Riemann Surfaces and Modular Function Field Extensions

Bryden R. Cais
cais@fas.harvard.edu
(617) 493–2628

Supervised by Noam D. Elkies

A thesis presented to the Department of Mathematics
in partial fulfillment of the requirements
for the degree of Bachelor of Arts with Honors

Harvard University
Cambridge, Massachusetts
April 1, 2002

Contents

0.1 Introduction

Given any Riemann surface X, we can consider the field of meromorphic functions on X, denoted $K(X)$. This is always an extension field of $\mathbb{C}$ and is isomorphic to $\mathbb{C}(x, y)$ with $F(x, y) = 0$ for some rational function in x and y. Now if $p : X \longrightarrow Y$ is any nonconstant holomorphic mapping of Riemann surfaces, then we can view $K(Y)$ as a subfield of $K(X)$ by the injective map $f \mapsto f \circ p$, for any $f \in K(Y)$. In fact, we have a correspondence between degree n covering maps $p : X \longrightarrow Y$ and degree n field extensions $K(X)/K(Y)$. The goal of Chapter 1 is to develop enough of the theory of compact Riemann surfaces to prove part of this correspondence.

In Chapter 2, we turn our attention to the group $\mathbf{SL}_2(\mathbb{Z})$ of two by two integer matrices with determinant 1, which has a natural action on the complex vector space $\mathbb{C}^2$. This action descends to an action on the Riemann sphere $\mathbb{P}^1$, which in turn descends to an action on the complex upper half plane $\mathbb{H}$, given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z := \frac{az + b}{cz + d}.$$

Clearly, any subgroup Γ of $\mathbf{SL}_2(\mathbb{Z})$ also acts on $\mathbb{H}$ by the same formula. We may therefore consider the quotient space $\mathbb{H}/\Gamma$. By the appropriate construction, this space can be made into a compact Riemann surface X. In particular, for each positive integer N, we study the Riemann surface $X(N)$ that arises as a quotient of the upper half plane by the *principal congruence subgroup of level N* of $\mathbf{SL}_2(\mathbb{Z})$. These are the normal subgroups $\Gamma(N)$ of $\mathbf{SL}_2(\mathbb{Z})$ given as the kernel of the reduction modulo N map $r_N : \mathbf{SL}_2(\mathbb{Z}) \longrightarrow \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$. There is a natural map $\pi : X(N) \longrightarrow X(1)$ that realizes $X(N)$ as a covering space of $X(1)$. We can then use the tools of the first chapter to show that we have a field extension $K(X(N))/K(X(1))$. The fact that $\Gamma(N)$ is normal in $\Gamma(1)$ for each N will enable us to show that this is in fact a Galois field extension with Galois group $\Gamma(1)/\pm\Gamma(N) \simeq \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})/\pm\{1\}$.

In the last chapter, we study the Galois field extensions $K(X(N))/K(X(1))$. For precisely 5 values of N, namely $1 \leq N \leq 5$, the Riemann surface $X(N)$ is of genus 0. We will show that this implies that the field $K(X(N))$ is generated by a single element over $\mathbb{C}$. We then turn to the theory of Elliptic functions and construct such a generator for each N of interest. Using this generator, the group $\Gamma(1)/\pm\Gamma(N)$ can be made to act on the Riemann sphere. We show how to interpret this action as the symmetries of an inscribed solid. Finally, we use the Galois correspondence to give a degree N field extension of $K(X(N))$ for each $2 \leq N \leq 5$ corresponding to an index N subgroup of $\Gamma(1)/\pm\Gamma(N)$. We know that such a field extension is given by adjoining a root of some degree N polynomial over $K(X(1))$. Using the tools of the first two chapters, we construct such a polynomial. As a consequence, when $N = 4$ or 5, we obtain a description of the field extension $K(X(N))/K(X(1))$ as the spliting field of a degree N polynomial.

Chapter 1

Riemann Surfaces

1.1 Definitions and Notations

A *Riemann surface* X is a one dimensional connected complex analytic manifold. That is, X is a connected, Hausdorff topological space S equipped with a *complex structure*. For every point $P \in S$, there exists a neighborhood U_P of P and a *complex chart* φ_P from U_P to the interior of the unit disc such that φ_P is a homeomorphism, and $\varphi_P(P) = 0$. These complex charts are required to be *holomorphically compatible*; that is, for any charts φ_P, φ_Q the map

$$\varphi_Q \varphi_P^{-1} : \varphi_P(U_P \cap U_Q) \longrightarrow \varphi_Q(U_P \cap U_Q)$$

is biholomorphic. Notice that it is immaterial whether we require a complex chart to map a neighborhood of a point homeomorphically to the interior of the unit disc or simply onto an open subset of $\mathbb{C}$ as we can easily convert the latter situation into the former by an appropriate Riemann mapping.

We remark that the usual topological terminology used in describing the space S carries over to the Riemann surface X. Thus, a point P of X is just the point P of S, a closed subset of X is a closed subset of S, and so on. In particular, if the space S is compact, we shall say that X itself is compact.

1.2 Maps Between Riemann Surfaces

Definition 1. Let X and X' be two Riemann surfaces with f any map from $U \subset X$ into X'. For a point P of X, set $Q = f(P)$ and let φ_P, φ'_Q be complex charts at P, Q respectively. Recall that f is said to be *analytic* at P if the function $\varphi'_Q f \varphi_P^{-1}$ which maps the interior of the unit disc to itself is complex analytic at the origin, and that if f is analytic at every point of U then it is an *analytic map* from U to X'. We also call f a *holomorphic mapping*.

Let X be any Riemann surface and $U \subset X$ a neighborhood of a point $P_0 \in X$. Let $f : U \longrightarrow \mathbb{C}$ be a one to one analytic map such that $f(P_0) = 0$. Then $t = f(P)$ is said to be a *locally uniformizing variable* at P_0. Since f is one to one, we will speak of a point t of X, by which we mean the point P of X such that $f(P) = t$. Locally uniformizing variables generalize the idea of complex charts, since for any point $P_0 \in X$, one may take $t = \varphi_{P_0}(P)$ as a locally uniformizing variable, where φ_{P_0} is a complex chart on X at P_0. This shows in particular that every point of X has a locally uniformizing variable. Any two locally uniformizing variables are related to each other by a power series, as the following theorem makes evident.

Theorem 1. *Suppose that $t = f(P)$ and $s = g(P)$ are two locally uniformizing variables at some point P_0 of a Riemann surface X. Then for all P near P_0 one has*

$$s = a_1 t + a_2 t^2 + \cdots , \tag{1.1}$$

where $a_1 \neq 0$. Conversely, if t is any locally uniformizing variable at P_0, then any power series of the above form gives another locally uniformizing variable at P_0.

Proof. Let U, V be the domains where f, g are defined. Since both s, t are locally uniformizing variables, the functions f, g are one to one and analytic. Thus, the map $gf^{-1} : f(U \cap V) \longrightarrow g(U \cap V)$ is one to one and analytic. Since $gf^{-1}(0) = 0$, we have $s = gf^{-1}(t) = a_1 t + a_2 t^2 + \cdots$ for all P near P_0. Since this map is one to one, we must have $a_1 \neq 0$. Conversely, any power series of the form 1.1 gives an analytic one to one function of t and hence an analytic one to one function g of P near P_0 with $g(P_0) = 0$. $\square$

We now use Theorem 1 to prove a key result about any analytic map between two Riemann surfaces.

Theorem 2. *Let X, Y be any two Riemann surfaces and let $f : X \longrightarrow Y$ be analytic at P_0, with $f(P_0) = Q_0$ and f not identically 0. Then there exist locally uniformizing variables t at P_0 and s at Q_0 such that*

$$s = t^n, \quad n \geq 1. \tag{1.2}$$

Proof. Let $t = \varphi(P)$ and $t' = \varphi'(Q)$ for $P \in X$ near P_0 and $Q \in Y$ near Q_0 be locally uniformizing variables. The analyticity of f at P_0 is equivalent to the condition that $t' = \varphi' f \varphi^{-1}(t)$ be analytic at $t = 0$. This amounts to being able to write t' as a power series in t, viz.

$$t' = a_1 t + a_2 t^2 + \cdots .$$

Since f is not identically zero, there exists some $n \geq 1$ such that $a_n \neq 0$, so that

$$t' = a_n t^n (1 + \cdots).$$

Since any power series with leading term $a_n t^n$ has an analytic n^{th} root, there exists some power series

$$t_1 = b_1 t + b_2 t^2 + \cdots ,$$

with $b_1 \neq 0$ and $t' = t_1^n$. The converse of Theorem 1 tells us that t_1 is a locally uniformizing variable at P_0, and this completes the proof. □

The integer n such that $s = t^n$ is in fact independent of the choice of uniformizing variable, which follows from the prrof of Theorem 2.

Theorem 3 (Identity Theorem). *Let X, Y be Riemann surfaces, and $f : X \longrightarrow Y$ a holomorphic mapping. If $f = 0$ on a set A having a limit point $a \in X$, then $f \equiv 0$ on X.*

Proof. Let $U \subset X$ be the set of all points in X having a neighborhood W such that $f \equiv 0$ on W. By definition, U is open. We show that it is closed. Let b be a limit point of U. Since f is continuous, $f(b) = 0$. Let $\varphi_1 : U_1 \longrightarrow V_1$ and $\varphi_2 : U_2 \longrightarrow V_2$ be locally uniformizing variables at b and $f(b)$ with the property that U_1 is connected. The map

$$\varphi_2 \circ f \circ \varphi_1^{-1} : V_1 \longrightarrow V_2 \subset \mathbb{C}$$

is holomorphic since f is. Moreover, $U \cap U_1 \neq \emptyset$ since b is a limit point of U and U_1 is open. By the identity theorem for holomorphic functions on domains in $\mathbb{C}$ [10, pg. 228], we see that f is identically 0 on U_1 and that $a \in U$. Thus, $b \in U$ and U is closed. Since X is connected, we must have $U = \emptyset$ or $U = X$. Since $a \in U$, the first case is impossible and $f \equiv 0$ on X. □

1.3 Functions on a Riemann Surface

Recall that a holomorphic function on an open subset Y of a Riemann surface X is just a holomorphic mapping $f : Y \longrightarrow \mathbb{C}$. It is not difficult to see that the sum and product of any two holomorphic functions on an open subset Y of a Riemann surface X are again holomorphic. Moreover, the constant functions are holomorphic. Therefore the set of all holomorphic functions on Y, denoted $\mathcal{O}(Y)$, is endowed with the structure of a $\mathbb{C}$-algebra. Every locally uniformizing variable on Y is clearly holomorphic. Sometimes we will have a function f which is holomorphic on some deleted neighborhood of a point $a \in Y \subset X$ and we will want to extend f to a holomorphic function on the entire (undeleted) neighborhood. The following well known theorem tells us when this is possible:

Theorem 4 (Riemann's Removable Singularities Theorem). *Let Y be an open subset of a Riemann surface X and let $a \in Y$. The function $f \in \mathcal{O}(Y \setminus \{a\})$ may be uniquely continued to a function $\tilde{f} \in \mathcal{O}(Y)$ precisely when f is bounded on some deleted neighborhood of a.*

We do not prove this theorem, but note that it follows easily from the analogous principle for holomorphic functions on $\mathbb{C}$ [4, pg. 5].

Given a Riemann surface X, we would like for the $\mathbb{C}$-algebra $\mathcal{O}(X)$ to have the structure of a field, but this is not possible if we restrict ourselves to working only with holomorphic functions, as dividing by a function f with a zero at some point a introduces a pole at a. Moreover, we will be working only with compact Riemann surfaces, and as the following lemma shows, holomorphic functions on a compact Riemann surface are very uninteresting:

Lemma 1. *Every holomorphic function f on a compact Riemann surface X is constant.*

Proof. By definition, f is a holomorphic mapping $f : X \longrightarrow \mathbb{C}$. Suppose that f is nonconstant. First, $f(X)$ is open. This follows from Theorem 2, since given a point $a \in X$ with $f(a) = b$, there exist locally uniformizing variables $t,\ s$ at $a,\ b$ with $s = t^k$, so that f maps a neighborhood of a to a neighborhood of b. Second, $f(X)$ is compact since X is. Thus, $f(X)$ is a compact, open subset of $\mathbb{C}$, and therefore empty. This is a contradiction. □

The remedy to these problems is to allow functions to take the value ∞.

Definition 2. Let Y be an open subset of a Riemann surface X. A *meromorphic function* f on Y is a holomorphic mapping(other than the constant mapping $f \equiv \infty$) to the Riemann sphere: $f : Y \longrightarrow \mathbb{P}^1$.

Let $z = f(P)$ be a meromorphic function on $Y \subset X$ and suppose that f is analytic at the point $P_0 \in Y$. We then have two cases:

1. $f(P_0) = z_0 \neq \infty$. Then $z - z_0$ is a locally uniformizing variable at P_0. Now let t be any locally uniformizing variably at P_0. Then by Theorem 1 we may write $z = a_0 + a_1 t + a_2 t^2 + \cdots$. The least integer n such that $a_n \neq 0$ is the *order* of f at P_0.

2. $f(P_0) = \infty$. Then $1/z$ is a locally uniformizing variable at P_0 and we have $1/z = a_1 t + a_2 t^2 + \cdots$. This enables us to write $z = b_{-n} t^{-n} + b_{-n+1} t^{-n+1} + \cdots$ for some positive integer n. The integer $-n$ is called the *order* of f at P_0.

Thus we see that a zero of f at P_0 corresponds to f having positive order at P_0 while a pole corresponds to f having negative order. A function that is identically 0 is defined to have order ∞. By the remark after the proof of Theorem 2, the order of a meromorphic function f at a point P_0 is well defined (that is, it does not depend on a choice of uniformizing variable).

Clearly, the sum and product of any two meromorphic functions on a Riemann surface X is again meromorphic. Moreover, by Theorem 3, we see that any meromorphic function that is not identically 0 has isolated zeroes so that its reciprocal is meromorphic. Thus, the set of all meromorphic functions on X is a field, denoted $K(X)$. Notice that every constant function $f : X \longrightarrow \mathbb{P}^1$ excluding $f \equiv \infty$ is a meromorphic function, so that $K(X)$ contains $\mathbb{C}$. We record this important fact here:

Theorem 5. *The set of all meromorphic functions on a Riemann surface X forms a field, denoted $K(X)$, and is an extension field of $\mathbb{C}$.*

The field $K(X)$ of meromorphic functions on a Riemann surface X is an important invariant of X (that is, under isomorphism) and, as we shall see, encodes many of the properties of X.

One might protest that we do not yet know that there are any nonconstant meromorphic functions on a given Riemann surface. In fact, given any n distinct points $P_1, P_2, \ldots, P_n$ in a Riemann surface Y, there exists a function $f \in K(Y)$ such that $f(P_j) \neq f(P_i)$ for $i \neq j$. We will prove this assuming the following:

Theorem 6. *[5, pg. 122] Let $P_1,\ P_2$ be any two distinct points on a Riemann surface X. Then there exists $f \in K(X)$ with a zero of order one at P_1 and a pole of order one at P_2.*

We now show by induction on n that there exists $f \in K(X)$ with $f(P_j)$ distinct for $j = 1, 2, \ldots, n$. By Theorem 6, we already know the result for $j = 2$, so assume that we have some $f \in K(X)$ such that $f(P_1), \ldots, f(P_k)$ are all distinct. Either $f(P_{k+1})$ is distinct from these values or we may reindex the P_j so that $f(P_k) = f(P_{k+1})$. Now we can find some fractional linear transformation which moves the values $f(P_j)$ for $j = 1, \ldots, k$ all away from ∞, with the resulting composition still being meromorphic. Thus, assume that $f(P_j)$ is finite for each j. By Theorem 6, we have some $g \in K(X)$ with $g(P_k) \neq g(P_{k+1})$. Now consider

$$\varphi_c \;=\; f + cg$$

with $c \in \mathbb{C}$. Clearly, $\varphi_c(P_i) = \varphi_c(P_j)$ if and only if

$$c \;=\; \frac{f(P_i) - f(P_j)}{g(P_i) - g(P_j)}.$$

Since there are a finite number of pairs $P_i,\ P_j$, we can find a $c \in \mathbb{C}$ such that $\varphi_c(P_i) \neq \varphi_c(P_j)$ for $i \neq j$, as claimed.

As an example of working with function fields on a Riemann surface, we give an explicit description of the field $K(\mathbb{P}^1)$.

Theorem 7. *[4, pg. 11] $K(\mathbb{P}^1) \simeq \mathbb{C}(x)$.*

Proof. Any $f \in K(\mathbb{P}^1)$ is a holomorphic mapping $f : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$. Since $\mathbb{P}^1$ is compact, we see that f has at most a finite number of poles, for if not, then the set of all the poles of f would have a limit point in $\mathbb{P}^1$ and by Theorem 3, f would be identically ∞, and hence not a meromorphic function. Without loss of generality, suppose that ∞ is not a pole of f (otherwise replace f by $1/f$). Now let $p_1, \ldots, p_n$ denote the poles of f and let

$$R_j(z) \;=\; \sum_{i=-k_j}^{-1} a_{i,j}(z - p_j)^i$$

be the principal part of f at p_j. The function $f-(R_1+R_2+\ldots+R_n)$ is then a pole-free meromorphic function; that is, a holomorphic function on $\mathbb{P}^1$. From Lemma 1, it must be a constant. Since each R_j is a rational function of z, we conclude that f must be rational also. $\square$

1.4 Branched Covers of a Riemann Surface

Definition 3. Let X, Y be topological spaces. A map $p : X \longrightarrow Y$ is a *covering map* if for any $y \in Y$, there exists a neighborhood U of y such that

$$p^{-1}(U) = \coprod_{i=1}^{n} V_i$$

where each V_i is mapped homeomorphically by p to U and the V_i are disjoint. If $X,\ Y$ are Riemann surfaces, we require that the map p be holomorphic.

As it turns out, the requirement that *every* point of Y should have n distinct inverses in X is rather rigid, and excludes most of the holomorphic mappings between Riemann surfaces that interest us. For example, if $n > 1$ then $x \mapsto x^n$ is a holomorphic self-mapping of $\mathbb{P}^1$, but it is not a covering map in the above sense since $0,\ \infty$ have only one inverse image each. Moreover, Theorem 2 tells us that any holomorphic mapping between Riemann surfaces looks locally like x^n for some n. This leads naturally to the idea of *branch points* and *branched coverings.*

Definition 4. Let $X,\ Y$ be Riemann surfaces, $p : X \longrightarrow Y$ a holomorphic mapping, and P_0, Q_0 points of X, Y respectively with $Q_0 = p(P_0)$. By Theorem 2, there exist locally uniformizing variables t at P_0 and s at Q_0 with $s = t^n$ for some $n \geq 1$. If $n \neq 1$, the point $P_0 \in X$ is called a *branch point* and $Q_0 \in Y$ a *branch value.* The integer $n-1$ is called the *branch number* of p at P_0 and will be denoted $b_p(P_0)$.

We shall also refer to branch points as *ramification points*, and will often say that P_0 has *order* n over Q_0 or that Q_0 is *ramified* of order n. We can now modify our definition of a covering map of Riemann surfaces to include branch points:

Definition 5. Let $X,\ Y$ be Riemann surfaces and $p : X \longrightarrow Y$ a holomorphic mapping. If there is a discrete set $A \subset Y$ such that

$$p : p^{-1}(Y \setminus A) \longrightarrow Y \setminus A$$

is a covering map, then the map $p : X \longrightarrow Y$ is a *branched covering map.*

With this definition, we readily have:

Theorem 8. *Let X, Y be compact Riemann surfaces and $p : X \longrightarrow Y$ a non-constant holomorphic mapping. Then p is a branched covering map.*

Proof. Let p be as above. We have shown in the proof of Lemma 1 that any holomorphic mapping between compact Riemann surfaces is surjective. Moreover, since p is non-constant, we see by Theorem 3 that the critical points of p form a discrete and therefore finite set $B \subset X$. Thus, p is a local homeomorphism on $X \setminus B$. Moreover, the inverse image under p of any point in Y must be finite since it is discrete. Let $y \in Y$ and put $p^{-1}(y) = \{x_1, \ldots, x_n\}$. Let U_j be a neighborhood of x_j and V_j a neighborhood of y such that $p : U_j \longrightarrow V_j$ is biholomorphic and the U_j are disjoint. Now let $V \subset \bigcap_{j=1}^{n} V_j$ such that $p^{-1}(V) \subset \bigcup_{j=1}^{n} U_j$ and put $W_j = U_j \cap P^{-1}(V)$. Then the W_j are disjoint and $p : W_j \longrightarrow V$ is biholomorphic for each j. $\square$

Theorem 9. *Let X, Y be compact Riemann surfaces and $f : X \longrightarrow Y$ a non-constant holomorphic mapping. There exists a positive integer m such that every point of Y is taken precisely m times by f, counting multiplicities.*

Proof. For each integer $n \geq 1$, define

$$B_n = \left\{ Q \in Y : \sum_{P \in f^{-1}(Q)} (b_f(P) + 1) \geq n \right\}.$$

By Theorem 2, we see that B_n is open. We show that it is also closed. Let $Q_k \in B_n$ and suppose that $\lim_{k\to\infty} Q_k = Q$. The proof of Theorem 8 shows that the set of branch points of f is finite. Thus, without loss of generality we may assume that $b_f(Q_k) = 0$ for all k, i.e. that $f^{-1}(Q_k)$ has cardinality n or greater. Let $P_{1k}, P_{2k}, \ldots, P_{nk}$ be n distinct points of $f^{-1}(Q_k)$. Again, since X is compact, for each $j = 1, \ldots, n$, there exists a subsequence of $\{P_{jk}\}$ that converges to a limit P_j, and $f(P_j) = Q$. Of course, the P_j might not be distinct, but since $f(P_{jk}) = Q_k$ for all k, we must have

$$\sum_{P \in f^{-1}(Q)} (b_f(P) + 1) \geq n.$$

Therefore, since B_n is both open and closed for any $n \geq 1$, we see that B_n is either empty of all of X for each n. Now pick some $Q_0 \in Y$ and set $\sum_{P \in f^{-1}(Q_0)} (b_f(P) + 1) = m$. Then by compactness again, m is finite and we have $B_m = X$ and B_n empty for all $n > m$. $\square$

Definition 6. Let X, Y, f, and m be as above. Since f is a branched covering map by Theorem 8, we call the integer m the *degree* of the cover f, or simply the *degree* of f. We also refer to f as an *m-sheeted* branched cover of Y by X.

Geometrically, we view X as consisting of m copies of Y, each copy mapping biholomorphically (after the branch points have been excluded) to Y. The branch points are intuitively where the copies of Y ("sheets") are glued together to form X.

1.5 Deck Transformations

Definition 7. Let X, Y be Riemann surfaces and $p : X \longrightarrow Y$ a branched covering map. A *deck transformation* is a fiber preserving biholomorphic map, that is, a map f such that the diagram

$$\begin{array}{ccc} X & \xrightarrow{f} & X \\ {\scriptstyle p}\big\downarrow & & \big\downarrow{\scriptstyle p} \\ Y & = & Y \end{array}$$

commutes. It is not difficult to see that the set of all deck transformations of the cover p forms a group under composition. We denote this group $\mathrm{Deck}(X/Y)$.

Definition 8. Let X, Y be Riemann surfaces, $p : X \longrightarrow Y$ a branched covering map, and $A \subset Y$ the set of branch values of p. The covering is *normal* if the Deck group $\mathrm{Deck}(X/Y)$ acts transitively on the fiber $p^{-1}(Q)$ for all points $Q \in Y \setminus A$.

The notation $\mathrm{Deck}(X/Y)$ is suggestive. Namely, the group of Deck transformations of a cover is intuitively very much like the Galois group of a field extension since mappings $f \in \mathrm{Deck}(X/Y)$ are required to "fix" the base space, Y, while permuting the points in any given fiber $p^{-1}(Q)$. The analogy in fact turns out to be a correspondence.

1.6 The Main Correspondence

Given any branched covering map $p : X \longrightarrow Y$, any function $f \in K(Y)$ gives rise to a function $g \in K(X)$ via the *pullback* $p^*(f) = fp = g$. We in fact have an injection $p^* : K(Y) \longrightarrow K(X)$ so that we may view $K(Y)$ as a subfield of $K(X)$, and we will often exploit this fact.

Theorem 10 (Main Theorem). *Let X, Y be compact Riemann surfaces and $p : X \longrightarrow Y$ an n-sheeted branched covering map. Then $K(X)/p^*K(Y)$ is a degree n field extension. Conversely, let Y be a Riemann surface and $L/K(Y)$ a degree n field extension. Then there exists a Riemann surface X, an n-sheeted branched covering map $p : X \longrightarrow Y$, and $f \in K(X)$ such that $L \simeq K(X) = p^*K(Y)(f)$. In both cases, the Deck group $\mathrm{Deck}(X/Y)$ is isomorphic to $\mathrm{Aut}(K(X)/p^*K(Y))$.*

We will only prove the first assertion. For proofs of the rest, see [4].

Proof. Let $f \in K(X)$ and denote the set of branch values of p by A. We know that A is finite and that the inverse image under p of a point $Q \in Y \setminus A$ consists of n distinct points $P_1, \dots P_n$. Since p is a branched covering map, for each such Q we have a neighborhood U_Q of Q such that

$$p^{-1}(U_Q) \ = \ \coprod_{i=1}^{n} V_i$$

with V_i a neighborhood of P_i, the neighborhoods V_i, V_j disjoint for $i \neq j$, and $p : V_i \longrightarrow U_Q$ biholomorphic. Denote the restriction of p to V_i by p_i. Since p_i^{-1} is well defined on Y,

$$(p_i^{-1})^* f \ = \ f p_i^{-1}$$

is a meromorphic function on U_Q for each i. Therefore, the elementary symmetric functions

$$s_j \ = \sum_{1 \leq i_1 < \cdots < i_j \leq n} \prod_{k=1}^{j} (p_{i_k}^{-1})^* f$$

for $j = 1, 2, \ldots, n$ are meromorphic functions on U_Q. Repeathing this argument for each $Q \in Y \setminus A$, we see that these functions piece together to form meromorphic functions s_j defined on all of $Y \setminus A$. Call these functions the *elementary symmetric functions* of f with respect to the covering p. We now show that the elementary symmetric functions of f may be continued meromorphically to all of Y. Let $a \in A$. Since X is compact, $p^{-1}(a)$ is finite, say $p^{-1}(a) = \{b_1, \ldots, b_k\}$. Since f is meromorphic, it has isolated poles, so that we may take a neighborhood U of a so that the only possible poles of f in $V = p^{-1}(U)$ occur at the b_j. Let t be a locally uniformizing variable on U at a. Then $t(a) = 0$, so that the function $p^*t = tp \in \mathcal{O}(V)$ vanishes at each b_j. Since f is meromorphic, it has *finite* order at each of the b_j so that we can find an integer k such that

$$(p^*t)^k f$$

is holomorphic, and hence bounded, on V. Thus, if r_j are the elementary symmetric functions of $(p^*t)^k f$, then the r_j are bounded on $U \setminus \{a\}$. By Theorem 4, the r_j can be holomorphically extended to all of U. Since p^*t is a meromorphic function on $V \subset X$, the elementary symmetric functions of p^*t are just $\binom{n}{j} t^j$, so that we have

$$r_j \ = \ t^{kj} s_j$$

for each j. Since t is a meromorphic function on $U \subset Y$ and r_j may be holomorphically continued to all of U, we see that the s_j can be meromorphically continued to U for all j. This shows that the elementary symmetric functions of f with respect to p are in $K(X)$. Therefore, for every point $P \in X$, we have

$$f^n - p^* s_1 f^{n-1} + p^* s_2 f^{n-1} + \cdots + (-1)^n p^* s_n \ = \ 0.$$

This shows that the minimal polynomial of f over $K(Y)$ has degree at most n.

We showed after Theorem 6 that given n distinct points in X, there exists some $f \in K(X)$ that separates them. Thus, for a point $Q_0 \in Y$ with n distinct preimages $p^{-1}(Q_0) = \{P_1, \ldots, P_n\}$, we have a $f \in K(X)$ with $f(P_i) \neq f(P_j)$ for $i \neq j$. Moreover, since f is continuous, there exists a neighborhood U of Q_0 such that f takes on n distinct values for every $P \in p^{-1}(U)$. Let $m < n$ and suppose that the minimal polynomial of f over $K(Y)$ is

$$f^m + c_1 f^{m-1} + \cdots + c_m$$

where $c_j \in K(Y)$ for each j. Then the polynomial $f^m + c_1(Q)f^{m-1} + \cdots + c_m(Q) \in \mathbb{C}[f]$ has n distinct roots for *every* $Q \in U$, which implies that the c_j are identically zero on U, and by Theorem 3, identically zero on $K(Y)$. This is a contradiction, and the minimal polynomial of f has degree n.

Since $K(Y)$ is of characteristic zero, the primitive element theorem applies and $K(X)$ is generated as a field extension of $K(Y)$ by a single element. We claim that

$$K(X) \;=\; K(Y)(f). \tag{1.3}$$

To see this, notice that since the minimal polynomial of f over $K(Y)$ has degree n, we have $[K(Y)(f) : K(Y)] = n$. Now let $g \in K(X)$. Again, by the primitive element theorem we have $K(Y)(f,g) = K(Y)(h)$ for some $h \in K(X)$. But we have shown that the minimal polynomial of h has degree at most n. We then have

$$n = [K(Y)(f) : K(Y)] \leq [K(Y)(f,g) : K(Y)] \leq n,$$

from which we conclude that $g \in K(Y)(f)$ already. This gives our claim and completes the proof. $\square$

Theorem 10 is a very deep result. From it we obtain

Theorem 11. *Let X, Y be compact Riemann surfaces. Then $X \simeq Y$ if and only if $K(X) \simeq K(Y)$.*

In the particular case of branched covering maps $p : X \longrightarrow \mathbb{P}^1$ from a compact Riemann surface to the Riemann Sphere, we see by 1.3 and Theorem 7 that

$$K(X) \;=\; K(\mathbb{P}^1)(f) \;\simeq\; \mathbb{C}(x,s) \tag{1.4}$$

with

$$F(x,s) \;=\; 0, \tag{1.5}$$

where $f \in K(X)$ and F is a rational function of two variables and degree $[K(X) : p^*K(Y)]$.

Chapter 2

The Group $\mathbf{SL}_2(\mathbb{Z})$ and its Subgroups

2.1 Definitions and Properties

Given a ring R, we let $\mathbf{SL}_2(R)$ denotes the group of two by two matrices with entries in R of determinant 1 in R. Now $\mathbf{SL}_2(\mathbb{Z})$ has a natural action on $\mathbb{C}^2$ which descends to $(\mathbb{C}^2 \setminus \{0\})/\mathbb{C}^*$. Since $(\mathbb{C}^2 \setminus \{0\})/\mathbb{C}^* \simeq \mathbb{P}^1$, we see that $\mathbf{SL}_2(\mathbb{Z})$ acts on $\mathbb{P}^1$, and the resulting action is called *fractional linear transformation.* Explicitly, write $z \in \mathbb{P}^1$ as x/y for $(x, y) \in \mathbb{C}^2 \setminus \{0\}$. For any $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbb{Z})$ we have

$$\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix},$$

which gives

$$\begin{aligned} \alpha z &= \frac{ax + by}{cx + dy} \\ &= \frac{az + b}{cz + d}. \end{aligned}$$

This action in fact descends to $\mathbb{H}$, which is verified by the formula

$$\Im(\alpha(z)) = \frac{\Im(z)}{|cz + d|^2}, \tag{2.1}$$

where α is as before. In fact, since

$$\frac{-az - b}{-cz - d} = \frac{az + b}{cz + d},$$

we see that $\mathbf{SL}_2(\mathbb{Z})/\pm 1 = \mathbf{PSL}_2(\mathbb{Z})$ acts on $\mathbb{H}$, where 1 denotes the identity matrix. Given any subgroup $\Gamma \subset \mathbf{SL}_2(\mathbb{Z})$, we will denote by $\bar{\Gamma}$ the image of Γ in $\mathbf{SL}_2(\mathbb{Z})/\{\pm 1\}$. We now define an important class of normal subgroups of $\mathbf{SL}_2(\mathbb{Z})$.

Definition 9. For any positive integer N define $\Gamma(N)$ to be the subgroup of $\mathbf{SL}_2(\mathbb{Z})$ consisting of those matrices that are congruent modulo N to the identity. That is,

$$\Gamma(N) \;=\; \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}) \,:\, a \equiv d \equiv 1,\; b \equiv c \equiv 0 \mod N \right\}.$$

We call $\Gamma(N)$ the *principal congruence subgroup of level* N. Notice that that $\Gamma(N)$ acts on $\mathbb{H}$ since it is a subgroup of $\Gamma(1) = \mathbf{SL}_2(\mathbb{Z})$.

2.2 The Structure of $\Gamma(N)$

Let $r_N : \mathbf{SL}_2(\mathbb{Z}) \longrightarrow \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$ denote the reduction modulo N map. Clearly, $\Gamma(N)$ is the kernel of r_N and is hence a normal subgroup of $\Gamma(1)$. More is true, however, as the following theorem makes evident:

Theorem 12. *[8, pg. 61] The sequence*

$$1 \longrightarrow \Gamma(N) \longrightarrow \mathbf{SL}_2(\mathbb{Z}) \xrightarrow{r_N} \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z}) \longrightarrow 1,$$

where the first two maps are the obvious inclusions and the last map is trivial, is exact.

Proof. The only thing that is not obvious is the surjectivity of r_N. Let

$$\alpha \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be an element of $\mathbf{GL}_2(\mathbb{Z})$ with $ad - bc \equiv 1 \mod N$. Recall [8, pg. 61] that there exist $\gamma,\ \delta \in \mathbf{PSL}_2(\mathbb{Z})$ such that $\gamma\alpha\delta$ is diagonal. We may therefore assume that α is diagonal since r_N is a homomorphism. Put

$$\alpha \;=\; \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$$

with $ad \equiv 1 \mod N$. Since d has an inverse modulo N, we have $(d, N) = 1$, and hence there exist integers $u,\ v$ such that

$$ud + vN = 1.$$

Let $ad = 1 + rN$ for some integer r. Then the matrix

$$\beta \;=\; \begin{pmatrix} a - ruN & rvN \\ N & d \end{pmatrix}$$

satisfies $r_N(\beta) = \alpha$ and has determinant 1 since $ad - ruNd - rvN^2 = 1 + rN(1 - ud - vN) = 1$. Thus r_N is surjective. $\square$

We can now determine the structure of $\Gamma(1)/\Gamma(N)$.

Corollary 1. *For any positive integer N, we have $\Gamma(1)/\Gamma(N) \simeq \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$.*

Proof. This follows from an application of the first isomorphism theorem to the exact sequence given in Theorem 12. $\square$

2.3 Fundamental Domains

Since $\Gamma(N)$ acts on $\mathbb{H}$, we may consider the quotient space $\mathbb{H}/\Gamma(N)$ consisting of the $\Gamma(N)$ orbits on $\mathbb{H}$.

Definition 10. Given a subgroup Γ of $\Gamma(1)$, a *fundamental domain* for Γ is a connected open subset $F \subset \mathbb{H}$ such that any point z of $\mathbb{H}$ is equivalent modulo Γ to some point of the closure of F and no two points of F are equivalent under Γ.

As an example, we determine a fundamental domain for $\Gamma(1)$.

Theorem 13. *A fundamental domain F for $\Gamma(1)$ is the open set in $\mathbb{H}$ bounded by the lines $\Re(z) = -\frac{1}{2}$, $\Re(z) = \frac{1}{2}$ and the unit circle $\{z : |z| = 1\}$. Moreover,*

$$T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \textit{and} \quad S := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

generate $\Gamma(1)$.

Proof. (Adapted from [12, pg. 16] and [8, pg. 30]) Let Γ' be the subgroup of $\Gamma(1)$ generated by S, T. First we show that every $z \in \mathbb{H}$ is equivalent under Γ' to some z' in the closure of F. Let $z \in \mathbb{H}$ and put $\sigma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma'$. Since the set

$$\{cz + d : c,\ d \in \mathbb{Z}\}$$

is a lattice in $\mathbb{C}$, the quantity $|cz + d|$ for $(c, d) \in \mathbb{Z}^2 \setminus \{0\}$ is bounded below. By 2.1, the set $\{\Im(\sigma(z)) : \sigma \in \Gamma'\}$ is bounded above, say by $\sigma_0(z) = w$. Since $\sigma_0,\ S \in \Gamma'$, we have

$$\Im(S\sigma_0(z)) = \Im\left(\frac{-1}{w}\right) = \frac{\Im(w)}{|w|^2} \leq \Im(w),$$

so that $|w| \geq 1$. Now there exists some $n \in \mathbb{Z}$ such that

$$-\frac{1}{2} \leq \Re(T^n w) \leq \frac{1}{2}.$$

and obviously $\Im(T^n w) = \Im(w)$, so that $z' = T^n\sigma_0(z)$ lies in the closure of F and is Γ' equivalent to z. Since $\Gamma' \subset \Gamma(1)$, we trivially have that every point of $\mathbb{H}$ is equivalent under $\Gamma(1)$ to a point in F.

Now we show that no two elements of F are equivalent under $\Gamma(1)$ (and hence under Γ'). Let $z,\ z'$ be distinct points of F and suppose that there exists $\sigma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma(1)$ with $z' = \sigma(z)$. Without loss of generality, assume that $\Im(z) \leq \Im(z') = \Im(z)/|cz+d|^2$. We therefore have

$$|c|\,\Im(z) \;\leq\; |cz+d| \;\leq\; 1.$$

Notice that if $c = 0$ then $d = \pm 1$ which forces $a = d$ and $z' = z \pm b$ which is impossible since b is a nonzero integer. Hence, $c \neq 0$. From the definition of F, we see that $\Im(z) \geq \sqrt{3}/2$. Therefore, $|c|\sqrt{3}/2 \leq 1$ so that $c = \pm 1$ and $|z+d| \leq 1$. Again from the definition of F we see that if $|d| \geq 1$ then $|z+d| > 1$ for any $z \in F$, so that $d = 0$ and $|z| \leq 1$. This contradicts $z \in F$. In fact, we have shown that if $z,\ z'$ in the closure of F are $\Gamma(1)$ equivalent, then $z' = T^{\pm 1}z$ or $z' = Sz$. In both cases, $z,\ z'$ lie on the boundary of F and are equivalent under Γ'.

We have shown that F is a fundamental domain for both Γ' and $\Gamma(1)$. Together with the fact that $\Gamma' \subset \Gamma(1)$, this implies that $\Gamma' = \Gamma(1)$. Indeed, let $\alpha \in \Gamma(1)$ and $z \in F$. Since F is a fundamental domain for Γ', there exists $\beta \in \Gamma'$ with $w = \beta\alpha(z)$ in the closure of F. In fact, we have shown above that w must be in F and hence $w = z$ (since F is a fundamental domain for $\Gamma(1)$). Therefore, $\beta\alpha = 1$ so that $\alpha \in \Gamma'$. $\square$

By Theorem 13, we see that

$$F \cup \{z \in \mathbb{H} : \Re(z) = -1/2,\ |z| \geq 1\} \cup \{z \in \mathbb{H} : -1/2 \leq \Re(z) \leq 0,\ |z| = 1\}$$

is a set of representatives for $\mathbb{H}/\Gamma(1)$. Similarly, one can find an explicit set of representatives for $\mathbb{H}/\Gamma(N)$. We would like to put a topology on the resulting set so that the space is compact. To do this, however, we need to add some points. The situation is completely analogous to that of the one point compactification of $\mathbb{C}$ by adding the point at infinity.

2.4 Cusps

It is obvious that any subgroup Γ of $\Gamma(1)$ acts on $\mathbb{Q} \cup \{\infty\} = \mathbb{P}^1(\mathbb{Q})$ by the same formula that gives its action on $\mathbb{H}$.

Definition 11. The *cusps* of a subgroup Γ of $\Gamma(1)$ are the Γ orbits of $\mathbb{P}^1(\mathbb{Q})$. We will denote by C_Γ any complete set of representatives of cusps of Γ. By abuse of terminology, will often refer to a single point as a cusp.

The following theorem describes the set $C_{\Gamma(N)}$ for each positive integer N.

Theorem 14. *[12, pg. 23] With the convention that $\infty = \pm 1/0$, two points $a/b,\ c/d$ of $\mathbb{Q}$ with $(a,b) = (c,d) = 1$ are equivalent under $\Gamma(N)$ iff $\pm\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \equiv \left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right] \mod N$.*

Proof. In one direction, suppose that $\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \equiv \left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right] \mod N$. Since $(c, d) = 1$, we have integers r, s such that $rc - ds = 1$. Then $\tau = \left[\begin{smallmatrix} c & s \\ d & r \end{smallmatrix}\right] \in \Gamma(1)$ satisfies $\tau\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] = \left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right]$, so that $\tau^{-1}\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \equiv \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] \mod N$. Thus, if we can find some $\sigma \in \Gamma(N)$ with $\sigma\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right] = \tau^{-1}\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$, then we have $\tau\sigma\tau^{-1}\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right] = \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$. Since $\Gamma(N)$ is normal in $\Gamma(1)$, this shows that a/b and c/d are equivalent under $\Gamma(N)$. It therefore remains to show the result when $\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right] = \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. Since $(a, b) = (a, N) = 1$, there exist integers p, q so that $ap + bq = (1 - a)/N$. Now let $\sigma = \left[\begin{smallmatrix} a & -Nq \\ b & 1+Np \end{smallmatrix}\right]$. Then clearly $\sigma \in \Gamma(N)$ and $\sigma\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right] = \left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right]$. This shows in any case that a/b and c/d are equivalent under $\Gamma(N)$. Conversely, suppose that there exists $\sigma = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right] \in \Gamma(N)$ with $a/b = \sigma(c/d)$. Then $a/b = (cp + dq)/(rp + sq)$, so that there exists some $\lambda = m/n \in \mathbb{Q}$ with $(m, n) = 1$ and $\lambda\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] = \left[\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right]\left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right]$. Equivalently, we have $ma = n(pc + qd)$ and $mb = n(rc + sd)$, from which we conclude (since $(m, n) = 1$) that $n|a$ and $n|b$. But $(a, b) = 1$ so that $n = \pm 1$. Similarly, since σ has determinant 1 and $(c, d) = 1$, we have $m = \pm 1$. Therefore, since $\sigma \in \Gamma(N)$ we have $\pm\left[\begin{smallmatrix} a \\ b \end{smallmatrix}\right] \equiv \left[\begin{smallmatrix} c \\ d \end{smallmatrix}\right] \mod N$. □

Using this theorem, we easily see that $C_{\Gamma(1)} = \{\infty\}$.

Definition 12. We shall let $\mathbb{H}^*$ denote the *extended complex plane*, that is, the subset of $\mathbb{P}^1$ given by $\mathbb{H}^* = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$.

From our remarks above, $\Gamma(N)$ acts on $\mathbb{H}^*$ for any N, so that the quotient $\mathbb{H}^*/\Gamma(N)$ makes sense. Note that since $\Gamma(N)$ preserves $\mathbb{P}^1(\mathbb{Q})$, we alwsys have $\mathbb{H}^*/\Gamma(N) = \mathbb{H}/\Gamma(N) \cup C_{\Gamma(N)}$. We now specify a topology on $\mathbb{H}^*/\Gamma(N)$ so that the resulting space is compact.

2.5 Topology

First, we topologize $\mathbb{H}^*$ by specifying a fundamental system of open neighborhoods of any point $z \in \mathbb{H}^*$. If $z \in \mathbb{H}$, then a fundamental system of open neighborhoods of z is just the usual one under the standard topology on $\mathbb{C}$. If $z \neq \infty$ is in $\mathbb{Q}$, we take as a fundamental system of open neighborhoods all sets of the form $\{z\} \cup S_z$, where S_z is a circle of radius $r > 0$ centered at $z + ir$. Finally, as a fundamental system of open neighborhoods of ∞, we take all sets of the form $\{\infty\} \cup \{z \in \mathbb{H} : \Im(z) > c\}$, where $c > 0$ is constant. Clearly, this specifies a topology on $\mathbb{H}^*$. Moreover, in this topology, every $\sigma \in \Gamma(N)$ is a homeomorphism $\sigma : \mathbb{H}^* \longrightarrow \mathbb{H}^*$.

We now endow $\mathbb{H}^*/\Gamma(N)$ with the quotient topology. That is, if $\pi : H^* \longrightarrow H^*/\Gamma(N)$ is the quotient map, then a set $X \subset \mathbb{H}^*/\Gamma(N)$ is open precisly when $\pi^{-1}(X)$ is open in $\mathbb{H}^*$. It can be shown [12, pg. 12] that $\mathbb{H}^*/\Gamma(N)$ with the above topology is Hausdorff and locally compact, though we will not prove these facts. From our discussion of a fundamental domain for $\mathbb{H}^*/\Gamma(1)$ above, it is not difficult to prove that $\mathbb{H}^*/\Gamma(1)$ is compact. Indeed, $\mathbb{H}^*/\Gamma(1) = (\mathbb{H}/\Gamma(1)) \cup \{\infty\}$ is the one-point compactification of $\mathbb{H}/\Gamma(1)$. Using this fact, we readily show that $\mathbb{H}^*/\Gamma(N)$ is compact for each N.

Theorem 15. *For every N, the topological space $\mathbb{H}^*/\Gamma(N)$ is compact.*

Proof. Since $\Gamma(N)$ is of finite index in $\Gamma(1)$, we may write

$$\Gamma(1) \;=\; \bigcup_{i=1}^{m} g_i\Gamma(N),$$

where $g_i \in \Gamma(1)$ and $m = [\Gamma(1) : \Gamma(N)]$. Therefore, we have

$$\mathbb{H}^*/\Gamma(N) \;=\; \bigcup_{i=1}^{m} g_i(\mathbb{H}^*/\Gamma(1)),$$

which is a finite union of compact spaces, and hence compact. □

2.6 Fixed Points

Now that we have made $\mathbb{H}^*/\Gamma(N)$ into a compact topological space, we would like to make it into a compact Riemann surface. Before we can endow $\mathbb{H}^*/\Gamma(N)$ with a smooth structure, however, we need to understand the points of $\mathbb{H}$ that have nontrivial $\Gamma(N)$ stabiliziers.

Definition 13. By the *isotropy subgroup* of a group Γ acting on a topological space S at a point $z \in S$ we mean the stabilizer of z in Γ, that is, the group $\Gamma_z = \{\sigma \in \Gamma : \sigma(z) = z\}$.

Theorem 16. *Every point $z \in \mathbb{H}$ with nontrivial stabilizer $\bar{\Gamma}(1)_z \subset \bar{\Gamma}(1)$ is equivalent under $\Gamma(1)$ to i or $\omega = e^{2\pi i/3}$. Every $s \in \mathbb{Q}$ is $\Gamma(1)$ equivalent to ∞. Moreover, the isotropy subgroups corresponding to i and ω are $\langle S\rangle$, $\langle ST\rangle$ respectively, and the isotropy subgroup of $\bar{\Gamma}(1)$ at ∞ is generated by T.*

Proof. Suppose that $z \in \mathbb{H}$ is fixed by some nontrivial $\alpha \in \bar{\Gamma}(1)$. By Theorem 13 there exists some $\sigma \in \Gamma(1)$ so that $\sigma(z) \in \bar{F}$. We then have $\sigma\alpha(z) = \sigma(z)$, so that $\sigma\alpha\sigma^{-1}$ fixes $\sigma(z) \in \bar{F}$. Thus, it is enough to consider the fixed points of $\bar{F}$. Let $\gamma = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \Gamma(1)$ fix $z \in \bar{F}$. Then from 2.1, we see that $|c|\,|z| \le |cz+d| = 1$. If $c = 0$ then $a = d = \pm 1$ and $z \in \mathbb{H}$ is not fixed by γ. Hence $c \ne 0$. Since $z \in \bar{F}$ we have $|z| \ge 1$ so that $|c| = |z| = 1$. If $d \ge 1$ then $|cz+d| > 1$ unless $d = 1$ and $z = \omega$ or $d = -1$ and $z = \omega + 1$. If $d = 0$ then $b = \mp c = \pm 1$. Since in this case we have $z = \pm a - 1/z$, we conclude $a = 1$ or 0 from the fact that $|z| = 1$. Thus we see that $z = i, \omega$ or $1 + \omega$ and that the isotropy subgroups are

$$\begin{aligned}
\bar{\Gamma}(1)_i \;&=\; \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle \;=\; \langle S\rangle, \quad \text{and} \\
\bar{\Gamma}(1)_\omega \;&=\; \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \right\rangle \;=\; \langle ST\rangle.
\end{aligned}$$

Now by Theorem 14, every $s \in \mathbb{Q}$ is $\Gamma(1)$ equivalent to ∞. Moreover if $z \mapsto (az+b)(cz+d)$ stabilizes ∞, we must have $c = 0$. This forces $a = d = \pm 1$, so that every $\gamma \in \bar{\Gamma}(1)$ fixing ∞ has the form $\gamma = \left(\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right)$ for some integer b. Since $T^b = \gamma$ and T also fixes ∞, we see that T generates $\bar{\Gamma}(1)_\infty$, as claimed. □

2.7 The Riemann Surface $\mathbb{H}^*/\Gamma(N)$

We can now specify a smooth structure on $\mathbb{H}^*/\Gamma(N)$. We will need the following useful lemma, which we do not prove:

Lemma 2. *Let $z \in \mathbb{H}^*$, and as before let $\Gamma(N)_z$ be the isotropy subgroup of $\Gamma(N)$ at z. Then there exists an open neighborhood U of z with*

$$\Gamma(N)_z \;=\; \{\gamma \in \Gamma(N) \,:\, \gamma(U) \cap U \neq \emptyset\}.$$

For a proof of this lemma, see [12, pg. 17]. Now let $z, \Gamma(N)_z, U$ be as in Lemma 2, and denote by π the natural projection (i.e. the quotient map)

$$\pi_N \,:\, \mathbb{H}^* \longrightarrow \mathbb{H}^*/\Gamma(N).$$

We know by Theorem 16 that $\bar{\Gamma}(N)_z$ is a finite cyclic subgroup of $\Gamma(N)$ if $z \in \mathbb{H}$, and that if z is a cusp, then there exists $g \in \Gamma(1)$ with $gz = \infty$ and hence $g\bar{\Gamma}(N)_z g^{-1} = \langle T^k \rangle$ for some $k \geq 1$. From the definition of U, it is easy to see that the natural map (given by inclusion)

$$U/\Gamma(N)_z \longrightarrow \mathbb{H}^*/\Gamma(N)$$

is injective. Moreover, $U/\Gamma(N)_z$ is an open neighborhood of $\pi_N(z)$ in $\mathbb{H}^*/\Gamma(N)$.

1. If $\bar{\Gamma}(N)_z$ is trivial, then $\pi_N \,:\, U \longrightarrow U/\Gamma(N)_z$ is a homeomorphism. We therefore use $U/\Gamma(N)_z$ as an open neighborhood of z and π_N^{-1} as a locally uniformizing variable at $\pi_N(z)$.

2. If $\bar{\Gamma}(N)_z$ is cyclic of order $n > 1$, (where in the cases that we are dealing with, $n = 2$ or 3), then let $\lambda \,:\, \mathbb{H} \longrightarrow \Delta$ be a biholomorphic mapping of the upper half plane to the unit disc with $\lambda(z) = 0$. Notice that $\lambda\bar{\Gamma}(N)_z\lambda^{-1}$ is a cyclic group of automorphisms of the disc preserving 0. As such, $\lambda\bar{\Gamma}(N)_z\lambda^{-1}$ consists of the mappings $w \longrightarrow \zeta_n^k w$, where $\zeta_n = e^{2\pi i/n}$ is a primitive n^{th} root of unity. Therefore, the map $p \,:\, U/\Gamma(N)_z \longrightarrow \mathbb{C}$ given by

$$p(\pi_N(z)) \;=\; \lambda^n(z)$$

is a homeomorphism of $U/\Gamma(N)_z$ with an open subset of $\mathbb{C}$, so we have an open neighborhood $U/\Gamma(N)_z$ and a locally uniformizing variable p for the point $\pi_N(z)$.

3. Finally, if z is a cusp, then we have seen in Theorem 16 that $\bar{\Gamma}(N)_z$ is $\Gamma(1)$ conjugate to $\bar{\Gamma}(N)_\infty$ and that $\bar{\Gamma}(N)_\infty$ as a subgroup of $\bar{\Gamma}(N)$ is generated by some power T^k. (It is, in fact, not difficult to see that $k = N$ since N is the smallest power of T such that $T^N \in \bar{\Gamma}(N)$). We therefore reduce everything to the case $z = \infty$ by letting $g \in \Gamma(1)$ be some element taking z to ∞. Now the map $p \,:\, U/\Gamma(N)_z \longrightarrow \mathbb{C}$ given by

$$p(\pi_N(z)) \;=\; e^{2\pi i g(z)/k} \tag{2.2}$$

certainly maps $U/\Gamma(N)_z$ onto an open neighborhood of $\mathbb{C}$. That it is an injection follows from the fact that $e^{2\pi i g(z_1)/k} = e^{2\pi i g(z_2)/k}$ if and only if $(T^k)^m(g(z_1)) = g(z_2)$ for some m, that is, if and only if $\pi_N(g(z_1)) = \pi_N(g(z_2))$. We therefore have an open neighborhood and a uniformizing variable corresponding to the cusp z.

It may readily be checked that the complex charts specified above are compatible, and that we have thus defined a Riemann surface.

Definition 14. We define the *modular curve* $X(N)$ to be the Riemann surface $\mathbb{H}^*/\Gamma(N)$.

2.8 The Natural Map $\mathbb{H}^*/\Gamma(N) \longrightarrow \mathbb{H}^*/\Gamma(1)$

Consider the diagram

$$\begin{array}{ccc} \mathbb{H}^* & = & \mathbb{H}^* \\ {\scriptstyle \pi_N}\downarrow & & \downarrow{\scriptstyle \pi_1} \\ X(N) & \xrightarrow{f} & X(1) \end{array}$$

where π_N, π_1 are the natural quotient maps and f is the natural map which makes the diagram commute. That is, for any point $\pi_N(z) \in X(N)$, the value $f(\pi_N(z))$ is defined to be $\pi_1(z)$. Using the complex structures on $X(N)$ and $X(1)$, it can be shown that f is in fact a holomorphic mapping of Riemann surfaces and is the quotient mapping

$$f : X(N) \longrightarrow X(N)/G, \tag{2.3}$$

where $G = \bar{\Gamma}(1)/\bar{\Gamma}(N)$. Theorem 8 tells us that f is a branched covering map. Moreover, we see that f is of degree $[\bar{\Gamma}(1) : \bar{\Gamma}(N)]$ since every point in $X(1)$ has $[\bar{\Gamma}(1) : \bar{\Gamma}(N)]$ inverse images under f (counting multiplicities). We are now in precisely the situation analyzed in Chapter 1, and we shall use the techniques developed there to study the function fields $K(X(N))$ for certain value of N. First, however, we determine the genus of the surface $X(N)$.

2.9 Genus

Let X, Y be compact Riemann surfaces of genus g, g' respectively, and $p : X \longrightarrow Y$ a degree n branched covering map. Let $b_p(P)$ denote the branch number of p at $P \in X$. Then the *Hurwitz formula* [12, pg. 19] tells us that

$$2g' - 2 = n(2g - 2) + \sum_{P \in X} b_p(P). \tag{2.4}$$

We have shown in the proof of Theorem 16 that the branch points of the covering $f : X(N) \longrightarrow X(1)$ are all equivalent under $\Gamma(1)$ to one of $\omega = e^{2\pi i/3}$, i, ∞. Moreover, it is not difficult to determine the branch number for each point. Explicitly, since $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $ST = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$, we see that $S, ST \notin \Gamma(N)$ for any $N > 1$. Therefore, $S, ST \in \bar{\Gamma}(1)/\bar{\Gamma}(N) = G$, so that by 2.3 we have:

1. Every point $P \in X(N)$ that is equivalent under $\Gamma(1)$ to ω has $b_f(P) = 2$.
2. Every point $P \in X(N)$ that is equivalent under $\Gamma(1)$ to i has $b_f(P) = 1$.

Now let $\mu_N = [\bar{\Gamma}(1) : \bar{\Gamma}(N)]$ be the degree of the branched covering map f. Then we clearly have

1. The number of *distinct* points $P \in X(N)$ that are equivalent under $\Gamma(1)$ to ω is $\mu_N/|\bar{\Gamma}(N)_\omega| = \mu_N/3$.
2. The number of *distinct* points $P \in X(N)$ that are equivalent under $\Gamma(1)$ to i is $\mu_N/|\bar{\Gamma}(N)_i| = \mu_N/2$.

Finally, we must compute the branch number at infinity and the size of $C_{\Gamma(N)}$. Notice that $T^N = \left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right)$ is the smallest power of T contained in $\bar{\Gamma}(N)$. Therefore, G_∞ as a subgroup of G is cyclic of order N so that $X(N)$ has precisely μ_N/N inequivalent cusps, that is, $\left|C_{\Gamma(N)}\right| = \mu_N/N$. Finally, since each cusp has stabilizer conjugate to G_∞, we see that $b_f(s) = N - 1$ for any cusp s. Let g_N denote the genus of the Riemann surface $X(N)$. Then putting all of our information together and using 2.4, we find

$$\begin{aligned} 2g_N - 2 &= \mu_N(2g_1 - 2) + 2\mu_N/3 + \mu_N/2 + (N-1)\mu_N/N \\ &= \mu_N\left(2g_1 + \frac{N-6}{6N}\right), \end{aligned}$$

so that

$$g_N = 1 + \mu_N\left(g_1 + \frac{N-6}{12N}\right). \tag{2.5}$$

In the next chapter, we will show that $g_1 = 0$.

Chapter 3

Modular Functions

3.1 Definitions

Definition 15. By a *modular function* of level N we shall mean a meromorphic function on the modular curve $X(N)$.

Clearly, any modular function f of level N may be extended to a meromorphic function $\tilde{f}$ on $\mathbb{H}^*$ by the pullback $\tilde{f} = \pi_N^* f$. We then see that $\tilde{f}$ is a meromorphic function on $\mathbb{H}^*$, invariant under the action $g\tilde{f}(z) = \tilde{f}(g\tau)$ for all $g \in \Gamma(N)$ and $\tau \in \mathbb{H}^*$. Such functions may be viewed as meromorphic functions on $\mathbb{H}$ invariant under $\Gamma(N)$ with the following additional property: For each $g \in \Gamma(1)$, the function $g\tilde{f}(\tau) = \tilde{f}(g\tau)$ admits a laurent expansion in the variable $q^{1/N} := e^{2\pi i\tau/N}$ with only finitely many negative powers of $q^{1/N}$. That $\tilde{f}$ admits such an expansion may be seen as follows:

1. $\tilde{f}$ is invariant under the action of $\Gamma(N)$, and in particular the transformation T^N, so that by standard results from fourier analysis, $f(\tau)$ has such an expansion for all $\tau \in \mathbb{H}$.

2. At the cusps, 2.2 tells us that such an expansion exists since $e^{2\pi i h(z)/N}$ is a uniformizing variable at z where h takes z to ∞.

3. The expansion has only finitely many negative powers of $q^{1/N}$ because $\tilde{f}$ is meromorphic on $\mathbb{H}^*$.

3.2 The Field Extension $K(X(N))/K(X(1))$

We have seen in section 2.8 that the natural map $f : X(N) \longrightarrow X(1)$ is a branched covering map of degree $[\bar{\Gamma}(1) : \bar{\Gamma}(N)]$, and therefore that $K(X(N))/K(X(1))$ is degree $[\bar{\Gamma}(1) : \bar{\Gamma}(N)]$ field extension. In fact, we have

Theorem 17. *The extension $K(X(N))/K(X(1))$ is Galois, with Galois group $\bar{\Gamma}(1)/\bar{\Gamma}(N)$.*

Proof. Let $G = \bar{\Gamma}(1)/\bar{\Gamma}(N)$. By 2.3, we have $X(1) = X(N)/G$. We show that G injects into $\mathrm{Aut}(K(X(N))/K(X(1)))$. Let $g, h \in G$ be distinct. Then there exists some point $z \in X(N)$ with $gz \neq hz$. By Theorem 6, there exists $f \in K(X(N))$ with $f(gz) \neq f(hz)$, that is, $gf \neq hf$. This completes the proof. □

3.3 Elliptic Functions

Let $\omega_1, \omega_2 \in \mathbb{C}$ be such that $\Im(\omega_1/\omega_2) > 0$, fix $L \subset \mathbb{C}^2$ to be the lattice generated by ω_1, ω_2, and put $L' = L \setminus \{0\}$ Recall that the *Weierstrass* function defined by

$$\wp(z, L) := \frac{1}{z^2} + \sum_{\omega \in L'} \left\{ \frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right\} \tag{3.1}$$

is a meromorphic doubly periodic function of z with periods ω_1 and ω_2. The function $\wp(z, L)$ admits the Laurent series expansion

$$\wp(z, L) = \frac{1}{z^2} + \frac{1}{20} g_2(L) z^2 + \frac{1}{28} g_3(L) z^4 + \cdots,$$

where

$$\begin{aligned} g_2(L) &:= 60 \sum_{\omega \in L'} \frac{1}{\omega^4} \quad \text{and} \\ g_3(L) &:= 140 \sum_{\omega \in L'} \frac{1}{\omega^6} \end{aligned}$$

are the *Eisenstein series* of weights 4 and 6 [8, pg. 10]. Obviously, $g_2(L)$ and $g_3(L)$ satisfy

$$\begin{aligned} g_2(\lambda L) &= \lambda^{-4} g_2(L) \\ g_3(\lambda L) &= \lambda^{-6} g_3(L) \end{aligned}$$

for any $\lambda \in \mathbb{C}^*$; that is, they are *homogenous* of degrees -4 and -6, respectively. The *discriminant*

$$\Delta(L) = g_2^3(L) - 27 g_3^2(L),$$

so named because it *is* the discriminant of the cubic polynomial $y^2 = 4x^3 - g_2 x - g_3$ satisfied by $(x, y) = (\wp(z, L), \wp'(z, L))$, is therefore homogenous of degree -12. It is furthermore true [8, pg. 11] that $\Delta(L) \neq 0$ for any lattice $L \subset \mathbb{C}^2$. This fact will be crucial when we define the modular function J. More than this is true, however. When L is the lattice generated by 1 and $\tau \in \mathbb{H}$, we in fact have the product expansion [8, pg. 249]

$$\Delta(L) \;=\; (2\pi i)^{12} q \prod_{n=1}^{\infty} (1-q^n)^{24}, \tag{3.2}$$

where $q = e^{2\pi i \tau}$.

Recall that the *Weierstrass sigma* and *zeta* functions are defined as [8, pg. 239]

$$\sigma(z, L) \;:=\; z \prod_{\omega \in L'} \left(1 - \frac{z}{\omega}\right) e^{z/\omega + \frac{1}{2}(z/\omega)^2} \tag{3.3}$$

and

$$\zeta(z, L) \;:=\; \frac{\sigma'(z, L)}{\sigma(z, L)}. \tag{3.4}$$

Taking the logarithm of the product 3.3 and differentiating twice with respect to z, it may be seen using 3.1 and 3.4 that

$$\zeta'(z, L) \;=\; -\wp(z, L).$$

It follows that

$$\zeta(z + \omega, L) \;=\; \zeta(z, L) + \eta(\omega, L),$$

for some constant $\eta(\omega, L)$ and any $\omega \in L$. In fact, $\eta(\omega, L)$ extends to a function $\eta(z, L)$ which is $\mathbb{R}$-linear in z [7, pg. 27]. Both $\zeta(z, L)$ and $\eta(z, L)$ are homogenous of degree -1, that is

$$\begin{aligned} \zeta(\lambda z, \lambda L) &= \lambda^{-1} \zeta(z, L), \\ \eta(\lambda z, \lambda L) &= \lambda^{-1} \eta(z, L), \end{aligned} \tag{3.5}$$

while $\sigma(z, L)$ is homogenous of degree 1:

$$\sigma(\lambda z, \lambda L) \;=\; \lambda \sigma(z, L), \tag{3.6}$$

for any $\lambda \in \mathbb{C}^*$. Furthermore, the sigma function satisfies [7, pg. 28]

$$\sigma(z + \omega, L) \;=\; \varepsilon(\omega) e^{\eta(\omega, L)(z + \omega/2)} \sigma(z, L), \tag{3.7}$$

where $\omega \in L$ is arbitrary and $\varepsilon(\omega)$ is defined by

$$\varepsilon(\omega) \;:=\; \begin{cases} 1 & \text{if} \quad \omega \in 2L \\ -1 & \text{otherwise} \end{cases}.$$

Since we will work exclusively with the lattice L generated by $1, \tau$ with $\tau \in \mathbb{H}$, set $W = \binom{\omega_1}{\omega_2} = \binom{\tau}{1}$. Further, let $a = (a_1, a_2) \in \mathbb{Q}^2$ and fix $z = a \cdot W = a_1\tau + a_2$. We now define the *Klein forms*

$$\kappa_a(W) \;=\; \kappa(z, L) \;:=\; e^{-\eta(z,L)z/2}\sigma(z, L). \tag{3.8}$$

For fixed a, it is evident that $\kappa_a(W)$ is a function only of τ, and we will often write $\kappa_a(\tau) = \kappa_a(W)$. The Klein forms satisfy several properties:

1. They are homogenous of degree 1:
$$\kappa_a(\lambda W) \;=\; \lambda\kappa_a(W) \tag{3.9}$$
for any $\lambda \in \mathbb{C}^*$. This follows from 3.5 and 3.6.

2. Using 3.7, it can be shown [7, pg. 28] that
$$\kappa_{a+b}(W) \;=\; \epsilon(a, b)\kappa_a(W), \tag{3.10}$$
where $b \in \mathbb{Z}^2$ is arbitrary and $\epsilon(a, b)$ is a d^{th} root of unity with d the least common multiple of the denominators of the components of a.

3. Finally, let $\alpha \in \mathbf{SL}_2(\mathbb{Z})$. Then
$$\kappa_a(\alpha W) \;=\; \kappa_{a\alpha}(W). \tag{3.11}$$
This last property follows from the Definition 3.8.

Our main use of Klein forms will be in the explicit construction of modular functions of level N. Let $a \in \frac{1}{N}\mathbb{Z}^2$ and suppose that $\alpha \in \Gamma(N)$. Then by 3.10 and 3.11, we have

$$\kappa_a(\alpha W) \;=\; \epsilon(\alpha)\kappa_a(W),$$

where $\epsilon(\alpha)$ is in fact a $(2N)^{\text{th}}$ root of unity. Thus, the Klein forms may be used as the "building blocks" of modular functions. In section 3.7, we will see how this explicit construction is carried out. We conclude this section by noting the q-product expansion for the Klein forms and a useful corollary of this formula: with $z = a \cdot W = a_1\tau + a_2$, put $q = e^{2\pi i\tau}$ as before and $q_z = e^{2\pi iz}$. We then have

$$\kappa_a(\tau) \;=\; -\frac{q^{(1/2)(a_1^2-a_1)}}{2\pi i}e^{\pi i a_2(a_1-1)}(1-q_z)\prod_{n=1}^{\infty}\frac{(1-q^nq_z)(1-q^n/q_z)}{(1-q^n)^2}. \tag{3.12}$$

Notice that if we change a to $a + b$ for some $b \in \mathbb{Z}^2$ then the q-series 3.12 changes by a root of unity. Therefore, from now on we consider a to be the representative of its class modulo $\mathbb{Z}^2$ such that

$$0 \le a_1 < 1 \quad \text{and} \quad 0 \le a_2 < 1.$$

Corollary 2. *Let a be as above. Then the Klein form $k_a(\tau)$ is holomorphic on $\mathbb{H}$ and the order of $k_a(\tau)$ at ∞ is*

$$\frac{1}{2}(a_1^2 - a_1). \tag{3.13}$$

Proof. This follows directly from the formula 3.12. □

We remark that we can use Corollary 2 to find the order of $\kappa_a(\tau)$ at any cusp s. Explicitly, let $\alpha_s \in \Gamma(1)$ take s to ∞. Then the order of $\kappa_a(\tau)$ at s is simply the order of $\kappa_{a\alpha_s}(\tau)$ at ∞.

3.4 The J Function

We now construct a modular function of level 1. As before, let $\tau \in \mathbb{H}$ and $L = \langle 1, \tau \rangle$ be the lattice generated by 1 and τ. Since $\Delta(L)$ and $g_2(L)^3$ are both homogenous of degree -12, the function

$$J(\tau) := \frac{g_2^3(L)}{\Delta} \tag{3.14}$$

$$= \frac{g_2^3(L)}{g_2^3(L) - 27g_3^2(L)} \tag{3.15}$$

is homogenous of degree 0. Therefore, since the lattices $L = \langle 1, \tau \rangle$ and $\mathcal{L} = \langle 1, -1/\tau \rangle$ satisfy the obvious relation $L = \tau\mathcal{L}$, we see that $J(\tau)$ is invariant under the transformation $\tau \to -1/\tau = S\tau$. Moreover, since $\langle 1, \tau + 1 \rangle = L$, we have $J(\tau + 1) = J(\tau)$. Finally, the nonvanishing of Δ for any $\tau \in \mathbb{H}$ tells us that $J(\tau)$ is holomorphic on $\mathbb{H}$. Expanding 3.15 as q-series, we see that

$$J(\tau) = \frac{1}{1728}\left(\frac{1}{q} + 744 + 196884q + \cdots\right). \tag{3.16}$$

Therefore, $J(\tau)$ has a pole of order 1 at ∞. Since J is invariant under S and T, by Theorem 13 we see that J defines a meromorphic function on $X(1)$, that is, $J \in K(X(1))$. We then have:

Theorem 18.

$$K(X(1)) = \mathbb{C}(J).$$

Proof. (See [8, pg. 63]) Let $f \in K(X(1))$. If f has a pole of order r at $z_0 \in \mathbb{H}$ then the function $f(J - J(z_0))^r$ is analytic at z_0. Therefore, there exists some polynomial $Q \in \mathbb{C}(J)$ such that Qf is holomorphic on $\mathbb{H}$. By Lemma 1, if Qf is not constant on $X(1)$, it has a pole at infinity (the only cusp) of order m, say. Since J has a simple pole at infinity, there exists a constant $c_0 \in \mathbb{C}$ so that $Qf - cJ^m$ has a pole of at most order $m - 1$. By descent, there exists some polynomial $P \in \mathbb{C}(J)$ so that $Qf - P$ has no pole at infinity and no pole in $\mathbb{H}$. But then $Qf - P$ is a holomorphic function on $X(1)$ and therefore constant by Lemma 1. We conclude that $f \in \mathbb{C}(J)$, as desired. □

Notice that this proof only uses two facts:

1. The surface $X(1)$ is compact.
2. The function $J \in K(X(1))$ has only a single pole, of order one at infinity.

Therefore, we see that the same proof will work for any compact Riemann surface $\mathcal{R}$ provided we can find a function $\mathcal{J} \in K(\mathcal{R})$ that has only a single pole of order one at infinity (or equivalently at any cusp).

3.5 Special Values of J

Since the only points in $\mathbb{H}^*$ with nontrivial stabilizer in $\bar{\Gamma}(1)$ are equivalent under $\Gamma(1)$ to i, ω, ∞, (Theorem 16), these points are in some sense "special" points. We therefore compute the value of J at each of them, since we will use these values frequently.

Proposition 1. *We have*

$$\begin{aligned} J(\omega) &= 0 \\ J(\infty) &= \infty \\ J(i) &= 1. \end{aligned}$$

Proof. The value at ∞ follows from the q-series 3.16. Let L_τ be the lattice generated by $1, \tau$. Then

$$\begin{aligned} g_3(L_i) &= \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0\}} \frac{1}{(m+ni)^6} \\ &= i^6 \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0\}} \frac{1}{(mi-n)^6} \\ &= -g_3(L_i), \end{aligned}$$

from which it follows that $g_3(L_i) = 0$. Similarly,

$$\begin{aligned} g_2(L_\omega) &= \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0\}} \frac{1}{(m+n\omega)^4} \\ &= \omega^4 \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0\}} \frac{1}{((m-n)\omega-n)^4} \\ &= \omega g_2(L_\omega), \end{aligned}$$

so that $g_2(L_\omega) = 0$. The proposed values of J now follow. $\square$

3.6 The Function Field $K(X(N))$

Theorem 18 tells us that $K(\mathbb{P}^1) \simeq K(X(1))$. By Theorem 10, we conclude that $X(1) \simeq \mathbb{P}^1$, as was alluded to in section 2.9. Then by 2.5, we have that the genus g_N of $X(N)$ is

$$g_N \;=\; 1 + \mu_N \frac{N-6}{12N}, \tag{3.17}$$

where $\mu_N = [\bar{\Gamma}(1) : \bar{\Gamma}(N)]$ is the degree of the natural covering map $f \,:\, X(N) \longrightarrow X(1)$ of section 2.8 and $N > 1$. Therefore, for $1 \leq N \leq 5$, $g_N = 0$ so that the field $K(X(N))$ is *rational*, that is, generated by a single element over $\mathbb{C}$. Moreover, it follows immediately from 3.17 that for $N > 6$, we have $g_N > 1$, while $g_6 = 1$. Therefore, the five values of N in the range $1 \leq N \leq 5$ are the *only* values of N for which $X(N)$ has genus 0 and the corresponding function field $K(X(N))$ is rational.

We have already constructed a generator for $K(X(1))$ over $\mathbb{C}$; namely, the function J of section 3.4. The goal of the next two sections is to determine such functions for the other four values of N cited above.

3.7 Products of Klein Forms

As mentioned in section 3.3, we will use the Klein forms to explicitly construct modular functions of level N. To do this, we need to know when a product of Klein forms (taken to both positive and negative integer exponents) is a modular function. The following theorem tells us when this is so:

Theorem 19. *Fix a positive integer $N \geq 1$ and let $A \subset \mathbb{Z}^2$ be a finite set consisting of pairs of integers not both divisible by N. Put*

$$\mathcal{A} \;=\; \left\{ \frac{1}{N} a \,:\, a \,\in\, A \right\}$$

and to $\alpha = (a_1/N, a_2/N) \,\in\, \mathcal{A}$ associate the integer $m(\alpha)$. Suppose that

$$\sum_{\alpha \in \mathcal{A}} m(\alpha) \;=\; 0,$$

and let

$$f(\tau) \;=\; \prod_{\alpha \in \mathcal{A}} \kappa_\alpha^{m(\alpha)}(\tau).$$

Then if N is odd, f is a modular function of level N if and only if

$$\sum_{\alpha \in \mathcal{A}} m(\alpha) a_1^2 \equiv \sum_{\alpha \in \mathcal{A}} m(\alpha) a_2^2 \;\equiv\; \sum_{\alpha \in \mathcal{A}} m(\alpha) a_1 a_2 \;\equiv\; 0 \mod N,$$

while if N is even, f is a modular function of level N if and only if

$$\sum_{\alpha\in\mathcal{A}} m(\alpha)a_1^2 \equiv \sum_{\alpha\in\mathcal{A}} m(\alpha)a_2^2 \equiv 0 \mod 2N \quad \textit{and}$$
$$\sum_{\alpha\in\mathcal{A}} m(\alpha)a_1a_2 \equiv 0 \mod N.$$

Proof. See [7, pg. 68]. □

3.8 Generators for $K(X(N))$, $N \le 5$

Theorem 20. *Set $\zeta_k = e^{2\pi i/k}$ and let the functions J_N for $2 \le N \le 5$ be given by*

$$J_2 = \frac{\kappa^4_{(0,\frac{1}{2})}}{\kappa^4_{(\frac{1}{2},0)}} = 16q^{1/2}\left(\prod_{n=1}^{\infty}\frac{1-q^{2n}}{1-q^{n/2}}\right)^8 \qquad J_3 = \frac{\kappa^3_{(\frac{1}{3},0)}}{\kappa^4_{(0,\frac{1}{3})}} = \frac{1}{i\sqrt{27}}q^{-1/3}\left(\prod_{n=1}^{\infty}\frac{1-q^{n/3}}{1-q^{3n}}\right)^3$$

$$J_4 = \frac{\kappa^3_{(0,\frac{1}{4})}\kappa_{(\frac{1}{2},\frac{1}{4})}}{\kappa^3_{(\frac{1}{4},0)}\kappa_{(\frac{1}{4},\frac{1}{2})}} = \zeta_8^3\sqrt{8}q^{1/4}\prod_{n=1}^{\infty}\frac{(1-q^{4n})^2(1-q^{n/2})}{(1-q^{n/4})^2(1-q^{2n})}$$

$$J_5 = \frac{\kappa_{(\frac{2}{5},0)}\kappa_{(\frac{2}{5},\frac{1}{5})}\kappa_{(\frac{2}{5},\frac{2}{5})}\kappa_{(\frac{2}{5},\frac{3}{5})}\kappa_{(\frac{2}{5},\frac{4}{5})}}{\kappa_{(\frac{1}{5},0)}\kappa_{(\frac{1}{5},\frac{1}{5})}\kappa_{(\frac{1}{5},\frac{2}{5})}\kappa_{(\frac{1}{5},\frac{3}{5})}\kappa_{(\frac{1}{5},\frac{4}{5})}} = \zeta_5 q^{-1/5}\prod_{n=1}^{\infty}\frac{(1-q^{5n-2})(1-q^{5n-3})}{(1-q^{5n-4})(1-q^{5n-1})}.$$

Then for $2 \le N \le 5$, $K(X(N)) = \mathbb{C}(J_N)$.

Proof. In section 3.4, we showed that any function on $X(N)$ with a single simple pole generates $K(X(N))$. Thus, it is enough to show that $J_N \in K(X(N))$ and that J_N has only a single pole of order one, for each N with $2 \le N \le 5$. That each J_N is a modular function of level N follows after a short calculation from Theorem 19. We then use Theorem 14 to compute the set of cusps $C_{\Gamma(N)}$ for the above values of N. We find:

$$\begin{aligned}
C_{\Gamma(2)} &= \{0, 1, \infty\}, \\
C_{\Gamma(3)} &= \{0, 1/2, 1, \infty\}, \\
C_{\Gamma(4)} &= \{0, 1/3, 1/2, 2/3, 1, \infty\}, \\
C_{\Gamma(5)} &= \{0, 2/9, 1/4, 2/7, 1/3, 2/5, 1/2, 5/8, 2/3, 3/4, 1, \infty\}.
\end{aligned}$$

Finally, we use Corollary 2 to compute the order of J_N at each cusp and find that J_N has only a single simple pole for each N with $2 \le N \le 5$. This completes the proof. □

As a consequence of this theorem, we are now able to give an explicit isomorphism $X(N) \longrightarrow \mathbb{P}^1$. In particular, since the parameter x of $\mathbb{C}(x)$ induces the trivial isomorphism $\mathrm{id} : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$, the map

$$\tau \longrightarrow J_N(\tau) \tag{3.18}$$

is an isomorphism of $X(N)$ with $\mathbb{P}^1$. We shall use this in the next section to show how $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ acts on $\mathbb{P}^1$ for $2 \leq N \leq 5$.

3.9 The Action of $\Gamma(1)$ on J_N

We showed in section 3.2 that the field extension $K(X(N))/K(X(1))$ is Galois with Galois group $\bar{\Gamma}(1)/\bar{\Gamma}(N)$. Since by Theorem 13 the transformations $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$ and $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ generate $\Gamma(1)$, we can determine the action of the Galois group $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ on J_N by determining the action of S and T on J_N.

Now if $f \in K(X(N))$ satisfies $K(X(N)) = \mathbb{C}(f)$, then we certainly have $gf = R(f)$ for any $g \in \Gamma(1)$, where $R \in \mathbb{C}(x)$. However, it is not difficult to see that f has only a single simple pole if and only if gf has only a single simple pole. Therefore, $\mathbb{C}(gf) = \mathbb{C}(f)$, so that we may write $f = Q(gf)$ where $Q \in \mathbb{C}(x)$. Then we have $f = Q \circ R(f)$ so that $Q \circ R = 1$ which implies that Q, R are degree one rational maps. Put more simply, the automorphism group of $\mathbb{P}^1$ is just $\mathbf{PSL}_2(\mathbb{C})$ [9, pg. 12], and since $f \mapsto gf$ induces an automorphism of $K(\mathbb{P}^1)$, it follows that $f = \phi_g(gf)$ for some fractional linear transformation $\phi_g \in \mathbf{PSL}_2(\mathbb{C})$. Moreover, it is not difficult to explicitly determine ϕ_g. Without loss of generality, we may suppose that f has a simple pole at infinity (if not, replace f by $1/(f - c_1)$ where c_1 is the constant term in the q-expansion of f). If gf has a simple pole at infinity, then there exists $r \in \mathbb{C}$ such that $f - r(gf)$ has no poles in $X(N)$ and is therefore constant. Otherwise, let c_g be the constant term in the q-expansion of gf. Then $1/(gf - c_g)$ has a simple pole at infinity and we are reduced to the previous case. We now use this process to determine the action of S and T on J_N for $2 \leq N \leq 5$.

With $\zeta_k = e^{2\pi i/k}$ as before, we have

$$T \circ J_2 = \frac{-J_2}{1 + J_2} \qquad S \circ J_2 = \frac{1}{J_2} \tag{3.19}$$

$$T \circ J_3 = \zeta_3 + \zeta_3^2 J_3 \qquad S \circ J_3 = \frac{-1}{J_3} \tag{3.20}$$

$$T \circ J_4 = \frac{\zeta_4 J_4}{1 - J_4} \qquad S \circ J_4 = \frac{1}{\zeta_4 J_4} \tag{3.21}$$

$$T \circ J_5 = \zeta_5^{-1} J_5 \qquad S \circ J_5 = \frac{\zeta_5^2 + (1 + \zeta_5 + \zeta_5^2) J_5}{J_5 - (1 + \zeta_5 + \zeta_5^2)} \tag{3.22}$$

Now by 3.18, the above action of $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ on J_N for $2 \leq N \leq 5$ induces an action on $\mathbb{P}^1$. That is for each N, for any point $z \in \mathbb{P}^1$, and any $g \in \bar{\Gamma}(1)/\bar{\Gamma}(N)$, we define

$$gz \;:=\; J_N(gJ_N^{-1}(z)).$$

We can give a simpler description of this action. We begin with $N = 2$, though this is in some senses the least intuitive of the 4 cases. The 3 cusps may be viewed as the 3 vertices of the equatorial triangle of a **double triangular pyramid** inscribed in the unit sphere. Moreover, the two preimages of 0, i.e. the $\bar{\Gamma}(1)/\bar{\Gamma}(2)$ orbit of ω, correspond to the two polar tips of this double pyramid. Projecting this solid to a triangulation of $\mathbb{P}^1$ by 6 triangles, it is evident that $\bar{\Gamma}(1)/\bar{\Gamma}(2) \simeq \mathbf{S}_3$ acts on $\mathbb{P}^1$ via symmetries of the double pyramid. The group is generated by rotations of $2\pi/3$ fixing polar points (corresponding to the order 3 stabilizer of ω generated by ST) and by the order 2 symmetry that interchanges the two tips (which corresponds to the transformation T).

Viewing the 4 cusps of $X(3)$ as the four vertices of a regular **tetrahedron** inscribed in $\mathbb{P}^1$, which has been projected to a triangulation of $\mathbb{P}^1$ by 4 triangles, we see that $\bar{\Gamma}(1)/\bar{\Gamma}(3)$ acts on $\mathbb{P}^1$ by symmetries of the tetrahedron: the group is generated by rotations of $2\pi/3$ about a vertex (corresponding to the transformation T, which stabilizes ∞) and by rotations of π about the midpoints of the edges (corresponding to S). We therefore see that the points in the orbit of i (i.e. the preimage of 1 under J) correspond to the midpoints of the vertices of this tetrahedron.

Similarly, we view the 6 cusps of $X(4)$ as the vertices of a regular **octrahedron** (projected to a triangulation of $\mathbb{P}^1$). Then $\bar{\Gamma}(1)/\bar{\Gamma}(4)$ acts on $\mathbb{P}^1$ by symmetries of the octahedron. As before, the group is generated by T (rotations by $\pi/2$ about a vertex) and S (rotations by π about the midpoint of an edge). The 12 midpoints of the vertices correspond to the points in $\mathbb{P}^1$ in the orbit of i.

Finally, the 12 cusps of $X(5)$ give us the 12 vertices of a regular **icosahedron**, and $\bar{\Gamma}(1)/\bar{\Gamma}(5)$ acts on $\mathbb{P}^1$ via the $\mathbf{A}_5$ action on the icosahedron. The group is generated by rotation by $2\pi/5$ about a vertex (again corresponding to T—a fact that is made clear by 3.22) and rotation through π about the midpoint of any edge. The 20 edge midpoints correspond, as before, to the points in the orbit of i.

3.10 Index N Subgroups of $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ and the Associated Curves

By Theorem 12, we have the isomorphism

$$\bar{\Gamma}(1)/\bar{\Gamma}(N) \;\simeq\; \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}.$$

For small values of N, we can easily determine the size and structure of $\mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$. Indeed, the description given in the previous section of the action of $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ on $\mathbb{P}^1$ realizes the group $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ as a subgroup of a permutation group. Moreover, the four groups $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ for $2 \leq N \leq 5$ enjoy a special property: namely, for each of the above N, the group $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ contains an index N subgroup. In summary, we have

1. $\mathbf{C}_3 \subset \mathbf{S}_3 \simeq \bar{\Gamma}(1)/\bar{\Gamma}(2)$.
2. $\mathbf{V}_4 \subset \mathbf{A}_4 \simeq \bar{\Gamma}(1)/\bar{\Gamma}(3)$.
3. $\mathbf{S}_3 \subset \mathbf{S}_4 \simeq \bar{\Gamma}(1)/\bar{\Gamma}(4)$.
4. $\mathbf{A}_4 \subset \mathbf{A}_5 \simeq \bar{\Gamma}(1)/\bar{\Gamma}(5)$.

Let G_N denote an index N subgroup of $\bar{\Gamma}(1)/\bar{\Gamma}(N)$. By Theorem 17 and the Galois correspondence, for each G_N we obtain a degree N extension of the field $\mathbb{C}(J)$ of rational functions in J. One way to describe these extensions is to use Theorem 10, which tells us that there is some Riemann surface $Y(N)$ and some $f \in K(Y(N))$ such that our degree N extension is just $K(Y(N)) = \mathbb{C}(J)(f)$ with $R_N(J, f) = 0$ for some degree N rational map R_N over $\mathbb{C}$. In order to give as simple and complete a description of these extensions as possible, we would like to explicitly find the maps R_N. That is the goal of the next section.

3.11 The Polynomials

As above, let $Y(N)$ be the Riemann surface corresponding to the index N subgroup G_N of $\bar{\Gamma}(1)/\bar{\Gamma}(N)$. Since $X(N)$ is rational over $\mathbb{C}$, so is $Y(N)$—that is, we have an isomorphism $f_N : Y(N) \longrightarrow \mathbb{P}^1$ with $K(Y(N)) = \mathbb{C}(f_N)$. Since $Y(N)$ is a degree N cover of $\mathbb{P}^1$, the map

$$J : Y(N) \longrightarrow \mathbb{P}^1 \tag{3.23}$$

is of degree N. Viewing $Y(N)$ as $\mathbb{P}^1$ by the isomorphism $\tau \longrightarrow f_N(\tau)$, we see that 3.23 is a degree N map of $\mathbb{P}^1$ to itself. By Theorem 7, there exists $R_N(x) \in \mathbb{C}(x)$ of degree N with $R_N(f_N) = J$. All this may be viewed in the following commutative diagram:

$$\begin{array}{ccc} Y(N) & \xrightarrow[f_N]{\sim} & \mathbb{P}^1 \\ {\scriptstyle J}\downarrow & & \downarrow{\scriptstyle R_N} \\ \mathbb{P}^1 & = & \mathbb{P}^1 \end{array}$$

Using that the automorphism group of $\mathbb{P}^1$ is $\mathbf{PSL}_2(\mathbb{C})$, it is not difficult to see that the map R_N is unique up to fractional linear transformation. Since $f_N : Y(N) \longrightarrow \mathbb{P}^1$ is an isomorphism, the map f_N is unbranched. However, by Theorem 16 and Proposition 1, we see that J as a map from $Y(N)$ to $\mathbb{P}^1$ *is* branched above $0, 1, \infty$ (and only these points). By composing f_N with a fractional linear transformation, we can ensure that f_N has a simple pole at ∞. Theorem 16 tells us that $\bar{\Gamma}(1)_\infty = \langle T \rangle$. Moreover, it is not difficult to see that T has order N in $\bar{\Gamma}(1)/\bar{\Gamma}(N)$. Since N divides the order of $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ only *once*, for each $2 \leq N \leq 5$, the group G_N cannot contain an element of order N, the upshot being that G_N has trivial intersection with $(\bar{\Gamma}(1)/\bar{\Gamma}(N))_\infty$ when N is prime.

That this is so for $N = 4$ follows from the fact that the square of any order 4 element in $\mathbf{S}_4$ is a double transposition and hence is not contained in $\mathbf{S}_3$.

Since $X(N)$ is a degree $[\bar{\Gamma}(1) : \bar{\Gamma}(N)]$ cover of $X(1)$, the preimage of ∞ under the map $J : X(N) \longrightarrow \mathbb{P}^1$ consists of $[\bar{\Gamma}(1) : \bar{\Gamma}(N)]/N$ points of order N. Since $G_N \cap (\bar{\Gamma}(1)/\bar{\Gamma}(N))_\infty = \{1\}$, the group G_N acts transitively on these points. Therefore, the preimage of ∞ under the map $J : Y(N) \longrightarrow \mathbb{P}^1$ consists of a single point of order N. Since f_N has a simple pole at ∞, we see that the map R_N has a pole of order N at ∞. Since R_N is a degree N rational map, we have shown that in fact R_N is a *polynomial* of degree N.

In the following analysis, especially for $N = 4, 5$, we shall need the following proposition:

Proposition 2. *Let S be a topological space and let G be a group acting on S. If $z_1, z_2 \in S$ are in the same G orbit then their stabilizers in G are conjugate.*

The proof follows directly from the definition of the stabilizer of a point, so we omit it. We now proceed to determine the degree N polynomial R_N for $2 \leq N \leq 5$. We shall implicitly use the fact from section 2.5 that $\bar{\Gamma}(1)_i \cap \Gamma(N) = \bar{\Gamma}(1)_\omega \cap \Gamma(N) = \{1\}$ for all $N \geq 2$

3.11.1 $N = 2$

Here, $\bar{\Gamma}(1)/\bar{\Gamma}(2) \simeq \mathbf{S}_3$ and the (unique, normal) index 2 subgroup of interest is $G_2 = \mathbf{C}_3$. We first compute the preimages of the points $0, 1$ of the map $J : X(2) \longrightarrow \mathbb{P}^1$.

1. Since the stabilizer $\bar{\Gamma}(1)_\omega$ is generated by the order 3 element ST, we see that the preimage of 0 consists of 2 points of order 3.

2. Similarly, we have $\bar{\Gamma}(1)_i = \langle S \rangle$, which is of order 2. Therefore, the preimage of 1 consists of 3 points of order 2.

With this information, we can then determine the preimages of each of the points $0, 1$ in $Y(2)$. In particular,

1. The group $\bar{\Gamma}(1)_\omega$ is isomorphic to $\mathbf{C}_3$. Therefore, the preimages of 0 in $X(2)$ are branch points of order 3 above the preimages of 0 in $Y(2)$. Thus, the preimage of 0 in $Y(2)$ consists of two single points.

2. Since $\bar{\Gamma}(1)_i \simeq \mathbf{C}_2$ has trivial intersection with $\mathbf{C}_3$, the latter acts transitively on the preimages of 1 under J in $X(2)$. Therefore, the preimage of 1 in $Y(2)$ is a single double point.

Item 1 above enables us to write $R_2(x) = c(x - a)(x - b)$ for some $a, b, c \in \mathbb{C}$, while item 2 tells us that $R_2(x) - 1$ has a double root. Therefore, $R_2(x) - 1$ has a root in common with $R_2'(x)$. We have

$$\begin{aligned} R_2'(x) &= c(x-a)(x-b)\left(\frac{1}{x-a} + \frac{1}{x-b}\right) \\ &= c(2x - a - b), \end{aligned}$$

so that

$$-c\left(\frac{a-b}{2}\right)^2 = 1.$$

Since a fractional linear transformation exists taking any three points to any other three points, we may suppose that $b = 0$ and $a = 2$ (since we have already specified that the preimage of ∞ under f_2 should be ∞). This gives $c = -1$ and we find that f_N satisfies the polynomial

$$-x(x-2) = J,$$

which may be rewritten as

$$-(x-1)^2 = J - 1,$$

that is, the degree two extension $K(Y(2))/K(X(1))$ is generated by a root of the polynomial

$$Z^2 - (J-1). \tag{3.24}$$

This is, of course, expected: since $\mathbb{C}$ contains all roots of unity and $K(Y(2))/K(X(1))$ is a Galois extension with cyclic Galois group (since $\mathbf{C_3}$ is normal in $\mathbf{S}_3$ and the quotient is $\mathbf{C}_2$), we know that $K(Y(2)) = K(X(1))(\sqrt{g})$, for some $g \in K(X(1))$ which is not a square in $K(X(1))$. In fact, using 3.15 and 3.2, we see that

$$\begin{aligned}
J - 1 &= \frac{g_2^3}{g_2^3 - 27g_3^2} \\
&= \frac{27g_3^2}{g_2^3 - 27g_3^2} \\
&= \frac{27g_3^2}{(2\pi i)^{12} q \prod_{n=1}^{\infty}(1-q^n)^{24}},
\end{aligned}$$

where $q = e^{2\pi i\tau}$ as usual. We now see explicitly that we can exrtract a square root of $J - 1$ to obtain the function

$$\sqrt{J-1} = \frac{3\sqrt{3}}{(2\pi i)^6}\frac{g_3}{q^{1/2}\prod_{n=1}^{\infty}(1-q^n)^{12}},$$

which generates $K(Y(2))$.

3.11.2 $N = 3$

Since $\bar{\Gamma}(1)/\bar{\Gamma}(3) \simeq \mathbf{A}_4$, we have the unique, normal index 3 subgroup $G_3 = \mathbf{V}_4$. As for $N = 2$, we first determine the preimages in $X(3)$ of 0 for the map $J : X(3) \longrightarrow \mathbb{P}^1$ and then use this information to find the preimage of 0 in $Y(3)$.

1. The stabilizer $\bar{\Gamma}(1)_\omega$ is generated by an element of order 3 corresponding to a 3-cycle in $\mathbf{A}_4$. Therefore, the preimage of 0 in $X(3)$ consists of 4 points of order 3.

Since every nontrivial element of $\mathbf{V}_4$ has order 2, we see that $\bar{\Gamma}(1)_\omega \simeq \mathbf{C}_3$ has trivial intersection with G_3. We have:

1. The 4 order 3 points in $X(3)$ above 0 form a single orbit under G_3. Therefore, the preimage of 0 in $Y(3)$ is a single point of order 3.

Thus, we can write $R_3(x) = c(x-a)^3$. Therefore, since $\mathbb{C}$ is algebraically closed (and hence the value of c is immaterial), we find that the degree 3 field extension $K(Y(3))/K(X(1))$ is the splitting field (since it is a Galois extension) of

$$Z^3 - J. \tag{3.25}$$

As for $N = 2$, this could have been predicted. The Galois group of $K(Y(3))/K(X(1))$ is $\mathbf{C}_3$ and therefore cyclic. Since $\mathbb{C}$ contains all roots of unity, the extension is obtained by extracting a cube root, in this case, of J. As for $N = 2$, we find

$$\sqrt[3]{J} = \frac{1}{(2\pi i)^4} \frac{g_2}{q^{1/3}\prod_{n=1}^{\infty}(1-q^n)^8}.$$

3.11.3 $N = 4$

The situation for $N = 4$ and 5 is somewhat different. Most notably, in these cases the group G_N is not normal in $\bar{\Gamma}(1)/\bar{\Gamma}(N)$ so that the field extension $K(Y(N))/K(X(1))$ is not Galois. We proceed as above.

1. We have $\bar{\Gamma}(1)/\bar{\Gamma}(4) \simeq \mathbf{S_4}$, so that the preimage of 0 in $X(4)$ consists of 8 points of order 3.

2. Similarly, over 1 we have 12 points of order 2.

By considering the irreducible two dimensional representation of $\mathbf{S}_4$, it can be seen that the stabilizer $(\bar{\Gamma}(1)/\bar{\Gamma}(4))_i$ corresponds to a 2-cycle in $\mathbf{S}_4$ and not a double transposition. Moreover, from Proposition 2, it follows that each 2-cycle in $\mathbf{S}_4$ (there are six in total) generates the stabilizer of two of the 12 points above 1. Fixing a copy of $\mathbf{S}_3 = G_3$ in $\mathbf{S}_4$ shows that for 6 of these 12 points, the stabilizer is contained in G_3, while G_3 acts transitively on the other 6. Since in any case the stabilizer has order 2, the six points whose stabilizer is contained in G_3 break up into two orbits under G_3.

The situation for 0 is similar. The stabilizer $(\bar{\Gamma}(1)/\bar{\Gamma}(4))_\omega$ viewed as a subgroup of $\mathbf{S}_4$ is generated by a 3-cycle. It follows from Proposition 2 that the 8 points of order 3 above 0 have isotropy subgroups generated by the 8 elements of order 3 in $\mathbf{S}_4$. Since our particular copy G_3 of $\mathbf{S}_3$ contains precisely 2 of the 3-cycles in $\mathbf{S}_4$, we see that 2 of the 8 points in the preimage of 0 have stabilizer contained in G_3, while G_3 acts transitively on the other 6 points. We have therefore shown that

1. The preimage of 0 in $Y(2)$ consists of a single point and a triple point.
2. The preimage of 1 consists of a double point and two single points.

Therefore, the polynomial $R_4(x)$ is

$$c(x-a)(x-b)^3,$$

for some $a, b, c \in \mathbb{C}$. Now since the preimage of 1 contains a double point, the polynomials

$$\begin{aligned} R_3(x) - 1 &= c(x-a)(x-b)^3 - 1 \quad \text{and} \\ R_3'(x) &= R_3(x)\left(\frac{1}{x-a} + \frac{3}{x-b}\right) \\ &= c(x-b)^2(4x-b-3a) \end{aligned}$$

have a common root. Since b is a root of $R_3(x)$, it cannot be a root of $R_3(x)-1$, and therefore we see that $(b+3a)/4$ must be a root of $R_3(x)-1$. We are then led to the equation

$$-\frac{27c}{2^8}(b-a)^4 = 1.$$

As before, we can make an affine change of variable so that a, b are any values we like, as long as they are not equal. So let $a = 3, b = -1$. Then we must have $c = -1/27$, so that the field extension $K(Y(4))/K(X(1))$ is generated by a root of

$$(Z-3)(Z+1)^3 + 27J. \tag{3.26}$$

The fact that the smallest normal subgroup of $\mathbf{S}_4$ containing (not necessarilly properly) $\mathbf{S}_3$ is $\mathbf{S}_4$ itself tells us that the normal closure of the extension $K(Y(4))/K(X(1))$ is the field $K(X(4))$ and consequently that $K(X(4))/K(X(1))$ may be described as the splitting field of the degree 4 polynomial 3.26.

3.11.4 $N = 5$

In this case we have $\bar{\Gamma}(1)/\bar{\Gamma}(5) \simeq \mathbf{A}_5$ and $G_5 = \mathbf{A}_4$. We at once see that

1. The preimage of 0 in $X(5)$ consists of 20 points of order 3. Moreover, by Proposition 2, the 20 isotropy subgroups at these points are generated by the 20 3-cycles in $\mathbf{A}_5$.
2. The preimage of 1 consists of 30 points of order 2. The stabilizer of any point is generated by a double transposition (since $\mathbf{A}_5$ contains no other elements of order 2). Since there are 15 double transpositions in $\mathbf{A}_5$, each double transposition generates the isotropy subgroup at precisely two points in the preimage of 1.

Now any copy of $\mathbf{A}_4$ inside $\mathbf{A}_5$ contains exactly 8 3-cycles and 3 double transpositions. Therefore, the 20 points above 0 break up into 3 orbits under G_5: one orbit consists of 12 points (unramified above the preimages of 0 in $Y(5)$) and the other two orbits consist of 4 points of order 3 above the preimages of 0 in $Y(5)$. Similarly, the 30 points in $X(5)$ above 1 form 3 orbits under G_5. Six of the 30 points have their stabilizers contained in our copy of $\mathbf{A}_4$, so that they form a single orbit of points of order 2 over $Y(5)$. The action of G_5 on the remaining 24 points is therefore transitive, so that we obtain 2 orbits of points of order 1 over $Y(5)$. Thus, we see that

1. The preimage of 0 in $Y(5)$ consists of 2 single points and one triple point.

2. The preimage of 1 in $Y(5)$ consists of 2 double points and one single point.

Therefore, the desired polynomial has the form

$$d(x-a)(x-b)(x-c)^3$$

for some $a, b, c, d \in \mathbb{C}$. As before, we can make an affine change of variable to ensure that $a = 1, b = -1$, say. The condition at 1 above then tells us that the polynomial

$$d(x^2-1)(x-c)^3 - 1$$

has two double roots, that is, that is shares two roots with its derivative

$$d(x-c)^2(5x^2-2cx-3).$$

We therefore see that

$$x_1 = \frac{c+\sqrt{c^2+15}}{5} \quad \text{and} \quad x_2 = \frac{c-\sqrt{c^2+15}}{5}$$

must both be roots of

$$d(x^2-1)(x-c)^3 - 1.$$

This gives two equations in two unknowns, which we readily solve to find

$$\begin{aligned} c &= -\frac{1}{3^2} i\sqrt{15} \\ d &= \frac{3^4 \cdot 5^2}{2^{11}} i\sqrt{15}. \end{aligned}$$

Therefore, the field extension $K(Y(5))/K(X(1))$ is generated by a root of

$$\frac{3^4 \cdot 5^2}{2^{11}} i\sqrt{15}\,(x^2-1)\left(x + \frac{1}{3^2} i\sqrt{15}\right)^3 - J.$$

We can simplify this equation by absorbing a factor of $i\sqrt{15}$ into x to obtain the polynomial

$$(15Z^2+1)(9Z-1)^3+\left(\frac{2^{11}}{5^4}\right)J. \tag{3.27}$$

As for $N=4$, since the smallest normal subgroup of $\mathbf{A}_5$ is $\mathbf{A}_5$ itself, we see that the Galois closure of $K(Y(5))/K(X(1))$ is $K(X(5))$ and therefore that $K(X(5))$ is the splitting field of the degree 5 polynomial 3.27.

Bibliography

[1] Elkies, N.D. *The Klein Quartic in Number Theory.* "The Eightfold Way." Levy, S., Ed. MSRI vol. 35, 1998.

[2] Farkas, H., Kra, I. *Riemann Surfaces.* 2nd ed. Springer-Verlag, New York 1992.

[3] Farkas, H., Kra, I. *Theta Constants, Riemann Surfaces, and the Modular Group.* GSM vol. 37. American Mathematical Society, Providence RI. 2001.

[4] Forster, O. *Lectures on Riemann Surfaces.* Springer-Verlag, New York 1981.

[5] Iwasawa, K. *Algebraic Functions.* Translations of Mathematical Monographs vol. 118. American Mathematical Society, Providence RI. 1991.

[6] Knapp, A.W. *Elliptic Curves.* Mathematical Notes 40. Princeton University Press, Princeton NJ. 1992.

[7] Kubert, D.S., Lang, S. *Modular Units.* Springer-Verlag, New York 1981.

[8] Lang, S. *Elliptic Functions.* Addison-Wesley, Reading MA. 1973.

[9] McKean, H., Moll, V. *Elliptic Curves.* Cambridge University Press, Cambridge 1997.

[10] Remmert, R. *Theory of Complex Functions.* Springer-Verlag, New York 1991.

[11] Serre, J.P. *A Course in Arithmetic.* Springer-Verlag, New York 1973.

[12] Shimura, G. *Introduction to the Arithmetic Theory of Automorphic Functions.* Princeton University Press, Princeton NJ. 1994.

[13] Whittaker, E.T., Watson, G.N. *A Course of Modern Analysis.* Cambridge University Press, Cambridge 1996.

[14] Wickelgren, K. *Galois Theory of Riemann Surfaces.* Harvard University, 2002. (Unpublished)