Suppose that $U = P_1 + P_2 - 2 \cdot \infty$, and $V = P_3 + P_4 - 2 \cdot \infty$, are variables on $J(C)$. We will alternately consider u or v fixed and the other variable. This should provide(a) little confusion. For example, when we speak of the zeroes of a function in u and v we encompass "the zeroes of u with v fixed" and vice versa. For the following we will always take u and v off θ so that the P_i are finite, $i = 1, \ldots, 4$. We continue to assume $P_i \neq P_j$ for $i \neq j$. For such u, v, $u+v$ is on θ if and only if a is zero. The poles of a are given by the zeroes of D which are just $x_i = x_j$, $i \neq j$ (D is Vandermonde). That is, a has a pole if $\overline{P}_i = \overline{P}_j$ for some $i = 1, 2$, $j = 3, 4$. We can't have $P_1 = \overline{P}_2$ or $P_3 = \overline{P}_4$ since we took u, v off θ. Both D and aD are zero when $P_i = P_j$. Therefore for u, v not on θ, $P_i \neq P_j$, for $i \neq j$:

$$\theta[\delta](u+v) = 0 \iff \frac{\begin{vmatrix} y_1 & x_1^2 & x_1 & 1 \\ y_2 & x_2^2 & x_2 & 1 \\ y_3 & x_3^2 & x_3 & 1 \\ y_4 & x_4^2 & x_4 & 1 \end{vmatrix}}{D} = 0$$

Switching v to $-v$ (note v is on θ, if any only if $-v$ is)

$$\theta[\delta](u-v) = 0 \iff \frac{\begin{vmatrix} y_1 & x_1^2 & x_1 & 1 \\ y_2 & x_2^2 & x_2 & 1 \\ -y_3 & x_3^2 & x_3 & 1 \\ -y_4 & x_4^2 & x_4 & 1 \end{vmatrix}}{D}$$

where the right-hand side has poles for $P_i = P_j$, $i = 1, 2$; $j = 3, 4$.

Therefore

$$A = \frac{\theta[\delta](u+v)\ \theta[\delta](u-v)}{\theta[\delta]^2(u)\ \theta[\delta]^2(v)}$$

has the same zeroes as

$$B = \frac{\begin{vmatrix} y_1 & x_1^2 & x_1 & 1 \\ y_2 & x_2^2 & x_2 & 1 \\ y_3 & x_3^2 & x_3 & 1 \\ y_4 & x_4^2 & x_4 & 1 \end{vmatrix} \begin{vmatrix} y_1 & x_1^2 & x_1 & 1 \\ y_2 & x_2^2 & x_2 & 1 \\ -y_3 & x_3^2 & x_3 & 1 \\ -y_4 & x_4^2 & x_4 & 1 \end{vmatrix}}{D^2}$$

for these u and v, and A has these zeroes to order 1. B, however, has poles when $P_i = P_j$ or $\overline{P}_j$ for $i = 1, 2$, $j = 3, 4$ (in short, when $x_i = x_j$, $i = 1, 2$; $j = 3, 4$). To cancel these poles we consider $B(x_1-x_3)(x_1-x_4)(x_2-x_3)(x_2-x_4)$, which is nothing more than

$$C = \frac{F(x_1,x_2) - 2y_1y_2}{(x_1-x_2)^2} - \frac{F(x_3,x_4) - 2y_3y_4}{(x_3-x_4)^2} + (x_1+x_2)x_3x_4 - x_1x_2(x_3+x_4)$$

$$= 16\left(\frac{p_{11}(u)}{4} - \frac{p_{11}(v)}{4} - p_{22}(u)\,p_{12}(v) + p_{22}(v)\,p_{12}(u)\right) .$$

Therefore, A and C have poles of precisely order 2 on θ (in u and v), and nowhere else. Further C has zeroes for $u \pm v$ on θ off the support of $(x_1-x_3)(x_1-x_4)(x_2-x_3)(x_2-x_4)$. Therefore C/A is analytic in u and so is a constant depending only on v – but is analytic in v so is a constant independent of v – and is therefore, just a constant.

To evaluate this constant, let

(1.8.3)

$$\frac{\theta[\delta](u+v)\,\theta[\delta](u-v)}{\theta[\delta]^2(u)\,\theta[\delta]^2(v)} = h\left(\frac{p_{11}(u)}{4} - \frac{p_{11}(v)}{4} - p_{22}(u)p_{12}(v) + p_{22}(v)p_{12}(u)\right)$$

$$= \frac{h}{16}\left(\frac{F(x_1,x_2) - 2y_1y_2}{(x_1-x_2)^2} - \frac{F(x_3,x_4) - 2y_3y_4}{(x_3-x_4)^2}\right.$$

$$\left. + (x_1+x_2)(x_3x_4) - x_1x_2(x_3+x_4)\right)$$

We will plug in $(x_1, y_1) = (a_1, 0), (x_2, y_2) = (a_2, 0)$, hence $u = \tau(\eta_1+\eta_2)' + (\eta_1+\eta_2)''$, and $(x_3, y_3) = (a_1, 0)$, $(x_4, y_4) = (a_3, 0)$, hence $v = \tau(\eta_1+\eta_3)' + (\eta_1+\eta_3)''$. To calculate the right-hand side of (1.8.3) we will use

$$\frac{F(x_1,x_2) - 2y_1y_2}{(x_1-x_2)^2} = \frac{f(x_1) + f(x_2) - 2y_1y_2}{(x_1-x_2)^2}$$

$$- (x_1+x_2)^3 + (x_1+x_2)x_1x_2 - b_1(x_1+x_2)^2 - b_2(x_1+x_2) - b_3$$

which follows directly from the definitions and long division.

For P_i, $i = 1, \ldots, 4$, as above we find

$$\frac{F(x_1,x_2)-2y_1y_2}{(x_1-x_2)^2}-\frac{F(x_3,x_4)-2y_3y_4}{(x_3-x_4)^2}$$

$$= -(a_1+a_2)^3+(a_1+a_2)a_1a_2-b_1(a_1+a_2)^2-b_2(a_1+a_2)-b_3$$

$$+(a_1+a_3)^3-(a_1+a_3)a_1a_3+b_1(a_1+a_3)^2+b_2(a_1+a_3)+b_3$$

$$= (a_2-a_3)[-3a_1^2-3a_1(a_2+a_3)-a_2^2-a_2a_3-a_3^2+a_1^2$$

$$+a_1(a_2+a_3)-b_1(2a_1+a_2+a_3)-b_2]$$

$$= (a_2-a_3)\left[-2a_1^2-2a_1a_2-2a_1a_3-a_2^2-a_3^2-a_2a_3\right.$$

$$\left.+\left(\sum_{i=1}^{5}a_i\right)(2a_1+a_2+a_3)-\sum_{1\le i\ne j\le 5}a_ia_j\right]$$

$$= (a_2-a_3)[a_1a_4+a_1a_5-a_4a_5]$$

So the right-hand side of (1.8.3) is

$$(1.8.4)\quad \frac{h}{16}((a_2-a_3)(a_1a_4+a_1a_5-a_4a_5)+(a_1+a_2)a_1a_3-a_1a_2(a_1+a_3))$$

$$= -\frac{h}{16}(a_2-a_3)(a_1-a_4)(a_1-a_5)$$

As for the left hand side of (1.8.3), we note that from the definition of θ:

$$(1.8.5)\quad \theta\begin{bmatrix}a\\b\end{bmatrix}(\tau c+d)=\theta\begin{bmatrix}a+c\\b+d\end{bmatrix}e^{-\pi i\,{}^tc\tau c-2\pi i\,{}^tc(b+d)},\ \text{for } c,d\in\mathbb{Q}^2$$

so the left hand side of (1.8.3) reduces to

$$\frac{\theta[\delta+2\eta_1+\eta_2+\eta_3](0)\,\theta[\delta+\eta_2-\eta_3](0)}{\theta[\delta+\eta_1+\eta_2]^2(0)\,\theta[\delta+\eta_1+\eta_3]^2(0)}\; e^{4\pi i\, {}^t(\eta_1+\eta_3)'\,\delta''}$$

and using (1.2.1) this becomes

$$\frac{\theta[\delta+\eta_2+\eta_3]^2(0)}{\theta[\delta+\eta_1+\eta_2]^2(0)\,\theta[\delta+\eta_1+\eta_3]^2(0)}\; e^{4\pi i[(\eta_1''+\eta_3'')(\delta'+\eta_2'+\eta_3')+(\eta_1+\eta_3)'\delta'']}$$

the exponential is just -1.

We can determine this by Thomae's formula (1.7.2):

$$\frac{\theta\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)}{\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\\\frac{1}{2}\end{bmatrix}^2(0)\;\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)}$$

$$= \frac{-e(\delta)\left((a_1-a_4)(a_1-a_5)(a_2-a_3)\right)^{1/2}}{d^{1/2}\left((a_1-a_2)(a_3-a_4)(a_3-a_5)(a_1-a_3)(a_2-a_4)(a_2-a_5)(a_4-a_5)\right)^{1/2}}$$

Combining (1.8.4) and (1.8.5) we get miraculously:

Proposition (1.8.7):

$$e(\delta)\,d^{1/2}\Delta(C)^{1/4}\,\xi_o(u,v) = 16\left(\frac{p_{11}(u)}{4}-\frac{p_{11}(v)}{4}-p_{22}(u)p_{12}(v)+p_{22}(v)p_{12}(u)\right)$$

We will now apply this to find $p_{11}(u) - p_{11}(v)$.

§9: Long Calculations, continued

In order to utilize proposition (1.8.7) in our search for $p_{11}(u) - p_{11}(v)$, we're going to have to calculate:

$$(1.9.0) \qquad -16(p_{22}(v)\,p_{12}(u) - p_{22}(u)\,p_{12}(v))$$

$$= ((x_1 + x_2)(u)\,x_1x_2(v) - (x_1 + x_2)(v)\,x_1x_2(u))$$

$$= \left(\sum_k a_k x_k(u) - \sum_k a_k\right)\left(\sum_k a_i a_j x_k(v)\right)$$

$$- \left(\sum_k a_k x_k(v) - \sum_k a_k\right)\left(\sum_k a_i a_j x_k(u)\right)$$

$$= E + D$$

where

$$E = \left(\sum_k a_k x_k(u)\right)\left(\sum_k a_i a_j x_k(v)\right) - \left(\sum_k a_k x_k(v)\right)\left(\sum_k a_i a_j x_k(u)\right)$$

and

$$D = \left(\sum_k a_k\right)\,(x_1x_2(u) - x_1x_2(v))$$

All summations are over the odd branch points. We have

$$(1.9.1)\quad E = \sum_{i<j} (x_i(u)x_j(v) - x_i(v)x_j(u))(a_i^2 - a_j^2)\,a_k$$

$$= \sum_{i<j} \theta[\eta_i]^2(0)\,\theta[\eta_j]^2(0)(\theta[\delta+\eta_i]^2(u)\,\theta[\delta+\eta_j]^2(v) - \theta[\delta+\eta_i]^2(v)\,\theta[\delta+\eta_j]^2(u))$$

$$\cdot \frac{(-1)^{4\,{}^t\delta\cdot(\eta_i'' + \eta_j'')}\,(a_i^2 - a_j^2)\,a_k}{\theta[0]^4(0)\;\theta[\delta]^2(u)\;\theta[\delta]^2(v)}$$

This is clearly a case for Riemann's Formula! We apply (1.7.0) with $\begin{bmatrix} a \\ e \end{bmatrix} = 2\delta + \eta_i + \eta_j$, $\begin{bmatrix} b \\ f \end{bmatrix} = \eta_i - \eta_j$, $\begin{bmatrix} c \\ g \end{bmatrix} = \begin{bmatrix} d \\ h \end{bmatrix} = 0$, $u+v$ and $u-v$ plugged in for x and y, and zero plugged in for u and v. This gives us:

$$(1.9.2)\quad \theta[\delta+\eta_i]^2(u)\;\theta[\delta+\eta_j]^2(v)$$

$$= \frac{1}{4} \sum_{\alpha,\beta \in \frac{1}{2}\mathbb{Z}^2} e^{-2\pi i\,{}^t\beta(2\delta' + 2\eta_i')}\,\theta\left[2\delta+\eta_i+\eta_j+\begin{matrix}\alpha\\ \beta\end{matrix}\right](u+v)$$

$$\cdot\,\theta\left[\eta_i - \eta_j + \begin{matrix}\alpha\\ \beta\end{matrix}\right](u-v)\;\theta\left[\begin{matrix}\alpha\\ \beta\end{matrix}\right]^2(0)$$

But

$$\theta\left[2\delta+\eta_i+\eta_j+\begin{matrix}\alpha\\ \beta\end{matrix}\right] = e^{2\pi i(2\delta'')(\eta_i' + \eta_j' + \alpha)}\,\theta\left[\eta_i+\eta_j+\begin{matrix}\alpha\\ \beta\end{matrix}\right]$$

and

$$\theta\left[\eta_i+\eta_j-2\eta_j+\begin{matrix}\alpha\\ \beta\end{matrix}\right] = e^{2\pi i(-2\eta_j'')(\eta_i' + \eta_j' + \alpha)}\,\theta\left[\eta_i+\eta_j+\begin{matrix}\alpha\\ \beta\end{matrix}\right]$$

So the sign (1.9.2) synthesizes to

$$\text{sign}(i,j,\alpha,\beta) = e^{-4\pi i\, {}^t\beta(\delta' + \eta_i') + 4\pi i(\delta'' - \eta_j'')(\eta_i' + \eta_j' + \alpha)}$$

We are interested in calculating

$$(1.9.3)\quad \theta[\delta+\eta_i]^2(u)\,\theta[\delta+\eta_j]^2(v) - \theta[\delta+\eta_i]^2(v)\,\theta[\delta+\eta_j]^2(u)$$

$$= \frac{1}{2}\sum \text{sign}(i,j,\alpha,\beta)$$

$$\cdot\, \theta\begin{bmatrix}\eta_i+\eta_j+\alpha\\ \beta\end{bmatrix}(u+v)\;\theta\begin{bmatrix}\eta_i+\eta_j+\alpha\\ \beta\end{bmatrix}(u-v)\;\theta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}^2(0)$$

where the only $\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ in the sum (1.9.3) are those for which $\theta\begin{bmatrix}\alpha\\ \beta\end{bmatrix}^2(0)$ is non-zero, and those for which $\theta\begin{bmatrix}\eta_i+\eta_j+\alpha\\ \beta\end{bmatrix}(u-v)$ changes signs when u and v are transposed, i.e., $\eta_i+\eta_j+\begin{matrix}\alpha\\ \beta\end{matrix}$ is odd.

We need another chart:

Chart 9.1	$\begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ even, $\eta_i + \eta_j + \begin{bmatrix}\alpha\\ \beta\end{bmatrix}$ odd	sign (i, j, α, β)
$i, j = 1, 3$ $\eta_i + \eta_j = \begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\\ \frac{1}{2}\\ 0\end{bmatrix}$	$\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$	1
	$\begin{bmatrix}\frac{1}{2}\\0\\0\\\frac{1}{2}\end{bmatrix}$	1
	$\begin{bmatrix}0\\\frac{1}{2}\\0\\0\end{bmatrix}$	-1
	$\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$	-1
$i, j = 1, 5$ $\eta_i + \eta_j = \begin{bmatrix}\frac{1}{2}\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$	$\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$	-1
	$\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\\0\end{bmatrix}$	1
	$\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}$	1
	$\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}$	-1
$i, j = 3, 5$ $\eta_i + \eta_j = \begin{bmatrix}0\\\frac{1}{2}\\0\\\frac{1}{2}\end{bmatrix}$	$\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$	1
	$\begin{bmatrix}0\\0\\\frac{1}{2}\\0\end{bmatrix}$	-1
	$\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}$	1
	$\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}$	-1

Therefore: $2E\theta[\delta]^2(u)\,\theta[\delta]^2(v)\,\theta[0]^4(0)$

$$= -(a_1^2 - a_3^2)a_5\,\theta[\eta_1]^2(0)\,\theta[\eta_3]^2(0)\left[-\theta[\delta](u+v)\,\theta[\delta](u-v)\,\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\right.$$

$$+\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)\,\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0) - \theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)\,\theta\begin{bmatrix}0\\\frac{1}{2}\\0\\0\end{bmatrix}^2(0)$$

$$\left.+\theta[\delta+\eta_5](u+v)\,\theta[\delta+\eta_5](u-v)\,\theta[0]^2(0)\right]$$

$$+(a_1^2 - a_5^2)a_3\,\theta[\eta_1]^2(0)\,\theta[\eta_5]^2(0)\left[-\theta[\delta](u+v)\,\theta[\delta](u-v)\,\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\right.$$

$$+\theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\\0\end{bmatrix}^2(0) - \theta[\delta+\eta_3](u+v)\,\theta[\delta+\eta_3](u-v)\theta[0]^2(0)$$

$$\left.+\theta[\delta+\eta_4]^2(u+v)\,\theta[\delta+\eta_4](u-v)\,\theta\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)\right]$$

$$-(a_3^2 - a_5^2)\,a_1\,\theta[\eta_3]^2(0)\,\theta[\eta_5]^2(0)\left[-\theta[\delta](u+v)\,\theta[\delta](u-v)\,\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\right.$$

$$+\theta[\delta+\eta_1](u+v)\,\theta[\delta+\eta_1](u-v)\,\theta[0]^2(0) - \theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)\,\theta\begin{bmatrix}0\\0\\\frac{1}{2}\\0\end{bmatrix}^2(0)$$

$$\left.+\theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)\,\theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\-\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\right]$$

$$= \theta[\delta](u+v)\,\theta[\delta](u-v)$$

$$\cdot\ ((a_1^2-a_3^2)a_5 - (a_1^2-a_5^2)a_3 + (a_3^2-a_5^2)a_1)\ \theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)$$

$$-\ \theta[\delta+\eta_1](u+v)\,\theta[\delta+\eta_1](u-v)(a_3^2-a_5^2)a_1\ \theta[0]^2(0)\ \theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)$$

$$+\ \theta[\delta+\eta_2](u+v)\,\theta[\delta+\eta_2](u-v)\Bigg[-(a_1^2-a_3^2)a_5\ \theta\begin{bmatrix}\frac{1}{2}\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)\ \theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)$$

$$+(a_1^2-a_5^2)a_3\ \theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)$$

$$+(a_3^2-a_5^2)a_1\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\Bigg]$$

$$-\theta[\delta+\eta_3](u+v)\,\theta[\delta+\eta_3](u-v)(a_1^2-a_5^2)a_3\ \theta[0]^2(0)\ \theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)$$

$$+\ \theta[\delta+\eta_4](u+v)\,\theta[\delta+\eta_4](u-v)\Bigg[(a_1^2-a_3^2)a_5\ \theta\begin{bmatrix}0\\\frac{1}{2}\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)$$

$$+(a_1^2-a_5^2)a_3\ \theta\begin{bmatrix}0\\0\\0\\\frac{1}{2}\end{bmatrix}^2(0)\ \theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)$$

$$-(a_3^2-a_5^2)a_1\ \theta\begin{bmatrix}\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)\ \theta\begin{bmatrix}0\\0\\\frac{1}{2}\\\frac{1}{2}\end{bmatrix}^2(0)\Bigg]$$

$$-\theta[\delta+\eta_5](u+v)\,\theta[\delta+\eta_5](u-v)(a_1^2-a_3^2)a_5\;\theta[0]^2(0)\;\theta\begin{bmatrix}\frac{1}{2}\\0\\0\\0\end{bmatrix}^2(0)\;\theta\begin{bmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{bmatrix}^2(0)$$

Let's calculate some of these coefficients. The coefficient of $\theta[\delta](u+v)\;\theta[\delta](u-v)$ is

$$-(a_1-a_3)(a_1-a_5)(a_3-a_5)\;\theta[\eta_1]^2(0)\;\theta[\eta_3]^2(0)\;\theta[\eta_5]^2(0)$$

$$= -e(\delta)\,d^{1/2}\,\Delta(C)^{1/4}\;\theta[0]^4(0)\ ,$$

using Thomae's formula.

The coefficient of $\theta[\delta+\eta_1](u+v)\,\theta[\delta+\eta_1](u-v)$ is

$$-\frac{e(\delta+\eta_1)\,a_1(a_3+a_5)\;\theta[0]^4\,d^{1/2}\,\Delta(C)^{1/4}}{\prod_{j\neq 1}\langle a_1-a_j\rangle^{1/2}}$$

The coefficients for η_3, η_5 are the same as for η_1 with the indices $\{1,3,5\}$ permuted. The calculation of the coefficient of $\theta[\delta+\eta_2](u+v)\theta[\delta+\eta_2](u-v)$ is nastier. After using Thomae's formula and simplifying we get:

$$\frac{e(\delta+\eta_2)\,d^{1/2}\,\Delta(C)^{1/4}\;\theta[0]^4(0)\;a_2(a_1+a_3+a_5-a_2)}{\prod_{j\neq 2}\langle a_2-a_j\rangle^{1/2}}$$

The coefficient for η_4 is the same as for η_2 with the indices 2 and 4 transposed.

So E can now be written as:

$$\left[- \frac{d^{1/2}\Delta(C)^{1/4}}{2} \sum_{i=1}^{5} a_i(a_1+a_3+a_5-a_i)\,\xi_i(u,v) \right] - \frac{e(\delta)}{2} d^{1/2}\Delta(C)^{1/4}\,\xi_o(u,v)$$

$$= \frac{d^{1/2}\Delta(C)^{1/4}}{2} \sum_{i=1}^{5} a_i^2\,\xi_i(u,v) - (a_1+a_3+a_5)(x_1x_2(u) - x_1x_2(v))$$

$$- 8\left(\frac{p_{11}(u)}{4} - \frac{p_{11}(v)}{4} - p_{22}(u)\,p_{12}(v) + p_{22}(v)\,p_{12}(u) \right)$$

by (1.7.11) and (1.8.7). Therefore, reinserting D clears a term and (1.9.0) becomes:

$$8\left(\frac{p_{11}(u)}{4} - \frac{p_{11}(v)}{4} \right) - 8(p_{22}(v)p_{12}(u) - p_{22}(u)p_{12}(v))$$

$$= \frac{1}{2} d^{1/2}\,\Delta(C)^{1/4} \sum_{i=1}^{5} a_i^2\,\xi_i(u,v)$$

we now add this to (1.8.7) multiplied by one-half to obtain:

Proposition (1.9.4):

$$p_{11}(u) - p_{11}(v) = \frac{1}{8} d^{1/2}\Delta(C)^{1/4} \left[e(\delta)\,\xi_o(u,v) + \sum_{i=1}^{5} a_i^2\,\xi_i(u,v) \right]$$

$$p(u,v) = \frac{d\Delta(C)^{1/2}}{64} \begin{vmatrix} e(\delta)\,\xi_o(u,v) + \sum a_i^2\,\xi_i(u,v) & -\sum a_i\,\xi_i(u,v) \\ -\sum a_i\,\xi_i(u,v) & \sum \xi_i(u,v) \end{vmatrix}$$

where the sums extend over $i = 1, \ldots, 5$.

A most satisfying simplification occurs after expansion of the determinant and infinite application of Thomae's formula. The determinant of (1.9.4) is

$$(1.9.5)\quad e(\delta)\xi_o(u,v)\sum_{i=1}^{5}\xi_i(u,v)+\sum_{i=1}^{5}a_i^2\,\xi_i(u,v)\sum_{i=1}^{5}\xi_i(u,v)$$

$$-\left(\sum_{i=1}^{5}a_i\,\xi_i(u,v)\right)^2$$

$$=e(\delta)\xi_o(u,v)\sum_{i=1}^{5}\xi_i(u,v)+\sum_{1\le i\le j\le 5}(a_i-a_j)^2\,\xi_i(u,v)\,\xi(u,v)$$

We want to plug in the definitions of ξ_o, ξ_i into (1.9.5) and multiply by $d\Delta(C)^{1/2}$. We'd better take this in pieces. Write

$$64p(u,v)=\sum_{i=1}^{5}A_i+\sum_{1\le i<j\le 5}B_{ij}$$

We first note:

$$\frac{B_{ij}}{\theta[\delta+\eta_i](u+v)\,\theta[\delta+\eta_i](u-v)\,\theta[\delta+\eta_j](u+v)\,\theta[\delta+\eta_j](u-v)}$$

$$=\frac{(-1)^{i+j}\,e(\delta+\eta_i)\,e(\delta+\eta_j)(a_i-a_j)^2\,d\prod_{i<j}\langle a_i-a_j\rangle}{\left(\prod_{k\neq i,j}\langle a_i-a_k\rangle^{1/2}\langle a_j-a_k\rangle^{1/2}\right)\langle a_i-a_j\rangle}$$

$$= (-1)^{i+j} e(\eta_i+\eta_j) d \left(\langle a_i - a_j \rangle^{1/2} \prod_{\substack{\ell,k \neq i,j \\ \ell \neq k}} \langle a_\ell - a_k \rangle^{1/2} \right)$$

$$\cdot \prod_{k \neq i,j} (\langle a_i-a_j\rangle\langle a_i-a_k\rangle\langle a_j-a_k\rangle\langle a_\ell-a_m\rangle)^{1/2}; \ \ell, m \neq i,j,k, \ \ell \neq m$$

$$= (-1)^{i+j} e(\eta_i + \eta_j) d^{-1} \prod_{q=1}^{4} e(\eta_{s(ij)_q}) \theta[\eta_{s(ij)_q}]^2(0)$$

where

$$s(ij)_q = \begin{cases} U \circ \{i,j,k\} & q = 1,2,3 \\ k \neq i,j & \\ U \circ \{\ell,m,n\} & q = 4 \\ \ell,m,n \neq i,j & \end{cases}$$

So

$$\prod_{q=1}^{4} e(\eta_{s(ij)_q}) = e\left(\sum_{k \neq i,j} (\delta+\eta_i+\eta_j+\eta_k) + \delta + \sum_{k \neq i,j} \eta_k \right) = e(\eta_i + \eta_j)$$

Therefore:

$$B_{ij} = (-1)^{i+j} d^{-1} \prod_{q=1}^{4} [\eta_{s(ij)_q}](0) \, \theta[\delta+\eta_i](u+v) \, \theta[\delta+\eta_i](u-v)$$

$$\cdot \theta[\delta+\eta_j](u+v) \, \theta[\delta+\eta_j](u-v)$$

We subsequently note:

$$\frac{A_i}{\theta[\delta+\eta_i](u+v)\;\theta[\delta+\eta_i](u-v)\;\theta[\delta+\eta_j](u+v)\;\theta[\delta+\eta_j](u-v)}$$

$$= \frac{(-1)^i\, e(\delta)\, e(\delta+\eta_i)\, d \prod_{i<j} \langle a_i - a_j \rangle}{\prod_{j\neq i} \langle a_i - a_j \rangle^{1/2}}$$

$$= (-1)^i\, e(\eta_i)\, d \prod_{j\neq i} \left(\langle a_i - a_j \rangle^{1/2} \prod_{\substack{k,\ell\neq i,j \\ k\neq\ell}} \langle a_k - a_\ell \rangle^{1/2} \right)$$

$$= (-1)^i\, e(\eta_i)\, d^{-1} \prod_{q=1}^{4} e(\eta_{s(i)_q})\, \theta[\eta_{s(i)_q}]^2(0)$$

where

$$s(i)_q = \{ U \circ \{j,k,\ell\} \}\,, \quad q = 1,2,3,4\,. \qquad j,k,\ell\neq i$$

Therefore

$$\eta_{s(i)_q} = \sum_{j\neq i} \eta_j = \eta_i \text{ modulo } 1\,.$$

Hence

$$\prod_{q=1}^{4} e(\eta_{s(i)_q}) = e(\eta_i)\,.$$

A rather pleasant thing has occurred. Letting $\eta_0 = 0$, we find $A_i = B_{oi}$! That is, $(-1)^{i+0} = (-1)^i$; $e(\eta_o + \eta_i) = e(\eta_i)$ (since we defined $d^{1/2}$ by $e(0) = 1$); and $s(i)_q = s(oi)_q$ once we note

$\eta_{U \circ \{i,j,k\}} = \eta_{\overline{U \circ \{i,j,k\}}} = \eta_{U \circ (\ell, m, o)}$ (where the bar denotes taking the complement in the set of indices $\{o, i, j, k, \ell, m\}$). Therefore the asymmetry of fixing a point at ∞ in (1.1.0) has disappeared from the "numerator" of $p(u,v)$ — as has our need to determine the signs $e(\eta_s)$. Completing our calculation we have:

Proposition (1.9.7):

$$64 dp(u,v) = \frac{N(u,v)}{D(u,v)} =$$

$$\frac{\sum_{0 \le i < j \le 5} (-1)^{i+j} \theta[\delta+\eta_i](u+v)\theta[\delta+\eta_i](u-v)\theta[\delta+\eta_j](u+v)\theta[\delta+\eta_j](u-v) \prod_{q=1}^{4} \theta[\eta_{s(ij)_q}]^2(0)}{\theta[\delta]^4(u)\ \theta[\delta]^4(v)}$$

where $N(u,v)$ and $D(u,v)$ are the numerator and denominator of the fraction on the right.

It is also worth noting that by (1.7.1):

$$dp(u,v) = \frac{\pm 1}{4\pi^2} (\det \omega)^2 p(u,v) \ .$$

§10: Modular Properties of $p(u, v)$ and Discriminants

We are finally ready to study the modular properties of $64dp(u, v) = \frac{N(u, v)}{D(u, v)}$. Both $N(u, v)$ and $D(u, v)$ are analytic, and being expressed by theta functions, are defined on all of $\mathfrak{H}^{(2)}$. We will even be able to adjust D and N so that each is modular in its own right.

For what follows, we need to have a better understanding of the factor ζ of absolute value 1 in the transformation formula for $\theta\begin{bmatrix} a \\ b \end{bmatrix}(z, \tau)$. In particular, if $z = 0$, $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma$ then (1.5.1) becomes [II, p. 176]

$$(1.10.1) \qquad \theta\begin{bmatrix} a^* \\ b^* \end{bmatrix}(0, \gamma \circ \tau) = \zeta \det(C\tau + D)^{1/2}\, \theta\begin{bmatrix} a \\ b \end{bmatrix}(0, \tau)$$

where

$$\zeta(\gamma) = \rho(\gamma)\, e^{-\pi i[{}^t a\, {}^t B D a - 2\, {}^t a\, {}^t B C b + {}^t b\, {}^t A C b + ({}^t a\, {}^t D - {}^t b\, {}^t C)(A\, {}^t B)_0]},$$

$\rho(\gamma)$ is an eighth root of unity,

and $$\begin{bmatrix} a^* \\ b^* \end{bmatrix} = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} + \frac{1}{2}((C\, {}^t D)_0, (A\, {}^t B)_0)$$

If $\gamma \in \Gamma(2)$ then $\rho(\gamma)$ is a fourth root of unity [M, I, pp. 194, 207]. Also note that by (1.2.1), if $a, b \in \frac{1}{p}\mathbb{Z}^2$, then $\theta\begin{bmatrix} a \\ b \end{bmatrix}^p(z, \tau)$ depends only on $\begin{bmatrix} a \\ b \end{bmatrix}$ modulo 1. Therefore from (1.10.1) we get:

Lemma (1.10.2). For $a, b \in \frac{1}{2p}\mathbb{Z}^2$, p odd:

i) $\theta\begin{bmatrix} a \\ b \end{bmatrix}^4 (0, \tau)$ is a modular form of level $2p^2$ and weight 2.

ii) $\theta\begin{bmatrix} a \\ b \end{bmatrix}^{4p} (0, \tau)$ is a modular form of level 2p and weight 2p.

Proof: i) Since $\Gamma(2p^2) \subset \Gamma(2)$, ρ^4 will be 1 in the transformation, so it's easy to see that ζ^4 will be 1, too. Further, $\begin{bmatrix} a^* \\ b^* \end{bmatrix} \equiv \begin{bmatrix} a \\ b \end{bmatrix} \bmod p$, so by (1.2.1), the characteristic is left unchanged.

ii) Here we only find that ζ^4 is a pth-root of unity, but this is why we raise to the pth-power. Also $\begin{bmatrix} a^* \\ b^* \end{bmatrix} \equiv \begin{bmatrix} a \\ b \end{bmatrix} \bmod 1$, but that's all that matters since we have raised to the pth-power. Q.E.D.

Let p be odd, $\gamma \in \Gamma(2p)$. We showed in §6 that $(\det \omega)^2 p(u, v)$ is a modular function of level 2p when u, v are points of order p in $\mathbb{C}^2/\Lambda$, and τ is restricted to being a period matrix. The meromorphic continuation of $(\det \omega)^2 p(u, v)$ to all of $\mathfrak{H}^{(2)}$ provided by (1.9.7) lends us a verification of its modular properties directly from those of the theta function. If we consider the action of γ on (1.9.7) we see for starters that since $\gamma \in \Gamma(2)$, it preserves the theta characteristics. In addition, if we write

$$(1.10.3) \quad u = \tau\frac{\alpha}{p} + \frac{\beta}{p}, \quad v = \tau\frac{\varepsilon}{p} + \frac{\varphi}{p}, \quad \alpha, \beta, \varepsilon, \varphi \in \mathbb{Z}^2$$

then the action of $\gamma \in \Gamma(p)$ on u and v doesn't change $\alpha, \beta, \varepsilon, \varphi$ modulo p (1.6.8), and therefore doesn't affect $p(u, v)$ which is periodic with respect to Λ in both u and v. The coup de grace is

delivered by the realization that $\gamma \in \Gamma(2)$ implies that $\rho(\gamma)$ is a fourth root of unity, as is $\zeta(\gamma)$.

Remark 1. We can go further along this line. If we now only assume $\gamma \in \Gamma(p)$, γ no longer fixed the denominator of $64dp(u, v)$ but it leaves the numerator "fixed." This is because of the symmetry in the summation of (1.9.7) and the weighted sign $(-1)^{i+j}$. To be precise, if we let $D\begin{bmatrix}a\\b\end{bmatrix}(u, v)$ denote $\theta\begin{bmatrix}a\\b\end{bmatrix}^4(u)\,\theta\begin{bmatrix}a\\b\end{bmatrix}^4(v)$, then for $\gamma \in \Gamma(p)$,

$$\frac{N(u, v)}{D[\delta](u, v)}(\gamma \circ \tau) = \pm \frac{N(u, v)}{D[\gamma \circ \delta](u, v)}(\tau)$$

where $\gamma \circ \delta$ denotes the action on the theta characteristic. We have a rather unilluminating proof of this fact and will therefore omit it.

Now $D\begin{bmatrix}a\\b\end{bmatrix}(u, v)$ $\left(\begin{bmatrix}a\\b\end{bmatrix}\right.$ a theta characteristic$\left.\right)$ is almost modular on its own. In fact

$$D'\begin{bmatrix}a\\b\end{bmatrix}(u, v) = W\left(\begin{bmatrix}a\\b\end{bmatrix}, \alpha, \beta, \varepsilon, \varphi\right) D\begin{bmatrix}a\\b\end{bmatrix}(u, v)$$

$$= \theta\begin{bmatrix}a + \alpha/p\\ b + \beta/p\end{bmatrix}^4(0)\,\theta\begin{bmatrix}a + \varepsilon/p\\ b + \varphi/p\end{bmatrix}^4(0)$$

is modular of level $2p^2$ by (1.10.2), where u, v are as in (1.10.3) and

$$(1.10.5)\quad W\left(\begin{bmatrix}a\\b\end{bmatrix}, \alpha, \beta, \varepsilon, \varphi\right)$$

$$= e^{(4\pi i/p^2)({}^t\alpha\tau\alpha + {}^t\varepsilon\tau\varepsilon + 2{}^t\alpha(b+\beta) + 2{}^t\varepsilon(b+\varphi)}$$

which is never zero.

Therefore, $N'(u,v) = N(u,v)\, W([\delta], \alpha, \beta, \epsilon, \varphi)$ is a modular form of weight 6 and level $2p^2$, and $64d\wp(u,v) = \dfrac{N'(u,v)}{D'(u,v)}$.

Our goal is to build up a "discriminant-like" object by multiplying $64d\wp(u,v)$ over all pairs $(u,v) \in \frac{1}{p}\Lambda$, u and $v \neq 0$ (to avoid points where it has poles identically in τ) and $u \neq \pm v$ (points where it is identically zero). This product will be of level 2 — we are in effect taking a norm from the field of modular functions of level $2p$ to those of level 2. It's difficult to nail down what modular function we obtain in general, but we can gauge what part of it is made up of discriminants of our curve.

We need to gather some facts about the lone (up to constants) modular form of weight 10. In terms of theta functions, it is given by [I2]

$$\Delta(\tau) = \prod_{\left[\begin{smallmatrix}a\\b\end{smallmatrix}\right]\text{even}} \theta\left[\begin{smallmatrix}a\\b\end{smallmatrix}\right]^2 (0,\tau)$$

The reason for the Δ symbol is that this function is essentially the discriminant of the curve for which τ is the period matrix. Thomae's formula gives us this precisely:

$$\Delta(\tau) = d^5 \prod_{i<j} (a_i - a_j)^2 = d^5 \Delta(C)$$

since $\prod_{\left[\begin{smallmatrix}a\\b\end{smallmatrix}\right]\text{even}} e\left(\left[\begin{smallmatrix}a\\b\end{smallmatrix}\right]\right) = 1$.

Hammond [H] shows $\Delta(\tau') = 0 \iff \tau' = \gamma \circ \tau$ for some $\gamma \in \Gamma$, $\tau = \begin{pmatrix} \tau_{11} & 0 \\ 0 & \tau_{22} \end{pmatrix}$, and that $\Delta(\tau)$ has a zero of precisely order 2 along these divisors.

Therefore, if $g(\tau)$ $\left(\tau = \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{12} & \tau_{22} \end{pmatrix}\right)$ is a modular form of level m and weight k, and $g(\gamma \circ \tau)$ has a zero of order 2n along $\tau_{12} = 0$ for all $\gamma \in \Gamma/\Gamma(m)$, then $g(\tau)/\Delta(\tau)^n$ is a modular form of level m and weight k-10n.

To see how many discriminants are "lurking" within modular forms built out of theta functions, we will use the following identity among power series, which follows directly from (1.2.0):

$$(1.10.6)\quad \theta\begin{bmatrix} e \\ f \\ g \\ h \end{bmatrix}(z, \tau)\Bigg|_{\tau_{12}=0} = \theta_{eg}(z_1, \tau_{11})\,\theta_{fh}(z_2, \tau_{22})$$

where

$$e, f, g, h \in \mathbb{Q}\,,\quad z = (z_1, z_2),\quad \tau = \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{12} & \tau_{22} \end{pmatrix}\,,$$

and θ_{eg} is the genus 1 theta function:

$$\theta_{eg}(z, \tau) = \sum_{n \in \mathbb{Z}} e^{\pi i (n+e)^2 \tau + 2\pi i (n+e)(z+g)}$$

$$z, \tau \in \mathbb{C}\,,\ \operatorname{im}\tau > 0$$

We need the following fact from genus 1 theta functions. Let $[1, \tau]$ denote the lattice generated in $\mathbb{C}$ by 1 and τ, then

(1.10.7) $\theta_{eg}(z,\tau) = 0 \iff z + e\tau + g \equiv \frac{1}{2} + \frac{1}{2}\tau$ modulo $[1,\tau]$.

Theorem (1.10.8): Let $\begin{bmatrix} a \\ b \end{bmatrix}$ be an odd theta characteristic

i) $D'_p \begin{bmatrix} a \\ b \end{bmatrix} = \prod_{\substack{u \neq \pm v \in \frac{1}{p}\Lambda \\ u,v \neq 0}} D'\begin{bmatrix} a \\ b \end{bmatrix}^p (u,v)$ is a modular form of weight $(p^4-1)(p^4-3)4p$ and level 2.

ii) $D'_p \begin{bmatrix} a \\ b \end{bmatrix}$ has a zero of order at least $8p(p^4-3)(p^2-1)$ along $\tau_{12} = 0$.

iii) $D'_p \begin{bmatrix} a \\ b \end{bmatrix} = \Delta(\tau)^{4p(p^4-3)(p^2-1)} f_p$, where f_p is a modular form of weight $4p(p^4-3)[(p^4-1) - 10(p^2-1)]$ and level 2.

Proof: i) If $\gamma \in \Gamma$, and $u, v \neq 0$, $u \neq \pm v$, then writing u, v as in (1.10.3), we see by (1.6.8) that $\gamma \circ u$, $\gamma \circ v \neq 0$, and $\gamma \circ u \neq \pm \gamma \circ v$. So γ acts merely to permute the p-division values. Since each $D'\begin{bmatrix} a \\ b \end{bmatrix}^p (u,v)$ is of level $2p$, the symmetrized product is of level 2. There are $(p^4-1)(p^4-3)$ terms in the product, each of weight $4p$.

ii) By (1.10.6), $\theta \begin{bmatrix} a + \begin{bmatrix} e \\ f \end{bmatrix} \\ b + \begin{bmatrix} g \\ h \end{bmatrix} \end{bmatrix} (0,\tau)$ has a zero along $\tau_{12} = 0$ only when either $\{a_1 + e, b_1 + g\}$ or $\{a_2 + f, b_2 + h\} \equiv \{1/2, 1/2\}$ modulo 1, where $a = (a_1, a_2)$, $b = (b_1, b_2)$, $a_i, b_i, e, f, g, h \in \mathbb{Q}$. For each odd $\begin{bmatrix} a \\ b \end{bmatrix}$, precisely one of $\{a_i, b_i\} \equiv \{1/2, 1/2\}$ modulo 1. Then there are precisely p^2-1 non-zero $u \in \frac{1}{p}\Lambda$, $u = \tau\begin{pmatrix} e \\ f \end{pmatrix} + \begin{pmatrix} g \\ h \end{pmatrix}$, such

that $\{e,g\}$ or $\{f,h\} = \{0,0\}$. Since each $\theta\begin{bmatrix} a+\begin{bmatrix} e \\ f \end{bmatrix} \\ b+\begin{bmatrix} g \\ h \end{bmatrix} \end{bmatrix}(0,\tau)$ appears in the product $8p(p^4-3)$ times, we see that $D'_p\begin{bmatrix} a \\ b \end{bmatrix}$ has a zero of order at least $8p(p^4-3)(p^2-1)$ along $\tau_{12} = 0$.

iii) To see how many times $\Delta(\tau)$ divides $D'_p\begin{bmatrix} a \\ b \end{bmatrix}$, we must count its order of zero along $\tau_{12} = 0$ and all its translates under $\gamma \in \Gamma$. But $D'_p\begin{bmatrix} a \\ b \end{bmatrix}(\tau)$ has a zero to a given order along $\gamma \circ (\tau_{12} = 0)$ if and only if $D'_p\begin{bmatrix} a \\ b \end{bmatrix}(\gamma \circ \tau)$ has a zero of the same order along $\tau_{12} = 0$. But $D'_p\begin{bmatrix} a \\ b \end{bmatrix}(\gamma \circ \tau)$ differs by a non-zero function from $D'_p\left[\gamma^{-1} \circ \begin{bmatrix} a \\ b \end{bmatrix}\right](\tau)$, and we know $\gamma^{-1} \circ \begin{bmatrix} a \\ b \end{bmatrix}$ is another odd theta characteristic, so it also has a zero there of order at least $8p(p^4-3)(p^2-1)$ by (ii). Q.E.D.

Corollary (1.10.9) $D'_3\begin{bmatrix} a \\ b \end{bmatrix} = K\Delta^{74880}$ for some constant K and is therefore of level 1.

Proof: Indeed the only modular form of weight 0 and level 2 is a constant.

Remark 2: Since $D'_3[\delta] = \left(\prod_{0 \neq u \in \frac{1}{p}\Lambda} \theta[\delta](u)\right)^{1872}$ we observe that $D'_3[\delta]$ being a power of $\Delta(\tau)$ is a modular affirmation that there are no 3-division points along the theta-divisor of the Jacobian of any curve of genus two. In arithmetic applications, the proof of the Manin-Mumford conjecture shows that for any curve of genus 2 defined over a number field, there are only finitely many torsion points

which lie upon the theta divisor of its Jacobian [Ra].

We can also tell when $N'(u,v)$ is zero along $\tau_{12} = 0$. Writing $u = (u_1, u_2)$, $v = (v_1, v_2)$ we claim that this happens whenever, $u_1, u_2, v_1, v_2, u_1 \pm v_1$, or $u_2 \pm v_2$ is zero. To show this is a pleasant — yet lengthy — exercise in the manipulation of elliptic theta function identities, so we will not present the proof. We only note that the "invariance" of $N(u,v)$ under the action of $\Gamma(p)$ (remark 1) implies that the test for the number of times $\Delta(\tau)$ divides $\prod_{\substack{u \neq \pm v \in \frac{1}{p}\Lambda \\ u,v \neq 0}} N'(u,v)$ need only be done along $\tau_{12} = 0$.

Chapter 2

Notation

We let $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ denote the integral, rational, real, and complex numbers. If R is a ring, $R^{\times}$ denotes its multiplicative group of invertible elements. If A is a group (ring, field) we let $\mathrm{End}(A)$, $\mathrm{Aut}(A)$ stand for its ring of endomorphisms or automorphisms, respectively.

If A/B is an extension of number fields, $\mathfrak{D}(A/B)$ will denote its different, $D(A/B)$ its discriminant, and $D_p(A/B)$ its "p-part" or local discriminant at p, where p is a prime of B. We will let $h(A)$ denote the class number of A. $\mathcal{O}_A$ will denote the ring of integers of A, and if p is a prime of A, $\mathcal{O}_{A,p}$ and A_p will denote the p-adic completions of $\mathcal{O}_A$ and A. We will drop the subscript "A" when the reference field is clear. If G is a finite group, we let $\hat{G}$ denote its group of one-dimensional characters, $\mathrm{Hom}(G, \mathbb{C}^{\times})$.

If A is an abelian variety defined over a number field F, we let A_m denote its group of points of order m, and $A(K)$ its group of K-rational points for any $F \subseteq K$. We denote addition on the variety by $\oplus$, subtraction by $\sim$, and any complex multiplications by $*$. We let O denote the origin on the variety.

We let ζ_p denote a primitive pth-root of unity. If f is a divisor of a number field K (formal product of finite and real primes, the latter with multiplicity zero or one); we let $I(f)$ denote the group of fractional ideals of K prime to f, and $P(f)$ the subgroup of

principal ideals with a generator α, $\alpha \equiv 1 \bmod^{\times} f$. If B is a class-field of A of conductor f, and H is the subgroup of I(f) corresponding to B by class field theory (in particular, $Gal(B/A) \simeq I(f)/H$), we say H "belongs to B."

When we talk of adjoining to a field the coordinates of a point in projective space, we mean adjoining the ratios of the coordinates.

§1: Introduction

Our goal is to emulate results that are well known about towers of fields of division points of an elliptic curve with complex multiplication. The situation is rather more complicated in dimensions greater than one. Fortuitously, we have been able to isolate a class which embodies many of the special properties of the elliptic case, namely that of an abelian variety A of dimension 2 with complex multiplication by a number field K which satisfies:

(2.1.0) i) $[K:\mathbb{Q}]$ is a cyclic extension, necessarily of order 4.

ii) $h(K) = 1$.

iii) The only roots of unity of K are ± 1.

iv) The endomorphisms of A are the full ring of integers of K.

We want to make certain calculations about towers of fields generated by division values of a point of infinite order in the Mordell-Weil group of A over K, $A(K)$. The corresponding calculations in the elliptic case were powerfully exploited by Coates and Wiles in their work on the conjectures of Birch and Swinnerton-Dyer [CW]. The jump from varieties of dimension 1 to dimension 2 carries with it technical difficulties which we will have to tackle in turn. The worst problem is that fields of division points "collapse," that is, they are not as large as we would have liked, had we our druthers. But we have lost our druthers long ago, and the speciality of our

choices (2.1.0) reflect the concessions we will make in order to render our calculations feasible. Indeed, ii), iii), iv) of (2.1.0) were assumed by Coates and Wiles, and i) is attractive to our purposes for a plethora of reasons. Prominant among these is a desire to have rational primes that remain inert in K. This will allow us to mimic much of the approach of Stark [St] and Gupta [G], who utilized rational primes in their versions of the tower-of-fields-of-division-points calculations used by Coates and Wiles. After we accumulate some basic facts about the special properties of our selected K's -- not the least of which is their existence -- we will develop that part of the theory of abelian varieties with complex multiplication that we need for our calculations. Our chief source for this is Lang's, Complex Multiplication [Lang 3], to which we will continually refer the reader for further details. Our notation will be Lang's. We then start our calculations, closely following the methods which proved fruitful in the elliptic case.

§2: Biquadratic Cyclic Fields: Class number one $\mathcal{O}_K$

We will now establish certain properties of the fields K which satisfy (2.1.0). These K are totally complex quadratic extensions of a totally real subfield — in the general parlance, a CM-field. There is a rich literature concerning the class numbers of such fields, and we will now exploit some of it.

Let $\mathrm{Gal}(K/\mathbb{Q}) = \langle \sigma \rangle$. We will also denote complex conjugation $= \sigma^2$, by ρ.

Proposition 2.2.0: Let K be a totally complex, cyclic, class number one extension of $\mathbb{Q}$ of degree 4, which contains only the roots of unity ± 1.

i) The units of K are of the form $\pm \zeta^i$, $i \in \mathbb{Z}$, where $\zeta \in K^+$, the real quadratic subfield of K.

ii) $h(K^+) = 1$, and $N_{K^+/\mathbb{Q}}(\zeta) = -1$.

iii) The subfields of index 3, 7, and 13 of $\mathbb{Q}(\zeta_{13})$, $\mathbb{Q}(\zeta_{29})$ and $\mathbb{Q}(\zeta_{53})$ are examples of such K.

Note: Uchida has shown that only finitely many such K exist [U].

Proof: i) Since $[K:\mathbb{Q}] = 4$, K totally complex, the units are of the form $\pm\zeta^i$, $i \in \mathbb{Z}$, where ζ is a unit in $\mathcal{O}_K$. We need only show ζ is real. We have $\sigma(\zeta) = (-1)^a \zeta^i$, where $a = 0, 1$, and $i \in \mathbb{Z}$. Therefore $\rho(\zeta) = \sigma^2(\zeta) = (-1)^{a(1+i)} \zeta^{i^2}$, and $\zeta = \sigma^4(\zeta) = (-1)^{a(1+i+i^2+i^3)} \zeta^{i^4}$; so $i^4 = 1$. That is, $i = \pm 1$ and $\rho(\zeta) = \zeta$;

hence ζ is in K^+.

ii) We lift a lemma from [Wa, p. 184]. Suppose the extension of number fields E/F contains no unramified abelian subextensions L/F with $L \neq F$. Then $h(F)$ divides $h(E)$. In our case, K/K^+ is a totally ramified extension (at least at the archimedean places of K^+) so the lemma applies. Therefore $h(K^+) \mid h(K)$; hence $h(K^+) = 1$.

Now suppose $N_{K^+/\mathbb{Q}}\zeta = 1$. Letting $\infty_1\infty_2$ denote the divisor corresponding to the product of the two real places of K^+, we see that $|I_{\infty_1\infty_2} / P_{\infty_1\infty_2}| = 2$, since all ideals are principal, and "half" are ideals all whose generators have different signs at each of the two real embeddings. So by class-field theory, there is a corresponding quadratic extension $L \mid K^+$, unramified at all finite places. We deduce that $KL \mid K$ is an abelian extension unramified at all finite places. But $h(K) = 1$, and K is totally complex, therefore $KL = K$, and $L \subseteq K$. It now suffices to show K/K^+ is ramified at a finite place. For then we would have $K^+ = L$, a contradiction forcing $N_{K^+/\mathbb{Q}}\zeta$ to be -1.

That K/K^+ is ramified at a finite place follows from the fact that $K^+/\mathbb{Q}$ is (totally!) ramified at a finite place, and the following lemma:

Lemma (2.2.1): Suppose K/E is a cyclic extension of number fields, and F a subextension which is totally ramified over a prime p of E, and such that $[K:F] \mid [F:E]$. Then K/E is totally ramified at p.

Proof: Let G_o be the inertia group of p for K/E, and L its fixed field. By assumption, $[F:E] \mid |G_o|$; a fortiori $[K:F] \mid |G_o|$, so $L \subseteq F$ since K/E is cyclic. But L/E is unramified at p, while F/E is totally ramified at p. Therefore $L = E$ and p is totally ramified in K/E.

iii) In each of these cases, we have $K \subseteq \mathbb{Q}(\zeta_p)$, p a prime. $\mathbb{Q}(\zeta_p)/K$ is totally ramified, so $h(K) \mid h(\mathbb{Q}(\zeta_p))$ by our aforementioned lifted lemma. If we can show $h(K) = 1$, then our K clearly fit the criteria; for we would have $K/\mathbb{Q}$ cyclic of order 4, and since the indices 3, 7, and 13 are odd, K would be totally complex. Further, we took $p > 5$ so each of these K would have only ± 1 as roots of unity.

The following data on cyclotomic fields is taken from [Wa, p. 353]. (All these fields have $h(\mathbb{Q}(\zeta_p)) = "h^-" = h(\mathbb{Q}(\zeta_p)^+)$).

I) $p = 13$. $h(\mathbb{Q}(\zeta_p)) = 1$ so $h(K) = 1$.

II) $p = 29$. $h(\mathbb{Q}(\zeta_p) = 2^3$. It suffices to show $2 \nmid h(K)$. We lift another lemma: If $\mathrm{Gal}(L/\mathbb{Q})$ is a p-group, and at most one finite prime is ramified, then $p \nmid h(L)$ [Wa, p. 185]. In our case, $K/\mathbb{Q}$

is ramified at only the finite prime (29), and $K/\mathbb{Q}$ is a 2-group. Hence $2 \nmid h(K)$.

III) $p = 53$. $h(\mathbb{Q}(\zeta_p)) = 4889$, a prime. If $h(K) \neq 1$, then $h(K)$ would be ridiculously large. We now apply a general theorem on CM-fields [Wa, p. 42].

$$h^-(K) \equiv h(K)/h(K^+) = Qw \prod_{\chi \text{ odd}} \left(-\frac{1}{2} B_{1,\chi}\right)$$

where: $Q = |\text{units of } (K)/\text{units of } (K^+)| = 1$ by (i)

$w = \#$ roots of unity of $K = 2$

$\chi \in \widehat{\text{Gal}}(K/\mathbb{Q})$, here of conductor 53

and

$$|B_{1,\chi}| = \left|\frac{1}{53} \sum_{a=1}^{53} \chi(a)\, a\right| \leq \frac{53+1}{2} = 27 \quad \text{by } |\chi(a)| \leq 1$$

There are two such odd χ so we get:

$$h^-(K) = h(K)/h(K^+) \leq \frac{1}{2} \cdot 27^2 < 4889$$

The result now follows from $h(K^+) = h(\mathbb{Q}(\sqrt{53})) = 1$ [BS, p. 422].

§3: Abelian Varieties with Complex Multiplication

Let A be an abelian variety of dimension n defined over the complex numbers. Then A is analytically isomorphic to $\mathbb{C}^n$ modulo a lattice Λ of dimension $2n$ over $\mathbb{R}$. If α is an endomorphism of $\mathbb{C}^n/\Lambda$, then α lifts to a $\mathbb{C}$-linear map $\tilde{\alpha}$ of $\mathbb{C}^n \to \mathbb{C}^n$. The mapping $\alpha \longmapsto \tilde{\alpha}$ gives a representation $R_{\mathbb{C}}(\alpha)$ of $\mathrm{End}(\mathbb{C}^n/\Lambda)$ as $n \times n$ complex matrices known as the complex representation. However, α must also map Λ into itself, giving a representation $R_{\mathbb{Q}}(\alpha)$ (the rational representation) of α by $2n \times 2n$ integer matrices. The rational representation is equivalent to the direct sum of the complex representation and its conjugate representation $\bar{R}_{\mathbb{C}}$. These representations extend by $\mathbb{Q}$-linearity to $\mathrm{End}(A)_{\mathbb{Q}} \equiv \mathrm{End}(A) \otimes \mathbb{Q}$. Let F be a CM field of dimension $2n$ over $\mathbb{Q}$. If we have an embedding $i : F \hookrightarrow \mathrm{End}(A)_{\mathbb{Q}}$, then the complex and rational representations of $\mathrm{End}(A)_{\mathbb{Q}}$ restrict to representations of $i(F)$. Since F is abelian, $R_{\mathbb{Q}}$ is equivalent to the direct sum of $2n$ one-dimensional representations of $i(F)$. These consist of the embeddings $\varphi_i : F \hookrightarrow \mathbb{C}$ that preserve $\mathbb{Q}$. Since $R_{\mathbb{Q}}(i(F))$ is $\mathbb{Q}$-valued, we must have that $R_{\mathbb{Q}}$ is the direct sum of each of the $2n$ distinct φ_i. However, since $R_{\mathbb{Q}}$ is equivalent to $R_{\mathbb{C}} \oplus \overline{R_{\mathbb{C}}}$, we must have that $R_{\mathbb{C}}$ is equivalent to $\bigoplus_{j=1}^{n} \varphi_{i_j}$ where the φ_{i_j} are distinct and no two are complex conjugates of each other. Let $\Phi = \{ \varphi_{i_j}, j = 1, \ldots, n \}$. In this guise we say that the abelian variety A with the embedding i admits

complex multiplication by F with CM-type Φ. We will abbreviate this by saying that the pair (A,i) is of type (F, Φ).

We call an abelian variety simple if it has no abelian subvarieties. If (A,i) is of type (F, Φ), then when F is Galois over $\mathbb{Q}$ we can tell easily whether A is simple [Lang 3, p. 13]. To wit, A is simple if and only if the only $\sigma \in \mathrm{Gal}(F/\mathbb{Q})$ such that $\Phi\sigma = \Phi$ (that is, the set $\{\varphi_{i_j}\}$ equals the set $\{\varphi_{i_j}\sigma\}$) is the identity. We call such an (F, Φ) a simple type.

Note: In particular, if (A,i) is of type (K,Φ) where K satisfies (2.1.0), then A is simple. This follows since $\Phi = \langle 1,\sigma \rangle, \langle 1, \sigma^3 \rangle$, $\langle \rho,\sigma \rangle$ or $\langle \rho, \sigma^3 \rangle$ and it is easily verified that $\Phi\sigma^i = \Phi$ only for $i = 0$. In what follows we will always use F to denote a general CM-field, reserving K for one which satisfies (2.1.0).

Associated to a type Φ is its type norm N_Φ, and type trace, T_Φ, defined (for F/Q Galois) by:

$$N_\Phi(x) = \prod_{\varphi \in \Phi} \varphi(x) \qquad \text{for } x \in F \tag{2.3.0}$$

$$T_\Phi(x) = \sum_{\varphi \in \Phi} \varphi(x)$$

The images of these maps will be of great importance. Attached to a type Φ (for F/Q Galois) is its reflex type, defined by $\Phi' = \{\varphi \mid \varphi^{-1} \in \Phi\}$. Associated to a CM-field F and type Φ we now define its reflex field F' by:

(2.3.1) $$F' = \mathbb{Q}(T_\Phi, (F))$$

Remark: For $F/\mathbb{Q}$ non-Galois, we use the Galois closure of F to extend these definitions.

Note: $K/\mathbb{Q}$ is Galois, so $K' \subseteq K$. Therefore K' is Galois over $\mathbb{Q}$, and $K'' \subseteq K' \subseteq K$. The statement that $\{K, \Phi\}$ is simple is equivalent to $K'' = K$ [Lang 3, p. 24], so we have $K = K'$. This will greatly simplify the general theorems of complex multiplication.

We shall always assume for technical simplicity that $i(F) \cap \mathrm{End}(A) = i(\mathcal{O}_F)$ (recall that the range of i is $\mathrm{End}(A)_{\mathbb{Q}}$). Such an (A,i) is called principal. Condition (2.1.0)(iv) simply states that we require any (A,i) of type (K,Φ) to be principal. In practice we will usually identify $\mathcal{O}_F$ with $\mathrm{End}(A)$.

Given an (A,i) of type (F, Φ), there is a lattice λ in F so that the analytic expression of A as a complex torus is given by $\theta : \mathbb{C}^n/\Phi(\lambda) \to A$, where $\Phi(a) = (\varphi_{i_1}(a), \ldots, \varphi_{i_n}(a))$ in $\mathbb{C}^n$, $a \in \lambda$; and θ commutes with the action of i, that is:

(2.3.2) $$i(a) \circ \theta = \theta \circ \Phi(a)$$

We then say that (A,i) is of type (F, Φ, λ). The reason for incorporating the lattice λ into our framework is that (F, Φ) determines (A,i) up to isogeny (isogeny for the pair being an isogeny of A which commutes with the action of i; for A simple, all isogenies of A are isogenies of (A,i)). However, (A,i) and (B,j) of types

(F, Φ, λ) and (F, Φ, μ) are isomorphic if and only if $\lambda = \mu\gamma$ for some $\gamma \in F$. Therefore there are $h(F)$ pairwise non-isomorphic classes of isogenous (principal) (A,i) of type (F, Φ). This result is implicitly used in lemma (2.4.2) [Lang 3, pp. 17, 59, 60], and this is why we restrict ourselves to K of class number one.

One of our main goals is to study fields generated by points of finite order on A. Our main tool will be to use the explicit characterization of the class fields generated by the images of points of finite order in a certain quotient variety of A. To discuss that quotient, we are going to have to consider polarized abelian varieties, that is, a triplet $(A, \mathcal{C}, i)$ where $\mathcal{C}$ is a polarization of A. What basic definition we take for polarization will not prove critical to our needs. But for specificity, we will take as our definition of polarization a class of divisors on A, $\mathcal{C} = \mathcal{C}(X)$, where X is an ample divisor on A, and such that $Y \in \mathcal{C}$ if and only if there are integers m and n so that mX is algebraically equivalent to nY. If $\alpha \in \text{End } A$ then we say α is an endomorphism of $(A, \mathcal{C})$ if $\alpha^{-1}(\mathcal{C}) \subseteq \mathcal{C}$, where α^{-1} is the induced map on divisors. The beauty of introducing the polarization is that while $\text{Aut}(A)$ is generally infinite (the units of F), $\text{Aut}(A, \mathcal{C})$ is finite [Lang 3, p. 71]. So $\text{Aut}(A, \mathcal{C})$ is contained within the roots of unity of F.

What is important for our needs (for details, see [Lang 3]) is that to every polarization is associated a Riemann Form $E_{\mathcal{C}}$ on

$\mathbb{C}^n \times \mathbb{C}^n$. Two polarizations are the same if and only if they determine the same Riemann form. There is an explicit characterization of the change of a Riemann form E under an endomorphism α of A, when $\alpha\rho(\alpha)$ is rational. Letting E' denote the Riemann form on $\alpha(A)$, we have [Lang 3, p. 74]

$$E' = \alpha\rho(\alpha)E$$

Since all roots of unity have absolute value one, they all preserve the Riemann form. In short, we have precisely that $\mathrm{Aut}(A, \mathbb{C})$ corresponds to the image of the roots of unity in $\mathrm{End}(A) \cap i(\mathcal{O}_F)$. For (A,i) principal, $\mathrm{Aut}(A, \mathbb{C})$ is precisely the set of roots of unity of F. This is the chief technical convenience afforded by our restriction to K containing only ± 1 as roots of unity.

Given a Riemann form E determined by a polarization $\mathbb{C}$ we say the triplet $(A, \mathbb{C}, i)$ is of type (F, Φ, λ, E) (with respect to the map θ: we needed it to define E). An isogeny of $(A, \mathbb{C}, i)$ into $(B, \mathfrak{D}, j)$ is an isogeny of (A,i) into (B,j) whose induced map on divisors maps $\mathfrak{D}$ into $\mathbb{C}$. We included E into our type data because (F, Φ, λ, E) determines the triplet $(A, \mathbb{C}, i)$ up to isomorphism.

There is one more technical point to worry about. A Riemann form determines an involution of $\mathrm{End}(A)_{\mathbb{Q}}$, and the theorems of complex multiplication demand that the set $i(F)$ be stable under this involution. For A simple this condition is always met.

§4: Fields of Definition

Since we are restricting ourselves to principal (A,i) of type (F, Φ, λ), we can say that (A,i) is defined over a field $L \subseteq \mathbb{C}$ whenever A is, and when all $i(a)$ are defined over L $(a \in F)$. When A is simple and defined over L, then (A,i) is defined over the compositum LF' [Lang 3, pp. 54-55]. If A is defined over L, we say a polarization $\mathcal{C}$ is defined over L if $\sigma(\mathcal{C}) = \mathcal{C}$ for every $\sigma \in \mathrm{Gal}(\mathbb{C}/L)$. We say $(A, \mathcal{C}, i)$ is defined over L whenever (A,i) and $\mathcal{C}$ are. For A simple, $\mathcal{C}$ is defined over any field of definition for (A,i). So if $(A, \mathcal{C}, i)$ is of type (K, Φ, λ, E), then $(A, \mathcal{C}, i)$ is defined over LK, where L is the field of definition of A (recall $K = K'$).

Let σ be any automorphism of $\mathbb{C}$. Then σ acts on A, (A,i) or $(A, \mathcal{C}, i)$ by $A \to A^{\sigma}$, $(A,i) \to (A^{\sigma}, i^{\sigma})$, or $(A, \mathcal{C}, i) \to (A^{\sigma}, \mathcal{C}^{\sigma}, i^{\sigma})$, where i^{σ} is defined by the commutivity of

$$\begin{array}{ccc} F & \xrightarrow{\ i\ } & \mathrm{End}(A) \\ & \searrow^{i^{\sigma}} & \downarrow \\ & & \mathrm{End}(A^{\sigma}) \end{array}$$

The downward arrow being the map induced by σ. We define the field of moduli of A, (A,i), or $(A, \mathcal{C}, i)$ (denoted by $M(A)$, $M(A,i)$, or $M(A, \mathcal{C}, i)$) as the fixed field of precisely those $\sigma \in \mathrm{Aut}(\mathbb{C})$ for which there exists an isomorphism s over $\mathbb{C}$; such that:

$$s : A \to A^{\sigma}$$

$$s : (A,i) \to (A,i)^{\sigma}$$

or $$s : (A, \mathbb{C}, i) \to (A, \mathbb{C}, i)^{\sigma}$$

These fields of moduli always exist [Lang 3, p. 123]. When A is simple, $M(A, \mathbb{C}, i) = M(A,i)$ (but $M(A) = M(A, \mathbb{C})$ is not always true). For example, when A is an elliptic curve, $M(A) = \mathbb{Q}(j(A))$, where j is the j-invariant of A. It is always the case (complex multiplication or not) that an elliptic curve is defined over its field of moduli. This is not always so for an abelian variety. However, for certain abelian varieties with complex multiplication (of which those of dimension 2 are just a subset), Shimura has found [Sh]:

Theorem (2.4.0): Let (F, Φ) be a simple type for F a CM-field, $[F : \mathbb{Q}] = 4$. Then any $(A, \mathbb{C}, i)$ of type (F, Φ, λ, E) has a model defined over its field of moduli, $M(A, \mathbb{C}, i)$.

In studying the fields generated by points of finite order on A, we must first introduce the Kummer Variety, the quotient of A by $\mathrm{Aut}(A, \mathbb{C}, i)$. For A simple, $\mathrm{Aut}(A, \mathbb{C}) = \mathrm{Aut}(A, \mathbb{C}, i)$ [Lang 3, p. 135]. For $(A, \mathbb{C}, i)$ of type (K, Φ, λ, E), we determined in the last section that $\mathrm{Aut}(A, \mathbb{C}) = \pm 1$. The Kummer variety is obtained therefore by identifying a point on A with its negative. Let W be the resulting variety; $h : A \to W$ the natural projection. Then the field of functions of W is just the subfield of even functions on A.

From now on we shall restrict ourselves to the case in question: $(A, \mathbb{C}, i)$ of type (K, Φ, λ, E), where K satisfies (2.1,0) and $i(K) = \mathrm{End}(A)$ (that is, (A, i) is principal). Then we have already noted:

(2.4.1) i) A is simple.

ii) $(A, \mathbb{C}, i)$ is defined over its field of moduli $M(A, \mathbb{C}, i)$.

iii) $K = K'$, its reflex field.

iv) $\mathrm{Aut}(A, \mathbb{C}, i) = \pm 1$.

We will now lift an important application of the main theorems of complex multiplication [Lang 3, p. 137].

Lemma (2.4.2): If A is principal, $M(A, \mathbb{C}, i)$ is contained in the Hilbert class field of K'.

Corollary (2.4.3): $M(A, \mathbb{C}, i) = K$.

Proof: Indeed $K = K'$, and the class number of K is 1. (This is the great simplification from the assumption $h(K) = 1$.) Therefore $M(A, \mathbb{C}, i) \subseteq K$. It's always the case that $K' \subseteq M(A, \mathbb{C}, i)$ [Lang 3, p. 125], so we have $M(A, \mathbb{C}, i) = K$.

§5: Fields Generated by Division Values

In this section, as in the sequel, p will be an odd rational prime which remains inert in K, and such that A has good reduction over K at p. There are infinitely many such p. For the sake of specificity, we will take $\Phi = \langle 1, \sigma \rangle$ where $\mathrm{Gal}(K/\mathbb{Q}) = \langle \sigma \rangle$. Then $\Phi' = \{ 1, \sigma^{-1} \}$. This choice is of no consequence except to simplify discussion.

Let m be any integer. Then A_m, the group of points of order m on A, is isomorphic to the additive group $\mathcal{O}_K/m\mathcal{O}_K$ [Lang 3, p. 138]. Let t_m be a primitive point, i.e. one such that $\mathcal{O}_K * t_m$ encompasses all the m-division points. Recall that h is the map from A onto the Kummer variety. We care to study the field generated over K by the coordinates of $h(A_m)$. (That is, the field generated by the even functions of A evaluated at points of order m.) Since all the endomorphisms of A are defined over K, this will be identical to the field $K(h(t_m))$, i.e., the field obtained by adjoining to K the coordinates of the point $h(t_m)$.

Our main tool will be another application of the main theorems of complex multiplication: the adjunction of division values generates class fields. We quote the following from [Lang 3, p. 138], tailoring the result to adjust for the conditions (2.4.1) and corollary (2.4.3):

Theorem (2.5.0): $K(h(t_m))$ is a classfield over K of conductor dividing m. It belongs to the group of ideals $H(m)$ consisting of all ideals $\mathfrak{B}$ prime to (m), such that:

There exists a $\beta \in K$ satisfying;

i) $N_{\Phi'}(\mathfrak{B}) = \mathfrak{B}\sigma^{-1}(\mathfrak{B}) = (\beta)$.

ii) $N_{K/\mathbb{Q}}(\mathfrak{B}) = \beta\rho(\beta)$.

iii) $\beta \equiv 1 \bmod(m)$.

Let's massage these conditions a bit. First of all, $\mathfrak{B}$ is principal, say equal to (a). Then (i) reduces to $(a\sigma^{-1}a) = (\beta)$ or $a\sigma^{-1}(a) = \pm\zeta^i\beta$, for some $i \in \mathbb{Z}$. Then (ii) implies $N_{K/\mathbb{Q}}((a)) = \rho(a\sigma^{-1}a)a\sigma^{-1}a = \beta\rho(\beta)$, which holds only when $(\pm\zeta^i)\,\rho(\pm\zeta^i) = (\pm\zeta^i)^2 = 1$. So we must have $a\sigma^{-1}(a) = \pm\beta$. Finally, condition (iii) translates into $a\sigma^{-1}(a) = \pm 1$ (m). We note that $N_{K^+/\mathbb{Q}}(\zeta) = -1$ implies $\zeta\sigma^{-1}(\zeta) = -1$ (m), so an ideal $\mathfrak{B}$ has a generator a such that $a\sigma^{-1}(a) = \pm 1$ (m) if and only if it has a generator a such that $a\sigma^{-1}(a) = \mp 1$ (m). Moreover, $(\pm\zeta^i a)\sigma^{-1}(\pm\zeta^i a) = (-1)^i a\sigma^{-1}(a)$. Summarizing we have:

Corollary (2.5.1): $K(h(t_m))$ is a classfield of K, of conductor dividing m, that belongs to the group $H(m)$, consisting of those ideals $\mathfrak{B}$ prime to (m), with generator a such that $a\sigma^{-1}(a) \equiv 1 \pmod m$. Moreover, if $\mathfrak{B}$ has one generator a with the property: $a\sigma^{-1}(a) \equiv \pm 1 \pmod m$, then all its generators have that property.

Recall that p is a fourth degree prime in $\mathcal{O}_K$. Therefore $\mathcal{O}_K/p\mathcal{O}_K$ is a fourth degree extension of $\mathbb{Z}/p\mathbb{Z}$ with Galois group $\langle\sigma\rangle$ and a lone intermediate quadratic extension, $\mathcal{O}_{K^+}/p\mathcal{O}_{K^+}$. We will now investigate the classfields obtained when $m = p^n$. We denote $K(h(t_{p^n}))$ by $K(E_n)$.

Theorem (2.5.2): $K(E_n)$ is an extension of degree $\frac{1}{2}p^{3n-3}((p^2+1)\cdot(p-1))$ over K.

Proof: n = 1. Here we have a surjection

$$\pi : (\mathcal{O}_K/p\mathcal{O}_K)^\times \longrightarrow I(p)/H(p)$$

whose kernel consists of those $a \in (\mathcal{O}_K/p\mathcal{O}_K)^\times$ such that $a\sigma^{-1}(a) = \pm 1$ (p). We note $a\sigma^{-1}(a) = \pm 1$ (p) implies $\sigma^{-1}(a) = \pm a^{-1}$ (p) so $\rho(a) = \sigma^{-2}(a) = \pm\sigma^{-1}(a^{-1}) = \pm(\pm a) = a \pmod p$. So a is congruent to a residue class in $(\mathcal{O}_{K^+}/p\mathcal{O}_{K^+})^\times$. We first count those a such that $a\sigma^{-1}(a) = a\sigma(a) = 1$ (p). These are those a in the cyclic extension $\mathcal{O}_{K^+}/p\mathcal{O}_{K^+}$ over $\mathbb{Z}/p\mathbb{Z}$ of norm 1, so by Hilbert's Theorem 90, $a = b(\sigma(b))^{-1}$ for some b in $(\mathcal{O}_{K^+}/p\mathcal{O}_{K^+})^\times$, which has order $p^2 - 1$. But $b(\sigma(b))^{-1} = b'(\sigma(b'))^{-1}$ if and only if $b'b^{-1}$ is in $(\mathbb{Z}/p\mathbb{Z})^\times$, so there are $p+1$ such a. Therefore there are also $p+1$ such ζa satisfying $(\zeta a)\,\sigma^{-1}(\zeta a) = -1 \pmod p$, so $|\ker \pi| = 2(p+1)$, and $[K(E_1): K] = \frac{1}{2}(p^2+1)(p-1)$.

$n > 1$: First we need a lemma

Lemma (2.5.3): Let $G = \mathrm{Gal}((\mathcal{O}_K/p\mathcal{O}_K)/(\mathbb{Z}/p\mathbb{Z})) = \langle \sigma \rangle$, cyclic of order 4. Then for $b \in \mathcal{O}_K/p\mathcal{O}_K$,

$$b - \sigma(b) + \sigma^2(b) - \sigma^3(b) = 0 \iff \exists\ c \in \mathcal{O}_K/p\mathcal{O}_K \text{ such that } b = c + \sigma(c).$$

Proof: The group algebra $\mathbb{Z}[G] \simeq \dfrac{\mathbb{Z}[t]}{t^4 - 1}$, where t is an indeterminant. Clearly for $\alpha \in \mathbb{Z}[G]$

$$(1+\sigma)\alpha = 0 \iff \alpha = (1 - \sigma + \sigma^2 - \sigma^3)\beta \quad \text{for some } \beta$$

$$(1 - \sigma + \sigma^2 - \sigma^3)\alpha = 0 \iff \alpha = (1 + \sigma)\beta \quad \text{for some } \beta$$

Therefore the following is a projective resolution of $\mathbb{Z}$ as a trivial $\mathbb{Z}[G]$-module:

$$0 \leftarrow \mathbb{Z} \xleftarrow{\varepsilon} \mathbb{Z}[G] \xleftarrow{D} \mathbb{Z}[G] \xleftarrow{N} \mathbb{Z}[G] \xleftarrow{D}$$

where D is multiplication by $1+\sigma$, N is multiplication by $1 - \sigma + \sigma^2 - \sigma^3$, and the augmentation ε is defined by $\varepsilon : \sigma^i \to (-1)^i$. Therefore if A is any $\mathbb{Z}[G]$-module,

$$H^1(G, A) \simeq \mathrm{Ann}_A N/DA \quad \text{where} \quad \mathrm{Ann}_A N = \{x \in A \mid Nx = 0\}.$$

Now G acts additively on $\mathcal{O}_K/p\mathcal{O}_K$, and (2.5.3) is just the statement that $H^1(G, \mathcal{O}_K/p\mathcal{O}_K) = 0$. But this is true for any Galois extension [Se1, p. 150].

Back to proof:

We have the following exact diagram:

$$(2.5.4)\quad \begin{array}{ccccccccc} 0 & \to & \{a \mid a\sigma^{-1}(a) = \pm 1 \ (\mathfrak{p}^n)\} & \to & (\mathfrak{O}_K/\mathfrak{p}^n\mathfrak{O}_K)^\times & \to & I(\mathfrak{p})/H(\mathfrak{p}^n) & \to & 0 \\ & & \downarrow \pi_1 & & \downarrow \pi_2 & & \downarrow \pi_3 & & \\ 0 & \to & \{a \mid a\sigma^{-1}(a) = \pm 1 \ (\mathfrak{p}^{n-1})\} & \to & (\mathfrak{O}_K/\mathfrak{p}^{n-1}\mathfrak{O}_K)^\times & \to & I(\mathfrak{p})/H(\mathfrak{p}^{n-1})) & \to & 0 \end{array}$$

where the maps π_i are induced by just considering everything mod $\mathfrak{p}^{n-1}$.

We will now show π_1 is surjective. Since again $(\zeta a)\sigma^{-1}(\zeta a) = -a\sigma^{-1}(a)$, we need only show that if

$$a\sigma^{-1}(a) = 1 + b\mathfrak{p}^{n-1} \ (\mathfrak{p}^n) \quad \text{for} \quad b \in \mathfrak{O}_K/\mathfrak{p}\mathfrak{O}_K \ ,$$

Then

$$(a + c\mathfrak{p}^{n-1})\sigma^{-1}(a + c\mathfrak{p}^{n-1}) = 1 \ (\mathfrak{p}^n) \quad \text{for some} \quad c \in \mathfrak{O}_K/\mathfrak{p}\mathfrak{O}_K \ .$$

But this last equation holds if and only if

$$a\sigma^{-1}(a) + (c\sigma^{-1}(a) + a\sigma^{-1}(c))\mathfrak{p}^{n-1} \equiv 1 \ (\mathfrak{p}^n)$$

i.e. $$b \equiv -c\sigma^{-1}(a) - a\sigma^{-1}(c) \pmod{\mathfrak{p}}$$

We note the following identity

$$(2.5.5)\qquad \frac{a\sigma^{-1}(a) \cdot \sigma^2(a\sigma^{-1}(a))}{\sigma(a\sigma^{-1}(a)) \cdot \sigma^3(a\sigma^{-1}(a))} = 1$$

which implies $b - \sigma^{-1}(b) + \sigma^{-2}(b) - \sigma^{-3}(b) \equiv 0 \pmod{\mathfrak{p}}$. So by lemma (2.5.3), (with σ^{-1} replaced by σ), we have

$$b \equiv d + \sigma^{-1}(d) \pmod{p} \quad \text{for some } d \in \mathcal{O}_K/p\mathcal{O}_K$$

Therefore we are searching for a c such that

$$-c\sigma^{-1}(a) - a\sigma^{-1}(c) \equiv d + \sigma^{-1}(d) \pmod{p}$$

But as before, $a\sigma^{-1}(a) \equiv 1 \ (p)$ implies $\sigma^{-1}(a) \equiv \sigma(a) \ (p)$ so $c = \sigma(-d/a)$ does the trick. Therefore, coker $(\pi_1) = 0$. We will now find ker (π_1).

If $a \equiv 1 \ (p^{n-1})$ and $a\sigma^{-1}(a) \equiv \pm 1 \ (p^n)$ then we have

$$a \equiv 1 + bp^{n-1} \ (p^n) \quad \text{for some } b \in \mathcal{O}_K/p\mathcal{O}_K$$

and $$a\sigma^{-1}(a) \equiv 1 + (b + \sigma^{-1}(b))p^{n-1} \ (p^n)$$

i.e., $b + \sigma^{-1}(b) = 0$. But then $\sigma^{-1}(b) + \sigma^{-2}(b) = 0$ and consequently $\sigma^2(b) = \sigma^{-2}(b) = b$, so $b \in \mathcal{O}_{K^+}/p\mathcal{O}_{K^+}$. Therefore we seek to find $V = \{ b \mid b + \sigma(b) = 0 \}$, the kernel of the trace map

$$\mathrm{Tr}: \mathcal{O}_{K^+}/p\mathcal{O}_{K^+} \longrightarrow \mathbb{Z}/p\mathbb{Z} \ .$$

By the additive version of Hilbert's Theorem 90, $\mathrm{Tr}(b) = 0$ precisely when $b = c - \sigma^{-1}(c)$ for some $c \in \mathcal{O}_{K^+}/p\mathcal{O}_{K^+}$. But c is unique only up to translation by an element of $\mathbb{Z}/p\mathbb{Z}$, so $|v| = p$. Hence, $|\ker \pi_1| = p$.

Note we easily have that ker $\pi_2 = (\mathbb{Z}/p\mathbb{Z})^4$ and coker $\pi_2 = 0$. Now let us apply the snake lemma to the diagram (2.5.4):

$$(2.5.6) \quad 0 \to \ker \pi_1 \to \ker \pi_2 \to \ker \pi_3 \to \mathrm{coker}\ \pi_1 \to \mathrm{coker}\ \pi_2 \to \mathrm{coker}\ \pi_3 \to 0$$

By our results above, this sequence is just:

$$0 \to \ker \pi_1 \to \ker \pi_2 \to \ker \pi_3 \to 0 \to 0 \to \operatorname{coker} \pi_3 \to 0$$

So $\operatorname{coker} \pi_3 = 0$, and $|\ker \pi_3| = |\ker \pi_2| \,/\, |\ker \pi_1| = p^4/p = p^3$.

Hence $|I(p)/H(p^n)| = p^3 \, |I(p)/H(p^{n-1})|$ and the result follows by induction.

Corollary to Proof (2.5.7): $\operatorname{Gal}(K(E_n)/K(E_{n-1})) \simeq (\mathbb{Z}/p\mathbb{Z})^3$.

Proof: Indeed

$$\operatorname{Gal}(K(E_n)/K(E_{n-1}) \simeq (I(p)/H(p^n)) \,/\, (I(p)/H(p^{n-1}))$$
$$\simeq \ker \pi_3 \simeq \ker \pi_2 / \ker \pi_1$$

And $\ker \pi_2 \simeq (\mathbb{Z}/p\mathbb{Z})^4$; $\ker \pi_1 = \mathbb{Z}/p\mathbb{Z}$.

Remark: This corollary is in essence an affirmation that the type Φ has "rank" 3. See [Lang 3, pp. 148-155].

An important observation is that since $h(K) = 1$, $K(E_n)$ is totally ramified over K. We will let $\mathfrak{B}_n$ denote the unique prime of $K(E_n)$ above p.

Lemma (2.5.8): If $K \subseteq L \subseteq K(E_n)$ and L has conductor at most p^i, then $L \subseteq K(E_i)$.

Proof: First of all, it suffices to show this for $i = n-1$. For $n = 1$, we note that if $K \subsetneq L \subseteq K(E_1)$, then L/K is totally ramified over p so has a non-trivial conductor. Hence if L has conductor at most $p^{n-1} = p^0 = 1$, then $L \subseteq K$, and by convention we'll consider $K =$ "$K(E_o)$." Now let $n > 1$. Then the classgroup H belonging to L/K contains those $x \in (\mathcal{O}_K/p^n\mathcal{O}_K)^\times$ such that $x\sigma^{-1}(x) \equiv \pm 1\ (p^n)$ as well as those $x \equiv 1\ (p^{n-1})$. We seek to show that it contains all those x such that $x\sigma^{-1}(x) \equiv \pm 1\ (p^{n-1})$. But for such an x:

$$x\sigma^{-1}(x) \equiv \pm 1 + bp^{n-1}\ (p^n) \quad \text{for some} \quad b \in \mathcal{O}_K/p\mathcal{O}_K$$

and by the identity (2.5.5):

$$b - \sigma(b) + \sigma^2(b) - \sigma^3(b) = 0$$

so $b = c + \sigma^{-1}(c)$ for some $c \in \mathcal{O}_K/p\mathcal{O}_K$ by lemma (2.5.3). Further, $y = 1 + cp^{n-1}\ (p^n)$ is in H as is:

$$\sigma^{-1}(y) = 1 + \sigma^{-1}(c)\, p^{n-1}\ (p^n).$$

Therefore $xy^{-1}\sigma^{-1}(xy^{-1}) = \pm 1\ (p^n)$, and so xy^{-1} — and consequently x — is in H.

Remark: In particular, the conductor of $K(E_n)$ is precisely p^n. The "collapsing" of the p-division values is not so pronounced as one might have first thought.

Lemma (2.5.9):

$$D_{K(E_n)/K} = p^{\left[\frac{1}{2}(p^2+1)(p-1)\left(\frac{np^{3n} - (n+1)p^{3n-3} + 1}{p^3 - 1}\right) - 1\right]}$$

Proof: Since the extension has conductor a power of p, the discriminant contains no primes outside p .

n = 1: $K(E_1)/K$ is a totally and tamely ramified extension of degree $\frac{1}{2}(p^2+1)(p-1)$. Therefore the power of $\mathfrak{P}_1$ in the different $\mathfrak{D}(K(E_1)/K)$ is $\frac{1}{2}(p^2+1)(p-1) - 1$, and the result follows from $N_{K(E_1)/K}(\mathfrak{P}_1) = p$.

n > 1: We proceed by induction. Let $H_n = \mathrm{Gal}(K(E_n)/K)$.

If $\chi \in \hat{H}_n$, χ is the identity on the subgroup

$$\mathrm{Gal}(K(E_n)/K(E_{n-1}))$$

if and only if χ lies in $\hat{H}_{n-1}$ under the natural inclusion. For any $\chi \in \hat{H}_n$, we let K_χ denote the fixed field of its kernel = $\{\sigma \in H_n \mid \chi(\sigma) = 1\}$, and $f(\chi)$ the conductor of K_χ over K. By lemma (2.5.8)

$$\chi \in \hat{H}_{n-1} \iff K_\chi \subseteq K(E_{n-1}) \iff f(\chi) \leq p^{n-1}$$

Now by the conductor-discriminant formula:

$$D_{K(E_n)/K} = \prod_{\chi \in \hat{H}_n} f(\chi)$$

$$D_{K(E_{n-1})/K} = \prod_{\chi \in \hat{H}_{n-1}} f(\chi)$$

So

$$D_{K(E_n)/K} \Big/ D_{K(E_{n-1})/K} = \prod_{\substack{\chi \in \hat{H}_n \\ \chi \notin \hat{H}_{n-1}}} f(\chi) = \prod_{f(\chi) = p^n} f(\chi)$$

$$= (p^n)^{\frac{1}{2}(p^2+1)(p-1)(p^{3n-3} - p^{3n-6})}$$

since $|\hat{H}_i| = |H_i| = \frac{1}{2}(p^2+1)(p-1)p^{3i-3}$ by Theorem (2.5.2). By the induction hypothesis (letting $\log_p$ stand for taking a logarithm to the base p)

$$\log_p D_K(K(E_n)/K) = \frac{1}{2}(p^2+1)(p-1)\left(\frac{(n-1)p^{3n-3} - np^{3n-6} + 1}{p^3 - 1}\right) - 1$$

$$+ \frac{1}{2}(p^2+1)(p-1)(np^{3n-3} - np^{3n-6})$$

$$= \frac{1}{2}(p^2+1)(p-1)\left[\frac{(p^3-1)(np^{3n-3} - np^{3n-6}) + (n-1)p^{3n-3} - np^{3n-6} + 1}{p^3 - 1}\right] - 1$$

$$= \frac{1}{2}(p^2+1)(p-1)\left[\frac{np^{3n} - (n+1)p^{3n-3} + 1}{p^3 - 1}\right] - 1 \qquad \text{Q. E. D.}$$

Now that we've determined the structure of the fields generated by adjoining to K the values of all even functions of A evaluated at points of order p^n, we want to study the structure of the fields generated by adjoining to K the values of all functions of A evaluated at points of order p^n — namely, the fields of p^n-division points of A over K. We denote these fields by K_n when the dependence on p is clear. We note that the field of functions on A is a quadratic

extension of the field of even functions, (odd/odd = even; $(\text{odd})^2$ = even). Therefore $[K_n : K(E_n)] = 1 \text{ or } 2$.

Theorem (2.5.10):

i) $[K_n : K] = 2[K(E_n) : K] = (p^2+1)(p-1)p^{3n-3}$.

ii) p ramifies totally in K_n/K. We let p_n denote the unique prime of K_n over p.

iii) $$D_p(K_n/K) = (p^2+1)(p-1)\left[\frac{np^{3n} - (n+1)p^{3n-3} + 1}{p^3-1}\right] - 1.$$

Proof: i) Any $\sigma \in \text{Gal}(K_n/K)$ is determined by its action on t_{p^n}, a primitive point of order p^n. Indeed, $\sigma(t_{p^n}) = \alpha * t_{p^n}$ for some $\alpha \in (\mathcal{O}_K/p^n\mathcal{O}_K)^\times$, and the map $\sigma \mapsto \alpha$ defines an injection. This injection is compatible with the taking of inverse limits. That is, if $n \mid m$ and $\sigma \in \text{Gal}(K_m/K)$, the the following diagram commutes:

(2.5.11)
$$\begin{array}{ccc} \sigma & \longrightarrow & \alpha \\ \downarrow & & \downarrow \\ \sigma|_{K_n} & \longrightarrow & \bar{\alpha} \end{array}$$

where $\bar{\alpha}$ denotes the reduction of α mod p^n. So we get an embedding of $\text{Gal}\left(\left(\bigcup_n K_n\right)/K\right)$ into $\mathcal{O}_p^\times$ (the p-adic units, p a prime of K). Further, we can recover $\text{Gal}(K_n/K)$ by projecting the image of $\text{Gal}\left(\left(\bigcup_n K_n\right)/K\right)$ in $\mathcal{O}_p^\times$ onto $(\mathcal{O}_K/p^n\mathcal{O}_K)^\times$. With this in mind we quote a theorem from [Lang 3, p. 101]. (The theorem

simplfies when $h(K) = 1$ and $K = K'$.) It states that if A has good reduction at p, (our omnipresent assumption), then:

$$\mathrm{Gal}\left(\left(\bigcup_n K_n\right)/K\right) \simeq N_{\Phi'}(\mathcal{O}_p^{\times}) \subseteq \mathcal{O}_p^{\times}$$

where $N_{\Phi'}$, is the reflex norm we tackled in theorem (2.5.2). So we have:

$$\mathrm{Gal}(K_n/K) \simeq \text{Image of } N_{\Phi'}(\mathcal{O}_p^{\times})$$

in $(\mathcal{O}_K/p^n\mathcal{O}_K)^{\times}$. Since $N_{\Phi'}(p) = p\sigma^{-1}(p) = p^2$, $N_{\Phi'}(p^n) \subseteq (p^n)$, hence

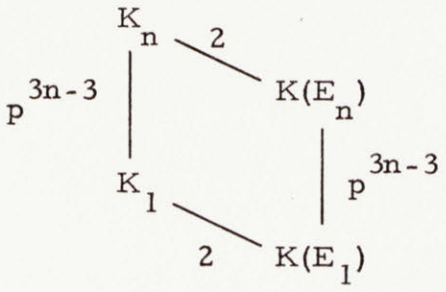
$$\mathrm{Gal}(K_n/K) \simeq N_{\Phi'}((\mathcal{O}_K/p^n\mathcal{O}_K)^{\times} \quad \text{in} \quad (\mathcal{O}_K/p^n\mathcal{O}_K)^{\times} \quad .$$

$N_{\Phi'}$, is surjective onto its image, so

$$\mathrm{Gal}(K_n/K) \simeq (\mathcal{O}_K/p^n\mathcal{O}_K)^{\times}/\ker N_{\Phi'},$$

where $\ker N_{\Phi'} = \{a \in (\mathcal{O}_K/p^n\mathcal{O}_K)^{\times} \mid a\sigma^{-1}(a) \equiv 1 \ (p^n)\}$. But in the proof of theorem (2.5.2) we calculated that this kernel has order $(p+1)p^{n-1}$, i.e., it consists of half the number of a such that $a\sigma^{-1}(a) \equiv \pm 1 \ (p^n)$. Therefore $[K_n:K] = 2[K(E_n):K]$.

ii) We will first show $\mathfrak{B}_1$ ramifies in K_1 to p_1^2. Then from the following diagram:



(*)

we have that the ramification index of $\mathfrak{B}_1$ in K_n is even, $[K(E_n):K(E_1)]$ is odd, so $\mathfrak{B}_n$ ramifies in K_n.

To show that $\mathfrak{B}_1$ ramifies in K_1, we note that the extension K_1/K is cyclic, since it embeds as a subgroup in $(\mathcal{O}_K/p\mathcal{O}_K)^\times$. We can now apply lemma (2.2.1) since

$$2 = [K_1:K(E_1)] \mid [K(E_1):K] = \frac{1}{2}(p^2+1)(p-1).$$

Therefore p ramifies totally in K_1.

iii) Since $[K_n:K(E_n)]$ is a quadratic extension and $p = N_{K(E_n)/K}(\mathfrak{B}_n)$ is odd, $\mathfrak{B}_n$ ramifies tamely in K_n. So $D_p(K_n/K(E_n)) = \mathfrak{B}_n$. Piecing the following formula with lemma (2.5.9); we have:

$$D_p(K_n/K) = N_{K(E_n)/K}(D(K_n/K(E_n))(D_{K(E_n)/K})^2$$

$$= p^{\left[\frac{1}{2}(p^2+1)(p-1)\left(\frac{np^{3n}-(n+1)p^{3n-3}+1}{p^3-1}\right)-1\right]^2}_p$$

$$= p^{(p^2+1)(p-1)\left[\frac{np^{3n}-(n+1)p^{3n-3}+1}{p^3-1}\right]-1}$$

We can glean even more information from our diagram (*) .

Corollary (2.5.12): No primes outside $\mathfrak{P}_1$ ramify in the infinite extension $\left(\bigcup_n K_n\right)/K_1$. A has good reduction everywhere over K_1.

Proof: Looking again at (*), we see that any prime $\mathfrak{S}$ of K_1 which ramifies in K_n has a factor of p in its ramification index. If c is the prime of $K(E_1)$ lying below $\mathfrak{S}$, then c must also ramify to K_n with a ramification index that is a multiple of p. However, since $[K_n : K(E_n)]$ is two, we must have that c ramifies in $K(E_n)$. But we know that only $\mathfrak{P}_1$ ramifies in that extension, so no such $\mathfrak{S} \neq p_1$ exists. Now by the criterion of Neron-Ogg-Shafarevich, A has good reduction over K_1 at all primes of K_1 not above p (considered as a rational prime). However, p remains prime in K, and p_1 is the lone prime of K_1 over p. Further, we chose p so that A over K had good reduction mod p, so A over K_1 has good reduction mod p_1 , too.

§6. Points of Infinite Order

Suppose that A has a Mordell-Weil group $A(K)$ of positive rank; that is, the K-rational points of A, modulo torsion, are generated freely over $\mathbb{Z}$ by some $P_1, \ldots, P_g$, $g > 0$. Let P be one of those generators. Then if $n \neq \pm 1$, there is no Q in $A(K)$ such that $n * Q = P$. However, there are n^4 such Q in $A(\overline{K})$, the points of A in the algebraic closure of K. We will let L_n denote the field obtained by adjoining to K_n the coordinates of all points Q so that $p^n * Q = P$. However, $p^n * Q = P$ and $p^n * Q' = P$ imply $Q \sim Q' \in A_{p^n}$. Since addition is defined over K, if the coordinates of one fixed Q such that $p^n * Q = P$ are adjoined to the K_n, we get all of L_n. We will make a permanent choice of such a Q and call it Q_n.

Lemma (2.6.0): i) L_n/K is a normal extension. ii) $\mathrm{Gal}(L_n/K_n)$ embeds into A_{p^n}, considered as an additive group.

Proof: i) Since multiplication is defined over K, if $\sigma \in \mathrm{Aut}(\mathbb{C}/K)$, $p^n * \sigma(Q_n) = \sigma(P) = P$, so $\sigma(Q_n)$ is a point whose coordinates already lie in L_n.

ii) Let $\sigma \in \mathrm{Gal}(L_n/K_n)$. Then as above,

$$p^n * (\sigma(Q_n) \sim Q_n) = O \quad \text{(the identity on } A\text{)}.$$

So $\sigma(Q_n) \sim Q_n \in A_{p^n}$. The map $\sigma \longmapsto \sigma(Q_n) \sim Q_n$ is independent

of the choice of Q_n, since any two such Q_n differ by a point in A_{p^n}, upon which σ acts as the identity.

A discussion of $\mathrm{Gal}(L_n/K)$ is facilitated by the following identifications. For any $\sigma \in \mathrm{Gal}(L_n/K)$

$$\sigma(Q_n) = Q_n \oplus \alpha_\sigma \quad \text{where} \quad \alpha_\sigma \in A_{p^n}$$

and for any $R \in A_{p^n}$

$$\sigma(R) = \beta_\sigma * R \quad \text{for some} \quad \beta_\sigma \in (\mathcal{O}_K/p^n\mathcal{O}_K)^\times$$

So we can embed $\mathrm{Gal}(L_n/K)$ into $GL_2(\mathcal{O}_K/p^n\mathcal{O}_K)$ by

$$\sigma \longmapsto \begin{pmatrix} 1 & \alpha_\sigma \\ 0 & \beta_\sigma \end{pmatrix} \tag{2.6.1}$$

Where we identify α_σ with its image under the isomorphism $A_{p^n} \simeq \mathcal{O}_K/p^n\mathcal{O}_K$. For the embedding, we will have to write group actions on the right. For any σ, τ in $\mathrm{Gal}(L_n/K)$ we verify:

$$\tau\sigma \longmapsto \begin{pmatrix} 1 & \alpha_\tau \\ 0 & \beta_\tau \end{pmatrix}\begin{pmatrix} 1 & \alpha_\sigma \\ 0 & \beta_\sigma \end{pmatrix} = \begin{pmatrix} 1 & \alpha_\sigma + \alpha_\tau\beta_\sigma \\ 0 & \beta_\tau\beta_\sigma \end{pmatrix} \tag{2.6.2}$$

which agrees with $\sigma(Q_n \oplus \alpha_\tau) = Q \oplus \alpha_\sigma \oplus \alpha_\tau * \beta_\sigma$. Under this scheme, $\mathrm{Gal}(L_n/K_n)$ consists of all matrices in $\mathrm{Gal}(L_n/K)$ of the form $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$.

The main question concerning the embedding of $\mathrm{Gal}(L_n/K_n)$ into A_{p^n} is whether it is surjective. To employ an overused phrase, is the Galois group "as big as possible?" (Kudos to the logician who finds content within that statement.) For elliptic curves the answer is almost always yes [Bas, Se2]. For abelian varieties with complex multiplication, Ribet [R] has shown that $\mathrm{Gal}(L_1/K_1)$ is as big as possible, for all but finitely many p, even when L_1 is obtained from K_1 by adjoining the division values of any number of independent generators of $A(K)$ (K any number field). Lang presents an argument to reduce the proof for L_n/K_n to that of L_1/K_1 in the elliptic case [Lang 4], and it carries over practically word for word for abelian varieties.

However, our case of K is sufficiently special that we can show directly that $\mathrm{Gal}(L_n/K_n) \simeq A_{p^n}$ for all n, and our p (inert in K, odd, good reduction).

We note that $\mathrm{Gal}(K_n/K)$ acts on $A_{p^n} \simeq \mathcal{O}_K/p^n\mathcal{O}_K$ by multiplication, and similarly acts on $A_p \simeq \mathcal{O}_K/p\mathcal{O}_K$ via its quotient $\mathrm{Gal}(K_1/K)$.

Lemma (2.6.3). $H^1(\mathrm{Gal}(K_n/K), A_p) = 0$.

Proof: $n = 1$. Note that A_p is a p-group, so $H^1(\mathrm{Gal}(K_1/K), A_p)$ is a p-group, too. However, $\mathrm{Gal}(K_1/K)$ is cyclic of order d $(= (p^2+1)(p-1))$ prime to p, so by [Se 1, p. 130] $H^1(\mathrm{Gal}(K_1/K), A_p)$ must by annihilated by d. Hence $H^1(\mathrm{Gal}(K_1/K), A_p) = 0$.

$n>1$: $\mathrm{Gal}(K_n/K)$ is isomorphic to $T_p \times N_p$, the product of its p-torsion subgroup T_p and its non-p-torsion subgroup N_p. N_p acts on A_p through its image under the projection onto $\mathrm{Gal}(K_1/K)$. In fact, N_p is isomorphic to $\mathrm{Gal}(K_1/K)$ under this projection since the groups have the same order and any element of $\mathrm{Gal}(K_n/K)$ that is congruent to 1 mod p is a p-torsion element. Hence, $H^1(N_p, A_p) \simeq H^1(\mathrm{Gal}(K_1/K), A_p) = 0$. Furthermore, the only element of A_p fixed under the action of N_p is the origin.

Corresponding to the exact sequence:

$$0 \to N_p \to \mathrm{Gal}(K_n/K) \to \mathrm{Gal}(K_n/K)/N_p \to 0$$

we get the restriction-inflation sequence [Se 1, p. 117]

$$0 \to H^1(\mathrm{Gal}(K_n/K)/N_p, A_p^{N_p}) \to H^1(\mathrm{Gal}(K_n/K), A_p) \to H^1(N_p, A_p)$$

where $A_p^{N_p}$ denotes the elements of A_p fixed under the action of N_p, and hence is trivial. So $H^1(\mathrm{Gal}(K_n/K), A_p)$ is sandwiched exactly between two zero terms, and therefore is zero.

Corollary (2.6.4): L_n/K_n is a non-trivial extension.

Proof: We chose P so that there were no Q in A(K) such that $p * Q = P$. Therefore the extension M/K obtained by adjoining to K the coordinates of any such Q is non-trivial. If L_1/K_1 were a trivial extension, then we would have $M \subseteq K_1$, and in fact it's easy

to see that K_1 would be equal to K adjoined with the coordinates of all the points Q such that $p * Q = P$. For any such Q we get a map

$$f_Q : \mathrm{Gal}(K_1/K) \to A_p$$

by $$f_Q(\sigma) = a_\sigma = \sigma(Q) \sim Q$$

By the relation (2.6.2), this is easily seen to be a one-cocycle, so $\sigma(Q) \sim Q = (\sigma - 1)a$ for some fixed $a = a(Q)$ in A_p, and any $\sigma \in \mathrm{Gal}(K_1/K)$. If $a = 0$ for some Q, then $Q \in A(K)$, a contradiction. But if $a \neq 0$, then every Q has $(p^2+1)(p-1) = [K_1 : K]$ conjugates over K, an impossibility since $(p^2+1)(p-1) \nmid p^4$. Therefore L_1/K_1 is non-trivial. Equivalently, there is an element $\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}$ of $\mathrm{Gal}(L_1/K_1)$, where $\alpha \neq 0$. Then for any $\begin{pmatrix} 1 & \alpha_\sigma \\ 0 & \beta_\sigma \end{pmatrix}$ in $\mathrm{Gal}(L_1/K)$

$$\begin{pmatrix} 1 & \alpha_\sigma \\ 0 & \beta_\sigma \end{pmatrix}\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \alpha + \alpha_\sigma \\ 0 & \beta_\sigma \end{pmatrix}$$

$$\begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & \alpha_\sigma \\ 0 & \beta_\sigma \end{pmatrix} = \begin{pmatrix} 1 & \alpha\beta_\sigma + \alpha_\sigma \\ 0 & \beta_\sigma \end{pmatrix}$$

Since $\alpha \neq 0$ and we can choose $\beta_\sigma \neq 1$, we see that $\mathrm{Gal}(L_1/K)$ is not abelian. As a result, we can now see that L_n/K_n is non-trivial. For $\mathrm{Gal}(L_1/K)$ is a quotient of $\mathrm{Gal}(L_n/K)$, the latter which would be abelian were $L_n \subseteq K_n$.

Lemma 2.6.5: Identifying $\mathrm{Gal}(K_n/K)$ with its image in $(\mathcal{O}_K/p^n\mathcal{O}_K)^\times$, there is an $\alpha \in \mathrm{Gal}(K_n/K)$ such that $\mathcal{O}_K/p^n\mathcal{O}_K = \mathbb{Z}/p^n\mathbb{Z}[\alpha]$.

Corollary 2.6.6. Every element of $\mathcal{O}_K/p^n\mathcal{O}_K$ can be written as a sum of elements of $\mathrm{Gal}(K_n/K)$.

Proof of lemma: Since $|\mathrm{Gal}(K_1/K)| = (p^2+1)(p-1) > (p^2-1) = |\mathcal{O}_{K^+}/p\mathcal{O}_{K^+}^\times|$, there is an $\bar{\alpha} \in \mathrm{Gal}(K_1/K)$ which is not in the "real subfield" $\mathcal{O}_{K^+}/p\mathcal{O}_{K^+}$ of $\mathcal{O}_K/p\mathcal{O}_K$. Since $\mathcal{O}_K/p\mathcal{O}_K$ is a cyclic extension of order 4 over $\mathbb{Z}/p\mathbb{Z}$ the real subfield is the only intermediate field, so we must have that $\mathcal{O}_K/p\mathcal{O}_K = \mathbb{Z}/p\mathbb{Z}[\bar{\alpha}]$. Let α be any element of $\mathrm{Gal}(K_n/K)$ which projects onto $\bar{\alpha}$, and let $\tilde{\alpha}$ be any element of $\mathcal{O}_{K,p}$ (the p-adic completion of $\mathcal{O}_K$, p a prime of K) such that $\tilde{\alpha}$ reduces to α modulo p^n. Then $\tilde{\alpha} \equiv \bar{\alpha} \bmod p$, and since $\mathcal{O}_{K,p}$ over $\mathbb{Z}_p$ is unramified, we must have that $\mathcal{O}_{K,p} = \mathbb{Z}_p[\tilde{\alpha}]$. Hence

$$\mathcal{O}_K/p^n\mathcal{O}_K = \mathbb{Z}_p[\tilde{\alpha}]/p^n\mathbb{Z}_p[\tilde{\alpha}] = \mathbb{Z}/p^n\mathbb{Z}[\alpha].$$

Following the approach in [G] we show:

Lemma (2.6.7): Let $H = \left\{\alpha \;\middle|\; \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \in \mathrm{Gal}(L_n/K_n)\right\}$. Then H is a non-zero ideal in $\mathcal{O}_K/p^n\mathcal{O}_K$.

Proof: H is non-zero by corollary (2.6.4).

Clearly H is an additive subgroup of $\mathcal{O}_K/p^n\mathcal{O}_K$, so we need only check that $\gamma\alpha \in H$ for any $\alpha \in H$, $\gamma \in \mathcal{O}_K/p^n\mathcal{O}_K$. But by corollary (2.6.6), it suffices to show $\gamma\alpha \in H$ just for those $\gamma \in \mathrm{Gal}(K_n/K)$. Given such a γ, there are elements of $\mathrm{Gal}(L_n/K)$ which project onto γ and γ^{-1}, say

$$M_1 = \begin{pmatrix} 1 & \beta_1 \\ 0 & \gamma^{-1} \end{pmatrix} \quad \text{and} \quad M_2 = \begin{pmatrix} 1 & \beta_2 \\ 0 & \gamma \end{pmatrix}$$

Computing M_1M_2 and M_2M_1 we find

$$\beta_2 + \beta_1\gamma\,, \quad \beta_1 + \beta_2\gamma^{-1} \in H \tag{2.6.8}$$

Note further that if $\alpha' \in H$,

$$\begin{pmatrix} 1 & \beta_1 \\ 0 & \gamma^{-1} \end{pmatrix}\begin{pmatrix} 1 & \alpha' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \alpha' + \beta_1 \\ 0 & \gamma^{-1} \end{pmatrix}$$

so the choice of β_1 (and β_2) in (2.6.8) can be translated by any element of H. Therefore, for any $\alpha \in H$ we have:

$$-\beta_1 - \beta_2\gamma^{-1} + \alpha \in H\,, \text{ and } \beta_2 + [(\beta_1 + (-\beta_1 - \beta_2\gamma^{-1} + \alpha)]\gamma = \alpha\gamma \in H\,.$$

We now have enough preliminaries to show:

Theorem (2.6.8): $\mathrm{Gal}(L_n/K_n) \simeq A_{p^n}$.

Proof: By the previous lemma, we have $\mathrm{Gal}(L_n/K_n) \simeq H \subseteq \mathcal{O}_K/p^n\mathcal{O}_K \simeq A_{p^n}$, where H is a non-zero ideal in $\mathcal{O}_K/p^n\mathcal{O}_K$.

For $n = 1$, we have immediately that $L_1/K_1 \simeq A_p$, since $\mathcal{O}_K/p\mathcal{O}_K$ is a field and hence has no proper ideals. For $n > 1$, $\mathcal{O}_K/p^n\mathcal{O}_K$ is a local ring with maximal ideal (p), so $H = (p)^m$ for some $m \geq 0$. We wish to show $m = 0$, so postulate contrarywise that $m \geq 1$. Then for any $\sigma \in \mathrm{Gal}(K_n/K)$, and any $\bar{\sigma}$, $\bar{\bar{\sigma}} \in \mathrm{Gal}(L_n/K)$ which project onto σ, we have the identifications:

$$\bar{\sigma} \longmapsto \begin{pmatrix} 1 & \alpha_{\bar{\sigma}} \\ 0 & \beta_\sigma \end{pmatrix}, \qquad \bar{\bar{\sigma}} \longmapsto \begin{pmatrix} 1 & \alpha_{\bar{\bar{\sigma}}} \\ 0 & \beta_\sigma \end{pmatrix}$$

and

$$\bar{\sigma}\bar{\bar{\sigma}}^{-1} \longmapsto \begin{pmatrix} 1 & (\alpha_{\bar{\sigma}} - \alpha_{\bar{\bar{\sigma}}})\,\beta_\sigma^{-1} \\ 0 & 1 \end{pmatrix}$$

Therefore $\alpha_{\bar{\sigma}} \equiv \alpha_{\bar{\bar{\sigma}}}$ (p) (recall $\beta_\sigma^{-1} \in (\mathcal{O}_K/p^n\mathcal{O}_K)^\times$). Hence $\beta_\sigma \longmapsto \bar{\alpha}_\sigma$ (p) is a well-defined map of $\mathrm{Gal}(K_n/K)$ into A_p. We denote any of the $\bar{\alpha}_\sigma$ modulo p by a_σ. This defines a one-cocycle $f : \mathrm{Gal}(K_n/K) \to A_p$ since (2.6.2) implies $\alpha_{\sigma\tau} = \alpha_\sigma + \beta_\sigma \alpha_\tau$ (recall that multiplication by β_σ is the action of σ on A_{p^n} and therefore, via its quotient, is the action on A_p). Now by lemma (2.6.3) we have a γ in A_p such that

$$\alpha_\sigma = \gamma - \beta_\sigma \gamma, \text{ hence}$$

$$(2.6.9) \qquad \beta_\sigma \equiv 1 \text{ (p)} \implies \alpha_\sigma \equiv 0 \pmod{p}$$

We have

$$\mathrm{Gal}(K_1/K) = \mathrm{Gal}(L_n/K) \Big/ \left\{ \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \Big| \beta \equiv 1 \ (p) \right\}$$

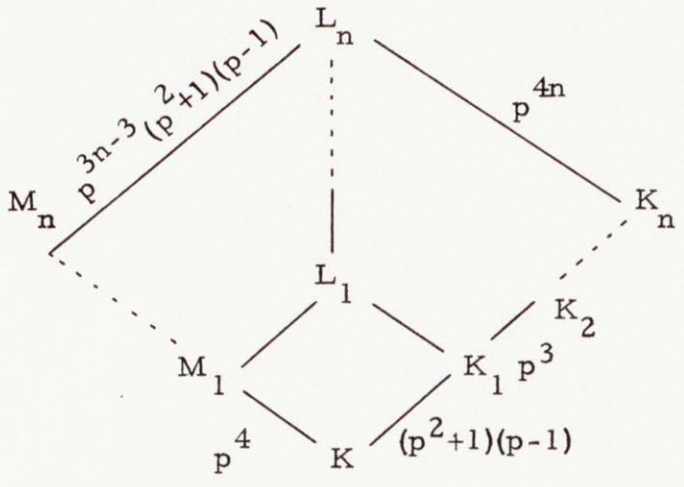
and $$\mathrm{Gal}(L_1/K) = \mathrm{Gal}(L_n/K) \Big/ \left\{ \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \Big| \beta \equiv 1, \ \alpha \equiv 0 \ (p) \right\}$$

So by (2.6.9), L_1/K_1 would be a trivial extension, a contradiction.

Therefore $m = 0$ and $\mathrm{Gal}(L_n/K_n) \simeq A_{p^n}$.

§7: Some Character Relations

Now that we know $\mathrm{Gal}(L_n/K_n) \simeq (\mathbb{Z}/p^n\mathbb{Z})^4$, we can make the following diagram of field extensions, letting $M_n = K(Q_n)$ (that is, K adjoined with the coordinates of Q_n).



In our identification (2.6.1), $\mathrm{Gal}(L_n/M_n)$ corresponds to matrices of the form $\begin{pmatrix} 1 & 0 \\ 0 & * \end{pmatrix}$. Our goal is to relate the conductors of characters on $\mathrm{Gal}(L_n/K_n)$ to relative discriminants in the towers of M's and K's.

First let's calculate the character on $\mathrm{Gal}(L_n/K)$ induced from the unit character on $\mathrm{Gal}(L_n/M_n)$; we will denote it by $1^*(L_n/M_n)$. Then

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

$$= \frac{1}{[L_n : M_n]} \left(\#\left\{ \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} \in \mathrm{Gal}(L_n/K) \,\middle|\, \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix}^{-1} \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} \in \mathrm{Gal}(L_n/M_n) \right\} \right)$$

We calculate that

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

so we need to find the number of $(u,v) \in (\mathcal{O}_K/p^n \mathcal{O}_K) \times \mathrm{Gal}(K_n/K)$ such that $(1-\beta)u + \alpha v \equiv 0 \ (p^n)$.

Let $0 \le m \le n$ be such that $p^m \| (1-\beta)$ (precisely divides). Then there are no solutions unless $p^m | \alpha$, in which case we are searching for solutions to

$$\left(\frac{1-\beta}{p^m}\right) u + \left(\frac{\alpha}{p^m}\right) v \equiv 0 \ (p^{n-m})$$

For any of the $[K_n:K]$ choices for v, there is a unique choice for u modulo p^{n-m}, and therefore p^{4m} choices for u modulo p^n. So for $p^m \| (1-\beta)$, $0 \le m \le n$,

$$1^*(L_n/M_n) \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} = \begin{cases} p^{4m} & p^m | \alpha \\ 0 & \text{otherwise} \end{cases}$$

Note that $\mathrm{Gal}(L_n/M_{n-1})$ consists of matrices of the form $\begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}$ where $\alpha \equiv 0 \ (p^{n-1})$. We will now investigate the induction of the unit character on $\mathrm{Gal}(L_n/M_{n-1})$ to all of $\mathrm{Gal}(L_n/K)$.

$$1^*(L_n/M_{n-1}) \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix}$$

$$= \frac{1}{[L_n:M_{n-1}]} \left\{ \# \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} \in \mathrm{Gal}(L_n/K) \left| \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix}^{-1} \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} \begin{pmatrix} 1 & u \\ 0 & v \end{pmatrix} \in \mathrm{Gal}(L_n/M_{n-1}) \right. \right\}$$

That is, we need count the number of (u,v) in $\mathcal{O}_K/p^n \mathcal{O}_K \times \mathrm{Gal}(K_n/K)$ such that $(1-\beta)u + \alpha v \equiv 0 \ (p^{n-1})$.

Let $0 \le m \le n-1$ be such that $p^m \| (1-\beta)$, and set $m = n-1$ if $\beta = 1$. Then there are no solutions unless $p^m \mid \alpha$, in which case we are searching for solutions to

$$\left(\frac{1-\beta}{p^m}\right)u + \left(\frac{\alpha}{p^m}\right)v \equiv 0 \quad (p^{n-m-1})$$

For every v there is a unique u modulo p^{n-m-1}, and therefore p^{4m+4} choices for u modulo p^n. So for $p^m \| (1-\beta)$, $0 \le m \le n-1$,

$$1^*(L_n/M_{n-1}) \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} = \begin{cases} p^{4m} & p^m \mid \alpha \\ 0 & \text{otherwise} . \end{cases}$$

Comparing these results, we note $1^*(L_n/M_{n-1})$ and $1^*(L_n/M_n)$ differ only when $\beta \equiv 1 \ (p^n)$. Let ψ be the character equal to their difference, $1^*(L_n/M_n) - 1^*(L_n/M_{n-1})$. Then

$$(2.7.0) \qquad \psi \begin{pmatrix} 1 & \alpha \\ 0 & \beta \end{pmatrix} = \begin{cases} p^{4n} - p^{4n-4} & \beta = 1, \ \alpha = 0 \\ -p^{4n-4} & \beta = 1, \ p^{n-1} \| \alpha \\ 0 & \text{otherwise} \end{cases}$$

Let χ be any of the $p^{4n} - p^{4n-4}$ first degree characters on $H = \mathrm{Gal}(L_n/K_n)$ such that $\chi^{p^{n-1}} \neq 1$. Let $\langle \ , \ \rangle_H$ denote the inner product of characters on the group H, and let $\psi\big|_H$ denote ψ restricted to H.

$$\langle \psi\big|_H, \chi \rangle = \frac{1}{p^{4n}} \left[(p^{4n} - p^{4n-4})\chi(I) + (-p^{4n-4}) \sum_{p^{n-1} \| \alpha} \chi \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} \right]$$

We rewrite α as $p^{n-1}b$ where $b \in (\mathcal{O}_K/p\mathcal{O}_K)^\times$ ($\alpha \neq 0$), and evaluate the sum

$$\sum_{b \in (\mathcal{O}_K/p\mathcal{O}_K)^\times} \chi\begin{pmatrix} 1 & bp^{n-1} \\ 0 & 1 \end{pmatrix} = \sum_{b \in (\mathcal{O}_K/p\mathcal{O}_K)^\times} \chi^{p^{n-1}}\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = -1$$

since $\chi^{p^{n-1}}$ was chosen to be nontrivial. Hence

$$\langle \psi\big|_H, \chi \rangle_H = \frac{1}{p^{4n}}[p^{4n} - p^{4n-4} + p^{4n-4}] = 1$$

We note that the degree of ψ is $p^{4n} - p^{4n-4}$, and there are precisely that many χ such that $\chi^{p^{n-1}} \neq 1$. Therefore we find

$$\psi\big|_H = \bigoplus_{\chi^{p^{n-1}} \neq 1} \chi$$

Letting "$*$" denote induction to all of $\mathrm{Gal}(L_n/K)$

$$(\psi\big|_H)^* = \bigoplus_{\chi^{p^{n-1}} \neq 1} \chi^* \tag{2.7.1}$$

But since $\psi(x) = 0$ for $x \notin H$, we verify for $x \in H$

$$(\psi\big|_H)^*(x) = \frac{1}{[L_n:K_n]} \sum_{g \in \mathrm{Gal}(L_n/K)} (\psi\big|_H)(gxg^{-1}) \tag{2.7.2}$$

$$= \frac{1}{[L_n:K_n]} \sum_{g \in \mathrm{Gal}(L_n/K)} \psi(x)$$

$$= [K_n:K]\psi(x)$$

And by Frobenius reciprocity

$$(2.7.3) \qquad \langle \psi , \chi^* \rangle_{Gal(L_n/K)} = \langle \psi \big|_H , \chi \rangle_H = 1$$

So we can conclude:

i) Any irreducible character on $Gal(L_n/K)$ contained in χ^* must be contained in ψ, ((2.7.1), (2.7.2)).

ii) Hence by (2.7.3), χ^* is irreducible.

iii) So again by (2.7.1) and (2.7.2), ψ is a sum of different χ^*'s, perhaps with multiplicities.

Comparing degrees: $\psi(1) = p^{4n} - p^{4n-4}$

$$\chi^*(1) = p^{3n-3}(p^2+1)(p-1)$$

We conclude $\psi = \Sigma(\chi^*)_i m_i$ where the $(\chi^*)_i$ denotes an indexed set of distinct induced characters, and $\Sigma m_i = p^{n-1}(p+1)$, $m_i > 0$. But we can calculate from (2.7.0)

$$\Sigma m_i^2 = \langle \psi , \psi \rangle_{Gal(L_n/K)} = \frac{1}{p^{7n-3}(p^2+1)(p-1)} [(p^{4n} - p^{4n-4})^2 + (-p^{4n-4})^2(p^4-1)]$$

$$= \frac{p^{8n} - p^{8n-4}}{p^{7n-2}(p^2+1)(p-1)} = p^{n-1}(p+1)$$

Hence all the $m_i = 1$. Summarizing

Theorem (2.7.4): The $(p^4-1)p^{4n-4}$ first degree characters χ on $Gal(L_n/K_n)$ each induce to a χ^* on $Gal(L_n/K)$ where

i) χ^* is irreducible, and there are $p^{3n-3}(p^2+1)(p-1)$ χ which induce to each of the distinct $(\chi^*)_i$, $1 \le i \le p^{n-1}(p+1)$.

ii) $1^*(L_n/M_n) - 1^*(L_n/M_{n-1}) = \bigoplus_{i=1}^{p^{n-1}(p+1)} (\chi^*)_i$.

Corollary (2.7.5):

$$D(M_n/K) = D(M_{n-1}/K)(D(K_n/K))^{p^{n-1}(p+1)} \prod_{i=1}^{p^{n-1}(p+1)} N_{K_n/K}(f(\chi_i))$$

where χ_i is any character on $\mathrm{Gal}(L_n/K_n)$ which induces to $(\chi^*)_i$.

Proof: We recall some standard properties of conductors, namely for $A \supset B \supset C$ number fields, A/C Galois, χ_i a character on Gal(A/B), χ_i^* its induction to Gal(A/C),

1) $f(\chi_1 + \chi_2) = f(\chi_1) f(\chi_2)$

2) $f(\chi^*) = D(B/C)^{\chi(1)} N_{B/C}(f(\chi))$

3) $f(1) = 1$

We now apply these in turn to calculate the conductor of both sides of part (ii) of Theorem (2.7.4).

Our calculations with corollary (2.7.5) will be local, involving only the p-part of the discriminants. This will result in no diminution in our knowledge of the $f(\chi)$, since we have the following:

Lemma (2.7.6): L_n/K_n is unramified outside p_n.

Proof: By corollary (2.5.12), A has good reduction at any prime b of K_n. Take $b \neq p_n$. If L_n/K_n were ramified at some -- and therefore every -- $\mathfrak{B}$ over b, and σ an element of the inertia group of b, then $\sigma(Q_n) \equiv Q_n(\mathfrak{B})$. But then the reduction of A mod b would have fewer than p^{4n} points $\bar{Q}$ such that $p^n * \bar{Q} = \bar{P}$ (a bar denotes reduction modulo b); impossible if A has good reduction at b. Q.E.D.

We hope to determine later whether for some n, L_n/K_n is ramified over p_n. We now note that it would suffice to show that the local extension $L_{n_q}/K_{n_{p_n}}$ is non-trivial for any (and therefore all) primes q of L_n above p_n. (In other words, the residue class degree of q over p_n is 1.) For were $L_{n_q}/K_{n_{p_n}}$ non-trivial and unramified, then $L_{n_q}/K_{n_{p_n}}$ would be the unique unramified extension of $K_{n_{p_n}}$ of degree $[L_{n_q} : K_{n_{p_n}}]$. But $[K_{n_{p_n}} : K_p]$ is totally ramified, and letting N be the unique unramified extension of K_p of degree $[L_{n_q} : K_{n_{p_n}}]$, we would have $L_{n_q} = K_{n_{p_n}} N$, the latter of which is abelian over K_p. However in the proof of corollary (2.6.4) we showed that no non-trivial element of $\mathrm{Gal}(L_n/K_n)$ could commute with all of $\mathrm{Gal}(K_n/K) = \mathrm{Gal}(K_{n_{p_n}}/K_p)$.

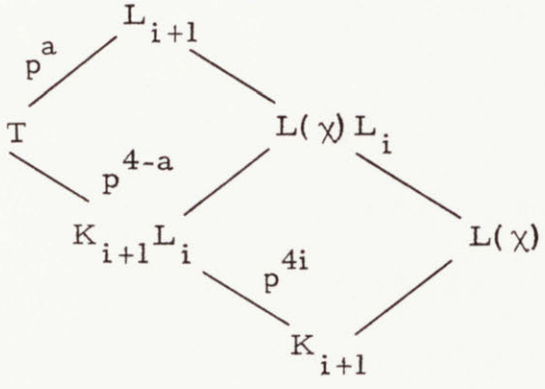
Under the supposition that L_n/K_n is ramified for some n, let i be the index such that $L_1/K_1, \ldots, L_i/K_i$ are unramified and L_{i+1}/K_{i+1} is ramified.

Theorem (2.7.7): $L_{i+1}/K_{i+1}L_i$ is totally ramified at any of the primes q in $K_{i+1}L_i$ above p_{i+1}.

Proof: The preliminary remarks guarantee that p_{i+1} splits into p^{4i} primes q_j in $K_{i+1}L_i$, and since L_{i+1}/K_{i+1} is abelian, if one of the q_j is totally ramified then all of them are. Fix one such q. By the assumption on i, the inertia group I of q in L_{i+1}/K_{i+1} is non-trivial.

$$\mathrm{Gal}(L_{i+1}/K_{i+1}L_i) \simeq \left\{\begin{pmatrix} 1 & \alpha \equiv 0 \ (p^i) \\ 0 & 1 \end{pmatrix}\right\} \simeq (\mathcal{O}_K/p\mathcal{O}_K) \simeq (\mathbb{Z}/p\mathbb{Z})^4 .$$

So we need to exclude the following three cases: $|I| = p^a$, $a = 1, 2, 3$.



Let T be the fixed field of I. Then $T/K_{i+1}L_i$, an extension of degree p^{4-a}, is unramified at q. We want to count the number of characters χ on $\mathrm{Gal}(L_{i+1}/K_{i+1})$, $\chi^{p^i} \neq 1$, such that $f(\chi) = 1$. If χ is any character on $\mathrm{Gal}(L_{i+1}/K_{i+1})$, $\chi^{p^i} \neq 1$, we have that $L(\chi)$, the fixed field of the kernel of χ, is not contained in $K_{i+1}L_i$, and that $L(\chi)L_i$ is the fixed field of the kernel of χ

restricted to $\mathrm{Gal}(L_{i+1}/K_{i+1}L_i)$. Since L_iK_{i+1}/K_{i+1} is unramified, $L(\chi)/K_{i+1}$ is unramified if and only if $L(\chi)L_i/K_{i+1}L_i$ is, and $L(\chi)L_i/K_{i+1}L_i$ is unramified if and only if it is contained in T.

Further, there are precisely p^{4i} χ which have a given restriction to $\mathrm{Gal}(L_{i+1}/K_{i+1}L_i)$, and $(p-1)$ restrictions which define the same $L(\chi)L_i$. Of these $L(\chi)L_i$, $(p^{4-a}-1)/(p-1)$ are contained in T. Therefore precisely $(p^{4-a}-1)p^{4i}$ of the characters χ on $\mathrm{Gal}(L_{i+1}/K_{i+1})$, $\chi^{p^i} \neq 1$, have $f(\chi) = 1$. But of the $p^{4i+4} - p^{4i}$ characters χ on $\mathrm{Gal}(L_{i+1}/K_{i+1})$, $\chi^{p^i} \neq 1$, there are $(p^2+1)(p-1)p^{3i}$ which induce the same χ^* on $\mathrm{Gal}(L_{i+1}/K)$, and therefore have the same conductor.

Since $(p^2+1)(p-1)p^{3i} \nmid p^{4i}(p^{4-a}-1)$ for $a = 1, 2, 3$, we have that q is totally ramified in L_{i+1}.

Lemma (2.7.8): Let L/K be a normal extension of number fields, totally ramified at a prime p of norm p^f, such that $\mathrm{Gal}(L/K) = (\mathbb{Z}/p\mathbb{Z})^n$. Then $p^2 \mid f(\chi)$ for every $\chi \neq 1$ in $\widehat{\mathrm{Gal}}(L/K)$.

Proof: Let $L(\chi)$ be the fixed field of χ; then $\mathrm{Gal}(L(\chi)/K) = \mathbb{Z}/p\mathbb{Z}$. The inertia group G_o of p in $L(\chi)/K$ is $\mathbb{Z}/p\mathbb{Z}$ by construction. Letting G_1 be the next ramification group; $(|G_o/G_1|, p) = 1$ so $G_1 = \mathbb{Z}/p\mathbb{Z}$, too. Hence $p^{2p-2} \mid D(L(\chi)/K) = f(\chi)^{p-1}$.

Corollary (2.7.9): Assume L_n/K_n is ramified for some n, and let $i+1$ be the first index at which ramification occurs. Then we have $p^2 \mid N_{K_{i+1}/K}(f(\chi))$, where $\chi \in \widehat{Gal}(L_{i+1}/K_{i+1})$, $\chi^{p^i} \neq 1$.

Proof: Let $L(\chi)$ be the fixed field of χ, and q any prime of $K_{i+1}L_i$ over p_{i+1}. By theorem (2.7.7), $L(\chi)L_i/K_{i+1}L_i$ is totally ramified over q, hence $L(\chi)/K_{i+1}$ is totally ramified over p_{i+1}. Therefore by lemma (2.7.8), $p_{i+1}^2 \mid f(\chi)$, and taking norms, $p^2 \mid N_{K_{i+1}/K}(f(\chi))$.

Lemma (2.7.10): Keeping the same definition of i, for $n \leq i$

$$D_p(M_n/K) = p^{\left[\frac{np^{4n+4} - (n+1)p^{4n} + 1}{p^4 - 1} - \frac{p^{4n} + (p^n - 1)p(p+1) - 1}{p^3 - 1}\right]}$$

Proof: For $n = 1$ this is just $D_p(K_1/K)^{p+1}$. It follows by induction from corollary (2.7.5).

References

[Bak] Baker, H. F.: "An Introduction to the Theory of Multiply Periodic Functions," Cambridge University Press, Cambridge, 1907.

[Bas] Bashmakov, M.: Un Théorème de Finitude sur la Cohomology des courbes elliptiques, C.R. Acad. Sci. Paris Sér. A-B 270, (1970) A999-A1000.

[BS] Borevich, Z. and Shafarevich, I.: "Number Theory," Academic Press, New York, 1966.

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

[Fr] Freitag, E.: Zur Theorie der Modulformen Zweiten Grades. Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II, (1965) 151-157.

[G] Gupta, R.: "Fields of Division Points of Elliptic Curves Related to Coates-Wiles," Ph.D. Thesis, M.I.T., 1983.

[H] Hammond, W.: On the Graded Ring of Siegel Modular Forms of Genus Two. Am. J. Math. 87, (1965) 502-506.

[I1] Igusa, J.: "Theta Functions," Die Grund. der math. Wiss. Band 194, Springer-Verlag, Berlin, 1972.

[I2] Igusa, J.: On Siegel Modular Forms of Genus Two. I, II. Am. J. Math. 84, (1962) 175-200; 86 (1964) 392-412.

[Land] Landau, E.: "Elementary Number Theory," Chelsea, New York, 1966.

[Lang1] Lang, S.: "Introduction to Algebraic and Abelian Functions," 2nd ed., Grad. Texts in Math. no. 89, Springer-Verlag, New York, 1982.

[Lang2] Lang, S.: "Elliptic Functions," Addison-Wesley, Reading, 1973.

[Lang3] Lang, S.: "Complex Multiplication," Die Grund. der math. Wiss. Band 255, Springer-Verlag, New York, 1983.

[Lang4] Lang, S.: "Elliptic Curves: Diophantine Analysis," Grund. der math. Wiss. Band 231, Springer-Verlag, Berlin, 1978.

[M1] Mumford, D.: "Tata Lectures on Theta. I, II," Prog. in Math. Vol. 28, 43, Birkhäuser, Boston, 1983, 1984.

[M2] Mumford, D.: "Curves and Their Jacobians," University of Michigan Press, Ann Arbor, 1976.

[Ma] Maass, H.: "Siegel Modular Forms and Dirichlet Series," Lect. Notes in Math. no. 216, Springer-Verlag, Berlin, 1971.

[R] Ribet, K.: Dividing Rational Points on Abelian Varieties of CM-type. Compositio Math. 33, Fasc. 1, (1976) 69-74.

[Ra] Raynaud, M.: Courbes sur une Variété Abélienne et Points de de Torsion. Invent. Math. 71, no. 1, (1983) 207-233.

[Se1] Serre, J.-P.: "Local Fields," Grad. Texts in Math. no. 67, Springer-Verlag, New York, 1979.

[Se2] Serre, J.-P.: Propriétés Galoisiennes des Points d'Ordre Fini des courbes elliptiques. Invent. Math. 15 (1972) 259-331.

[Sh] Shimura, G.: On the Zeta-Function of an Abelian Variety with Complex Multiplication, Ann. of Math. (2), 94 (1971) 504-533.

[St] Stark, H.: The Coates-Wiles Theorem Revisited, in "Number Theory Related to Fermat's Last Theorem," Prog. in Math., Vol. 26, Birkhäuser, Boston, 1982.

[U] Uchida, K.: Class Numbers of Imaginary Abelian Number Fields. I. Tôhoku Math. J. (2) 23 (1972) 97-104.

[Wa] Washington, L.: "Introduction to Cyclotomic Fields," Grad. Texts in Math. no. 83, Springer-Verlag, New York, 1982.

[Wo] Wolfe, T.: "Look Homeward, Angel," Charles Scribner's Sons, New York, 1929.