LECTURES ON THE COMBINATORIAL STRUCTURE OF THE MODULI SPACES OF RIEMANN SURFACES

MOTOHICO MULASE

Contents

0. Introduction

0.1. **Motivation.** The goal of these lectures is to determine the combinatorial orbifold structure of the moduli spaces of Riemann surfaces, and to compute the Euler characteristic of these moduli spaces. In establishing our results concerning the moduli spaces, we will encounter, quite unexpectedly, quantum field theory through the Feynman diagram expansion technique due to Feynman and 'tHooft ([8], [51]), quantum gravity through the work of Witten [54] and Kontsevich [20], random matrices and Selberg integral formulas through the analysis of matrix integrals ([1], [25], [38]), and number theory and arithmetic algebraic geometry through Belyi's theorem and Grothendieck's idea of *dessins d'enfants* ([4], [12], [43]). Our goal thus serves as a motivation for us to learn these other areas of mathematics.

Let us denote by $\mathfrak{M}_{g,n}$ the *moduli space* of Riemann surfaces of genus g with n marked points. (All these terms are defined in the lectures.) Set theoretically, the moduli space is just the set of isomorphism classes of Riemann surfaces with n reference points marked. If no marked points are chosen, we simply denote $\mathfrak{M}_g = \mathfrak{M}_{g,0}$.

Since Riemann introduced the notion of moduli spaces in his monumental paper [39] of 1857, there has been a continuous effort in mathematics toward the understanding of $\mathfrak{M}_{g,n}$. After the explosive development in superstring theory that started in 1984, the study of moduli spaces has also become one of the central

Date: November 1, 2001.

themes in modern mathematical physics. The developments have led to the discovery of new ideas such as *quantum cohomology theory* [21], [22].

Before the recent burst of research motivated by string theory, there were many attempts to understand the natural geometric structure of $\mathfrak{M}_g$. Riemann [39] discovered that $\mathfrak{M}_g$ is naturally a subset of the quotient space $\mathcal{A}_g = \mathcal{S}_g \big/ Sp(g,\mathbb{Z})$, where $\mathcal{S}_g$ is the *Siegel upper half space* of genus g that consists of complex symmetric matrices T of size $g \times g$ whose imaginary parts are positive definite, and $Sp(g,\mathbb{Z})$ is the *modular group* of $2g \times 2g$ symplectic matrices with integer entries. An element $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ of $Sp(g,\mathbb{Z})$ acts on $T \in \mathcal{S}_g$ by the fractional linear transformation

$$T \longmapsto \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot T = (AT+B)(CT+D)^{-1}.$$

We will examine the simplest case of this quotient space shortly. We note that the complex dimension of $\mathcal{A}_g$ is $g(g+1)/2$. Riemann showed, through a heuristic argument, that the complex dimension of $\mathfrak{M}_g$ is only $3g-3$ for $g \geq 2$. Thus if $g \geq 4$, the moduli space is a very thin subset of $\mathcal{A}_g$. The problem of determining the defining equations of $\mathfrak{M}_g$ as a subspace of $\mathcal{A}_g$ is known as the **Riemann-Schottky Problem** or the **Schottky Problem**, after Riemann and his student Schottky who worked on the case of $g = 4$ [45]. A surprising solution to this problem was established in the works done by many authors [3], [26], [27], and [48]. The results are unexpected because they use the Kadomtsev-Petviashvili equations, an integrable system of nonlinear partial differential equations that describes shallow water wave motions.

Although a complete and beautiful characterization of $\mathfrak{M}_g$ in $\mathcal{A}_g$ has been obtained in terms of nonlinear partial differential equations, we still do not know any systematic algebraic equations that define $\mathfrak{M}_g$. For more information and recent developments in this direction, see articles in [7], the newly added sections of the 1994 edition of [36], and the articles [23], [28] and [30].

A compact Riemann surface is a complete non-singular projective algebraic curve defined over $\mathbb{C}$. Thus the moduli space $\mathfrak{M}_g$ should be constructed through algebro-geometric technique. We note, however, that $\mathfrak{M}_g$ is not naturally an algebraic variety. The notion of *algebraic stack* was introduced by Grothendieck, and further developed by Deligne and others [6], to study such objects as the moduli spaces. Mumford's geometric invariant theory [36] is another approach to moduli theory. This approach has been particularly useful in the study of moduli spaces of algebraic vector bundles. The theory of algebraic stacks has become fashionable these days.

A third way toward the understanding of the moduli space $\mathfrak{M}_g$ is through Teichmüller theory. This approach, originated by Teichmüller in the 1930's, uses heavy real harmonic analysis and complex analysis [16], [37]. The idea is to avoid dealing with $\mathfrak{M}_g$ directly because it is not a complex manifold, but instead to construct a complex manifold $\mathcal{T}_g$ such that there is a ramified covering map

$$\mathcal{T}_g \longrightarrow \mathfrak{M}_g$$

that is determined by the action of the *mapping class group*, or the *Teichmüller modular group*. Teichmüller theory turns out to be a great success because of the fundamental result that $\mathcal{T}_g$ is a complex manifold that is topologically homeomorphic to the unit ball in the complex vector space $\mathbb{C}^{3g-3}$. Therefore, all topological information of $\mathfrak{M}_g$ is contained in the mapping class group action on the Teichmüller space.

The approach that we will develop in these lectures is different from any of the above methods. Although we will not be able to introduce complex analytic or algebraic structures to the moduli spaces, our method gives a concrete *orbifold* structure to the product space $\mathfrak{M}_{g,n} \times \mathbb{R}_+^n$. In particular, it enables us to compute the *orbifold Euler characteristic* of $\mathfrak{M}_{g,n}$.

We follow [34] for the concrete realization of the orbifold structure of $\mathfrak{M}_{g,n} \times \mathbb{R}_+^n$, which is based on the earlier works of Strebel [50], Harer [13], and others. The advantage of the method established in [34] is that one can see the striking relation with Grothendieck's idea of dessins d'enfants, based on Belyi's theorem [4], [43].

The computation of the Euler characteristic was originally carried out by Harer and Zagier [15], through the study of group cohomologies of the mapping class group acting on the Teichmüller space. Another computational method, a more direct calculation, was proposed in [38] using the idea of quantum field theory. The correct asymptotic analysis of the Hermitian matrix integrals was established in [32] along this line, and the computation of the Euler characteristic was carried out completely rigorously.

In these lectures we will prove

Theorem 0.1. *The orbifold Euler characteristic of the moduli space is given by*

$$\chi(\mathfrak{M}_{g,n}) = (-1)^{n-1} \frac{(2g+n-3)!}{(2g-2)!} \zeta(1-2g), \tag{0.1}$$

where $\zeta(z)$ is the Riemann zeta function.

Remark. 1. The formula (0.1) is valid for all values of $g \geq 0$ and $n > 0$. For $g = 0$, we use

$$\frac{\zeta(1-2g)}{(2g-2)!} = \frac{(2g)(2g-1)}{(2g)!} \zeta(1-2g) = -2g \cdot \zeta(1-2g) = 1.$$

Hence

$$\chi(\mathfrak{M}_{0,n}) = (-1)^{n-1}(n-3)!.$$

2. The Euler characteristic should be an integer if the space in question is a manifold. For an orbifold, Thurston introduced the notion of *orbifold Euler characteristic*, which takes rational values. We note that

$$\zeta(1-2g) = -\frac{b_{2g}}{2g},$$

where b_m is the m-th Bernoulli number, and therefore the value of (0.1) is indeed a rational number.

We follow Penner's idea and examine a model of 0-dimensional quantum field theory, or an example of the Hermitian random matrix model. There are two completely different methods to solve the model: one utilizing the *Selberg integral formula* that gives the exact analytic result, and the other using the Feynman diagram expansion due to Feynman [8] and 'tHooft [51] that leads to the topological result. Since both results should agree, we obtain the equality of Theorem 0.1.

Let us now begin our long excursion toward our goal, enjoying the scenery surrounding us.

1. Riemann Surfaces and Elliptic Functions

1.1. **Basic Definitions.** Let us begin with defining *Riemann surfaces* and their *moduli spaces.*

Definition 1.1 (Riemann surfaces)**.** A **Riemann surface** is a paracompact Hausdorff topological space C with an open covering $C = \bigcup_\lambda U_\lambda$ such that for each open set U_λ there is an open domain V_λ of the complex plane $\mathbb{C}$ and a homeomorphism

$$\phi_\lambda : V_\lambda \longrightarrow U_\lambda \tag{1.1}$$

that satisfies that if $U_\lambda \cap U_\mu \neq \emptyset$, then the **gluing map** $\phi_\mu^{-1} \circ \phi_\lambda$

$$V_\lambda \supset \phi_\lambda^{-1}(U_\lambda \cap U_\mu) \xrightarrow{\ \phi_\lambda\ } U_\lambda \cap U_\mu \xrightarrow{\ \phi_\mu^{-1}\ } \phi_\mu^{-1}(U_\lambda \cap U_\mu) \subset V_\mu \tag{1.2}$$

is a biholomorphic function.

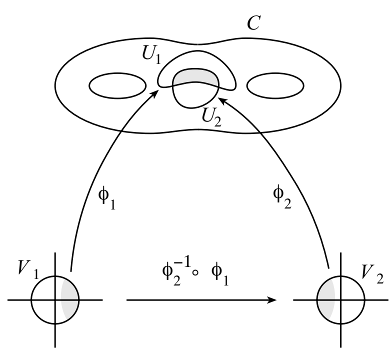

Figure 1.1. Gluing two coordinate charts.

Remark.
1. A topological space X is **paracompact** if for every open covering $X = \bigcup_\lambda U_\lambda$, there is a locally finite open cover $X = \bigcup_i V_i$ such that $V_i \subset U_\lambda$ for some λ. **Locally finite** means that for every $x \in X$, there are only finitely many V_i's that contain x. X is said to be **Hausdorff** if for every pair of distinct points x, y of X, there are open neighborhoods $W_x \ni x$ and $W_y \ni y$ such that $W_x \cap W_y = \emptyset$.
2. A continuous map $f : V \longrightarrow \mathbb{C}$ from an open subset V of $\mathbb{C}$ into the complex plane is said to be **holomorphic** if it admits a convergent Taylor series expansion at each point of $V \subset \mathbb{C}$. If a holomorphic function $f : V \longrightarrow V'$ is one-to-one and onto, and its inverse is also holomorphic, then we call it **biholomorphic**.
3. Each open set V_λ gives a **local chart** of the Riemann surface C. We often identify V_λ and U_λ by the homeomorphism ϕ_λ, and say "U_λ and U_μ are glued by a biholomorphic function." The collection $\{\phi_\lambda : V_\lambda \longrightarrow U_\lambda\}$ is called a **local coordinate system**.
4. A Riemann surface is a **complex manifold** of complex dimension 1. We call the Riemann surface structure on a topological surface a **complex structure**. The definition of complex manifolds of an arbitrary dimension can be given in a similar manner. For more details, see [19].

Definition 1.2 (Holomorphic functions on a Riemann surface)**.** A continuous function $f : C \longrightarrow \mathbb{C}$ defined on a Riemann surface C is said to be a **holomorphic function** if the composition $f \circ \phi_\lambda$

$$V_\lambda \xrightarrow{\ \phi_\lambda\ } U_\lambda \subset C \xrightarrow{\ f\ } \mathbb{C}$$

is holomorphic for every index λ.

Definition 1.3 (Holomorphic maps between Riemann surfaces)**.** A continuous map $h : C \longrightarrow C'$ from a Riemann surface C into another Riemann surface C' is a **holomorphic map** if the composition map $(\phi'_\mu)^{-1} \circ h \circ \phi_\lambda$

$$V_\lambda \xrightarrow{\ \phi_\lambda\ } U_\lambda \subset C \xrightarrow{\ h\ } C' \supset U'_\mu \xrightarrow{\ \phi'_\mu\ } V'_\mu$$

is a holomorphic function for every local chart V_λ of C and V'_μ of C'.

Definition 1.4 (Isomorphism of Riemann surfaces)**.** If there is a bijective holomorphic map $h : C \longrightarrow C'$ whose inverse is also holomorphic, then the Riemann surfaces C and C' are said to be **isomorphic**. We use the notation $C \cong C'$ when they are isomorphic.

Since the gluing function (1.2) is biholomorphic, it is in particular an orientation preserving homeomorphism. Thus each Riemann surface C carries the structure of an oriented topological manifold of real dimension 2. We call it the **underlying** topological manifold structure of C. The orientation comes from the the natural orientation of the complex plane. All local charts are glued in an orientation preserving manner by holomorphic functions.

A **compact** Riemann surface is a Riemann surface that is compact as a topological space without boundary. In these lectures we deal mostly with compact Riemann surfaces.

The classification of compact topological surfaces is completely understood. The simplest example is a 2-sphere S^2. All other oriented compact topological surfaces are obtained by attaching **handles** to an oriented S^2. First, let us cut out two small disks from the sphere. We give an orientation to the boundary circle that is compatible with the orientation of the sphere. Then glue an oriented cylinder $S^1 \times I$ (here I is a finite open interval of the real line $\mathbb{R}$) to the sphere, matching the orientation of the boundary circles. The surface thus obtained is a compact oriented surface of **genus** 1. Repeating this procedure g times, we obtain a **compact oriented surface of genus** g. The genus is the number of attached handles. Since the sphere has Euler characteristic 2 and a cylinder has Euler characteristic 0, the surface of genus g has Euler characteristic $2 - 2g$.

A set of **marked points** is an ordered set of distinct points $(p_1, p_2, \cdots, p_n)$ of a Riemann surface. Two Riemann surfaces with marked points $\big(C, (p_1, \cdots, p_n)\big)$ and $\big(C', (p'_1, \cdots, p'_n)\big)$ are **isomorphic** if there is a biholomorphic map $h : C \longrightarrow C'$ such that $h(p_j) = p'_j$ for every j.

Definition 1.5 (Moduli space of Riemann surfaces)**.** The **moduli space** $\mathfrak{M}_{g,n}$ is the set of isomorphism classes of Riemann surfaces of genus g with n marked points.

The goal of these lectures is to give an orbifold structure to $\mathfrak{M}_{g,n} \times \mathbb{R}_+^n$ and to determine its Euler characteristic for every genus and $n > 0$.

1.2. **Elementary Examples.** Let us work out a few elementary examples. The simplest Riemann surface is the complex plane $\mathbb{C}$ itself with the standard complex structure. The unit disk $D_1 = \{z \in \mathbb{C} \mid |z| < 1\}$ of the complex plane is another example. We note that although these Riemann surfaces are homeomorphic to one another, they are not isomorphic as Riemann surfaces. Indeed, if there was a biholomorphic map $f : \mathbb{C} \longrightarrow D_1$, then f would be a bounded (since $|f| < 1$) holomorphic function defined entirely on $\mathbb{C}$. From Cauchy's integral formula, one concludes that f is constant.

The simplest nontrivial example of a compact Riemann surface is the Riemann sphere $\mathbb{P}^1$. Let U_1 and U_2 be two copies of the complex plane, with coordinates z and w, respectively. Let us *glue* U_1 and U_2 with the identification $w = 1/z$ for $z \neq 0$. The union $\mathbb{P}^1 = U_1 \cup U_2$ is a compact Riemann surface homeomorphic to the 2-dimensional sphere S^2.

The above constructions give all possible complex structures on the 2-plane $\mathbb{R}^2$ and the 2-sphere S^2, which follows from the following:

Theorem 1.6 (Riemann Mapping Theorem)**.** *Let X be a Riemann surface with trivial fundamental group: $\pi_1(X) = 1$. Then X is isomorphic to either one of the following:*

1. *the entire complex plane $\mathbb{C}$ with the standard complex structure;*
2. *the unit disk $\{z \in \mathbb{C} \mid |z| < 1\}$ with the standard complex structure induced from the complex plane $\mathbb{C}$; or*
3. *the Riemann sphere $\mathbb{P}^1$.*

Remark. The original proof of the Riemann mapping theorem is due to Riemann, Koebe, Carathéodory, and Poincaré. Since the technique we need to prove this theorem has nothing to do with the topics we deal with in these lectures, we refer to [41], volume II, for the proof.

We note that $\mathbb{P}^1$ is a Riemann surface of genus 0. Thus the Riemann mapping theorem implies that $\mathfrak{M}_0$ consists of just one point.

A powerful technique to construct a new Riemann surface from a known one is the quotient construction via a group action on the old Riemann surface. Let us examine the quotient construction now.

Let X be a Riemann surface. An **analytic automorphism** of X is a biholomorphic map $f : X \longrightarrow X$. The set of all analytic automorphisms of X forms a group through the natural composition of maps. We denote by $\mathrm{Aut}(X)$ the group of analytic automorphisms of X. Let G be a group. When there is a group homomorphism $\phi : G \longrightarrow \mathrm{Aut}(X)$, we say the group G **acts on** X. For an element $g \in G$ and a point $x \in X$, it is conventional to write

$$g(x) = (\phi(g))(x),$$

and identify g as a biholomorphic map of X into itself.

Definition 1.7 (Fixed point free and properly discontinuous action)**.** Let G be a group that acts on a Riemann surface X. A point $x \in X$ is said to be a **fixed point** of $g \in G$ if $g(x) = x$. The group action of G on X is said to be **fixed point free** if no element of G other than the identity has a fixed point. The group action is said to be **properly discontinuous** if for every compact subsets Y_1 and Y_2 of X, the cardinality of the set

$$\{g \in G \mid g(Y_1) \cap Y_2 \neq \emptyset\}$$

is finite.

Remark. A finite group action on a Riemann surface is always properly discontinuous.

When a group G acts on a Riemann surface X, we denote by X/G the **quotient space**, which is the set of orbits of the G-action on X.

Theorem 1.8 (Quotient construction of a Riemann surface)**.** *If a group G acts on a Riemann surface X properly discontinuously and the action is fixed point free, then the quotient space X/G has the structure of a Riemann surface.*

Proof. Let us denote by $\pi : X \longrightarrow X/G$ the natural projection. Take a point $\widehat{x} \in X/G$, and choose a point $x \in \pi^{-1}(\widehat{x})$ of X. Since X is covered by local coordinate systems $X = \bigcup_\mu U_\mu$, there is a coordinate chart U_μ that contains x. Note that we can cover each U_μ by much smaller open sets without changing the Riemann surface structure of X. Thus without loss of generality, we can assume that U_μ is a disk of radius ϵ centered around x, where ϵ is chosen to be a small positive number. Since the closure $\overline{U}_\mu$ is a compact set, there are only finitely many elements g in G such that $g(\overline{U}_\mu)$ intersects with $\overline{U}_\mu$. Now consider taking the limit $\epsilon \rightarrow 0$. If there is a group element $g \neq 1$ such that $g(\overline{U}_\mu) \cap \overline{U}_\mu \neq \emptyset$ as ϵ becomes smaller and smaller, then $x \in \overline{U}_\mu$ is a fixed point of g. Since the G action on X is fixed point free, we conclude that for a small enough ϵ, $g(\overline{U}_\mu) \cap \overline{U}_\mu = \emptyset$ for every $g \neq 1$.

Therefore, $\pi^{-1}(\pi(U_\mu))$ is the disjoint union of $g(U_\mu)$ for all distinct $g \in G$. Moreover,

$$\pi : U_\mu \longrightarrow \pi(U_\mu) \tag{1.3}$$

gives a bijection between U_μ and $\pi(U_\mu)$. Introduce the **quotient topology** to X/G by defining $\pi(U_\mu)$ as an open neighborhood of $\widehat{x} \in X/G$. With respect to the quotient topology, the projection π is continuous and locally a homeomorphism. Thus we can introduce a holomorphic coordinate system to X/G by (1.3). The gluing function of $\pi(U_\mu)$ and $\pi(U_\nu)$ is the same as the gluing function of U_μ and $g(U_\nu)$ for some $g \in G$ such that $U_\mu \cap g(U_\nu) \neq \emptyset$, which is a biholomorphic function because X is a Riemann surface. This completes the proof. □

Remark.
1. The above theorem generalizes to the case of a manifold. If a group G acts on a manifold X properly discontinuously and fixed point free, then X/G is also a manifold.
2. If the group action is fixed point free but not properly discontinuous, then what happens? An important example of such a case is a free Lie group action on a manifold. A whole new theory of fiber bundles starts here.
3. If the group action is properly discontinuous but not fixed point free, then what happens? The quotient space is no longer a manifold. Thurston coined the name **orbifold** for such an object. We will study orbifolds in later sections.

Definition 1.9 (Fundamental domain)**.** Let G act on a Riemann surface X properly discontinuously and fixed point free. A region Ω of X is said to be a **fundamental domain** of the G-action if the disjoint union of $g(\Omega)$, $g \in G$, covers the entire X:

$$X = \coprod_{g \in G} g(\Omega).$$

The simplest Riemann surface is $\mathbb{C}$. What can we obtain by considering a group action on the complex plane? First we have to determine the automorphism group of $\mathbb{C}$. If $f : \mathbb{C} \longrightarrow \mathbb{C}$ is a biholomorphic map, then f cannot have an essential singularity at infinity. (Otherwise, f is not bijective.) Hence f is a polynomial in the coordinate z. By the fundamental theorem of algebra, the only polynomial that gives a bijective map is a polynomial of degree one. Therefore, $\mathrm{Aut}(\mathbb{C})$ is the group of affine transformations

$$\mathbb{C} \ni z \longmapsto az + b \in \mathbb{C},$$

where $a \neq 0$.

Exercise 1.1. Determine all subgroups of $\mathrm{Aut}(\mathbb{C})$ that act on $\mathbb{C}$ properly discontinuously and fixed point free.

Let us now turn to the construction of compact Riemann surfaces of genus 1. Choose an element $\tau \in \mathbb{C}$ such that $Im(\tau) > 0$, and define a free abelian subgroup of $\mathbb{C}$ by

$$\Lambda_\tau = \mathbb{Z} \cdot \tau \oplus \mathbb{Z} \cdot 1 \subset \mathbb{C}. \tag{1.4}$$

This is a **lattice of rank** 2. An **elliptic curve of modulus** τ is the quotient abelian group

$$E_\tau = \mathbb{C}/\Lambda_\tau. \tag{1.5}$$

It is obvious that the natural Λ_τ-action on $\mathbb{C}$ through addition is properly discontinuous and fixed point free. Thus E_τ is a Riemann surface. Figure 1.2 shows a fundamental domain Ω of the Λ_τ-action on $\mathbb{C}$. It is a parallelogram whose four vertices are 0, 1, $1+\tau$, and τ. It includes two sides, say the interval $[0, 1)$ and $[0, \tau)$, but the other two parallel sides are not in Ω.

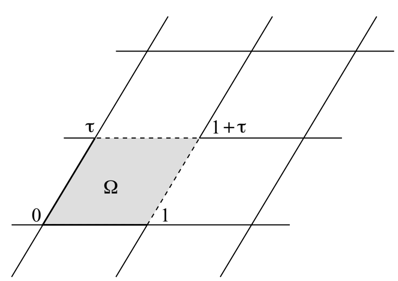

FIGURE 1.2. A fundamental domain of the Λ_τ-action on the complex plane.

Topologically E_τ is homeomorphic to a torus $S^1 \times S^1$. Thus an elliptic curve is a compact Riemann surface of genus 1. Conversely, one can show that if a Riemann surface is topologically homeomorphic to a torus, then it is isomorphic to an elliptic curve.

Exercise 1.2. Show that every Riemann surface of genus 1 is an elliptic curve. (Hint: Let Y be a Riemann surface and X its universal covering. Show that X has a natural complex structure such that the projection map $\pi : X \longrightarrow Y$ is locally biholomorphic.)

An elliptic curve $E_\tau = \mathbb{C}/\Lambda_\tau$ is also an abelian group. The group action by addition

$$E_\tau \times E_\tau \ni (x,y) \longmapsto x+y \in E_\tau$$

is a holomorphic map. Namely, for every point $y \in E_\tau$, the map $x \longmapsto x+y$ is a holomorphic automorphism of E_τ. Being a group, an elliptic curve has a privileged point, the origin $0 \in E_\tau$.

When are two elliptic curves E_τ and E_μ isomorphic? Note that since $Im(\tau) > 0$, τ and 1 form a $\mathbb{R}$-linear basis of $\mathbb{R}^2 = \mathbb{C}$. Let

$$\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \tau \\ 1 \end{bmatrix},$$

where

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2,\mathbb{Z}).$$

Then the same lattice Λ_τ is generated by (ω_1, ω_2):

$$\Lambda_\tau = \mathbb{Z}\cdot\omega_1 \oplus \mathbb{Z}\cdot\omega_2.$$

Therefore,

$$E_\tau = \mathbb{C}/(\mathbb{Z}\cdot\omega_1 \oplus \mathbb{Z}\cdot\omega_2). \tag{1.6}$$

To make (1.6) into the form of (1.5), we divide everything by ω_2. Since the division by ω_2 is a holomorphic automorphism of $\mathbb{C}$, we have

$$E_\tau = \mathbb{C}/(\mathbb{Z}\cdot\omega_1 \oplus \mathbb{Z}\cdot\omega_2) \cong E_\mu,$$

where

$$\tau \longmapsto \mu = \frac{\omega_1}{\omega_2} = \frac{a\tau+b}{c\tau+d}. \tag{1.7}$$

The above transformation is called a **linear fractional transformation**, which is an example of a **modular transformation**. Note that we do not allow the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to have determinant -1. This is because when the matrix has determinant -1, we can simply interchange ω_1 and ω_2 so that the net action is obtained by an element of $SL(2,\mathbb{Z})$.

Exercise 1.3. Show that the linear fractional transformation (1.7) is a holomorphic automorphism of the upper half plane $H = \{\tau \in \mathbb{C} \mid Im(\tau) > 0\}$.

Conversely, suppose we have an isomorphism

$$f : E_\tau \xrightarrow{\sim} E_\mu.$$

We want to show that μ and τ are related by a fractional linear transformation (1.7). By applying a translation of E_τ if necessary, we can assume that the isomorphism f maps the origin of E_τ to the origin of E_μ, without loss of generality. Let us denote by $\pi_\tau : \mathbb{C} \longrightarrow E_\tau$ the natural projection. We note that it is a **universal covering** of the torus E_τ. It is easy to show that the isomorphism f **lifts** to a homeomorphism $\widetilde{f} : \mathbb{C} \longrightarrow \mathbb{C}$. Moreover it is a holomorphic automorphism of $\mathbb{C}$:

$$\begin{array}{ccc} \mathbb{C} & \xrightarrow[\sim]{\widetilde{f}} & \mathbb{C} \\ {\scriptstyle\pi_\tau}\downarrow & & \downarrow{\scriptstyle\pi_\mu} \\ E_\tau & \xrightarrow[\sim]{f} & E_\mu. \end{array} \tag{1.8}$$

Since $\widetilde{f}$ is an affine transformation, $\widetilde{f}(z) = sz + t$ for some $s \neq 0$ and t. Since $f(0) = 0$, $\widetilde{f}$ maps Λ_τ to Λ_μ bijectively. In particular, $t \in \Lambda_\mu$. We can introduce a new coordinate in $\mathbb{C}$ by shifting by t. Then we have

$$\begin{bmatrix} \mu \\ 1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} s\tau \\ s \end{bmatrix} \tag{1.9}$$

for some element $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL(2,\mathbb{Z})$ (again after interchanging μ and 1, if necessary). The equation (1.9) implies (1.7).

It should be noted here that a matrix $A \in SL(2,\mathbb{Z})$ and $-A$ (which is also an element of $SL(2,\mathbb{Z})$) define the same linear fractional transformation. Therefore, to be more precise, the projective group

$$PSL(2,\mathbb{Z}) = SL(2,\mathbb{Z})\big/\{\pm 1\},$$

called the **modular group**, acts on the upper half plane H through holomorphic automorphisms. We have now established the first interesting result on the moduli theory:

Theorem 1.10 (The moduli space of elliptic curves)**.** *The moduli space of Riemann surfaces of genus* 1 *with one marked point is given by*

$$\mathfrak{M}_{1,1} = H\big/PSL(2,\mathbb{Z}),$$

where $H = \{\tau \in \mathbb{C} \mid Im(\tau) > 0\}$ *is the* ***upper half plane****.*

Since H is a Riemann surface and $PSL(2,\mathbb{Z})$ is a discrete group, we wonder if the quotient space $H\big/PSL(2,\mathbb{Z})$ becomes naturally a Riemann surface. To answer this question, we have to examine if the modular transformation has any fixed points. To this end, it is useful to know

Proposition 1.11 (Generators of the modular group)**.** *The group* $PSL(2,\mathbb{Z})$ *is generated by two elements*

$$T = \begin{bmatrix} 1 & 1 \\ & 1 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} & -1 \\ 1 & \end{bmatrix}.$$

Proof. Clearly, the subgroup $\langle S, T\rangle$ of $PSL(2\mathbb{Z})$ generated by S and T contains

$$\begin{bmatrix} 1 & n \\ & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & \\ n & 1 \end{bmatrix}$$

for an arbitrary integer n. Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an arbitrary element of $PSL(2,\mathbb{Z})$. The condition $ad - bc = 1$ implies a and b are relatively prime. The effect of the left and right multiplication of the above matrices and S to A is the **elementary transformation** of A:

1. Add any multiple of the second row to the first row and leave the second row unchanged;
2. Add any multiple of the second column to the first column and leave the second column unchanged;
3. Interchange two rows and change the sign of one of the rows;
4. Interchange two columns and change the sign of one of the columns.

The consecutive application of elementary transformations on A has an effect of performing the Euclidean algorithm to a and b. Since they are relatively prime, at the end we obtain 1 and 0. Thus the matrix A is transformed into $\begin{bmatrix} 1 & 0 \\ c' & 1 \end{bmatrix}$. It is in $\langle S, T \rangle$, hence so is A. This completes the proof. $\square$

We can immediately see that $S(i) = i$ and $(TS)(e^{\pi i/3}) = e^{\pi i/3}$. Note that $S^2 = 1$ and $(TS)^3 = 1$ in $PSL(2, \mathbb{Z})$. The system of equations

$$\begin{cases} \frac{ai+b}{ci+d} = i \\ ad - bc = 1 \end{cases}$$

shows that 1 and S are the only stabilizers of i, and

$$\begin{cases} \frac{ae^{\pi i/3}+b}{ce^{\pi i/3}+d} = e^{\pi i/3} \\ ad - bc = 1 \end{cases}$$

shows that 1, TS, and $(TS)^2$ are the only stabilizers of $e^{\pi i/3}$. Thus the subgroup $\langle S \rangle \cong \mathbb{Z}/2\mathbb{Z}$ is the stabilizer of i and $\langle TS \rangle \cong \mathbb{Z}/3\mathbb{Z}$ is the stabilizer of $e^{\pi i/3}$. In particular, the quotient space $H/PSL(2, \mathbb{Z})$ is not naturally a Riemann surface.

Since the $PSL(2, \mathbb{Z})$-action on H has fixed points, the fundamental domain cannot be defined in the sense of Definition 1.9. But if we allow overlap

$$X = \bigcup_{g \in G} g(\Omega)$$

only at the fixed points, an almost as good *fundamental domain* can be chosen. Figure 1.3 shows the popular choice of the fundamental domain of the $PSL(2, \mathbb{Z})$-action.

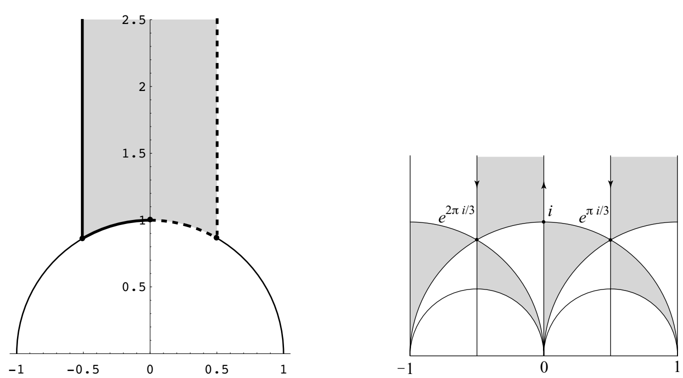

FIGURE 1.3. The fundamental domain of the $PSL(2, \mathbb{Z})$-action on the upper half plane H and the tiling of H by the $PSL(2, \mathbb{Z})$-orbits.

Since the transformation T maps $\tau \longmapsto \tau + 1$, the fundamental domain can be chosen as a subset of the vertical strip $\{\tau \in H \mid -1/2 \le Re(\tau) < 1/2\}$. The transformation $S : \tau \longmapsto -1/\tau$ interchanges the inside and the outside of the semicircle $|\tau| = 1$, $Im(\tau) > 0$. Therefore, we can choose the fundamental domain as in Figure 1.3. The arc of the semicircle from $e^{2\pi i/3}$ to i is included in the fundamental domain, but the other side of the semicircle is not. Actually, S maps the left-side segment of the semicircle to the right-side, leaving i fixed. Note that the union

$$H = \bigcup_{A \in PSL(2,\mathbb{Z})} A(\Omega)$$

of the orbits of the fundamental domain Ω by all elements of $PSL(2,\mathbb{Z})$ is not disjoint. Indeed, the point i is covered by Ω and $S(\Omega)$, and there are three regions that cover $e^{2\pi i/3}$.

The quotient space $H/PSL(2,\mathbb{Z})$ is obtained by gluing the vertical line $Re(\tau) = -1/2$ with $Re(\tau) = 1/2$, and the left arc with the right arc. Thus the space looks like Figure 1.4.

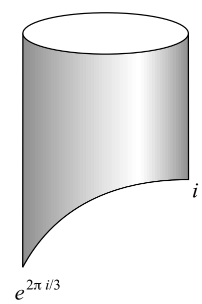

FIGURE 1.4. The moduli space $\mathfrak{M}_{1,1}$.

The moduli space $\mathfrak{M}_{1,1}$ is an example of an **orbifold**, and in algebraic geometry, it is an example of an **algebraic stack**. It has two **corner singularities**.

1.3. **Weierstrass Elliptic Functions.** Let ω_1 and ω_2 be two nonzero complex numbers such that $Im(\omega_1/\omega_2) > 0$. (It follows that ω_1 and ω_2 are linearly independent over the reals.) The **Weierstrass elliptic function**, or the Weierstrass $\wp$-function, of periods ω_1 and ω_2 is defined by

(1.10)
$$\wp(z) = \wp(z|\omega_1, \omega_2) = \frac{1}{z^2} + \sum_{\substack{m,n \in \mathbb{Z} \\ (m,n) \neq (0,0)}} \left(\frac{1}{(z - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2} \right).$$

Let $\Lambda_{\omega_1,\omega_2} = \mathbb{Z} \cdot \omega_1 + \mathbb{Z} \cdot \omega_2$ be the lattice generated by ω_1 and ω_2. If $z \notin \Lambda_{\omega_1,\omega_2}$, then the infinite sum (1.10) is absolutely and uniformly convergent. Thus $\wp(z)$ is a holomorphic function defined on $\mathbb{C} \setminus \Lambda_{\omega_1,\omega_2}$. To see the nature of the convergence of (1.10), fix an arbitrary $z \notin \Lambda_{\omega_1,\omega_2}$, and let N be a large positive number such that if $|m| > N$ and $|n| > N$, then $|m\omega_1 + n\omega_2| > 2|z|$. For such m and n, we have

$$\begin{aligned}
\left| \frac{1}{(z - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2} \right| &= \frac{|z| \cdot |2(m\omega_1 + n\omega_2) - z|}{|(z - m\omega_1 - n\omega_2)^2 (m\omega_1 + n\omega_2)^2|} \\
&< \frac{|z| \cdot \frac{5}{2} |m\omega_1 + n\omega_2|}{\frac{1}{4} |m\omega_1 + n\omega_2|^4} \\
&< C \frac{1}{|m\omega_1 + n\omega_2|^3}
\end{aligned}$$

for a large constant C independent of m and n. Since

$$\sum_{|m|>N,|n|>N} \frac{1}{|m\omega_1+n\omega_2|^3} < \infty,$$

we have established the convergence of (1.10). At a point of the lattice $\Lambda_{\omega_1,\omega_2}$, $\wp(z)$ has a double pole, as is clearly seen from its definition. Hence the Weierstrass elliptic function is globally meromorphic on $\mathbb{C}$. From the definition, we can see that $\wp(z)$ is an even function:

$$\wp(z) = \wp(-z). \tag{1.11}$$

Another characteristic property of $\wp(z)$ we can read off from its definition (1.10) is its double periodicity:

$$\wp(z+\omega_1) = \wp(z+\omega_2) = \wp(z). \tag{1.12}$$

For example,

$$\begin{aligned}
&\wp(z+\omega_1)\\
&= \frac{1}{(z+\omega_1)^2} + \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \left(\frac{1}{(z-(m-1)\omega_1-n\omega_2)^2} - \frac{1}{(m\omega_1+n\omega_2)^2}\right)\\
&= \frac{1}{(z+\omega_1)^2} + \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0),(1,0)}} \left(\frac{1}{(z-(m-1)\omega_1-n\omega_2)^2} - \frac{1}{((m-1)\omega_1+n\omega_2)^2}\right)\\
&\qquad + \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0),(1,0)}} \left(\frac{1}{((m-1)\omega_1+n\omega_2)^2} - \frac{1}{(m\omega_1+n\omega_2)^2}\right)\\
&\qquad + \frac{1}{z^2} - \frac{1}{\omega_1^2}\\
&= \frac{1}{z^2} + \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(1,0)}} \left(\frac{1}{(z-(m-1)\omega_1-n\omega_2)^2} - \frac{1}{((m-1)\omega_1+n\omega_2)^2}\right)\\
&= \wp(z).
\end{aligned}$$

Note that there are no holomorphic functions on $\mathbb{C}$ that are doubly periodic, except for a constant. The derivative of the $\wp$-function,

$$\wp'(z) = -2\sum_{m,n\in\mathbb{Z}} \frac{1}{(z-m\omega_1-n\omega_2)^3}, \tag{1.13}$$

is also a doubly periodic meromorphic function on $\mathbb{C}$. The convergence and the periodicity of $\wp'(z)$ is much easier to prove than (1.10).

Let us define two important constants:

$$g_2 = g_2(\omega_1,\omega_2) = 60 \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{(m\omega_1+n\omega_2)^4}, \tag{1.14}$$

$$g_3 = g_3(\omega_1,\omega_2) = 140 \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{(m\omega_1+n\omega_2)^6}. \tag{1.15}$$

These are the most fundamental examples of the **Eisenstein series**. The combination $\wp(z) - 1/z^2$ is a holomorphic function near the origin. Let us calculate its Taylor expansion. Since $\wp(z) - 1/z^2$ is an even function, the expansion contains only even powers of z. From (1.10), the constant term of the expansion is 0. Differentiating it twice, four times, etc., we obtain

$$\wp(z) = \frac{1}{z^2} + \frac{1}{20} g_2 z^2 + \frac{1}{28} g_3 z^4 + O(z^6). \tag{1.16}$$

It follows from this expansion that

$$\wp'(z) = -2\,\frac{1}{z^3} + \frac{1}{10} g_2 z + \frac{1}{7} g_3 z^3 + O(z^5).$$

Let us now compare

$$\begin{aligned} (\wp'(z))^2 &= 4\,\frac{1}{z^6} - \frac{2}{5} g_2 \frac{1}{z^2} - \frac{4}{7} g_3 + O(z^2), \\ 4(\wp(z))^3 &= 4\,\frac{1}{z^6} + \frac{3}{5} g_2 \frac{1}{z^2} + \frac{3}{7} g_3 + O(z^2). \end{aligned}$$

It immediately follows that

$$(\wp'(z))^2 - 4(\wp(z))^3 = -g_2 \frac{1}{z^2} - g_3 + O(z^2).$$

Using (1.16) again, we conclude that

$$f(z) \stackrel{\text{def}}{=} (\wp'(z))^2 - 4(\wp(z))^3 + g_2 \wp(z) + g_3 = O(z^2).$$

The equation means that

1. $f(z)$ is a globally defined doubly periodic meromorphic function with possible poles at the lattice $\Lambda_{\omega_1,\omega_2}$;
2. it is holomorphic at the origin, and hence holomorphic at every lattice point, too;
3. and it has a double zero at the origin.

Therefore, we conclude that $f(z) \equiv 0$. We have thus derived the **Weierstrass differential equation**:

$$(\wp'(z))^2 = 4(\wp(z))^3 - g_2 \wp(z) - g_3. \tag{1.17}$$

The differential equation implies that

$$z = \int dz = \int \frac{dz}{d\wp} d\wp = \int \frac{d\wp}{\wp'} = \int \frac{d\wp}{\sqrt{4(\wp)^3 - g_2\wp - g_3}}.$$

This last integral is called an **elliptic integral**. The Weierstrass $\wp$-function is thus the inverse function of an elliptic integral, and it explains the origin of the name *elliptic function*. Consider an ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \qquad b > a > 0.$$

From its parametric expression

$$\begin{cases} x = a\cos(\theta) \\ y = b\sin(\theta), \end{cases}$$

the arc length of the ellipse between $0 \leq \theta \leq s$ is given by

$$\int_0^s \sqrt{\left(\frac{dx}{d\theta}\right)^2 + \left(\frac{dy}{d\theta}\right)^2} d\theta = b\int_0^s \sqrt{1-k^2\sin^2(\theta)}d\theta, \tag{1.18}$$

where $k^2 = 1 - a^2/b^2$. This last integral is the Legendre-Jacobi second **elliptic integral**. Unless $a = b$, which is the case for the circle, (1.18) is not calculable in terms of elementary functions such as the trigonometric, exponential, and logarithmic functions. Mathematicians were led to consider the *inverse functions* of the elliptic integrals, and thus discovered the elliptic functions. The integral (1.18) can be immediately evaluated in terms of elliptic functions.

The usefulness of the elliptic functions in physics was recognized soon after their discovery. For example, the exact motion of a pendulum is described by an elliptic function. Unexpected appearances of elliptic functions have never stopped. It is an amazing coincidence that the Weierstrass differential equation implies that

$$u(x,t) = -\wp(x+ct) + \frac{c}{3}$$

solves the **KdV equation**

$$u_t = \frac{1}{4}u_{xxx} + 3uu_x,$$

giving a periodic wave solution traveling at the velocity $-c$. This observation was the key to the vast development of the 1980s on the Schottky problem and integrable systems of nonlinear partial differential equations called **soliton equations** [30].

1.4. **Elliptic Functions and Elliptic Curves.** A meromorphic function is a holomorphic map into the Riemann sphere $\mathbb{P}^1$. Thus the Weierstrass $\wp$-function defines a holomorphic map from an elliptic curve onto the Riemann sphere:

$$\wp : E_{\omega_1,\omega_2} = \mathbb{C}/\Lambda_{\omega_1,\omega_2} \longrightarrow \mathbb{P}^1 = \mathbb{C}\cup\{\infty\}. \tag{1.19}$$

To prove that the map is surjective, we must show that the Weierstrass $\wp$-function

$$\mathbb{C}\setminus\Lambda_{\omega_1,\omega_2} \longrightarrow \mathbb{C}$$

is surjective. To this end, let us first recall **Cauchy's integral formula**. Let

$$\sum_{n=-k}^{\infty} a_n z^n$$

be a power seires such that the sum of positive powers $\sum_{n\geq 0} a_n z^n$ converges absolutely around the oriign 0 with the radius of convergence $r > 0$. Then for any positively oriented circle $\gamma = \{z \in \mathbb{C} \mid |z| = \epsilon\}$ of radius $\epsilon < r$, we have

$$\frac{1}{2\pi i}\oint_\gamma \sum_{n=-k}^{\infty} a_n z^n = a_{-1}.$$

In this formulation, the integral formula is absolutely obvious. It has been generalized to the more familiar form that is taught in a standard complex analysis course.

Let Ω be the parallelogram whose vertices are 0, ω_1, ω_2, and $\omega_1 + \omega_2$. This is a fundamental domain of the $\Lambda_{\omega_1,\omega_2}$-action on the plane. Since the group acts by addition, the translation $\Omega + z_0$ of Ω by any number z_0 is also a fundamental domain. Now, choose the shift z_0 cleverly so that $\wp(z)$ has no poles or zeros on the boundary γ of $\Omega + z_0$. From the double periodicity of $\wp(z)$ and $\wp'(z)$, we have

$$\oint_\gamma \frac{\wp'(z)}{\wp(z)} dz = \oint_\gamma \frac{d}{dz} \log \wp(z) dz = 0. \tag{1.20}$$

This is because the integral along opposite sides of the parallelogram cancels. The function $\wp'(z)/\wp(z)$ has a simple pole of residue m where $\wp(z)$ has a zero of order m, and has a simple pole of residue $-m$ where $\wp(z)$ has a pole of order m. It is customary to count the number of zeros and poles with their multiplicity. Therefore, (1.20) shows that the number of poles and zeros of $\wp(z)$ are exactly the same on the elliptic curve E_{ω_1,ω_2}. Since we know that $\wp(z)$ has only one pole of order 2 on the elliptic curve, it must have two zeros or a zero of order 2. Here we note that the formula (1.20) is also true for

$$\frac{d}{dz} \log(\wp(z) - c)$$

for any constant c. This means that $\wp(z) - c$ has two zeros or a zero of order 2. It follows that the map (1.19) is surjective, and its inverse image consists of two points, generically.

Let e_1, e_2, and e_3 be the three roots of the polynomial equation

$$4X^3 - g_2 X - g_3 = 0.$$

Then except for the four points e_1, e_2, e_3, and ∞ of $\mathbb{P}^1$, the map $\wp$ of (1.19) is two-to-one. This is because only at the preimage of e_1, e_2, and e_3 the derivative $\wp'$ vanishes, and we know $\wp$ has a double pole at 0. We call the map $\wp$ of (1.19) a **branched double covering** of $\mathbb{P}^1$ **ramified** at e_1, e_2, e_3, and ∞.

It is quite easy to determine the preimages of e_1, e_2 and e_3 via the $\wp$-function. Recall that $\wp'(z)$ is an odd function in z. Thus for $j = 1, 2$, we have

$$\wp'\left(\frac{\omega_j}{2}\right) = \wp'\left(\frac{\omega_j}{2} - \omega_j\right) = \wp'\left(-\frac{\omega_j}{2}\right) = -\wp'\left(\frac{\omega_j}{2}\right).$$

Hence

$$\wp'\left(\frac{\omega_1}{2}\right) = \wp'\left(\frac{\omega_2}{2}\right) = \wp'\left(\frac{\omega_1 + \omega_2}{2}\right) = 0.$$

It is customary to choose the three roots e_1, e_2 and e_3 so that we have

$$\wp\left(\frac{\omega_1}{2}\right) = e_1, \qquad \wp\left(\frac{\omega_2}{2}\right) = e_2, \qquad \wp\left(\frac{\omega_1 + \omega_2}{2}\right) = e_3. \tag{1.21}$$

The quantities $\omega_1/2$, $\omega_2/2$, and $(\omega_1 + \omega_2)/2$ are called the **half periods** of the Weierstrass $\wp$-function.

The **complex projective space** $\mathbb{P}^n$ of dimension n is the set of equivalence classes of nonzero vectors $(x_0, x_1, \cdots, x_n) \in \mathbb{C}^{n+1}$, where $(x_0, x_1, \cdots, x_n)$ and $(y_0, y_1, \cdots, y_n)$ are equivalent if there is a nonzero complex number c such that $y_j = c x_j$ for all j. The equivalence class of a vector $(x_0, x_1, \cdots, x_n)$ is denoted by $(x_0 : x_1 : \cdots : x_n)$. We can define a map from an elliptic curve into $\mathbb{P}^2$,

$$(\wp, \wp') : E_{\omega_1,\omega_2} = \mathbb{C}/\Lambda_{\omega_1,\omega_2} \longrightarrow \mathbb{P}^2, \tag{1.22}$$

as follows: for $E_{\omega_1,\omega_2} \ni z \neq 0$, we map it to $(\wp(z) : \wp'(z) : 1) \in \mathbb{P}^2$. The origin of the elliptic curve is mapped to $(0 : 1 : 0) \in \mathbb{P}^2$. In terms of the global coordinate

$(X : Y : Z) \in \mathbb{P}^2$, the image of the map (1.22) satisfies a homogeneous cubic equation

$$Y^2Z - 4X^3 + g_2XZ^2 + g_3Z^3 = 0. \tag{1.23}$$

The zero locus C of this cubic equation is a **cubic curve**, and this is why the Riemann surface E_{ω_1,ω_2} is called a *curve*. The **affine part** of the curve C is the locus of the equation

$$Y^2 = 4X^3 - g_2X - g_3$$

in the (X, Y)-plane and its real locus looks like Figure 1.5.

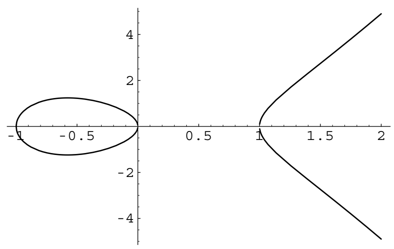

FIGURE 1.5. An example of a nonsingular cubic curve $Y^2 = 4X^3 - g_2X - g_3$.

We note that the association

$$\begin{cases} X = \wp(z) \\ Y = \wp'(z) \\ Z = 1 \end{cases}$$

is holomorphic for $z \in \mathbb{C} \setminus \Lambda_{\omega_1,\omega_2}$, and provides a local holomorphic parameter of the cubic curve C. Thus C is **non-singular** at these points. Around the point $(0 : 1 : 0) \in C \subset \mathbb{P}^2$, since $Y \neq 0$, we have an affine equation

$$\frac{Z}{Y} - 4\left(\frac{X}{Y}\right)^3 + g_2\frac{X}{Y}\left(\frac{Z}{Y}\right)^2 + g_3\left(\frac{Z}{Y}\right)^3 = 0.$$

The association

$$\begin{cases} \frac{X}{Y} = \frac{\wp(z)}{\wp'(z)} = -\frac{1}{2}z + O(z^5) \\ \frac{Z}{Y} = \frac{1}{\wp'(z)} = -\frac{1}{2}z^3 + O(z^7), \end{cases} \tag{1.24}$$

which follows from the earlier calculation of the Taylor expansions of $\wp(z)$ and $\wp'(z)$, shows that the curve C near $(0 : 1 : 0)$ has a holomorphic parameter $z \in \mathbb{C}$ defined near the origin. Thus the cubic curve C is everywhere non-singular.

Note that the map $z \longmapsto (\wp(z) : \wp'(z) : 1)$ determines a bijection from E_{ω_1,ω_2} onto C. To see this, take an arbitrary point $(X : Y : Z)$ on C. If it is the point at infinity, then (1.24) shows that the map is bijective near $z = 0$ because the relation can be solved for $z = z(X/Y)$ that gives a holomorphic function in X/Y. If $(X : Y : Z)$ is not the point at infinity, then there are two points z and z' on E_{ω_1,ω_2} such that

$$\wp(z) = \wp(z') = X/Z.$$

Since $\wp$ is an even function, actually we have $z' = -z$. Indeed, $z = -z$ as a point on E_{ω_1,ω_2} means $2z = 0$. This happens exactly when z is equal to one of the three half periods. For a given value of X/Z, there are two points on C, namely $(X : Y : Z)$ and $(X : -Y : Z)$, that have the same X/Z. If

$$(\wp(z) : \wp'(z) : 1) = (X : Y : Z),$$

then

$$(\wp(-z) : \wp'(-z) : 1) = (X : -Y : Z).$$

Therefore, we have

$$\begin{array}{ccccc} E_{\omega_1,\omega_2} & \xrightarrow[\text{bijection}]{(\wp:\wp':1)} & C & \xrightarrow{\subset} & \mathbb{P}^2 \\ \wp\downarrow & & \downarrow & & \\ \mathbb{P}^1 & = & \mathbb{P}^1, & & \end{array}$$

where the vertical arrows are $2 : 1$ ramified coverings.

Since the inverse image of the $2 : 1$ holomorphic mapping $\wp : E_{\omega_1,\omega_2} \longrightarrow \mathbb{P}^1$ is $\pm z$, the map $\wp$ induces a bijective map

$$E_{\omega_1,\omega_2}\big/\{\pm 1\} \xrightarrow[\text{bijection}]{} \mathbb{P}^1.$$

Because the group $\mathbb{Z}\big/2\mathbb{Z} \cong \{\pm 1\}$ acts on the elliptic curve E_{ω_1,ω_2} with exactly four fixed points

$$0, \quad \frac{\omega_1}{2}, \quad \frac{\omega_2}{2}, \quad \text{and} \quad \frac{\omega_1+\omega_2}{2},$$

the quotient space $E_{\omega_1,\omega_2}\big/\{\pm 1\}$ is not naturally a Riemann surface. It is $\mathbb{P}^1$ with **orbifold singularities** at e_1, e_2, e_3 and ∞.

1.5. **Degeneration of the Weierstrass Elliptic Function.** The relation between the coefficients and the roots of the cubic polynomial

$$4X^3 - g_2X - g_3 = 4(X - e_1)(X - e_2)(X - e_3)$$

reads

$$\begin{cases} 0 = e_1 + e_2 + e_3 \\ g_2 = -4(e_1e_2 + e_2e_3 + e_3e_1) \\ g_3 = 4e_1e_2e_3. \end{cases}$$

The **discriminant** of this polynomial is defined by

$$\triangle = (e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2 = \frac{1}{16}(g_2^3 - 27g_3^2).$$

We have noted that e_1, e_2, e_3, and ∞ are the branched points of the double covering $\wp : E_{\omega_1,\omega_2} \longrightarrow \mathbb{P}^1$. When the discriminant vanishes, these branched points are no longer separated, and the cubic curve (1.23) becomes singular.

Let us now consider a special case

$$\begin{cases} \omega_1 = ri \\ \omega_2 = 1, \end{cases} \tag{1.25}$$

where $r > 0$ is a real number. We wish to investigate what happens to $g_2(ri, 1)$, $g_3(ri, 1)$ and the corresponding $\wp$-function as $r \to +\infty$. Actually, we will see the Eisenstein series degenerate into a Dirichlet series. The **Riemann zeta function** is a **Dirichlet series** of the form

$$\zeta(s) = \sum_{n>0} \frac{1}{n^s}, \qquad Re(s) > 1. \tag{1.26}$$

We will calculate its special values $\zeta(2g)$ for every integer $g > 0$ later. For the moment, we note some special values:

$$\zeta(2) = \frac{\pi^2}{6}, \qquad \zeta(4) = \frac{\pi^4}{90}, \qquad \zeta(6) = \frac{\pi^6}{945}.$$

Since

$$\begin{aligned}\frac{1}{60}\, g_2(ri,1) &= \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{(mri+n)^4}\\ &= 2\sum_{n>0}\frac{1}{n^4} + 2\sum_{m>0}\frac{1}{(mr)^4} + 2\sum_{m>0}\sum_{n>0}\left(\frac{1}{(mri+n)^4} + \frac{1}{(mri-n)^4}\right)\\ &= 2\zeta(4) + 2\,\frac{1}{r^4}\zeta(4) + 2\sum_{m>0,n>0}\frac{(mri-n)^4+(mri+n)^4}{((mr)^2+n^2)^4},\end{aligned}$$

we have an estimate

$$\begin{aligned}\left|\frac{1}{60}\, g_2(ri,1) - 2\zeta(4)\right| &\leq 2\,\frac{1}{r^4}\zeta(4) + 2\sum_{m>0,n>0}\frac{2((mr)^2+n^2)^2}{((mr)^2+n^2)^4}\\ &= 2\,\frac{1}{r^4}\zeta(4) + 4\sum_{m>0,n>0}\frac{1}{((mr)^2+n^2)^2}\\ &< 2\,\frac{1}{r^4}\zeta(4) + 4\sum_{m>0,n>0}\frac{1}{(mr)^2((mr)^2+n^2)}\\ &< 2\,\frac{1}{r^4}\zeta(4) + 4\sum_{m>0,n>0}\frac{1}{(mr)^2n^2}\\ &= 2\,\frac{1}{r^4}\zeta(4) + 4\frac{1}{r^2}(\zeta(2))^2.\end{aligned}$$

Hence we have established

$$\lim_{r\to+\infty} g_2(ri,1) = 120\,\zeta(4).$$

Similarly, we have

$$\begin{aligned}\frac{1}{140}\, g_3(ri,1) &= \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{(mri+n)^6}\\ &= 2\zeta(6) - 2\frac{1}{r^6}\zeta(6) + 2\sum_{m,n>0}\left(\frac{1}{(mri+n)^6} + \frac{1}{(mri-n)^6}\right)\\ &= 2\zeta(6) - 2\frac{1}{r^6}\zeta(6) + 2\sum_{m,n>0}\frac{(mri-n)^6+(mri+n)^6}{((mr)^2+n^2)^6},\end{aligned}$$

hence

$$\begin{aligned}\left|\frac{1}{140}\, g_3(ri,1) - 2\zeta(6)\right| &< 2\frac{1}{r^6}\zeta(6) + 2\sum_{m,n>0}\frac{2((mr)^2+n^2)^3}{((mr)^2+n^2)^6}\\ &= 2\frac{1}{r^6}\zeta(6) + 4\sum_{m,n>0}\frac{1}{((mr)^2+n^2)^3}\\ &< 2\frac{1}{r^6}\zeta(6) + 4\,\frac{1}{r^4}\zeta(2)\zeta(4).\end{aligned}$$

Therefore,

$$\lim_{r\to+\infty} g_3(ri,1) = 280\,\zeta(6).$$

Note that the discriminant vanishes for these values:

$$(120\,\zeta(4))^3 - 27(280\,\zeta(6))^2 = 0.$$

Now let us study the degeneration of $\wp(z) = \wp(z|ri,1)$, the Weierstrass elliptic function with periods ri and 1, when $r \to +\infty$. Since one of the periods goes to ∞, the resulting function would have only one period, 1. It would have a double pole at each $n \in \mathbb{Z}$, and its leading term in its $(z-n)$-expansion would be $\frac{1}{(z-n)^2}$. There is such a function indeed:

$$\sum_{n=-\infty}^{\infty} \frac{1}{(z-n)^2} = \frac{\pi^2}{\sin^2(\pi z)}.$$

Thus we expect that the degeneration would have this limit, up to a constant term adjustment. From its definition, we have

$$\begin{aligned}
\wp(z|ri,1) &= \frac{1}{z^2} + \sum_{(m,n)\neq(0,0)} \left(\frac{1}{(z-mri-n)^2} - \frac{1}{(mri+n)^2}\right)\\
&= \sum_{n\in\mathbb{Z}} \frac{1}{(z-n)^2} - 2\sum_{n>0}\frac{1}{n^2} + \sum_{m>0}\sum_{n\in\mathbb{Z}} \left(\frac{1}{(z-mri-n)^2} - \frac{1}{(mri+n)^2}\right)\\
&\qquad + \sum_{m<0}\sum_{n\in\mathbb{Z}} \left(\frac{1}{(z-mri-n)^2} - \frac{1}{(mri+n)^2}\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\left|\wp(z|ri,1) - \frac{\pi^2}{\sin^2(\pi z)} + 2\,\zeta(2)\right| &\le \sum_{m>0}\left|\sum_{n\in\mathbb{Z}} \frac{1}{(z-mri-n)^2} - \sum_{n\in\mathbb{Z}}\frac{1}{(mri+n)^2}\right|\\
&\qquad + \sum_{m<0}\left|\sum_{n\in\mathbb{Z}} \frac{1}{(z-mri-n)^2} - \sum_{n\in\mathbb{Z}}\frac{1}{(mri+n)^2}\right|\\
&= \sum_{m>0}\left|\frac{\pi^2}{\sin^2(\pi z-\pi mri)} - \frac{\pi^2}{\sin^2(\pi mri)}\right|\\
&\qquad + \sum_{m<0}\left|\frac{\pi^2}{\sin^2(\pi z-\pi mri)} - \frac{\pi^2}{\sin^2(\pi mri)}\right|\\
&< \sum_{m>0}\frac{\pi^2}{|\sin^2(\pi z-\pi mri)|} + \sum_{m>0}\frac{\pi^2}{|\sin^2(\pi mri)|}\\
&\qquad + \sum_{m<0}\frac{\pi^2}{|\sin^2(\pi z-\pi mri)|} + \sum_{m<0}\frac{\pi^2}{|\sin^2(\pi mri)|}.
\end{aligned}$$

For $m > 0$, we have a simple estimate

$$\begin{aligned}
\sum_{m>0}\frac{1}{|\sin^2(\pi mri)|} &= \sum_{m>0}\frac{1}{\sinh^2(\pi mr)}\\
&< \frac{1}{\sinh^2(\pi r)} + \int_1^\infty \frac{dx}{\sinh^2(\pi rx)}
\end{aligned}$$

$$= \frac{1}{\sinh^2(\pi r)} + \frac{\coth(\pi r) - 1}{\pi r} \underset{r\to\infty}{\longrightarrow} 0.$$

The same is true for $m < 0$. To establish an estimate of the terms that are dependent on $z = x + iy$, let us impose the following restrictions:

$$0 \le Re(z) = x < 1, \qquad -\frac{r}{2} < Im(z) = y < \frac{r}{2}. \tag{1.27}$$

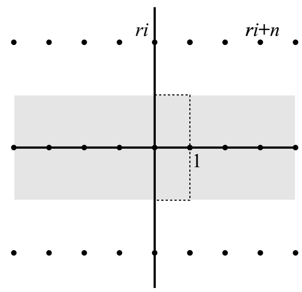

FIGURE 1.6. Degeneration of a lattice to the integral points on the real axis.

Since all functions involved have period 1, the condition for the real part is not a restriction. We wish to show that

$$\lim_{r\to\infty} \sum_{m>0} \frac{\pi^2}{|\sin^2(\pi z - \pi m r i)|} = 0$$

uniformly on every compact subset of (1.27).

$$\begin{aligned}
\sum_{m>0} \frac{1}{|\sin^2(\pi z - \pi m r i)|} &= \sum_{m>0} \frac{4}{|e^{\pi i z} e^{\pi m r} - e^{-\pi i z} e^{-\pi m r}|^2} \\
&= \sum_{m>0} e^{-2\pi m r} \frac{4}{|e^{\pi i z} - e^{-\pi i z} e^{-2\pi m r}|^2} \\
&\le \sum_{m>0} e^{-2\pi m r} \frac{4}{(e^{-\pi y} - e^{\pi y} e^{-2\pi m r})^2} \\
&< \sum_{m>0} e^{-2\pi m r} \frac{4}{(e^{-\pi y} - e^{-\pi r})^2} \\
&= \frac{e^{-2\pi r}}{1 - e^{-2\pi r}} \cdot \frac{4}{(e^{-\pi y} - e^{-\pi r})^2} \underset{r\to\infty}{\longrightarrow} 0.
\end{aligned}$$

A similar estimate holds for $m < 0$. We have thus established the convergence

$$\lim_{r\to\infty} \wp(z|ri, 1) = \frac{\pi^2}{\sin^2(\pi z)} - 2\,\zeta(2).$$

Let $f(z)$ denote this limiting function, $g_2 = 120\,\zeta(4)$, and $g_3 = 280\,\zeta(6)$. Then, as we certainly expect, the following differential equation holds:

$$(f'(z))^2 = 4f(z)^3 - g_2 f(z) - g_3 = 4\left(f(z) - \frac{2\pi^2}{3}\right) \cdot \left(f(z) + \frac{\pi^2}{3}\right)^2.$$

Geometrically, the elliptic curve becomes an infinitely long cylinder, but still the top circle and the bottom circle are glued together as one point. It is a singular algebraic curve given by the equation

$$Y^2 = 4\left(X - \frac{2\pi^2}{3}\right) \cdot \left(X + \frac{\pi^2}{3}\right)^2 .$$

We note that at the point $(-\pi^2/3, 0)$ of this curve, we cannot define the unique tangent line, which shows that it is a singular point.

1.6. **The Elliptic Modular Function.** The Eisenstein series g_2 and g_3 depend on both ω_1 and ω_2. However, the quotient

$$J(\tau) = \frac{g_2(\omega_1,\omega_2)^3}{g_2(\omega_1,\omega_2)^3 - 27g_3(\omega_1,\omega_2)^2} = \frac{\omega_2^{12}}{\omega_2^{12}} \cdot \frac{g_2(\tau,1)^3}{g_2(\tau,1)^3 - 27g_3(\tau,1)^2} \tag{1.28}$$

is a function depending only on $\tau = \omega_1/\omega_2 \in H$. This is what is called the **elliptic modular function**. From its definition it is obvious that $J(\tau)$ is invariant under the modular transformation

$$\tau \longmapsto \frac{a\tau + b}{c\tau + d} ,$$

where $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in PSL(2,\mathbb{Z})$. Since the Eisenstein series $g_2(\tau,1)$ and $g_3(\tau,1)$ are absolutely convergent, $J(\tau)$ is complex differentiable with respect to $\tau \in H$. Hence $J(\tau)$ is holomorphic on H, except for possible singularities coming from the zeros of the discriminant $g_2^3 - 27g_3^2$. However, we have already shown that the cubic curve (1.23) is non-singular, and hence $g_2^3 - 27g_3^2 \neq 0$ for any $\tau \in H$. Therefore, $J(\tau)$ is indeed holomorphic everywhere on H.

To compute a few values of $J(\tau)$, let us calculate $g_3(i,1)$. First, let

$$\Lambda = \{mi + n \mid m \geq 0, n > 0, (m,n) \neq (0,0)\}.$$

Since the square lattice $\Lambda_{i,1}$ has $90°$ rotational symmetry, it is partitioned into the disjoint union of the following four pieces:

$$\Lambda_{i,1} \setminus \{0\} = \Lambda \cup i\Lambda \cup i^2\Lambda \cup i^3\Lambda.$$

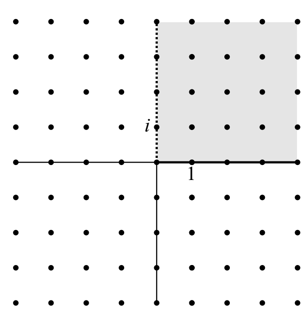

FIGURE 1.7. A partition of the square lattice $\Lambda_{i,1}$ into four pieces.

Thus we have

$$\frac{1}{140} g_3(i,1) = \sum_{\substack{m,n\in\mathbb{Z} \\ (m,n)\neq(0,0)}} \frac{1}{(mi+n)^6}$$

$$= \left(1 + \frac{1}{i^6} + \frac{1}{i^{12}} + \frac{1}{i^{18}}\right) \sum_{\substack{m\geq 0, n>0\\(m,n)\neq(0,0)}} \frac{1}{(mi+n)^6}$$
$$= 0\ .$$

Similarly, let $\omega = e^{\pi i/3}$. Since $\omega^6 = 1$, the honeycomb lattice $\Lambda_{\omega,1}$ has $60°$ rotational symmetry. Let

$$L = \{m\omega + n \mid m \geq 0, n > 0, (m,n) \neq (0,0)\}.$$

Due to the $60°$ rotational symmetry, the whole honeycomb is divided into the disjoint union of six pieces:

$$\Lambda_{\omega,1} \setminus \{0\} = L \cup \omega L \cup \omega^2 L \cup \omega^3 L \cup \omega^4 L \cup \omega^5 L.$$

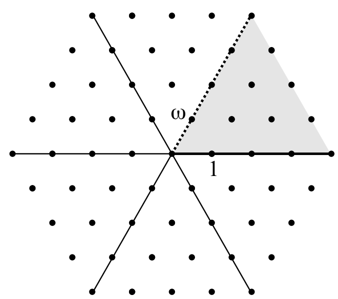

FIGURE 1.8. A partition of the honeycomb lattice $\Lambda_{\omega,1}$ into six pieces.

Therefore,

$$\frac{1}{60} g_2(\omega,1) = \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{(m\omega+n)^4}$$
$$= \left(1 + \frac{1}{\omega^4} + \frac{1}{\omega^8} + \frac{1}{\omega^{12}} + \frac{1}{\omega^{16}} + \frac{1}{\omega^{20}}\right) \sum_{\substack{m\geq 0, n>0\\(m,n)\neq(0,0)}} \frac{1}{(m\omega+n)^4}$$
$$= (1 + \omega^2 + \omega^4 + 1 + \omega^2 + \omega^4) \sum_{\substack{m\geq 0, n>0\\(m,n)\neq(0,0)}} \frac{1}{(m\omega+n)^4}$$
$$= 0.$$

We have thus established

$$J(i) = \frac{g_2(i,1)^3}{g_2(i,1)^3 - 27 g_3(i,1)^2} = 1,$$
$$J(e^{2\pi i/3}) = J(\omega) = \frac{g_2(\omega,1)^3}{g_2(\omega,1)^3 - 27 g_3(\omega,1)^2} = 0. \tag{1.29}$$

Moreover, we see that $J(\tau) - 1$ has a double zero at $\tau = i$, and $J(\tau)$ has a triple zero at $\tau = e^{2\pi i/3}$. This is consistent with the fact that i and $e^{2\pi i/3}$ are the fixed points of the $PSL(2,\mathbb{Z})$-action on H, with an order 2 stabilizer subgroup at i and an order 3 stabilizer subgroup at $e^{2\pi i/3}$.

Another value of $J(\tau)$ we can calculate is the value at the infinity $i\infty$:

$$J(i\infty) = \lim_{r\to+\infty} J(ri) = \lim_{r\to+\infty} \frac{g_2(ri,1)^3}{g_2(ri,1)^3 - 27g_3(ri,1)^2} = \infty.$$

The following theorem is a fundamental result.

Theorem 1.12 (Properties of J). 1. *The elliptic modular function*

$$J : H \longrightarrow \mathbb{C}$$

is a surjective holomorphic function which defines a bijective holomorphic map

$$H/PSL(2,\mathbb{Z}) \cup \{i\infty\} \longrightarrow \mathbb{P}^1. \tag{1.30}$$

2. *Two elliptic curves E_τ and $E_{\tau'}$ are isomorphic if and only if*

$$J(\tau) = J(\tau').$$

Remark. The bijective holomorphic map (1.30) is *not* biholomorphic. Indeed, as we have already observed, the expansion of J^{-1} starts with $\sqrt[3]{z}$ at $z = 0 \in \mathbb{P}^1$, and starts with $\sqrt{z-1}$ at $z = 1$. From this point of view, the moduli space $\mathfrak{M}_{1,1}$ is not isomorphic to $\mathbb{C}$.

In order to prove Theorem 1.12, first we parametrize the structure of an elliptic curve E_τ in terms of the branched points of the double covering

$$\wp : E_\tau \longrightarrow \mathbb{P}^1.$$

We then re-define the elliptic modular function J in terms of the branched points. The statements follow from this new description of the modular invariant.

Let us begin by determining the holomorphic automorphisms of $\mathbb{P}^1$. Since

$$\mathbb{P}^1 = \big(\mathbb{C}^2 \setminus (0,0)\big)/\mathbb{C}^\times,$$

where $\mathbb{C}^\times = \mathbb{C}\setminus\{0\}$ is the multiplicative group of complex numbers, we immediately see that

$$PGL(2,\mathbb{C}) = GL(2,\mathbb{C})/\mathbb{C}^\times$$

is a subgroup of $\mathrm{Aut}(\mathbb{P}^1)$. In terms of the coordinate z of $\mathbb{P}^1 = \mathbb{C} \cup \{\infty\}$, the $PGL(2,\mathbb{C})$-action is described again as linear fractional transformation:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot z = \frac{az+b}{cz+d}\ . \tag{1.31}$$

Let $f \in \mathrm{Aut}(\mathbb{P}^1)$. Since (1.31) can bring any point z to ∞, by composing f with a linear fractional transformation, we can make the automorphism fix ∞. Then this automorphism is an affine transformation, since it is in $\mathrm{Aut}(\mathbb{C})$. Therefore, we have shown that

$$\mathrm{Aut}(\mathbb{P}^1) = PGL(2,\mathbb{C}).$$

We note that the linear fractional transformation (1.31) brings 0, 1, and ∞ to the following three points:

$$\begin{cases} 0 \longmapsto \frac{b}{d} \\ 1 \longmapsto \frac{a+b}{c+d} \\ \infty \longmapsto \frac{a}{c}. \end{cases}$$

Since the only condition for a, b, c and d is $ad-bc \neq 0$, it is easy to see that 0, 1 and ∞ can be brought to any three distinct points of $\mathbb{P}^1$. In other words, $PGL(2,\mathbb{C})$ acts on $\mathbb{P}^1$ **triply transitively**.

Now consider an elliptic curve E defined by a cubic equation

$$Y^2 Z = 4X^3 - g_2 X Z^2 - g_3 Z^3$$

in $\mathbb{P}^2$, and the projection to the X-coordinate line

$$p : E \ni \begin{cases} (X:Y:Z) \longmapsto (X:Z) \in \mathbb{P}^1 & Z \neq 0, \\ (0:1:0) \longmapsto (1:0) \in \mathbb{P}^1. \end{cases}$$

Note that we are assuming that $g_2^3 - 27g_3^2 \neq 0$. As a coordinate of $\mathbb{P}^1$, we use $x = X/Z$. As before, let

$$4x^3 - g_2 x - g_3 = 4(x - e_1)(x - e_2)(x - e_3).$$

Then the double covering p is ramified at e_1, e_2, e_3 and ∞. Since these four points are distinct, we can bring three of them to 0, 1, and ∞ by an automorphism of $\mathbb{P}^1$. The fourth point cannot be brought to a prescribed location, so let λ be the fourth branched point under the action of this automorphism. In particular, we can choose

$$\lambda = \frac{e_3 - e_2}{e_3 - e_1}. \tag{1.32}$$

This is the image of e_3 via the transformation

$$x \longmapsto \frac{x - e_2}{x - e_1}.$$

This transformation maps

$$\begin{cases} e_1 \longmapsto \infty \\ e_2 \longmapsto 0 \\ e_3 \longmapsto \lambda \\ \infty \longmapsto 1 \, . \end{cases}$$

Noting the relation $e_1 + e_2 + e_3 = 0$, a direct calculation shows

$$\begin{aligned} J(\tau) &= \frac{g_2^3}{g_2^3 - 27g_3^2} \\ &= \frac{-64(e_1e_2 + e_2e_3 + e_3e_1)^3}{16(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2} \\ &= \frac{4}{27} \frac{\big(-3(e_1e_2 + e_2e_3 + e_3e_1)\big)^3}{(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2} \\ &= \frac{4}{27} \frac{\big((e_1 + e_2 + e_3)^2 - 3(e_1e_2 + e_2e_3 + e_3e_1)\big)^3}{(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2} \\ &= \frac{4}{27} \frac{\big(e_1^2 + e_2^2 + e_3^2 - (e_1e_2 + e_2e_3 + e_3e_1)\big)^3}{(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2} \\ &= \frac{4}{27} \frac{\big((e_3 - e_2)^2 - (e_3 - e_2)(e_3 - e_1) + (e_3 - e_1)^2)\big)^3}{(e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2} \\ &= \frac{4}{27} \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2}. \end{aligned}$$

Of course naming the three roots of the cubic polynomial is arbitrary, so the definition of λ (1.32) receives the action of the **symmetric group** $\mathfrak{S}_3$. We could have chosen any one of the following six choices as our λ:

$$\begin{aligned} \lambda = \frac{e_3 - e_2}{e_3 - e_1}, \quad \frac{1}{\lambda} = \frac{e_3 - e_1}{e_3 - e_2}, \quad 1 - \frac{1}{\lambda} = \frac{e_2 - e_1}{e_2 - e_3}, \\ \frac{\lambda}{\lambda - 1} = \frac{e_2 - e_3}{e_2 - e_1}, \quad 1 - \lambda = \frac{e_1 - e_2}{e_1 - e_3}, \quad \frac{1}{1-\lambda} = \frac{e_1 - e_3}{e_1 - e_2}. \end{aligned} \tag{1.33}$$

Since the rational map

$$\mu = j(\lambda) = \frac{4}{27} \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2} \tag{1.34}$$

is a symmetric function of e_1, e_2, e_3, it has the same value for any of the six choices (1.33). The rational map j has degree 6, and hence the inverse image $j^{-1}(\mu)$ of $\mu \in \mathbb{C}$ exactly coincides with the 6 values given above. The value $j(\lambda)$ of the elliptic curve is called the **j-invariant**.

Lemma 1.13. *Let E (resp. E') be an elliptic curve constructed as a double covering of $\mathbb{P}^1$ ramified at 0, 1, ∞, and λ (resp. λ'). Suppose $j(\lambda) = j(\lambda')$. Then E and E' are isomorphic.*

Proof. Since $j : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ is a ramified covering of degree 6, $j(\lambda) = j(\lambda')$ implies that λ' is one of the 6 values listed in (1.33). Now let us bring back the four ramification points 0, 1, ∞, and λ to e_1, e_2, e_3, and ∞ by solving two linear equations

$$\lambda = \frac{e_3 - e_2}{e_3 - e_1} \qquad \text{and} \qquad e_1 + e_2 + e_3 = 0. \tag{1.35}$$

The solution is unique up to an overall constant factor, which does not affect the value

$$j(\lambda) = \frac{4}{27} \frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(\lambda - 1)^2} = \frac{g_2^3}{g_2^3 - 27g_3^2}.$$

We also note that (e_1, e_2, e_3) and their constant multiple (ce_1, ce_2, ce_3) define an isomorphic elliptic curve. So choose a particular solution e_1, e_2 and e_3 of (1.35). Then the difference between λ and λ' is just a permutation of e_1, e_2 and e_3. In particular, the defining cubic equation of the elliptic curve, which is symmetric under permutation of e_1, e_2 and e_3, is exactly the same. Thus E and E' are isomorphic. □

We are now ready to prove Theorem 1.12.

Proof. First, take an arbitrary $\mu \in \mathbb{C}$, and let λ be a point in the inverse image of μ via the map j. We can construct a cubic curve E as a double cover of $\mathbb{P}^1$ ramified at 0, 1, ∞, and λ. Since it is a Riemann surface of genus 1, it is isomorphic to a particular elliptic curve E_τ for some $\tau \in H$. Realize E_τ as a cubic curve, and choose its ramification points 0, 1, ∞, and λ'. Here we have applied an automorphism of $\mathbb{P}^1$ to choose this form of the ramification point. Since E and E_τ are isomorphic, we have

$$J(\tau) = j(\lambda') = j(\lambda) = \mu.$$

This establishes that $J : H \longrightarrow \mathbb{C}$ is surjective. We have already established that $J(\tau) = J(\tau')$ implies the isomorphism $E_\tau \cong E_{\tau'}$, by translating the equation into λ-values. This fact also shows that the map

$$J : H\big/PSL(2,\mathbb{Z}) \longrightarrow \mathbb{C}$$

is one-to-one.

We have shown that $J(i\infty) = \infty$, but we have not seen how the modular function behaves at infinity. This is our final subject of this section, which completes the proof of Theorem 1.12.

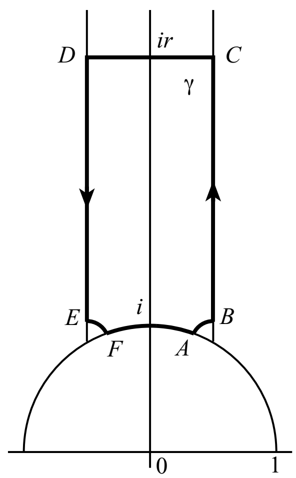

FIGURE 1.9. The boundary of the fundamental domain as an integration contour.

Let γ be the contour defined in Figure 1.9. It is the boundary of a fundamental domain of the $PSL(2,\mathbb{Z})$-action, except for an arc $\widehat{AB}$, a line segment $\overline{CD}$, and another arc $\widehat{EF}$. Since $J : H\big/PSL(2,\mathbb{Z}) \longrightarrow \mathbb{C}$ is a bijective holomorphic map, and since $J(e^{2\pi i/3}) = 0$, J does not have any other zeros in the fundamental domain. Therefore, $d \log J(\tau)$ is holomorphic everywhere inside the contour γ, and we have

$$\oint_\gamma d \log J(\tau) = 0.$$

By differentiating the equation $J(\tau) = J\big((a\tau + b)/(c\tau + d)\big)$, we obtain

$$J'(\tau) = J'\left(\frac{a\tau + b}{c\tau + d}\right)\frac{1}{(c\tau + d)^2}.$$

Hence

$$\begin{aligned} dJ(\tau) = J'(\tau)d\tau &= J'\left(\frac{a\tau + b}{c\tau + d}\right)\frac{1}{(c\tau + d)^2}d\tau = J'\left(\frac{a\tau + b}{c\tau + d}\right) d\left(\frac{a\tau + b}{c\tau + d}\right) \\ &= dJ\left(\frac{a\tau + b}{c\tau + d}\right). \end{aligned}$$

(The exterior differentiation d and the integer d should not be confused.) It follows that

$$d\log J(\tau) = d\log J\left(\frac{a\tau+b}{c\tau+d}\right).$$

Therefore, we have

$$\int_B^C d\log J(\tau) + \int_D^E d\log J(\tau) = 0$$
$$\int_F^i d\log J(\tau) + \int_i^A d\log J(\tau) = 0.$$

Next, since $J(e^{2\pi i/3}) = 0$ is a zero of order 3, the integral of $d\log J(\tau)$ around $e^{2\pi i/3}$ is given by

$$\oint d\log J(\tau) = 6\pi i.$$

The arcs $\widehat{AB}$ and $\widehat{EF}$ joined together form a third of a small circle going around $e^{2\pi i/3}$ clockwise. Therefore, we have

$$\int_A^B d\log J(\tau) + \int_E^F d\log J(\tau) = -\frac{1}{3}\oint d\log J(\tau) = -2\pi i.$$

Thus we are left with the integration along the line segment $\overline{CD}$.

In order to study the behavior of $J(\tau)$ as $\tau \longrightarrow i\infty$, we introduce a new varialbe $q = e^{2\pi i\tau}$. Since $J(\tau+1) = J(\tau)$, the elliptic modular function admits a **Fourier series expansion** in terms of $q = e^{2\pi i\tau}$. So let

$$f(q) = f(e^{2\pi i\tau}) = J(\tau)$$

be the Fourier expansion of $J(\tau)$. Note that

$$\frac{J'(\tau)}{J(\tau)}d\tau = d\log J(\tau) = d\log f(q) = \frac{f'(q)}{f(q)}dq.$$

The points $C = 1/2+ir$ and $D = -1/2+ir$ in τ-coordinate transform into $e^{\pi i}e^{-2\pi r}$ and $e^{-\pi i}e^{-2\pi r}$ in q-coordinate, respectively. Therefore, the path $\overline{CD}$ is a loop of radius $e^{-2\pi r}$ around $q = 0$ with the counter clockwise orientation in q-coordinate. Thus we have

$$\int_C^D d\log J(\tau) = \int_{e^{\pi i}e^{-2\pi r}}^{e^{-\pi i}e^{-2\pi r}} d\log f(q) = -\oint d\log f(q) = 2\pi i n,$$

where n is the order of the pole of $f(q)$ at $q = 0$.

Altogether, we have established

$$0 = \oint_\gamma d\log J(\tau) = 2\pi i n - 2\pi i = 2\pi i(n-1).$$

Therefore, we conclude $n = 1$. Hence $f(q)$ has a simple pole at $q = 0$, or $\tau = i\infty$. In other words, the map

$$J : H/PSL(2,\mathbb{Z}) \cup \{i\infty\} \longrightarrow \mathbb{P}^1$$

is holomorphic around the point $i\infty$. This completes the proof of Theorem 1.12. $\square$

The first few terms of the q-expansion of $J(\tau)$ are given by

$$J(\tau) = \frac{1}{1728}(q^{-1} + 744 + 196884q + 21493760q^2 + \cdots).$$

We refer to [9] for the story of these coefficients, the *Monstrous Moonshine*, and its final mathematical outcome.

1.7. **Compactification of the Moduli of Elliptic Curves.** We have introduced two different ways to parametrize the moduli space $\mathfrak{M}_{1,1}$ of elliptic curves. The first one is through the **period** $\tau \in H$ of an elliptic curve, and the other via the fourth **ramification point** $\lambda \in \mathbb{P}^1 \setminus \{0,1,\infty\}$ when we realize an elliptic curve as a double cover over $\mathbb{P}^1$ ramified at $0, 1, \infty$ and λ. The equality we have proven,

$$J(\tau) = \frac{g_2(\tau,1)^3}{g_2(\tau,1)^3 - 27g_3(\tau,1)^2} = \frac{4}{27}\,\frac{\lambda^2-\lambda+1}{\lambda^2(1-\lambda)^2},$$

gives two holomorphic fibrations over $\mathbb{C}$:

$$\begin{array}{ccc} & & \mathbb{P}^1 \setminus \{0,1,\infty\} \\ & & \quad j \Big\downarrow \mathfrak{S}_3\text{-action} \\ H & \xrightarrow[PSL(2,\mathbb{Z})\text{-action}]{J} & \mathbb{C}. \end{array} \tag{1.36}$$

We have also established that the function $j(\lambda)$ is invariant under the action of $\mathfrak{S}_3$, and the elliptic modular function $J(\tau)$ is invariant under the action of the modular group $PSL(2,\mathbb{Z})$.

It is intriguing to note the similarity of these two groups. In the presentation by generators and their relations, we have

$$\begin{aligned} PSL(2,\mathbb{Z}) &= \langle S, T \,|\, S^2 = (ST)^3 = 1\rangle, \\ \mathfrak{S}_3 &= \langle s, t \,|\, s^2 = t^2 = (st)^3 = 1\rangle. \end{aligned} \tag{1.37}$$

Therefore, there is a natural surjective homomorphism

$$h : PSL(2,\mathbb{Z}) \longrightarrow \mathfrak{S}_3 \tag{1.38}$$

defined by $h(S) = s$ and $h(T) = t$. The kernel $\mathrm{Ker}(h)$ is a normal subgroup of $PSL(2,\mathbb{Z})$ of index 6.

Proposition 1.14 (Congruence subgroup modulo 2)**.** *The kernel* $\mathrm{Ker}(h)$ *of the homomorphism* $h : PSL(2,\mathbb{Z}) \longrightarrow \mathfrak{S}_3$ *is equal to the* ***congruence subgroup*** *of* $PSL(2,\mathbb{Z})$ ***modulo*** *2:*

$$\mathrm{Ker}(h) = \Gamma(2) \underset{def}{=} \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in PSL(2,\mathbb{Z}) \;\middle|\; \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mod 2 \right\}.$$

In particular, we have an isomorphism

$$PSL(2\mathbb{Z})\big/\Gamma(2) \cong \mathfrak{S}_3.$$

Proof. Let

$$\begin{aligned} A = T^2 &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \\ B = ST^{-2}S &= \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}. \end{aligned}$$

Obviously A and B are elements of both $\mathrm{Ker}(h)$ and $\Gamma(2)$. First let us show that A and B generate $\Gamma(2)$:

$$\Gamma(2) = \langle A, B\rangle.$$

The condition $ad - bc = 1$ means that a and b are relatively prime, and the congruence condition

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mod 2$$

means that a and d are odd and b and c are even. Since the multiplication of the matrix A^n from the right to $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ changes b to $2na + b$, by a suitable choice of the power n, we can make $|b| < |a|$. (They cannot be equal because a is odd and b is even.) On the other hand, the multiplication of B^m from the right changes a to $a + 2mb$. Thus by a suitable choice of the power of B, we can make $|a| < |b|$. Hence by consecutive multiplications of suitable powers of A and B from the right, $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma(2)$ is brought to the form $\begin{bmatrix} 1 & b' \\ c' & d' \end{bmatrix}$, which is still an element of $\Gamma(2)$. Thus b' is even, and hence further application of $A^{-b'/2}$ from the right brings the matrix to $\begin{bmatrix} 1 & 0 \\ c' & * \end{bmatrix}$. The determinant condition dictates that $* = 1$. Since c' is also even, $\begin{bmatrix} 1 & 0 \\ c' & 1 \end{bmatrix} = B^{c'/2}$. Hence $\Gamma(2)$ is generated by A and B. In particular, $\Gamma(2) \subset \mathrm{Ker}(h)$.

Next let us determine the index of $\Gamma(2)$ in $PSL(2\mathbb{Z})$. The method of exhaustive listing works here. As a representative of the coset $PSL(2\mathbb{Z})/\Gamma(2)$, we can choose

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}.$$

Therefore, $\Gamma(2)$ is an index 6 subgroup of $PSL(2\mathbb{Z})$. It implies that $\Gamma(2) = \mathrm{Ker}(h)$. This completes the proof. □

Figure 1.10 shows a fundamental domain of the $\Gamma(2)$-action on the upper half plane H.

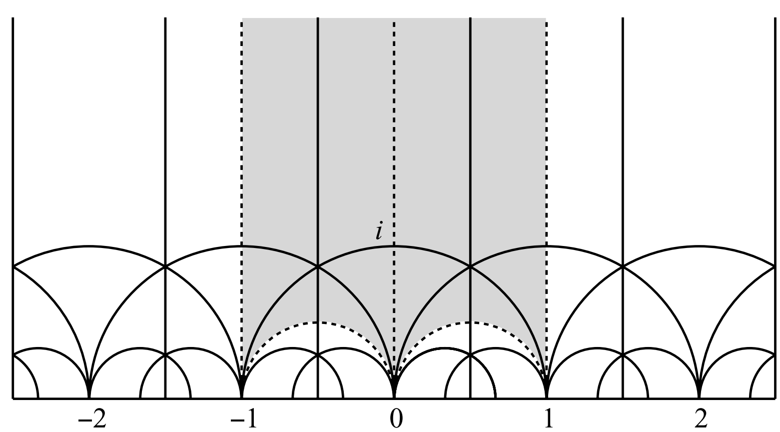

FIGURE 1.10. A fundamental domain of the action of the congruence subgroup $\Gamma(2) \subset PSL(2,\mathbb{Z})$.

We observe that the line $Re(\tau) = -1$ is mapped to the line $Re(\tau) = 1$ by $A = T^2 \in \Gamma(2)$, and the semicircle connecting -1, $\frac{-1+i}{2}$ and 0 is mapped to the semicircle connecting 1, $\frac{1+i}{2}$ and 0 by $B = ST^{-2}S \in \Gamma(2)$. Gluing these dotted lines and semicircles, we obtain a sphere minus three points. Because of the triple transitivity of $\mathrm{Aut}(\mathbb{P}^1)$, we know that $\mathfrak{M}_{0,3}$ consists of only one point:

$$\mathfrak{M}_{0,3} = \big\{\big(\mathbb{P}^1, (0, 1, \infty)\big)\big\}. \tag{1.39}$$

Therefore, we can identity

$$H\big/\Gamma(2) \cong \mathbb{P}^1 \setminus \{0, 1, \infty\}.$$

We now have a commutative diagram that completes (1.36).

$$\begin{array}{ccc} H & \xrightarrow{\Gamma(2)\text{-action}} & \mathbb{P}^1 \setminus \{0,1,\infty\} \\ \Big\| & & j \Big\downarrow \mathfrak{S}_3\text{-action} \\ H & \xrightarrow[PSL(2,\mathbb{Z})\text{-action}]{J} & \mathbb{P}^1 \setminus \{\infty\}. \end{array} \tag{1.40}$$

Let us study the geometry of the map

$$j : \mathbb{P}^1 \ni \lambda \longmapsto \frac{4}{27}\frac{(\lambda^2 - \lambda + 1)^3}{\lambda^2(1-\lambda)^2} \in \mathbb{P}^1.$$

We see that $j^{-1}(\infty) = \{0, 1, \infty\}$, and that each of the three points has multiplicity 2. To see the ramification of j at 0 and 1, let us consider the inverse image of the closed real interval $[0, 1]$ on the target $\mathbb{P}^1$ via j. Figure 1.11 shows $j^{-1}([0, 1])$. The shape is the union of two circles of radius 1 centered at 0 and 1, intersecting at $e^{\pi i/3}$ and $e^{-\pi i/3}$ with a 120° angle. The inverse image $j^{-1}([0, 1])$ also contains the vertical line segment $\overline{e^{\pi i/3}e^{-\pi i/3}}$. Each point of $j^{-1}(\infty) = \{0, 1, \infty\}$ is surrounded by a **bigon**, representing the fact that the ramification at each point is of degree 2. Each point of $j^{-1}(0) = \{e^{\pi i/3}, e^{-\pi i/3}\}$ has multiplicity 3, which can be seen by the **tri-valent** vertex of the **graph** $j^{-1}([0, 1])$ at $e^{\pi i/3}$ and $e^{-\pi i/3}$. And finally, each point of $j^{-1}(1) = \{-1, \frac{1}{2}, 2\}$ has multiplicity 2 and is located at the middle of an edge of the graph.

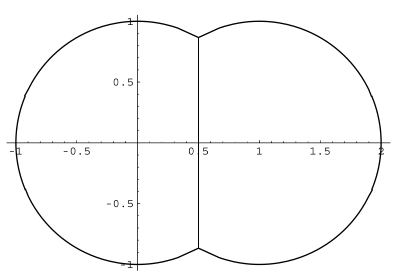

FIGURE 1.11. The inverse image of $[0, 1]$ via the map $j(\lambda) = \frac{4}{27}\frac{\lambda^2-\lambda+1}{\lambda^2(1-\lambda)^2}$. (Graphics produced by Josephine Yu.)

More geometrically, consider $\mathbb{P}^1$ as a sphere with its real axis as the equator, and $\omega = e^{\pi i/3}$ and $\omega^{-1} = e^{-\pi i/3}$ as the north and the south poles. Then we can see that the $\mathfrak{S}_3$-action on $\mathbb{P}^1$ is equivalent to the action of the **dihedral group** D_3 on the equilateral triangle $\triangle 01\infty$. It becomes obvious that $\omega = e^{\pi i/3}$ and $\omega^{-1} = e^{-\pi i/3}$ are stabilized by the action of the cyclic group $\mathbb{Z}/3\mathbb{Z}$ through the 120° rotations about the axis connecting the poles, and each of $0, 1, \infty$ and $-1, \frac{1}{2}, 2$ is invariant under the 180° rotation about a diameter of the equator.

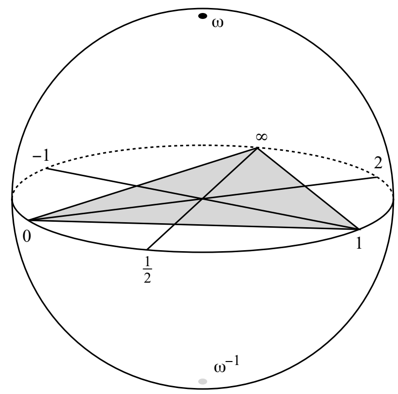

FIGURE 1.12. The $\mathfrak{S}_3$-action on $\mathbb{P}^1$ through the dihedral group action.

Since the $PSL(2,\mathbb{Z})$-action on H factors through the $\Gamma(2)$-action and the $\mathfrak{S}_3$-action, we have the equality

$$\mathfrak{M}_{1,1} = H/PSL(2,\mathbb{Z}) = \Big(H/\Gamma(2)\Big)\Big/\mathfrak{S}_3 = \Big(\mathbb{P}^1 \setminus \{0,1,\infty\}\Big)\Big/\mathfrak{S}_3.$$

At this stage, we can *define* a **compactification** of the moduli space $\mathfrak{M}_{1,1}$ by

$$\overline{\mathfrak{M}_{1,1}} = \mathbb{P}^1/\mathfrak{S}_3. \tag{1.41}$$

Since $\mathfrak{S}_3$ acts on $\mathbb{P}^1$ properly discontinuously, the quotient is again an orbifold. The stabilizer subgroup at $e^{\pi i/3}$ is $\mathbb{Z}/3\mathbb{Z}$, and the stabilizer subgroup at $\frac{1}{2}$ is $\mathbb{Z}/2\mathbb{Z}$. Therefore, the orbifold structure of $\mathbb{P}^1/\mathfrak{S}_3$ at its singular points $j = 0$ and $j = 1$ is exactly the same as we have observed before. (We refer to Chapter **??** for the definition of orbifolds and the terminology from the orbifold theory.)

However, there is a big difference in the singularity structure at ∞. From (1.41), the compactified moduli space has the quotient singularity modeled by the $\mathbb{Z}/2\mathbb{Z}$-action on $\mathbb{P}^1$ at ∞. On the other hand, as we have seen in the last section, the elliptic modular function $J(\tau)$ has a simple pole at $q = 0$ in terms of the variable $q = e^{2\pi i\tau}$. This shows that the moduli space has a compactification

$$\overline{\mathfrak{M}_{1,1}} = H/PSL(2\mathbb{Z}) \cup \{i\infty\} \xrightarrow[\text{holomorphic and bijective}]{J} \mathbb{C} \cup \{\infty\} = \mathbb{P}^1,$$

and that J^{-1} is a holomorphic map at $\infty \in \mathbb{P}^1$. How do we reconcile this difference?

This is due to the fact that **the upper half plane** H, which is isomorphic to the unit open disk $\{z \in \mathbb{C} \mid |z| < 1\}$ by the Riemann mapping theorem, **does not have any natural compactification as a Riemann surface**. Therefore we cannot take the compactification of H before taking the quotient by the modular group $PSL(2,\mathbb{Z})$. The point $\{i\infty\}$, called the **cusp point**, is added only after taking the full quotient. But if we take another route by first constructing the quotient by a normal subgroup such as the congruence subgroup $\Gamma(2)$, then we can add *three* points to compactify the quotient space. The moduli space in question is the quotient of this intermediate quotient space by the action of the factor group $PSL(2,\mathbb{Z})\big/\Gamma(2) = \mathfrak{S}_3$. In this second construction, we end up with a compact orbifold with a singularity at ∞. The moduli space $\mathfrak{M}_{1,1}$ is an infinite cylinder near ∞. Therefore, depending on when we compactify it, the point at infinity can be an orbifold singularity modeled by any $\mathbb{Z}\big/n\mathbb{Z}$-action on the complex plane.

Thus we note that **the moduli space $\mathfrak{M}_{1,1}$ does not have a canonical orbifold compactification**. The point at infinity can be added as a non-singular point, or as a $\mathbb{Z}\big/2\mathbb{Z}$-singular point, or in many other different ways. We also note that **if we wish to consider the compactified moduli space of elliptic curves as an algebraic variety, then the natural identification is**

$$\overline{\mathfrak{M}_{1,1}} \cong \mathbb{P}^1$$

without any singularities. Its complex structure is introduced by the modular function $J(\tau)$.

For a higher genus, the situation becomes far more complex. Compactification of $\mathfrak{M}_{g,n}$ as an algebraic variety is no longer unique, and compactification as an orbifold is even more non-unique. In the later chapters, we consider the **canonical orbifold structure of the non-compact moduli space $\mathfrak{M}_{g,n}$**. It is still an open question to find an orbifold compactification of $\mathfrak{M}_{g,n}$ with an orbifold cell-decomposition that restricts to the canonical orbifold cell-decomposition of the moduli space.

References

[1] George Andrews, Richard Askey, and Ranjan Roy, *Special functions*, Encyclopedia of Mathematics and its Applications **71**, Cambridge University Press (1999).

[2] Enrico Arbarello, M. Cornalba, Phillip Griffiths, and Joseph Harris, *Geometry of algebraic curves* volume 1, Grundlehren der mathematischen Wissenschaften **267**, Springer-Verlag (1985).

[3] Enrico Arbarello and C. De Concini, *On a set of equations characterizing the Riemann matrices*, Annals of Mathematics **120** (1984), 119–140.

[4] G. V. Belyi, *On galois extensions of a maximal cyclotomic fields*, Math. U.S.S.R. Izvestija **14** (1980), 247–256.

[5] D. Bessis, C. Itzykson and J. B. Zuber, *Quantum field theory techniques in graphical enumeration*, Advanced in Applied Mathematics **1** (1980), 109–157.

[6] Pierre Deligne and David Mumford, *The irreducibility of the space of curves of given genus*, Publ. Math. I.H.E.S. **86** (1969), 75–110.

[7] Leon Ehrenpreis and Robert C. Gunning, Editors, *Theta functions–Bowdoin 1987*, AMS Proceedings of Symposia in Pure Mathematics **49** (1989).

[8] Richard P. Feynman, *Space-time approach to quantum electrodynamics*, Physical Review **76** (1949), 769–789. Reprinted in [46].

[9] Igor Frenkel, James Lepowsky, and Arne Meurman, *Vertex operator algebras and the monster*, Academic Press (1988).

[10] Frederick P. Gardiner, *Teichmüller theory and quadratic differentials*, John Wiley & Sons (1987).

[11] Jonathan L. Gross and Thomas W. Tucker, *Topological graph theory*, John Wiley & Sons (1987).

[12] Alexander Grothendieck, *Esquisse d'un programme* (1984), reprinted in [44], 7–48.

[13] John L. Harer, *The virtual cohomological dimension of the mapping class group of an orientable surface*, Inventiones Mathematicae **84** (1986), 157–176.

[14] John L. Harer, *The cohomology of the moduli space of curves*, in Theory of Moduli, Montecatini Terme, 1985 (Edoardo Sernesi, ed.), Springer-Verlag, 1988, pp. 138–221.

[15] John L. Harer and Don Zagier, *The Euler characteristic of the moduli space of curves*, Inventiones Mathematicae **85** (1986), 457–485.

[16] Y. Imayoshi and M. Taniguchi, *An introduction to Teichmüller spaces*, Springer-Verlag (1992).

[17] Frances Kirwan, *Complex algebraic curves*, London Mathematical Society Student Text **23**, Cambridge University Press (1992).

[18] Neal Koblitz, *Introduction to Elliptic Curves and Modular Forms*, Graduate Texts in Mathematics, Vol. **97**, Second Edition, Springer-Verlag (1993).

[19] Kunihiko Kodaira, *Complex manifolds and deformation of complex structures*, Die Grundlehren der mathematischen Wissenschaften **283**, Springer-Verlag (1986).

[20] Maxim Kontsevich, *Intersection Theory on the Moduli Space of Curves and the Matrix Airy Function*, Communications in Mathematical Physics **147** (1992), 1–23.

[21] Maxim Kontsevich and Yuri Manin, *Gromov-Witten classes, quantum cohomology, and enumerative geometry*, Communications in Mathematical Physics **164** (1994), 525–562.

[22] Maxim Kontsevich and Yuri Manin, *Quantum cohomology of a product*, Inventiones Mathematicae **124** (1996), 313–339.

[23] Yingchen Li and Motohico Mulase *Prym varieties and integrable system*, Communications in Analysis and Geometry **5** (1997), 279–332.

[24] Eduard Looijenga, *Cellular decompositions of compactified moduli spaces of pointed curves*, in Moduli Space of Curves, R. H. Dijkgraaf et al., Editors, Birkhaeuser (1995), 369–400.

[25] Madan Lal Mehta, *Random matrices*, Second Edition, Academic Press (1991).

[26] Motohico Mulase, *Algebraic geometry of soliton equations*, Proceedings of the Japan Academy **59** (1983), 285–288.

[27] Motohico Mulase, *Cohomological structure in soliton equations and Jacobian varieties*, Journal of Differential Geometry **19** (1984), 403–430.

[28] Motohico Mulase, *KP equations, strings, and the Schottky problem*, in Algebraic Analysis **II**, Masaki Kashiwara and Takahiro Kawai, Editors, (1988) 473–492.

[29] Motohico Mulase, *Category of vector bundles on algebraic curves and infinite dimensional Grassmannians*, International Journal of Mathematics **1** (1990), 293–342.

[30] Motohico Mulase, *Algebraic theory of the KP equations*, in Perspectives in Mathematical Physics, R. Penner and S. T. Yau, Editors., Intern. Press Co. (1994) 157–223.

[31] Motohico Mulase, *Matrix integrals and integrable systems*, in Topology, geometry and field theory, K. Fukaya et al. Editors, World Scientific (1994), 111–127.

[32] Motohico Mulase, *Asymptotic analysis of a hermitian matrix integral*, International Journal of Mathematics **6** (1995), 881–892.

[33] Motohico Mulase, *Lectures on the Asymptotic Expansion of a Hermitian Matrix Integral*, in Supersymmetry and Integrable Models, Henrik Aratin et al., Editors, Springer Lecture Notes in Physics **502** (1998), 91–134.

[34] Motohico Mulase and Michael Penkava, *Ribbon graphs, quadratic differentials on Riemann surfaces, and algebraic curves defined over* $\overline{\mathbb{Q}}$, Asian Journal of Mathematics **2** (1998), 875–920.

[35] Motohico Mulase and Michael Penkava, *Periods of Strebel differentials and algebraic curves defined over the field of algebraic numbers*, UCD Preprint 2001.

[36] David Mumford, John Fogarty and Frances Kirwan, *Geometric Invariant Theory*, Third Edition, Springer-Verlag (1994).

[37] Subhashis Nag, *The complex analytic theory of Teichmüller spaces*, Wiley-Interscience Publications (1988).

[38] Robert C. Penner, *Perturbation series and the moduli space of Riemann surfaces*, Journal of Differential Geometry **27** (1988), 35–53.

[39] Bernhard Riemann, *Theorie der Abel'schen Functionen*, Journal für die reine und angewandte Mathematik **54** (1857), 115–155.

[40] Gerhard Ringel, *Map color theorem*, Die Grundlehren der mathematischen Wissenschaften **209**, Springer-Verlag (1974).

[41] Giovanni Sansone and Johan Gerretsen, *Lectures on the theory of functions of a complex variable*, Volume I and II, Wolters-Noordhoff Publishing, 1960 and 1969.

[42] Ichiro Satake, *The Gauss-Bonnet theorem for V-manifold*, Journal of the Mathematical Society of Japan **9** (1957), 464–492.

[43] Leila Schneps, *The grothendieck theory of dessins d'enfants*, London Mathematical Society Lecture Notes Series, vol. 200, 1994.

[44] Leila Schneps and Pierre Lochak, editors, *Geometric Galois actions: Around Grothendieck's esquisse d'un programme*, London Mathematical Society Lecture Notes Series, vol. 242, 1997.
[45] F. Schottky, *Zur Theorie der Abelschen Functionen von vier Variablen*, Journal für die reine und angewandte Mathematik **102** (1888), 304–352.
[46] Julian Schwinger, *Selected Papers on Quantum Electrodynamics*, Dover Publications (1958).
[47] Graeme Segal and George Wilson, *Loop groups and equations of KdV type*, Publ. Math. I.H.E.S. **61** (1985), 5–65.
[48] Takahiro Shiota, *Characterization of jacobian varieties in terms of soliton equations*, Inventiones Mathematicae **83** (1986), 333–382.
[49] Daniel D. Sleator, Robert E. Tarjan, and William P. Thurston, *Rotation distance, triangulations, and hyperbolic geometry*, Journal of the American Mathematical Society **1** (1988), 647–681.
[50] Kurt Strebel, *Quadratic differentials*, Springer-Verlag, 1984.
[51] Gerard 'tHooft", *A Planer Diagram Theory for Strong Interactions*, Nuclear Physics B **72** (1974), 461–473.
[52] William Thurston, *Three-dimensional geometry and topology, volume 1 and 2*, Princeton University Press, 1997, (volume 2 to be published).
[53] E. T. Whittaker and G. N. Watson, *A course of modern analysis*, Fourth Edition, Cambridge University Press (1927), Reprint (1999).
[54] Edward Witten, *Two dimensional gravity and intersection theory on moduli space*, Surveys in Differential Geometry **1** (1991), 243–310.
[55] L. Zapponi, *Dessins d'enfants et actions galoisiennes*, Thèse de Doctorat, Besançon, 1998.

Department of Mathematics, University of California, Davis, CA 95616–8633
E-mail address: `mulase@math.ucdavis.edu`