LECTURES ON THE COMBINATORIAL STRUCTURE OF THE MODULI SPACES OF RIEMANN SURFACES AND FEYNMAN DIAGRAM EXPANSION OF MATRIX INTEGRALS

MOTOHICO MULASE

Contents

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** $\phi_\mu^{-1} \circ \phi_\lambda$

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

Date: March 14, 2009.

is a biholomorphic function.

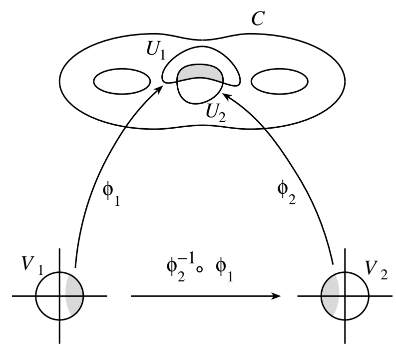

FIGURE 1.1. Gluing two coordinate charts.

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 $\epsilon \to 0$. 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.

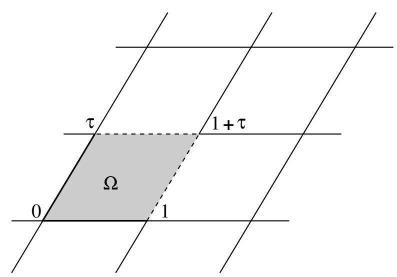

FIGURE 1.2. A fundamental domain of the Λ_τ-action on the complex plane.

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,

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

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

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

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 $H = \{\tau \in \mathbb{C} \bigm| Im(\tau) > 0\}$.

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 $\mathbb{C}$:

$$\begin{tikzcd} \mathbb{C} \arrow[r, "\widetilde{f}", "\sim"'] \arrow[d, "\pi_\tau"'] & \mathbb{C} \arrow[d, "\pi_\mu"] \\ E_\tau \arrow[r, "f", "\sim"'] & E_\mu. \end{tikzcd} \tag{1.8}$$

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*

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

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. □

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.

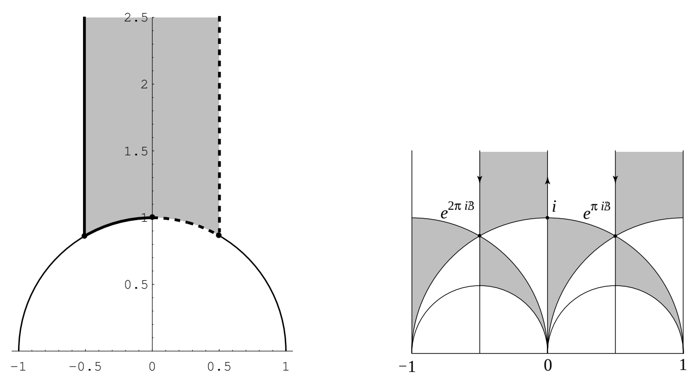

FIGURE 1.3. The fundamental domain of the $PSL(2,\mathbb{Z})$-action on the upper half plane H and the tiling of H by the $PSL(2,\mathbb{Z})$-orbits.

Since the transformation T maps $\tau \longmapsto \tau + 1$, the fundamental domain can be chosen as a subset of the vertical strip $\{\tau \in H \mid -1/2 \leq Re(\tau) < 1/2\}$. 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.

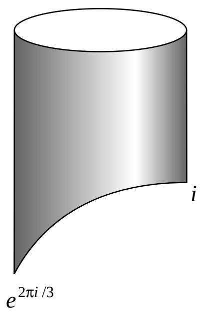

FIGURE 1.4. The moduli space $\mathfrak{M}_{1,1}$.

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

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

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 $0 \le \theta \le s$ 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:

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

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 $\sum_{n\ge 0} a_n z^n$ 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 $\wp'(z)\big/\wp(z)$ 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

$$4X^3-g_2X-g_3=0.$$

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 $y_j=cx_j$ 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.

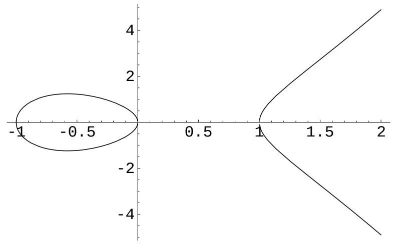

FIGURE 1.5. An example of a nonsingular cubic curve $Y^2 = 4X^3 - g_2X - g_3$.

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

$$\begin{tikzcd} E_{\omega_1,\omega_2} \arrow[r, "(\wp:\wp':1)", "\text{bijection}"'] \arrow[d, "\wp"'] & C \arrow[r, "\subset"] \arrow[d] & \mathbb{P}^2 \\ \mathbb{P}^1 \arrow[r, equal] & \mathbb{P}^1, & \end{tikzcd}$$

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

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

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,

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

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

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

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:

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

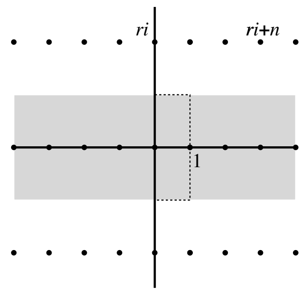

FIGURE 1.6. Degeneration of a lattice to the integral points on the real axis.

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:

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

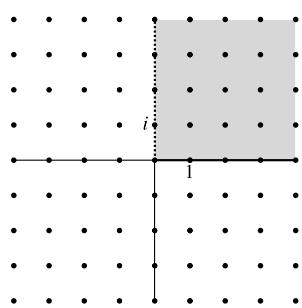

FIGURE 1.7. A partition of the square lattice $\Lambda_{i,1}$ into four pieces.

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:

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

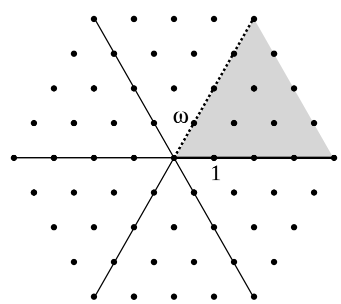

FIGURE 1.8. A partition of the honeycomb lattice $\Lambda_{\omega,1}$ into six pieces.

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

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

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

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

(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

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

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

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

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

$$Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$$

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

$$4x^3 - g_2x - g_3 = 4(x-e_1)(x-e_2)(x-e_3).$$

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.

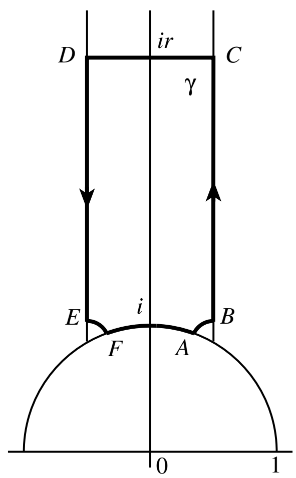

FIGURE 1.9. The boundary of the fundamental domain as an integration contour.

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

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

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 $\mathbb{C}$:

$$\begin{tikzcd} & \mathbb{P}^1 \setminus \{0,1,\infty\} \arrow[d, "j"', "\mathfrak{S}_3\text{-action}"] \\ H \arrow[r, "J", "PSL(2,\mathbb{Z})\text{-action}"'] & \mathbb{C}. \end{tikzcd} \tag{1.36}$$

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

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

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

$$A = T^2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

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

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 $PSL(2\mathbb{Z})\big/\Gamma(2)$, 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. $\square$

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

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).

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

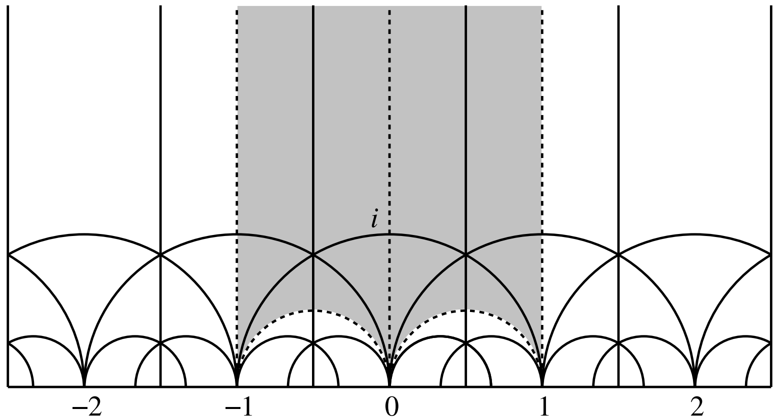

FIGURE 1.10. A fundamental domain of the action of the congruence subgroup $\Gamma(2) \subset PSL(2, \mathbb{Z})$.

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.

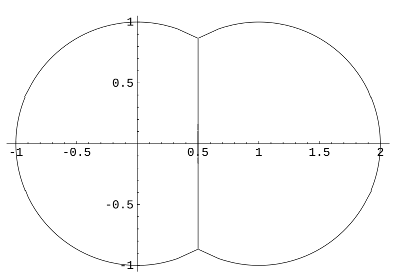

FIGURE 1.11. The inverse image of $[0, 1]$ via the map $j(\lambda) = \frac{4}{27} \frac{\lambda^2 - \lambda + 1}{\lambda^2(1-\lambda)^2}$. (Graphics produced by Josephine Yu.)

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.

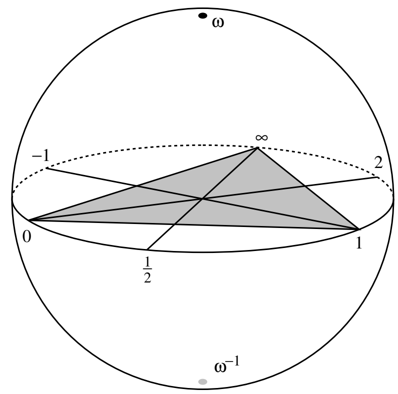

FIGURE 1.12. The $\mathfrak{S}_3$-action on $\mathbb{P}^1$ through the dihedral group action.

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

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

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

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

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

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

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 $PSL(2,\mathbb{Z})/\Gamma(2) = \mathfrak{S}_3$. 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 $\mathbb{Z}/n\mathbb{Z}$-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 $\mathbb{Z}/2\mathbb{Z}$-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.

2. Matrix Integrals and Feynman Diagram Expansion

This chapter is devoted to the study of the asymptotic analysis of various matrix integrals. We investigate symmetric, Hermitian, and quarternionic self-adjoint matrices separately. These integrals can be thought of as 0-dimensional models of Quantum Field Theory. QFT produces many interesting and useful mathematical tools. In this chapter, we deal with QFT as a machinery of **counting formula**. QFT provides us with a clever method of counting the order of certain finite groups.

Often QFT is not well-defined mathematically, but all our models lead to finite dimensional integrals and therefore they are well-defined. We will develop two different methods for calculating some of the QFT integrals. Since the original integral is well-defined, the two methods should provide the same answer. This apparent equality turns out to be an interesting equality in mathematics.

Let us begin by reviewing asymptotic analysis of holomorphic functions.

2.1. Asymptotic Expansion of Analytic Functions.

A holomorphic function admits a convergent Taylor series expansion at each point of the domain of definition. What happens if we try to expand the function into a power series at a boundary point of the domain? We investigate this question in this section. Since our goal is the asymptotic analysis of matrix integrals, we focus our study on the techniques used in matrix integrals, instead of developing the most general theory of asymptotic series.

When we are first introduced to complex analysis, perhaps the most surprizing thing may have been the fact that complex differentiability implies complex analyticity. Let $h(z)$ be a continuous function defined on an open domain $U \subset \mathbb{C}$. If $h(z)$ is continuously differentiable everywhere in U, then it satisfies Cauchy's Theorem of Integration:

$$\oint_\gamma h(z)dz = 0,$$

where γ is a closed loop in U. One can then show that $h(z)$ satisfies the Cauchy Integral Formula

$$h(w) = \frac{1}{2\pi i}\oint_\gamma \frac{h(z)}{z-w}dz, \tag{2.1}$$

where γ is a simple loop in U that goes around $w \in U$ once counter-clockwise. But (2.1) immediately implies that $h(z)$ has Taylor expansion everywhere in U. What happens if $h(z)$ is continuously differentiable not on an entire neighborhood of a point, say 0, but only a part of the neighborhood? This motivates us to introduce the following definition.

Definition 2.1 (Asymptotic Expansion)**.** Let h be a holomorphic function defined on a wedge-shaped domain Ω:

$$\Omega = \{z \in \mathbb{C} \;\big|\; \alpha < \arg(z) < \beta, \quad |z| < r.\}$$

A power series $\sum_{n\ge 0} a_n z^n$ is said to be an **asymptotic expansion** of h at the origin $0 \in \partial\Omega$ if

$$\lim_{\substack{z\to 0\\ z\in\Omega}} \frac{h(z) - \sum_{n=0}^{m-1} a_n z^n}{z^m} = a_m \tag{2.2}$$

holds for every $m \ge 0$. When an asymptotic expansion exists, we say h has an asymptotic expansion on Ω at its boundary point 0.

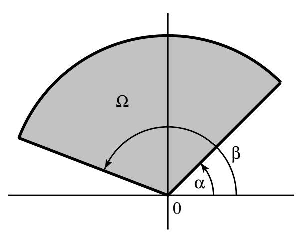

Figure 2.1. A wedge-shaped domain.

Let us examine the implications of (2.2). For $m = 0$, it requires the convergence of $h(z)$ as $t \longrightarrow 0$ while in Ω. Thus $h(z)$ is continuous at 0 when approaching from inside Ω. We can *define* the value of $h(z)$ at 0 by $h(0) = a_0$. For $m = 1$, the existence of

$$\lim_{\substack{z\to 0\\ z\in\Omega}} \frac{h(z)-h(0)}{z} = a_1$$

implies that $h(z)$ is differentiable at 0 when approaching from inside Ω. Let us call the situation **Ω-differentiable**. Since $h(z)$ is holomorphic on Ω, we can differentiate the numerator and the denominator of (2.2) $(m-1)$-times and obtain the same limit. The existence of the limit

$$\lim_{\substack{z\to 0\\ z\in\Omega}} \frac{h(z) - \sum_{n=0}^{m-1} a_n z^n}{z^m} = \lim_{\substack{z\to 0\\ z\in\Omega}} \frac{h^{(m-1)}(z) - (m-1)!a_{m-1}}{m!z} = a_m$$

thus implies that $h^{(m-1)}(z)$ is Ω-differentiable, and that $h^{(m)}(0) = m!a_m$. Therefore, the existence of an asymptotic expansion at $0 \in \overline{\Omega}$ simply means the function $h(z)$ is infinitely many times continuously Ω-differentiable at 0.

The above consideration immediately implies

Proposition 2.2 (Uniqueness of asymptotic expansion)**.** *If a holomorphic function h on Ω has an asymptotic expansion at $0 \in \partial\Omega$ as above, then it is unique.*

The simplest example of an asymptotic expansion is the Taylor expansion when h is holomorphic at 0. Since h is infinitely many times continuously differentiable in a neighborhood of 0, $h^{(m)}(0)$ is well-defined for all $m \geq 0$, and the Taylor expansion

$$h(z) = \sum_{n\geq 0} \frac{h^{(m)}(0)}{m!} z^m$$

gives the asymptotic expansion of $h(z)$ at $z = 0$.

The technique of asymptotic expansion is developed to study the behavior of a holomorphic function at its **essential singularity**. When a holomorphic function $h(z)$ has an essential singularity at 0, often we can find a wedge-shaped domain Ω with 0 as its vertex such that the function is infinitely many times continuously differentiable on $\overline{\Omega}$. We can then expand the function into its asymptotic series and study its properties. The existence of such a domain is significant because $h(z)$ can take arbitrary values except for up to two excluded values in any neighborhood of 0 (Picard's Theorem). If h is defined on a larger domain Ω' that contains Ω and has an asymptotic expansion on Ω', then h has an asymptotic expansion also on Ω and the asymptotic series are exactly the same. In general, however, the existence depends on the choice of the domain Ω.

Example 2.1. Consider $h(z) = e^{1/z}$. It is holomorphic on $\mathbb{C} \setminus \{0\}$. If we choose

$$\Omega = \{z \in \mathbb{C} \mid \frac{\pi}{2} + \epsilon < \arg(z) < \frac{3\pi}{2} - \epsilon\}, \tag{2.3}$$

then it has an asymptotic expansion on Ω at 0, and its asymptotic series is the zero series. However, if we choose a wedge-shaped domain contained in the right half plane $Re(z) > 0$, then $e^{1/z}$ does not have any asymptotic expansion.

This example also shows that two different holomorphic functions may have the same asymptotic expansion on the same domain. From this point of view, the holomorphic function h and its asymptotic series $\sum_{n\geq 0} a_n z^n$ are not *equal*. We use the notation

$$\mathcal{A}(h) = \sum_{n\geq 0} a_n z^n \tag{2.4}$$

to indicate that the series of the right hand side is the asymptotic expansion of $h(z)$. We also use

$$h(z) \equiv g(z)$$

if $h(z)$ and $g(z)$ have the same asymptotic expansion on the same domain. Thus $0 \equiv e^{1/z}$ on the domain of (2.3).

Proposition 2.3 (Properties of the asymptotic expansion)**.** *Let $f(z)$ and $h(z)$ be holomorphic functions on a domain Ω and have asymptotic expansions at its* *boundary point* $0 \in \partial\Omega$*:*

$$\mathcal{A}(f) = \sum_{n\geq 0} a_n z^n, \qquad \mathcal{A}(h) = \sum_{n\geq 0} b_n z^n.$$

Then

$$\mathcal{A}(f+h) = \mathcal{A}(f) + \mathcal{A}(h) \tag{2.5}$$

$$\mathcal{A}(f\cdot h) = \mathcal{A}(f)\cdot\mathcal{A}(h) \tag{2.6}$$

Proof. For every $m \geq 0$, we have

$$\begin{aligned}\lim \frac{f(z)+h(z)-\sum_{n=0}^{m-1}(a_n+b_n)z^n}{z^m} &= \lim \frac{f(z)-\sum_{n=0}^{m-1} a_n z^n}{z^m} \\ &\quad + \lim \frac{h(z)-\sum_{n=0}^{m-1} b_n z^n}{z^m} \\ &= a_m + b_m.\end{aligned}$$

This proves (2.5).

Since we know

$$\lim \frac{f(z)h(z)-f(z)\sum_{n=0}^{m-1} b_k z^k}{z^m} = a_0 b_m$$

and

$$\lim \frac{f(z)\sum_{k=0}^{m-1} b_k z^k - \sum_{n=0}^{m-1} a_n z^n \sum_{k=0}^{m-1} b_k z^k}{z^m} = a_m b_0,$$

adding the above two equations, we have

$$\lim \frac{f(z)h(z)-\sum_{n=0}^{m-1} a_n z^n \sum_{k=0}^{m-1} b_k z^k}{z^m} = a_0 b_m + a_m b_0.$$

Note that

$$\begin{aligned}&\sum_{n=0}^{m-1} a_n z^n \sum_{k=0}^{m-1} b_k z^k \\ &\qquad = \sum_{n+k\leq m-1} a_n b_k z^{n+k} + (a_1 b_{m-1} + a_2 b_{m-2} + \cdots a_{m-1} b_1) z^m + O(z^{m+1}).\end{aligned}$$

Therefore, we obtain

$$\lim \frac{f(z)h(z)-\sum_{n+k\leq m-1} a_n b_k z^{n+k}}{z^m} = \sum_{n+k=m} a_n b_k.$$

This proves (2.6). □

Let us now consider a simple example

$$Z_4(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} e^{\frac{t}{4!}x^4}\, dx.$$

The integral $Z_4(t)$ is a holomorphic function in t for $Re(t) < 0$ and continuous for $Re(t) \leq 0$. Let

$$\Omega = \{t \in \mathbb{C} \mid 2\pi/3 < \arg(t) < 4\pi/3\}.$$

We wish to find the asymptotic expansion of Z_4 on Ω at $t = 0$. First we note that if $t \in \Omega$, then

$$\left| e^{\frac{t}{4!}x^4} \right| = e^{\frac{Re(t)}{4!}x^4} \leq 1.$$

We claim:

$$\mathcal{A}\big(Z_4(t)\big) = \sum_{n\geq 0} \frac{t^n}{(4!)^n n!}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} x^{4n} dx\right). \tag{2.7}$$

Indeed, we have

$$\begin{aligned}
&\lim_{\substack{t\to 0\\ t\in\Omega}} \frac{1}{t^m}\left(\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} e^{\frac{t}{4!}x^4} dx - \sum_{n=0}^{m-1} \frac{t^n}{(4!)^n n!}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2} x^{4n} dx\right)\\
&\quad= \lim_{\substack{t\to 0\\ t\in\Omega}} \frac{1}{t^m}\left(\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}\sum_{n\geq 0}\frac{t^n}{(4!)^n n!}x^{4n}dx - \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}\sum_{n=0}^{m-1}\frac{t^n}{(4!)^n n!}x^{4n}dx\right)\\
&\quad= \lim_{\substack{t\to 0\\ t\in\Omega}} \frac{1}{t^m}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}\sum_{n=m}^{\infty}\frac{t^n}{(4!)^n n!}x^{4n}dx\\
&\quad= \lim_{\substack{t\to 0\\ t\in\Omega}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}\sum_{n=0}^{\infty}\frac{t^n}{(4!)^{n+m}(n+m)!}x^{4(n+m)}dx\\
&\quad= \lim_{\substack{t\to 0\\ t\in\Omega}} \frac{1}{m!}\frac{d^m}{dt^m}Z_4(t)\\
&\quad= \frac{1}{(4!)^m m!}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}x^{4m}dx,
\end{aligned}$$

where we have used the uniform continuity of $Z_4^{(m)}(t)$ on $\overline{\Omega}$ for every $m\geq 0$.

To evaluate this last integral, let us consider

$$Z(J) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2+Jx}dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-J)^2}e^{\frac{1}{2}J^2}dx = e^{\frac{1}{2}J^2}. \tag{2.8}$$

If we put $J=\sqrt{-1}p$, then (2.8) simply says that the Fourier transform of $e^{-1/2x^2}$ is $e^{-1/2p^2}$. The multiplication by x^{4n} to the function $e^{-1/2x^2}$ changes into the $4n$-th order differentiation of its Fourier transform, and we have

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}x^{4n}dx &= \left.\frac{d^{4n}}{dJ^{4n}}Z(J)\right|_{J=0}\\
&= \left.\frac{d^{4n}}{dJ^{4n}}\sum_{m\geq 0}\frac{1}{2^m m!}J^{2m}\right|_{J=0}\\
&= \frac{(4n)!}{2^{2n}(2n)!}\\
&= \frac{(4n)(4n-1)(4n-2)\cdots 4\cdot 3\cdot 2\cdot 1}{(4n)(4n-2)\cdots 4\cdot 2}\\
&= (4n-1)(4n-3)\cdots 3\cdot 1\\
&\underset{\text{def}}{=} (4n-1)!!.
\end{aligned} \tag{2.9}$$

Thus the final result of the asymptotic expansion is given by

$$\mathcal{A}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}x^2}e^{\frac{t}{4!}x^4}dx\right) = \sum_{n\geq 0}\frac{(4n-1)!!}{(4!)^n n!}t^n. \tag{2.10}$$

The above method works for other **potential** terms such as x^6, x^8, or more general

$$V(x) = \sum_{j=1}^{2m} \frac{t_j}{j!} x^j.$$

But unfortunately the asymptotic expansion becomes unappealingly complicated. The technique developed in the next section provides an amazingly beautiful interpretation of the asymptotic formula.

2.2. **Feynman Diagram Expansion.** The key technique of the computation of the asymptotic expansion (2.10) is the introduction of the **source** term Jx in (2.8) and the fact that the integration changes into the differentiation through **Fourier transform**, as we have seen in (2.9). Now, instead of calculating the Taylor expansion of $Z(J) = e^{J^2/2}$, let us find a combinatorial interpretation of the mechanism.

The simplest case, $n = 1$, is illustrative. We have

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} x^4 dx &= \left(\frac{d}{dJ}\right)^4 e^{\frac{1}{2}J^2}\bigg|_{J=0} \\
&= \left(\frac{d}{dJ}\right)^3 J e^{\frac{1}{2}J^2}\bigg|_{J=0} \\
&= \left(\frac{d}{dJ}\right)^2 \left(e^{\frac{1}{2}J^2} + J^2 e^{\frac{1}{2}J^2}\right)\bigg|_{J=0} \\
&= \left(\frac{d}{dJ}\right) \left(J e^{\frac{1}{2}J^2} + 2J e^{\frac{1}{2}J^2} + J^3 e^{\frac{1}{2}J^2}\right)\bigg|_{J=0} \\
&= \left(e^{\frac{1}{2}J^2} + 2e^{\frac{1}{2}J^2} + 6J^2 e^{\frac{1}{2}J^2} + J^4 e^{\frac{1}{2}J^2}\right)\bigg|_{J=0} = 3.
\end{aligned}$$

We see that a differentiation creates a factor of J in front of $e^{\frac{1}{2}J^2}$, which is evaluated at $J = 0$ in the end. Thus unless another differentiation annihilates the factor J, the contribution of this term is 0. If we name each operator a, b, c, d, then there are three different pairs $\{ab, cd\}$, $\{ac, bd\}$, and $\{ad, bc\}$. In other words, the answer 3 of the integral represents the number of ways of making two pairs of differential operators out of four. In general, the differentiation $Z^{(4n)}(0)$ gives **the number of ways of making $2n$-pairs out of the $4n$ objects**. Indeed, we have

$$\begin{aligned}
Z^{(4n)}(0) &= \frac{\binom{4n}{2}\binom{4n-2}{2}\cdots\binom{4}{2}\binom{2}{2}}{(2n)!} \\
&= \frac{(4n)(4n-1)(4n-2)(4n-3)\cdots 4\cdot 3\cdot 2\cdot 1}{2^{(2n)}(2n)!} \\
&= \frac{(4n)!}{2^{(2n)}(2n)!} \\
&= (4n-1)!! \ ,
\end{aligned}$$

and this coincides with the calculation of (2.9). In order to visualize the situation, let us provide n sets of 4 dots. Each dot represents a differential operator $\partial/\partial J$. We connect two dots with a line when the corresponding operators are paired

(Figure 2.2). Let us call the dots and the lines connecting paired dots a **pairing scheme**.

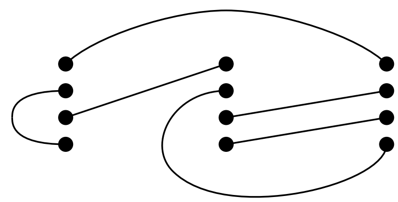

FIGURE 2.2. A pairing scheme of n sets of 4 dots.

What follows is an ingenious idea of Richard Feynman. He replaces the set of four dots with a **vertex** of **valence** four. Then the paring scheme changes into a **graph** (Figure 2.3).

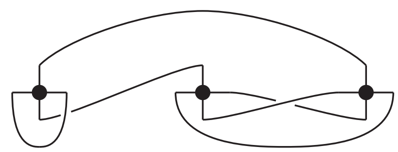

FIGURE 2.3. A 4-valent graph.

The formula (2.9) thus gives the number of pairing schemes. Then what does the coefficient

$$\frac{(4n-1)!!}{(4!)^n n!}$$

of the asymptotic expansion (2.10) represent? We can view the $4n$ dots as the **total space** $\mathcal{D}$ of a **fiber bundle** defined over a finite set $\mathcal{V}$ of n elements as the **base space**, with a **fiber** F_p at a base point $p \in \mathcal{V}$ consisting of 4 dots:

$$\begin{array}{ccc} F_p & \longrightarrow & \mathcal{D} \\ \Big\downarrow & & {\scriptstyle\pi}\Big\downarrow \\ \{p\} & \longrightarrow & \mathcal{V}. \end{array}$$

The symmetric group $\mathfrak{S}_{4n}$ acts on the total space $\mathcal{D}$ by permutation. Let $G \subset \mathfrak{S}_{4n}$ be a maximal subgroup that preserves the fiber bundle structure. In other words, G consists of those permutations that map each fiber onto another fiber. Clearly, every element of G induces a transformation of $\mathcal{V}$, and the kernel of the homomorphism $G \longrightarrow \mathfrak{S}_n$ is $\mathfrak{S}_4^n$, which acts on each fiber as permutation of 4 elements. Thus we have an **exact sequence** of groups

$$\mathfrak{S}_4^n \longrightarrow G \longrightarrow \mathfrak{S}_n,$$

and hence

$$\mathfrak{S}_4^n \ltimes \mathfrak{S}_n \cong G \subset \mathfrak{S}_{4n}.$$

The passage from the paring scheme P as in Figure 2.2 to the graph Γ as in Figure 2.3 is the projection of the pairing scheme onto the base space $\mathcal{V}$. From this point of view, let us denote

$$\pi(P) = \Gamma.$$

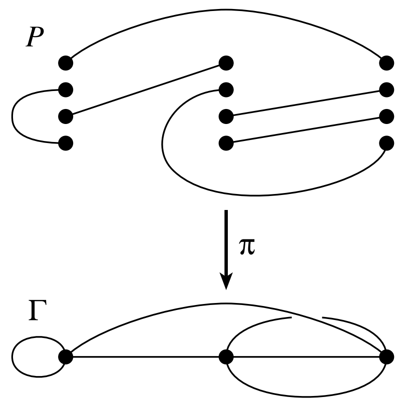

FIGURE 2.4. From a pairing scheme to a graph through the projection of the fiber bundle.

The group G also acts on the set of pairing schemes $\mathcal{P}$. If this action is fixed point free, then we can identify the orbit space $\mathcal{P}/G$ with the set of all 4-valent graphs with n vertices. Note that if two pairing schemes P and P' are on the same G-orbit, then their stabilizer subgroups are isomorphic:

$$G_P \cong G_{P'}.$$

Let $\Gamma = \pi(P)$ denote the graph obtained from a pairing scheme P, and $\Gamma' = \pi(P')$. Then these graphs should be defined to be **isomorphic**, and their **automorphism group** can be defined by

$$\mathrm{Aut}(\Gamma) = G_P.$$

Since the G-orbit $G \cdot P$ is related to the stabilizer G_P by

$$G \cdot P \cong G/G_P,$$

we have the **counting formula**

$$\frac{|\mathcal{P}|}{|G|} = \frac{1}{|G|} \sum_{\pi(P)\in \mathcal{P}/G} |G \cdot P| = \frac{1}{|G|} \sum_{\pi(P)\in \mathcal{P}/G} |G/G_P| = \sum_{\Gamma} \frac{1}{\mathrm{Aut}(\Gamma)},$$

where Γ runs all 4-valent graphs consisting of n vertices. We have thus established a desired interpretation of the coefficient of the asymptotic expansion:

$$\frac{(4n-1)!!}{(4!)^n n!} = \sum_{\substack{\Gamma \text{ 4-valent graph} \\ \text{with } n \text{ vertices}}} \frac{1}{|\mathrm{Aut}(\Gamma)|}. \tag{2.11}$$

In order to proceed further to more complicated integrals, we need to give the precise definition of graphs and their automorphisms here.

2.3. **Preparation from Graph Theory.**

Definition 2.4 (Graph)**.** A **graph** is a collection

$$\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$$

consisting of a finite set $\mathcal{V}$ of **vertices**, a finite set $\mathcal{E}$ of **edges**, and their **incidence relation**

$$\mathcal{I} : \mathcal{E} \longrightarrow (\mathcal{V} \times \mathcal{V})/\mathfrak{S}_2$$

that maps the set of edges to the set of symmetric pairs of vertices. A vertex V and an edge E of a graph Γ is said to be **incident** if $\mathcal{I}(E) = (V, V')$ for a vertex V'.

Remark. A graph is a visual object. We place the vertices in the space, and connect a pair of vertices with a line if there is an edge incident to them. If an edge is incident to the same vertex twice, then it forms a **loop** starting and ending at the vertex.

Let V and V' be two vertices of a graph Γ. The quantity

$$a_{VV'} = |\mathcal{I}^{-1}(V, V')|$$

gives the number of edges that connect these vertices. The **valence**, or the **degree**, of a vertex V is the number

$$j(V) = \sum_{\substack{V' \in \mathcal{V} \\ V' \neq V}} a_{VV'} + 2a_{VV}.$$

This is the number of edges that are incident to V. Note that when an edge is incident to V twice, forming a loop, then it contributes 2 to the valence of V.

Remark. **To avoid unnecessary complexity, we assume that all graphs we deal with in these lectures have no vertices of valence less than 3, unless otherwise stated.**

Definition 2.5 (Graph isomorphism)**.** Two graphs $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$ and $\Gamma' = (\mathcal{V}', \mathcal{E}', \mathcal{I}')$ are said to be **isomorphic** if there are bijections $\alpha : \mathcal{V} \xrightarrow{\sim} \mathcal{V}'$ and $\beta : \mathcal{E} \xrightarrow{\sim} \mathcal{E}'$ that are compatible with the incidence relations:

$$\begin{array}{ccc} \mathcal{E} & \xrightarrow{\mathcal{I}} & (\mathcal{V} \times \mathcal{V})/\mathfrak{S}_2 \\ {\scriptstyle \beta}\big\downarrow & & \big\downarrow{\scriptstyle \alpha \times \alpha} \\ \mathcal{E}' & \xrightarrow{\mathcal{I}'} & (\mathcal{V}' \times \mathcal{V}')/\mathfrak{S}_2. \end{array}$$

For example, the graph of Figure 2.3 and the graph at the bottom of Figure 2.4 are isomorphic. The notion of isomorphism of graphs *should* naturally lead to the notion of graph automorphisms. However, we immediately see that there is a big difference between what we need in Feynman diagram expansion and the notion of graph automorphisms in a more traditional sense. Let us consider the case of $n = 1$ in (2.11). We have a 4-valent graph with only one vertex. There is only one such graph, which has two loops attached to the vertex. In terms of traditional graph theory, the automorphism group should be $\mathfrak{S}_2$, which interchanges the two loops. But the formula we have established gives

$$\frac{3!!}{4! \times 1} = \frac{1}{8} = \frac{1}{|\mathrm{Aut}(\Gamma)|},$$

or $|\mathrm{Aut}(\Gamma)| = 8$. This example illustrates that we have to define the graph automorphism in a quite different way from the usual graph theory. To establish the right notion of graph automorphisms for our purpose, we need to consider directed graphs and the edge refinement of a graph.

A **directed edge** is an edge $E \in \mathcal{E}$ of a graph with an arrow assigned from the vertex at one end of E to the other. There are two distinct directions for every edge. A **directed graph** is a graph whose edges are all directed. There are $2^{|\mathcal{E}|}$

different directed graphs for each graph. For every directed edge $\overrightarrow{E}$ of a graph Γ that is incident to vertices V and V' (allowing the case $V = V'$), we can choose a midpoint V_E of it, and separate the edge E into two **half edges** E_- and E_+, such that the order (E_-, E_+) is consistent with the direction of the edge. Thus E_- is incident to (V, V_E), and E_+ is incident to (V', V_E). V_E is a new vertex of valence 2. The incidence relation of a directed graph is a map

$$\mathcal{I} : \mathcal{E} \ni E \longmapsto (V, V') \in \mathcal{V} \times \mathcal{V}$$

without taking the symmetric product, where V is the **initial vertex** of $\overrightarrow{E}$ and V' is its **terminal vertex**.

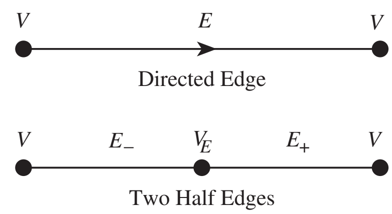

FIGURE 2.5. Creating two half edges from a directed edge.

Definition 2.6 (Edge refinement)**.** Let $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$ be a graph with no vertices of valence less than 3. The **edge refinement** of Γ is a graph obtained by adding a midpoint on each edge of Γ. More precisely, choose a direction on Γ. The edge refinement is a graph

$$\Gamma_{\mathcal{E}} = (\mathcal{V} \cup \mathcal{V}_{\mathcal{E}}, \mathcal{E}_- \cup \mathcal{E}_+, \mathcal{I}_{\mathcal{E}})$$

consisting of the set of vertices $\mathcal{V} \cup \mathcal{V}_{\mathcal{E}}$, the set of edges $\mathcal{E}_- \cup \mathcal{E}_+$, and an incidence relation $\mathcal{I}_{\mathcal{E}} : \mathcal{E}_- \cup \mathcal{E}_+ \longrightarrow \mathcal{V} \times \mathcal{V}_{\mathcal{E}}$ subject to the following conditions:

(1) $\mathcal{V}_{\mathcal{E}} = \mathcal{E}$ is the set of edges of the original graph that is identified with the set of midpoints of edges;
(2) $\mathcal{E}_- \cup \mathcal{E}_+$ is the set of half edges;
(3) the incidence relation $\mathcal{I}_{\mathcal{E}}$ is consistent with the original incidence relation, namely

$$\begin{tikzcd} \mathcal{E}_- \arrow[r, "\mathcal{I}_{\mathcal{E}}"] \arrow[d] & \mathcal{V} \times \mathcal{V}_{\mathcal{E}} \arrow[r] & \mathcal{V} \arrow[d, equal] \\ \mathcal{E} \arrow[r, "\mathcal{I}"] & \mathcal{V} \times \mathcal{V} \arrow[r, "pr_1"] & \mathcal{V}, \end{tikzcd} \qquad \begin{tikzcd} \mathcal{E}_+ \arrow[r, "\mathcal{I}_{\mathcal{E}}"] \arrow[d] & \mathcal{V} \times \mathcal{V}_{\mathcal{E}} \arrow[r] & \mathcal{V} \arrow[d, equal] \\ \mathcal{E} \arrow[r, "\mathcal{I}"] & \mathcal{V} \times \mathcal{V} \arrow[r, "pr_2"] & \mathcal{V}. \end{tikzcd}$$

Remark. (1) The edge refinement is independent of the choice of a direction of Γ. Indeed, let $\overrightarrow{E}$ be a directed edge of Γ connecting the initial vertex V_i and the terminal vertex V_t. Flipping the direction results in *renaming* the half edges E_- and E_+ and the vertices V_i and V_t, without altering the actual set of vertices, half edges, and the incidence relation.

(2) Since we are not allowing any vertices of valence less than 3 in Γ, the original graph can be recovered from its edge refinement $\Gamma_{\mathcal{E}}$ uniquely. Indeed, Γ is obtained by throwing away all 2-valent vertices from $\Gamma_{\mathcal{E}}$, and connecting half edges together when they meet.

(3) The valence of a vertex $V \in \mathcal{V}$ of Γ is the number of half edges of the edge refinement $\Gamma_{\mathcal{E}}$ that are incident to V.

Definition 2.7 (Graph automorphism)**.** Let $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$ be a graph with no vertices of valence less than 3. A **graph automorphism** of Γ is a triple $(\alpha, \alpha_{\mathcal{E}}, \beta)$ of bijections $\alpha : \mathcal{V} \xrightarrow{\sim} \mathcal{V}$, $\alpha_{\mathcal{E}} : \mathcal{V}_{\mathcal{E}} \xrightarrow{\sim} \mathcal{V}_{\mathcal{E}}$, and $\beta : \mathcal{E}_- \cup \mathcal{E}_+ \xrightarrow{\sim} \mathcal{E}_- \cup \mathcal{E}_+$ that are compatible with the incidence relation of the edge refinement $\Gamma_{\mathcal{E}} = (\mathcal{V} \cup \mathcal{V}_{\mathcal{E}}, \mathcal{E}_- \cup \mathcal{E}_+, \mathcal{I}_{\mathcal{E}})$ of Γ:

$$\begin{tikzcd} \mathcal{E}_- \cup \mathcal{E}_+ \arrow[r, "\mathcal{I}_{\mathcal{E}}"] \arrow[d, "\beta"'] & \mathcal{V} \times \mathcal{V}_{\mathcal{E}} \arrow[d, "\alpha \times \alpha_{\mathcal{E}}"] \\ \mathcal{E}_- \cup \mathcal{E}_+ \arrow[r, "\mathcal{I}_{\mathcal{E}}"] & \mathcal{V} \times \mathcal{V}_{\mathcal{E}}. \end{tikzcd}$$

The **group of graph automorphisms** of a graph Γ is denoted by $\mathrm{Aut}(\Gamma)$.

Example 2.2. There is only one $2j$-valent graph Γ with one vertex. Since every edge is a loop, Γ has j edges (Figure 2.6). There are $2j$ half edges in the edge refinement of Γ. Thus $\mathrm{Aut}(\Gamma)$ is a subgroup of $\mathfrak{S}_{2j}$ that acts on the set of half edges $\mathcal{E}_- \cup \mathcal{E}_+$ through permutation. Since a graph automorphism induces a permutation of midpoints $\mathcal{V}_{\mathcal{E}}$, we have an exact sequence

$$(\mathfrak{S}_2)^j \longrightarrow \mathrm{Aut}(\Gamma) \longrightarrow \mathfrak{S}_j.$$

Therefore, $\mathrm{Aut}(\Gamma) = (\mathfrak{S}_2)^j \ltimes \mathfrak{S}_j \subset \mathfrak{S}_{2j}$. In particular, it has $2^j j!$ elements. We note that from the point of view of traditional graph theory, there are only $j!$ automorphisms.

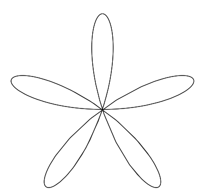

FIGURE 2.6. The unique $2j$-valent graph with 1 vertex.

If a graph Γ has no loops and its vertices have valance at least 3, then our $\mathrm{Aut}(\Gamma)$ is the same as the traditional automorphism group. Historically, one of the greatest motivations of graph theory came from applications to electric circuits and communication networks. In a context of electric circuits or networks, it is absolutely important to have 2-valent vertices, but loops are not welcome. In fact, a loop is a short circuit in an electrical circuit, and there is no need for a loop in a communication network. We need an alternative definition of automorphisms because our intent of application is different. Now we are ready to show that our definition is indeed the right notion for the Feynman diagram expansion we are considering.

Definition 2.8 (Pairing scheme)**.** Let $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$ be a graph. For each vertex $V \in \mathcal{V}$, we denote by $j(V)$ the valence of V. Note that

$$\sum_{V \in \mathcal{V}} j(V) = 2|\mathcal{E}|$$

is equal to the number of half edges of the edge refinement of Γ. A **pairing scheme associated with** graph Γ is a triple $(\mathcal{D}, \pi, \mathcal{I}_P)$ consisting of a collection $\mathcal{D}$ of a total of $2|\mathcal{E}|$ **dots**, a projection

$$\pi : \mathcal{D} \longrightarrow \mathcal{V}$$

whose fiber at V consists of $j(V)$ dots, and a bijection

$$\mathcal{I}_P : \mathcal{E}_- \cup \mathcal{E}_+ \xrightarrow{\sim} \mathcal{D}$$

satisfying the compatibility condition of incidence

$$\begin{array}{ccc} \mathcal{E}_- \cup \mathcal{E}_+ & \xrightarrow{\mathcal{I}_\mathcal{E}} & \mathcal{V} \times \mathcal{V}_\mathcal{E} \\ {\scriptstyle \mathcal{I}_P}\big\downarrow & & \big\downarrow{\scriptstyle pr_1} \\ \mathcal{D} & \xrightarrow{\pi} & \mathcal{V}. \end{array}$$

Two dots $D, D' \in \mathcal{D}$ are **connected** in the pairing scheme $(\mathcal{D}, \pi, \mathcal{I}_P)$ if there is an edge $E \in \mathcal{E}$ such that $D = \mathcal{I}_P(E_-)$ and $D' = \mathcal{I}_P(E_+)$, whereE_- and E_+ are the two half edges belonging to E with an appropriate choice of a direction of E.

A pairing scheme associated with a graph Γ is not unique. Indeed, an automorphism of the fibration $\pi : \mathcal{D} \longrightarrow \mathcal{V}$ transforms one pairing scheme to another.

Definition 2.9 (Automorphism of fibration)**.** An **automorphisms of the fibration** $\pi : \mathcal{D} \longrightarrow \mathcal{V}$ is a permutation f of the dots $\mathcal{D}$ that preserves the fibration:

$$\begin{array}{ccc} \mathcal{D} & \xrightarrow{f} & \mathcal{D} \\ {\scriptstyle \pi}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ \mathcal{V} & \xrightarrow{\overline{f}} & \mathcal{V}. \end{array} \tag{2.12}$$

Theorem 2.10 (Graph automorphisms and stabilizers of a pairing scheme)**.** *Let $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$ be a graph and $(\mathcal{D}, \pi, \mathcal{I}_P)$ the pairing scheme associated with Γ. By $St(\mathcal{D}, \pi, \mathcal{I}_P)$ we denote the stabilizer subgroup of the group of automorphisms of the fibration $\pi : \mathcal{D} \longrightarrow \mathcal{V}$ that fixes the given pairing scheme $(\mathcal{D}, \pi, \mathcal{I}_P)$. Then we have a natural group isomorphism*

$$\phi : St(\mathcal{D}, \pi, \mathcal{I}_P) \xrightarrow{\sim} \mathrm{Aut}(\Gamma).$$

Proof. Take an element $f \in St(\mathcal{D}, \pi, \mathcal{I}_P)$. It induces a bijection $\overline{f} : \mathcal{V} \longrightarrow \mathcal{V}$ as in (2.12). Let $E \in \mathcal{E}$ be an edge of Γ. It determines a pair of dots $(\mathcal{I}_P(E_-), \mathcal{I}_P(E_+))$ that are connected. Since f stabilizes the pairing scheme $(\mathcal{D}, \pi, \mathcal{I}_P)$, the pair of dots $\big(f(\mathcal{I}_P(E_-)), f(\mathcal{I}_P(E_+))\big)$ are again connected in $(\mathcal{D}, \pi, \mathcal{I}_P)$. Therefore, it determines an edge $\hat{f}(E)$ of Γ. More precisely, we have a bijection

$$\hat{f} : \mathcal{E}_- \cup \mathcal{E}_+ \longrightarrow \mathcal{E}_- \cup \mathcal{E}_+.$$

The fact that f is an automorphism of the fibration $\pi : \mathcal{D} \longrightarrow \mathcal{V}$ implies that the pair of bijections $(\overline{f}, \hat{f})$ is an automorphism of the graph Γ. This association defines the homomorphism ϕ. Clearly, the kernel of this homomorphism is trivial.

Conversely, let

$$(\alpha : \mathcal{V} \longrightarrow \mathcal{V}, \alpha_\mathcal{E} : \mathcal{V}_\mathcal{E} \longrightarrow \mathcal{V}_\mathcal{E}, \beta : \mathcal{E}_- \cup \mathcal{E}_+ \longrightarrow \mathcal{E}_- \cup \mathcal{E}_+)$$

be an automorphism of Γ. Through the bijection $\mathcal{I}_P : \mathcal{E}_- \cup \mathcal{E}_+ \longrightarrow \mathcal{D}$, β induces an automorphism f of the fibration π that is compatible with the other data:

$$\begin{tikzcd}
\mathcal{V}_\mathcal{E} \arrow[r, equal] \arrow[d, "\alpha_\mathcal{E}"'] & \mathcal{E} & \mathcal{E}_- \cup \mathcal{E}_+ \arrow[l] \arrow[r, "\mathcal{I}_P"] \arrow[d, "\beta"'] & \mathcal{D} \arrow[r, "\pi"] \arrow[d, "f"] & \mathcal{V} \arrow[d, "\alpha"] \\
\mathcal{V}_\mathcal{E} \arrow[r, equal] & \mathcal{E} & \mathcal{E}_- \cup \mathcal{E}_+ \arrow[l] \arrow[r, "\mathcal{I}_P"] & \mathcal{D} \arrow[r, "\pi"] & \mathcal{V}.
\end{tikzcd}$$

The automorphism f permutes the pairs of dots in $\mathcal{D}$, stabilizing the pairing scheme $(\mathcal{D}, \pi, \mathcal{I}_P)$. Thus the homomorphism ϕ is surjective. This completes the proof. $\square$

Definition 2.11 (Connectivity of a graph)**.** Two vertices V and V' of a graph Γ are said to be **connected in** Γ if there is a sequence of vertices

$$V = V_0, V_1, V_2, \cdots, V_n = V'$$

in $\mathcal{V}$ such that (V_i, V_{i+1}) are incident to an edge $E_i \in \mathcal{E}$ for every $i = 0, 1, 2, \cdots, n-1$. If every pair of vertices of Γ are connected in Γ, then we say the graph itself is **connected**.

2.4. **Asymptotic Analysis of 1×1 Matrix Integrals.** We are now ready to calculate the asymptotic expansion of

$$Z(t,m) = \frac{1}{\sqrt{2\pi}} \int_\mathbb{R} \exp\left(-\frac{1}{2}x^2\right) \exp\left(\sum_{j=3}^{2m} \frac{t_j}{j!} x^j\right) dx, \tag{2.13}$$

where $m > 0$ is an integer and

$$t = (t_3, t_4, \cdots, t_{2m-1}, t_{2m}) \in \mathbb{C}^{2m-3} \times \Omega_\epsilon.$$

From now on we use the domain

$$\Omega_\epsilon = \{t_{2m} \in \mathbb{C} \bigm| \pi - \epsilon < \arg(t_{2m}) < \pi + \epsilon\} \tag{2.14}$$

for the asymptotic expansion in t_{2m}, where ϵ is a small positive real number. Since $Re(t_{2m}) < 0$, the integral (2.13) is absolutely convergent on $\mathbb{C}^{2m-3} \times \Omega_\epsilon$. Choose $t_{2m} \in \Omega_\epsilon$ and fix it. Then $Z\big((t_3, t_4, \cdots, t_{2m-1}, t_{2m}), m\big)$ is absolutely and uniformly convergent in $(t_3, t_4, \cdots, t_{2m-1})$ on any compact subset of $\mathbb{C}^{2m-3}$. In particular, it has an absolutely convergent Taylor series expansion around the origin of $\mathbb{C}^{2m-3}$:

$$\begin{aligned}
Z(t,m) &= \frac{1}{\sqrt{2\pi}} \int_\mathbb{R} \exp\left(-\frac{1}{2}x^2\right) \exp\left(\sum_{j=3}^{2m-1} \frac{t_j}{j!} x^j\right) \exp\left(\frac{t_{2m}}{(2m)!} x^{2m}\right) dx \\
&= \sum_{v_3 \ge 0, v_4 \ge 0, \cdots, v_{2m-1} \ge 0} \prod_{j=3}^{2m-1} \frac{t_j^{v_j}}{(j!)^{v_j} v_j!} \\
&\qquad \cdot \frac{1}{\sqrt{2\pi}} \int_\mathbb{R} \exp\left(-\frac{1}{2}x^2\right) x^{\sum_{j=3}^{2m-1} j v_j} \exp\left(\frac{t_{2m}}{(2m)!} x^{2m}\right) dx.
\end{aligned}$$

For a fixed $(v_3, v_4, \cdots, v_{2m-1})$, the integral of the last line of the above and its all t_{2m} derivatives are uniformly continuous on Ω_ϵ. Therefore, as $t_{2m} \longrightarrow 0$ while in Ω_ϵ, we have an asymptotic expansion

$$\mathcal{A}\left(\frac{1}{\sqrt{2\pi}} \int_\mathbb{R} \exp\left(-\frac{1}{2}x^2\right) x^{\sum_{j=3}^{2m-1} j v_j} \exp\left(\frac{t_{2m}}{(2m)!} x^{2m}\right) dx\right)$$

$$= \sum_{v_{2m}\geq 0} \frac{t_{2m}^{v_{2m}}}{((2m)!)^{v_{2m}} v_{2m}!} \cdot \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}x^2\right) x^{\sum_{j=3}^{2m} j v_j} dx.$$

Let us denote by $\mathcal{A}\big(Z(t,m)\big)$ the Taylor expansion in $(t_3, t_4, \cdots, t_{2m-1}) \in \mathbb{C}^{2m-3}$ and the asymptotic expansion in $t_{2m} \in \Omega_\epsilon$ of $Z(t,m)$. We have thus established

$$\mathcal{A}\big(Z(t,m)\big) = \sum_{v_3\geq 0, v_4 \geq 0, \cdots, v_{2m}\geq 0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} x^{\sum_{j=3}^{2m} j v_j} dx \prod_{j=3}^{2m} \frac{t_j^{v_j}}{(j!)^{v_j} v_j!}. \tag{2.15}$$

It is worth noting that the function $Z\big((t_3, t_4, \cdots, t_{2m}), m\big)$ is not continuous as $t_{2m} \longrightarrow 0$ in Ω_ϵ. This is why we have to use the above argument to establish the asymptotic expansion of $Z(t,m)$.

We still have to calculate the coefficients of the expansion. To this end, recall the function

$$Z(J) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2 + Jx} dx = e^{\frac{1}{2}J^2}$$

of (2.8). It is easy to see that

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} x^{\sum_{j=3}^{2m} j v_j} dx = \left(\frac{d}{dJ}\right)^{\sum_{j=3}^{2m} j v_j} Z(J)\Bigg|_{J=0}.$$

Now consider the collection of dots $\mathcal{D}(v_3, \cdots, v_{2m})$, consisting of v_j sets of j dots for $j = 3, 4, \cdots, 2m$. Since only the paired differentiation contributes 1 to the answer, we have

$$\left(\frac{d}{dJ}\right)^{\sum_{j=3}^{2m} j v_j} Z(J)\Bigg|_{J=0} = \text{the number of pairing schemes on } \mathcal{D}(v_3, \cdots, v_{2m}).$$

Let $\mathcal{V}(v_3, \cdots, v_{2m})$ be the set of vertices consisting of v_j vertices of valence j, $j = 3, 4, \cdots, 2m$. Then there is a natural fibration

$$\pi : \mathcal{D}(v_3, \cdots, v_{2m}) \longrightarrow \mathcal{V}(v_3, \cdots, v_{2m}).$$

The automorphism group G of this fibration is given by

$$G = \prod_{j=3}^{2m} \mathfrak{S}_j^{v_j} \ltimes \prod_{j=3}^{2m} \mathfrak{S}_{v_j} \subset \mathfrak{S}_{|\mathcal{D}(v_3,\cdots,v_{2m})|}.$$

Let us denote by $\mathcal{P}(v_3, \cdots, v_{2m})$ the collection of all pairing schemes on $\mathcal{D}(v_3, \cdots, v_{2m})$. Then

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}x^2} x^{\sum_{j=3}^{2m} j v_j} dx \prod_{j=3}^{2m} \frac{1}{(j!)^{v_j} v_j!} &= \frac{|\mathcal{P}(v_3, \cdots, v_{2m})|}{|G|} \\
&= \frac{1}{|G|} \sum_{[P]\in \mathcal{P}(v_3,\cdots,v_{2m})\big/G} \big|G\cdot P\big| \\
&= \frac{1}{|G|} \sum_{[P]\in \mathcal{P}(v_3,\cdots,v_{2m})\big/G} \big|G/G_P\big| \\
&= \sum_{[P]\in \mathcal{P}(v_3,\cdots,v_{2m})\big/G} \frac{1}{|G_P|}
\end{aligned}$$

$$= \sum_{\Gamma} \frac{1}{|\mathrm{Aut}(\Gamma)|},$$

where Γ runs all graphs whose vertex set is equal to $\mathcal{V}(v_3, \cdots, v_{2m})$. We have thus proved

$$\mathcal{A}\left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}x^2\right) \exp\left(\sum_{j=3}^{2m} \frac{t_j}{j!}x^j\right) dx\right) = \sum_{\substack{\Gamma \text{ graph with vertices} \\ \text{of valence } j=3,4,\cdots,2m}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2m} t_j^{v_j(\Gamma)} \in \mathbb{Q}[[t_3, t_4, \cdots, t_{2m}]], \tag{2.16}$$

where $v_j(\Gamma)$ denotes the number of j-valent vertices of the graph Γ. The expansion result is a divergent formal power series in $\mathbb{Q}[[t_3, t_4, \cdots, t_{2m}]]$ with rational coefficients.

The number m chosen in the integral is artificial. Since the asymptotic expansion makes sense only when we have a holomorphic function, we placed it so that the integral $Z(t, m)$ converges. As a result, we obtained an artificial constraint in (2.16) that the graph Γ cannot have any vertices of valence greater than $2m$. In the rest of this section, let us investigate the limit

$$\lim_{m\to\infty} \mathcal{A}\big(Z(t, m)\big).$$

First let us recall the **Krull topology** of the formal power series ring $K[[t]]$ with coefficients in a field K. Let $\mathcal{J}_n = t^n K[[t]]$ be the ideal generated by t^n. The Krull topology is introduced to the ring $K[[t]]$ by defining the collection $\{\mathcal{J}_n\}_{n\geq 0}$ as the basis for open neighborhoods of $0 \in K[[t]]$. Since $\mathcal{J}_{n+1} \subset \mathcal{J}_n$, we have a **projective system**

$$\cdots \longrightarrow K[[t]]/\mathcal{J}_{n+1} \xrightarrow{p_{n+1}} K[[t]]/\mathcal{J}_n \longrightarrow \cdots .$$

Note that

$$\bigcap_{n\geq 0} \mathcal{J}_n = \{0\}.$$

Therefore, the natural homomorphism

$$K[[t]] \longrightarrow \varprojlim_n K[[t]]/\mathcal{J}_n$$

is injective, and hence they are canonically isomorphic.

In the same spirit, let us define the ring $K[[t_1, t_2, t_3, \cdots]]$ of formal power series in infinitely many variables as follows. We introduce the degree of each variable by

$$\deg(t_n) = n, \qquad n = 1, 2, 3, \cdots . \tag{2.17}$$

There is a natural inclusion

$$K[[t_1, t_2, \cdots, t_m]] \subset K[[t_1, t_2, \cdots, t_m, t_{m+1}]]. \tag{2.18}$$

Let $\mathcal{J}_n^m$ denote the ideal of $K[[t_1, t_2, \cdots, t_m]]$ generated by all polynomials of homogeneous degree n. Note that if $m \geq n$, then the natural inclusion (2.18) induces

$$K[[t_1, t_2, \cdots, t_m]]/\mathcal{J}_n^m = K[[t_1, t_2, \cdots, t_m, t_{m+1}]]/\mathcal{J}_n^{m+1}.$$

We also have a natural projection

$$K[[t_1, t_2, \cdots, t_m, t_{m+1}]] \longrightarrow K[[t_1, t_2, \cdots, t_m, t_{m+1}]]/(t_{m+1})$$

$$\xrightarrow{\sim} K[[t_1, t_2, \cdots, t_m]],$$

which induces a projective system

$$\cdots \longrightarrow K[[t_1, t_2, \cdots, t_n, t_{n+1}]]\big/\mathcal{J}_{n+1}^{n+1} \longrightarrow K[[t_1, t_2, \cdots, t_n]]\big/\mathcal{J}_n^n \longrightarrow \cdots .$$

We can now *define*

$$K[[t_1, t_2, t_3, \cdots]] \underset{\text{def}}{=} \varprojlim_n K[[t_1, t_2, \cdots, t_n]]\big/\mathcal{J}_n^n. \tag{2.19}$$

Let $\mathcal{J}_n$ denote the ideal of $K[[t_1, t_2, t_3, \cdots]]$ generated by polynomials of homogeneous degree n. This ideal is generated by a finite number of monomials of degree n. By definition,

$$K[[t_1, t_2, t_3, \cdots]]\big/\mathcal{J}_n = K[[t_1, t_2, \cdots, t_n]]\big/\mathcal{J}_n^n,$$

and we have

$$\bigcap_{n\geq 0} \mathcal{J}_n = \{0\}.$$

The Krull topology of $K[[t_1, t_2, t_3, \cdots]]$ is defined by identifying the collection $\{\mathcal{J}_n\}_{n\geq 0}$ as the basis for open neighborhoods of $0 \in K[[t_1, t_2, t_3, \cdots]]$. Since

$$\mathcal{J}_n \cap K[[t_1, t_2, \cdots, t_m]] = \mathcal{J}_n^m,$$

the induced topology on the subring

$$K[[t_1, t_2, \cdots, t_m]] \subset K[[t_1, t_2, t_3, \cdots]]$$

agrees with the canonical Krull topology of $K[[t_1, t_2, \cdots, t_m]]$.

With these preparations, let us go back to the asymptotic expansion (2.16). For a graph $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$, let us denote

$$\begin{aligned} v(\Gamma) &= |\mathcal{V}| = \text{the number of vertices of } \Gamma, \\ e(\Gamma) &= |\mathcal{E}| = \text{the number of edges of } \Gamma. \end{aligned} \tag{2.20}$$

As before, $v_j(\Gamma)$ denotes the number of j-valent vertices. It is easy to see that

$$v(\Gamma) = \sum_j v_j(\Gamma), \qquad e(\Gamma) = \frac{1}{2}\sum_j j v_j(\Gamma). \tag{2.21}$$

Therefore, the degree of the monomial in (2.16) is given by

$$\deg\left(\prod_{j=3}^{2m} t_j^{v_j(\Gamma)}\right) = 2e(\Gamma),$$

which takes only even values. Although bounding $v = v(\Gamma)$ does not bound the set of graphs with v vertices, if we fix the number $e(\Gamma)$, then there are only finitely many graphs with $e(\Gamma)$ edges. Hence every coefficient of a monomial in (2.16) is a finite sum. In particular, we can rearrange the summation of the asymptotic series as

$$\mathcal{A}\Big(Z\big((t_3, t_4, \cdots, t_{2m}), m\big)\Big) = \sum_{n\geq 0} \sum_{\substack{\Gamma \text{ graph with } e(\Gamma)=n \\ \text{and valence } j=3,\cdots,2m}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2m} t_j^{v_j(\Gamma)},$$

and for every $n \geq 0$, the graph sum

$$\sum_{\substack{\Gamma \text{ graph with } e(\Gamma)=n \\ \text{and valence } j=3,\cdots,2m}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2m} t_j^{v_j(\Gamma)}$$

is a weighted homogeneous polynomial of degree $2n$ if there is a graph Γ with exactly n edges. We also note that the maximum of $jv_j(\Gamma)$ for every given graph Γ does not exceed $2e(\Gamma)$. Therefore, for a fixed n and an arbitrary $m \geq n$, the polynomial

(2.22)

$$\begin{aligned}
\mathcal{A}\Big(Z\big((t_3, t_4, \cdots, t_{2m}), m\big)\Big) \mod \mathcal{J}_{2n+1}^{2m} &= \sum_{\substack{\Gamma \text{ graph with } e(\Gamma)\leq n \\ \text{and valence } j=3,\cdots,2n}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2n} t_j^{v_j(\Gamma)} \\
&\in \mathbb{Q}[[t_3, t_4, \cdots, t_{2m}]]\big/\mathcal{J}_{2n+1}^{2m} \\
&= \mathbb{Q}[[t_3, t_4, t_5, \cdots]]\big/\mathcal{J}_{2n+1}
\end{aligned}$$

is **stable**, i.e., it does not depend on m as long as it is larger than n. In the light of this stability, let us consider a sequence of polynomials

$$\mathcal{A}\Big(Z\big((t_3, t_4, \cdots, t_{2m}), m\big)\Big) \mod \mathcal{J}_m^{2m} \in \mathbb{Q}[[t_3, t_4, t_5, \cdots]]\big/\mathcal{J}_m$$

for $m \geq 0$. This defines an element of the projective system

$$\cdots \longrightarrow \mathbb{Q}[[t_3, t_4, t_5, \cdots]]\big/\mathcal{J}_{m+1} \longrightarrow \mathbb{Q}[[t_3, t_4, t_5, \cdots]]\big/\mathcal{J}_m \longrightarrow \cdots .$$

Definition 2.12. We define the limit of $\mathcal{A}\big(Z(t,m)\big)$ as m goes to ∞ as an element of the projective limit of the above projective system:

$$\begin{aligned}
&\lim_{m\to\infty} \mathcal{A}\Big(Z\big((t_3, t_4, \cdots, t_{2m}), m\big)\Big) \\
&\quad = \left\{\mathcal{A}\Big(Z\big((t_3, t_4, \cdots, t_{2m}), m\big)\Big) \mod \mathcal{J}_m\right\}_{m\geq 0} \\
&\quad \in \varprojlim_m \mathbb{Q}[[t_3, t_4, t_5, \cdots]]\big/\mathcal{J}_m \\
&\quad = \mathbb{Q}[[t_3, t_4, t_5, \cdots]].
\end{aligned}$$

Theorem 2.13 (Asymptotic expansion of the scalar integral)**.** *The asymptotic expansion $\mathcal{A}\big(Z(t,m)\big)$ of (2.16) has a well-defined limit as m goes to ∞, and the limiting formal power series as an element of $\mathbb{Q}[[t_3, t_4, t_5, \cdots]]$ is given by*

$$\begin{aligned}
&\lim_{m\to\infty} \mathcal{A}\left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}x^2\right) \exp\left(\sum_{j=3}^{2m} \frac{t_j}{j!} x^j\right) dx\right) \\
&\qquad\qquad = \sum_{n\geq 0} \sum_{\substack{\Gamma \text{ graph} \\ \text{with } e(\Gamma)=n}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2n} t_j^{v_j(\Gamma)}.
\end{aligned}$$

For every fixed $n \geq 0$, the graph sum is a finite sum, and the product $\prod_j t_j^{v_j(\Gamma)}$ is a monomial of degree $2n$.

Remark. If we set $t_j = 0$ for all $j \geq 3$, then the integral has value 1. This corresponds to the homogeneous degree 0 term of the formal power series in the right hand side. Since it means $e(\Gamma) = 0$, and since we are not allowing any vertex to have valence less than 3, the graph $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$ is an empty object. Therefore, we *define*

$$\sum_{\substack{\Gamma \text{ graph} \\ \text{with } e(\Gamma)=0}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{0} t_j^{v_j(\Gamma)} = 1$$

to make the equality hold for all cases.

Proof. The only remaining thing we have to check is that for every $n \geq 0$, the weighted homogeneous polynomial

$$\sum_{\substack{\Gamma \text{ with} \\ e(\Gamma)=n}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2n} t_j^{v_j(\Gamma)} \tag{2.23}$$

of degree $2n$ appears in the element

$$\left\{ \mathcal{A}\Big(Z\big((t_3, t_4, \cdots, t_{2m}), m\big)\Big) \mod \mathcal{J}_m \right\}_{m \geq 0} \tag{2.24}$$

of the projective limit, and that it is stable as m tends to ∞. From (2.22), if we choose $m \geq 2n+1$, then the homogeneous polynomial (2.23) appears in the sequence (2.24) and is stable for sufficiently large m. This completes the proof. $\square$

2.5. **The Logarithm and the Connectivity of Graphs.** For our purpose of using graph theory in the study of the moduli spaces of Riemann surfaces, we need to consider connected graphs. In the asymptotic expansion of Theorem 2.13, all graphs, connected or non-connected, appear in the right hand side. How can we restrict the sequence to have only connected graphs?

As we have seen, the power series in an infinite number of variables

$$f(t) = f(t_3, t_4, t_5, \cdots) = \sum_{n \geq 0} \sum_{\substack{\Gamma \text{ graph} \\ \text{with } e(\Gamma)=n}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2n} t_j^{v_j(\Gamma)} \in \mathbb{Q}[[t_3, t_4, t_5, \cdots]] \tag{2.25}$$

is a well-defined element. Therefore, its subseries

$$h(t) = h(t_3, t_4, t_5, \cdots) = \sum_{n > 0} \sum_{\substack{\Gamma \text{ connected} \\ \text{graph with } e(\Gamma)=n}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2n} t_j^{v_j(\Gamma)} \tag{2.26}$$

is also well-defined.

Remark. We considered the case when the graph Γ was an empty object in the last section, and gave the value 1 to the leading term of (2.25). This means that an empty set is counted as a graph. However, **we do not consider an empty graph to be connected.** This is consistent with the definition of a connected topological space, which dictates that an empty set is *not* connected. This is the reason our series $h(t)$ of (2.26) does not have the constant term.

Theorem 2.14 (Sequence of connected graphs)**.** *Let $f(t)$ and $h(t)$ be as above. Then*

$$f(t) = e^{h(t)} = \sum_{m=0}^{\infty} \frac{1}{m!}\big(h(t)\big)^m.$$

Proof. The **order** of a formal power series in $\mathbb{Q}[[t_3, t_4, t_5, \cdots]]$ is the degree of the lowest degree non-zero homogeneous polynomial (called the **leading term**) that appears in the series. Thus $h(t)$ has order 4, and the leading term corresponds to the unique graph with one vertex and two edges. In particular,

$$\big(h(t)\big)^m \in \mathcal{J}_k$$

if $4m \geq k$. Therefore,

$$\left(e^{h(t)} \mod \mathcal{J}_k\right) \in \mathbb{Q}[t_3, t_4, \cdots, t_{k-1}]$$

is a polynomial with rational coefficients for every $k \geq 0$.

Now consider the graph expansion

$$\frac{1}{m!}\big(h(t)\big)^m \equiv \frac{1}{m!} \sum_{\substack{\Gamma_1, \Gamma_2, \cdots, \Gamma_m \\ e(\Gamma_i) < k}} \prod_{i=1}^{m} \frac{1}{|\mathrm{Aut}(\Gamma_i)|} \prod_{j \geq 3} t_j^{v_j(\Gamma_i)} \mod \mathcal{J}_{2k}. \tag{2.27}$$

Let

$$\Gamma = \Gamma_1 \cup \Gamma_2 \cup \cdots \cup \Gamma_m$$

be the graph with m connected components Γ_i, $i = 1, 2, \cdots, m$. Obviously, we have

$$\mathrm{Aut}(\Gamma) = \left(\prod_{i=1}^{m} \mathrm{Aut}(\Gamma_i)\right) \ltimes \mathfrak{S}_m.$$

In particular,

$$\frac{1}{|\mathrm{Aut}(\Gamma)|} = \frac{1}{m!} \prod_{i=1}^{m} \frac{1}{|\mathrm{Aut}(\Gamma_i)|}. \tag{2.28}$$

Because of the construction of Γ, we also have

$$\prod_{j \geq 3} t_j^{v_j(\Gamma)} = \prod_{j \geq 3} t_j^{v_j(\Gamma_1 \cup \cdots \cup \Gamma_m)} = \prod_{i=1}^{m} \prod_{j \geq 3} t_j^{v_j(\Gamma_i)}. \tag{2.29}$$

From (2.27), (2.28) and (2.29), we obtain

$$\frac{1}{m!}\big(h(t)\big)^m \equiv \sum_{\substack{\Gamma \text{ graph with } m \text{ connected} \\ \text{components and } e(\Gamma) < k}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j \geq 3} t_j^{v_j(\Gamma)} \mod \mathcal{J}_{2k}.$$

Since the bound on $e(\Gamma)$ also bounds the number of connected components in Γ, we have

$$\begin{aligned} e^{h(t)} &\equiv \sum_{\substack{\Gamma \text{ graph with} \\ e(\Gamma) < k}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j \geq 3} t_j^{v_j(\Gamma)} \mod \mathcal{J}_{2k} \\ &\equiv f(t) \mod \mathcal{J}_{2k} \end{aligned}$$

for every $k \geq 0$. This establishes $f(t) = e^{h(t)}$. $\square$

The formal power series

$$f_{>0}(t) = f(t) - f(0) = f(t) - 1$$

has a positive order. Therefore,

$$\log f(t) \underset{\text{def}}{=} -\sum_{m\geq 1} \frac{(-1)^m}{m} \big(f_{>0}(t)\big)^m \in \mathbb{Q}[[t_3, t_4, t_5, \cdots]]$$

is well-defined. Of course it is $h(t)$:

$$\log f(t) = h(t).$$

In the same way as in the previous section, we can establish the equality for the asymptotic series that contains only connected graphs:

Theorem 2.15 (Asymptotic expansion with connected graphs)**.** *As an element of* $\mathbb{Q}[[t_3, t_4, t_5, \cdots]]$*, we have an equality*

$$\lim_{m\to\infty} \log \mathcal{A}\left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{1}{2}x^2\right) \exp\left(\sum_{j=3}^{2m} \frac{t_j}{j!} x^j\right) dx\right)$$
$$= \sum_{n>0} \sum_{\substack{\Gamma \text{ connected} \\ \text{graph with } e(\Gamma)=n}} \frac{1}{|\mathrm{Aut}(\Gamma)|} \prod_{j=3}^{2n} t_j^{v_j(\Gamma)}.$$

Remark. The factor $1/j!$ accompanying t_j in the integral is introduced so that the asymptotic expansion has a natural interpretation as the **generating function** of the reciprocal of the orders of graph automorphism groups.

2.6. **Ribbon Graphs and Oriented Surfaces.** We have found in Theorem 2.15 the generating function of the orders of the automorphism groups of connected graphs. Our next challenge is to restrict the graphs to be drawn on a surface, in particular a Riemann surface. The shape of the right hand side of the asymptotic formula inevitably becomes more complicated, because it should contain information of the genus of the surface on which a connected graph is drawn, while the left hand side has amazingly simple generalizations. In the next two sections we develop the extensions of the left hand side of the formula to deal with graphs on surfaces. In this section, we identify the conditions we have to impose on the graphs so that they are placed on a surface. For historical remarks on the research on graphs embedded in surfaces, we refer to Ringel [40].

Suppose we have a graph drawn on an oriented surface. The orientation of the surface determines a **cyclic ordering** of the edges incident to each vertex. This consideration motivates our definition of **ribbon graphs**.

Definition 2.16 (Cyclic ordering)**.** Consider a set X of j labeled objects, and let G be a subgroup of the symmetric group $\mathfrak{S}_j$. A G**-ordering** on X is a coset of the quotient space $\mathfrak{S}_j/G$. When $G = \mathbb{Z}/j\mathbb{Z}$ is the cyclic group of order j, we simply say the G-ordering a **cyclic ordering**.

Remark. An element of $\mathfrak{S}_j/G$ gives an ordering of the j elements of X that is invariant under the G-action. Therefore, when $G = \mathfrak{S}_j$, the G-ordering means no ordering. The $\{1\}$-ordering is thus the same as the ordering in the usual sense, and there are $j!$ different ways of ordering the elements of X.

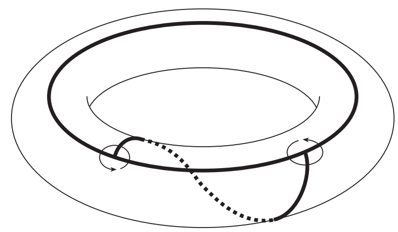

FIGURE 2.7. A graph drawn on an oriented surface. At each vertex, the orientation of the surface determines a cyclic ordering of the edges incident to the vertex.

Definition 2.17 (Ribbon graphs)**.** Let $\Gamma = (\mathcal{V}, \mathcal{E}, \mathcal{I})$ be a graph, and $\Gamma_{\mathcal{E}}$ its edge refinement. A **cyclic ordering** of edges at a vertex $V \in \mathcal{V}$ means a cyclic ordering of the set of half edges incident to V. A **ribbon graph structure** $\mathcal{C}$ on Γ is the collection of cyclic ordering at every vertex of Γ. A **ribbon graph** is a graph with a ribbon graph structure. We use the notation $\Gamma^R = (\Gamma, \mathcal{C})$ to indicate a ribbon graph. The graph Γ is the **underlying graph** of a ribbon graph Γ^R.

Remark. The terminology **ribbon graph** was first used by Kontsevich in [20]. Earlier, the same object was called a **fatgraph** by Penner [38], but the notion was well-known to the graph theory community for long time and called by different names, such as a **map of a surface**. A cyclic ordering is commonly referred to as a **rotation** or a **rotation system** in the literature ([11], [40]). We adopt the new terminology due to Kontsevich, that best represents the nature of the object.

Definition 2.18 (Ribbon graph isomorphism)**.** Let $(\Gamma, \mathcal{C})$ and $(\Gamma', \mathcal{C}')$ be two ribbon graphs. A graph isomorphism of the edge refinement

$$\phi : \Gamma_{\mathcal{E}} \xrightarrow{\sim} \Gamma'_{\mathcal{E}'}$$

induces the **pull-back** ribbon graph structure $\phi^*(\mathcal{C}')$. The ribbon graphs $(\Gamma, \mathcal{C})$ and $(\Gamma', \mathcal{C}')$ are said to be **isomorphic** if $\mathcal{C} = \phi^*(\mathcal{C}')$.

To visualize a ribbon graph Γ^R, let us provide an oriented plane (with the orientation represented by a counter clockwise rotation), and place every vertex on the plane so that the cyclic ordering of half edges at each vertex is drawn in the order of clockwise rotation. Since the half edges incident to a j-valent vertex V are cyclically ordered, let us prepare indices $i_1, i_2, \cdots, i_j$ to name each half edge. But instead of using a single index to name a half edge, we use double indices $i_1 i_2, i_2 i_3, \cdots, i_j i_1$ for the j half edges incident to V. In this way we can keep track of the cyclic ordering better. Since we use double indices, we can also use double lines to represent half edges. Now a half edge looks like a *ribbon*. This ribbon is a subset of the oriented plane, and hence it inherits the natural orientation. The orientation of each ribbon can be presented by an arrow on its boundary that is consistent with the orientation of the ribbon. Thus a j-valent vertex V looks like one in Figure 2.8.

When two vertices V and V' are connected by an edge, it is done in a way that the orientation of the half edges are consistent (Figure 2.8). Thus we have an oriented topological surface with boundary as a visualization of a ribbon graph Γ^R.

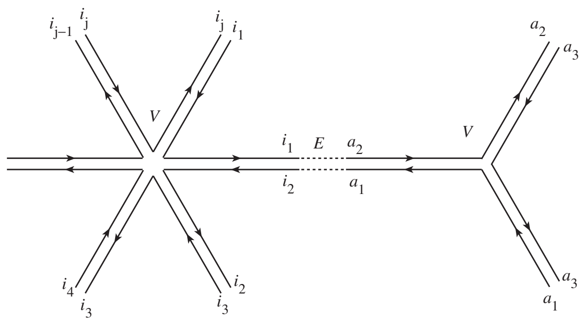

FIGURE 2.8. Two vertices with cyclic ordering connected to one another with a consistent orientation. A half edge labeled by $i_1 i_2$ is connected to a half edge labeled by $a_1 a_2$. In this example, the outward line i_1 from V is connected with the inward line a_2 going into V', and the outward line a_1 form V' with the inward line i_2 going to V.

Definition 2.19 (Boundary circuit of a ribbon graph)**.** Let $\Gamma^R = (\Gamma, \mathcal{C})$ be a ribbon graph, and $\overrightarrow{E}_1$ and $\overrightarrow{E}_2$ be two directed edges of Γ. The edge $\overrightarrow{E}_2$ is said to be the **successor** of the edge $\overrightarrow{E}_1$ at vertex V if the half edges E_{1+} and E_{2-} are incident to V, and E_{2-} is right after E_{1+} with respect to the cyclic ordering at the vertex V. A sequence $(\overrightarrow{E}_0, V_0, \overrightarrow{E}_1, V_1, \cdots, \overrightarrow{E}_n, V_n)$ of directed edges and vertices is said to be a **boundary circuit** of Γ^R if

(1) the directed edge $\overrightarrow{E}_{i+1}$ is the successor of $\overrightarrow{E}_i$ at the vertex V_i for $i = 0, 1, \cdots, n-1$;
(2) $(\overrightarrow{E}_n, V_n) = (\overrightarrow{E}_0, V_0)$; and
(3) $\overrightarrow{E}_0, \overrightarrow{E}_1, \cdots, \overrightarrow{E}_{n-1}$ are distinct as directed edges (i.e., the same edge can appear in a sequence up to twice with opposite directions).

In the topological visualization of the ribbon graph Γ^R, the boundary circuit $(\overrightarrow{E}_0, V_0, \overrightarrow{E}_1, V_1, \cdots, \overrightarrow{E}_n, V_n)$ is an oriented circle with n segments. We denote by $b(\Gamma^R)$ the number of boundary circuits of a ribbon graph Γ^R.

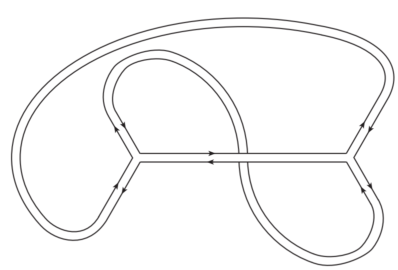

FIGURE 2.9. An example of a ribbon graph with two vertices, three edges and one boundary circuit.

As we have observed in Figure 2.7, a graph drawn on an oriented surface is naturally a ribbon graph. Conversely, a connected ribbon graph has a **canonical**

embedding into an oriented surface. Let Γ^R be a ribbon graph. Since its boundary circuit is an oriented circle, we can *glue* an oriented disk to each boundary circuit with consistent orientation. Then we obtain a compact oriented topological surface on which the underlying graph Γ is drawn. Let us denote by Σ_{Γ^R} the compact oriented topological surface thus obtained. The topological type of Σ_{Γ^R} minus $b(\Gamma^R)$ points is the same as the graph Γ. The genus of the surface is determined by the equation

$$2 - 2g(\Sigma_{\Gamma^R}) = \chi(\Sigma_{\Gamma^R}) = v(\Gamma) - e(\Gamma) + b(\Gamma^R). \tag{2.30}$$

By comparing Figure 2.7 and Figure 2.9, we see that the compact surface Σ_{Γ^R} for the ribbon graph Figure 2.9 is a torus. Indeed, we have

$$2 - 2g = 2 - 3 + 1 = 0,$$

hence $g = 0$.

Definition 2.20 (Ribbon graph automorphism)**.** A **ribbon graph automorphism** is an automorphism of a graph that preserves the cyclic ordering at each vertex. The group of automorphisms of a ribbon graph Γ^R is denoted by $\mathrm{Aut}(\Gamma^R)$. The group of ribbon graph automorphisms that fix each boundary circuit is denoted by $\mathrm{Aut}_b(\Gamma^R)$.

Remark. We have a natural subgroup inclusion relation

$$\mathrm{Aut}_b(\Gamma^R) \subset \mathrm{Aut}(\Gamma^R) \subset \mathrm{Aut}(\Gamma),$$

where Γ is the underlying graph of a ribbon graph Γ^R. When we study the orbifold structure of moduli spaces of Riemann surfaces, we use the more restricted automorphism group $\mathrm{Aut}_b(\Gamma^R)$. On the other hand, in the Feynman diagram expansion of Hermitian matrix integrals, it is the group $\mathrm{Aut}(\Gamma^R)$ that naturally occurs.

We have thus established an **intrinsic** condition for the graph to be drawn on an oriented surface. We can write down a generating function of the ribbon graph automorphism groups. Our next attention is the analysis counterpart of this generating function.

2.7. **Hermitian Matrix Integrals.** The goal of this section is to identify the asymptotic expansion of a Hermitian matrix integral

$$Z_{\mathcal{H}}(t, N; m) = \frac{1}{C_N} \int_{\mathcal{H}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right) \exp\left(\sum_{j=3}^{2m} \frac{t_j}{j}\mathrm{trace}(X^j)\right) dX \tag{2.31}$$

in terms of ribbon graphs. Here $\mathcal{H}_N$ denotes the space of all $N \times N$ hermitian matrices, and for $X = \left[x_{ij}\right] \in \mathcal{H}_N$, dX is the standard Lebesgue measure on $\mathcal{H}_N = \mathbb{R}^{N^2}$:

$$dX = \bigwedge_{i=1}^{N} dx_{ii} \wedge \bigwedge_{i<j} \Big(dRe(x_{ij}) \wedge dIm(x_{ij})\Big).$$

We note that

$$\mathrm{trace}(X^2) = \mathrm{trace}(X^\dagger X) = \sum_i (x_{ii})^2 + 2\sum_{i<j} (Re(x_{ij}))^2 + 2\sum_{i<j} (Im(x_{ij}))^2$$

is a positive definite quadratic form. The overall normalization constant is chosen to be

$$C_N = \int_{\mathcal{H}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right) dX$$

$$(2.32) \qquad = \int_{\mathbb{R}^{N^2}} \exp\left(-\frac{1}{2}\sum_{i,j}\overline{x}_{ij}x_{ij}\right)\prod_{i=1}^{N} dx_{ii}\prod_{i<j} dRe(x_{ij})\wedge dIm(x_{ij})$$

$$= \left(\sqrt{2\pi}\,\right)^N \pi^{N(N-1)/2} = \left(\sqrt{2}\,\right)^N \left(\sqrt{\pi}\,\right)^{N^2}.$$

The integral $Z_{\mathcal{H}}(t, N; m)$ is absolutely convergent for $Re(t_{2m}) < 0$ and arbitrary $t_3, t_4, \cdots, t_{2m-1}$. Therefore, $Z_{\mathcal{H}}(t, N; m)$ is a holomorphic function in

$$t = (t_3, t_4, \cdots, t_{2m-1}, t_{2m}) \in \mathbb{C}^{2m-3} \times \Omega_\epsilon,$$

where Ω_ϵ is the same domain as in (2.14). In exactly the same way as in the case of the scalar integral (2.15), we obtain the expansion of $Z_{\mathcal{H}}(t, N; m)$ as a Taylor series in $t_3, t_4, \cdots, t_{2m-1}$ and an asymptotic series in t_{2m}:

$$(2.33) \quad \mathcal{A}\big(Z_{\mathcal{H}}(t, N; m)\big)$$

$$= \frac{1}{C_N}\sum_{v_3\geq 0, v_4\geq 0,\cdots, v_{2m}\geq 0}\prod_{j=3}^{2m}\frac{t_j^{v_j}}{j^{v_j}v_j!}\int_{\mathcal{H}_N}\exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)\prod_{j=3}^{2m}\big(\mathrm{trace}(X^j)\big)^{v_j} dX.$$

We use the following lemmas to calculate this last integral.

Lemma 2.21 (Hermitian matrix differentiation)**.** *Let $J = \big[y_{ij}\big]_{1\leq i,j\leq N}$ be a Hermitian matrix valued variable, and let*

$$\frac{\partial}{\partial J} = \left[\tfrac{\partial}{\partial y_{ij}}\right]_{1\leq i,j\leq N}.$$

Then

$$\left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \exp\big(\mathrm{trace}(X^t J)\big)\bigg|_{J=0} = \big(\mathrm{trace}(X^j)\big)^n.$$

Proof. A simple calculation shows

$$(2.34) \quad \mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j \exp\big(\mathrm{trace}(X^t J)\big)\bigg|_{J=0}$$

$$= \left(\sum_{a_1,a_2,\cdots,a_j}\frac{\partial}{\partial y_{a_1a_2}}\frac{\partial}{\partial y_{a_2a_3}}\cdots\frac{\partial}{\partial y_{a_ja_1}}\right)\exp\left(\sum_{a,b}x_{ab}y_{ab}\right)\Bigg|_{J=0}$$

$$= x_{a_1a_2}x_{a_2a_3}\cdots x_{a_ja_1} = \mathrm{trace}(X^j).$$

Here we have used the fact that

$$\frac{\partial}{\partial y_{ij}}y_{ab} = \delta_{ia}\delta_{jb},$$

and in particular, if $i \neq j$,

$$\frac{\partial}{\partial y_{ij}}y_{ji} = \frac{\partial}{\partial y_{ij}}\overline{y}_{ij} = 0.$$

The desired formula follows from repeating (2.34) n-times. $\square$

Lemma 2.22 (Source term for Hermitian matrix integral)**.** *Let* $J = \big[y_{ij}\big]_{1\le i,j\le N}$ *and* $\frac{\partial}{\partial J} = \Big[\frac{\partial}{\partial y_{ij}}\Big]_{1\le i,j\le N}$ *be as above. Then*

$$\frac{1}{C_N}\int_{\mathcal{H}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)\big(\mathrm{trace}(X^j)\big)^n dX \\ = \left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \exp\left(\frac{1}{2}\mathrm{trace}(J^2)\right)\Bigg|_{J=0}.$$

Proof. Since the integral is absolutely convergent, we can interchange the integration and the differentiation with respect to a parameter, and we obtain

$$\begin{aligned}
&\frac{1}{C_N}\int_{\mathcal{H}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)\big(\mathrm{trace}(X^j)\big)^n dX\\
=&\frac{1}{C_N}\int_{\mathcal{H}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)\left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \exp\big(\mathrm{trace}(X^tJ)\big)\Bigg|_{J=0} dX\\
=&\left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \frac{1}{C_N}\int_{\mathcal{H}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)\exp\big(\mathrm{trace}(X^tJ)\big)\,dX\Bigg|_{J=0}\\
=&\left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \frac{1}{C_N}\int_{\mathcal{H}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^t-J)^2\right)\exp\left(\frac{1}{2}\mathrm{trace}(J^2)\right)dX\Bigg|_{J=0}\\
=&\left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \exp\left(\frac{1}{2}\mathrm{trace}(J^2)\right)\Bigg|_{J=0},
\end{aligned}$$

where we used the translational invariance of the Lebesgue measure dX on $\mathcal{H}_N$ and the fact that $\mathrm{trace}(X^2) = \mathrm{trace}\big((X^t)^2\big)$ and $\mathrm{trace}(X^tJ) = \mathrm{trace}(JX^t)$. ☐

As in the case of scalar integral, the quantity

$$\prod_{j=3}^{2m}\left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^{v_j} \exp\left(\frac{1}{2}\mathrm{trace}(J^2)\right)\Bigg|_{J=0}$$

is non-zero only when a pair of $\frac{\partial}{\partial y_{ij}}$'s in the differential operator operates on the exponential function $\exp\big(\frac{1}{2}\mathrm{trace}(J^2)\big)$. Note that

$$\begin{aligned}
\frac{\partial}{\partial y_{ij}}\cdot\frac{\partial}{\partial y_{k\ell}}\exp\left(\frac{1}{2}\mathrm{trace}(J^2)\right)\Bigg|_{J=0} &= \frac{\partial}{\partial y_{ij}}\cdot\frac{\partial}{\partial y_{k\ell}}\exp\left(\frac{1}{2}\sum_{a,b}y_{ab}\,y_{ba}\right)\Bigg|_{J=0}\\
&=\frac{1}{2}\sum_{a,b}(\delta_{ia}\delta_{jb}\delta_{kb}\delta_{\ell a}+\delta_{ib}\delta_{ja}\delta_{ka}\delta_{\ell b})\\
&=\frac{1}{2}(\delta_{i\ell}\delta_{jk}+\delta_{i\ell}\delta_{jk})\\
&=\delta_{i\ell}\delta_{jk}.
\end{aligned}\tag{2.35}$$

As before, let us introduce a set of dots $\mathcal{D}$ grouped into v_j sets of j dots for given indices $v_3, v_4, \cdots, v_{2m}$. Thus

$$|\mathcal{D}| = \sum_{j=3}^{2m} j v_j.$$

Since we are dealing with the differentiation by a matrix variable, a group of j dots are labeled with double indices like

$$\bullet_{a_1 a_2} \qquad \bullet_{a_2 a_3} \qquad \cdots \qquad \bullet_{a_j a_1}$$

and these labels introduce a **cyclic ordering** of the dots. From (2.35), we know that each pair of dots $(\bullet_{ij}, \bullet_{k\ell})$ in the differential operator contributes $\delta_{i\ell}\delta_{jk}$ in the computation of the derivative. Thus we have

$$\prod_{j=3}^{2m} \left(\operatorname{trace} \left(\frac{\partial}{\partial J} \right)^j \right)^{v_j} \exp\left(\frac{1}{2} \operatorname{trace}(J^2) \right) \Bigg|_{J=0} = \sum_{(P \text{ pairing scheme})} \sum_{(\text{all indices})} \prod_{\substack{\bullet_{ij} \text{ and } \bullet_{k\ell} \\ \text{are paired in } P}} \delta_{i\ell}\delta_{jk}.$$

Example 2.3. For $j = 4$ and $v_j = 1$, we have

$$\begin{aligned}
&\operatorname{trace} \left(\frac{\partial}{\partial J} \right)^4 \exp\left(\frac{1}{2} \operatorname{trace}(J^2) \right) \Bigg|_{J=0} \\
=& \sum_{i,j,k,\ell} \frac{\partial}{\partial J_{ij}} \frac{\partial}{\partial J_{jk}} \frac{\partial}{\partial J_{k\ell}} \frac{\partial}{\partial J_{\ell i}} \exp\left(\frac{1}{2} \sum_{a,b} J_{ab} J_{ba} \right) \Bigg|_{J=0} \\
=& \sum_{i,j,k,\ell} \frac{\partial}{\partial J_{ij}} \frac{\partial}{\partial J_{jk}} \frac{\partial}{\partial J_{k\ell}} \left(J_{i\ell} \cdot \exp\left(\frac{1}{2} \sum_{a,b} J_{ab} J_{ba} \right) \right) \Bigg|_{J=0} \\
=& \sum_{i,j,k,\ell} \frac{\partial}{\partial J_{ij}} \frac{\partial}{\partial J_{jk}} \big(\delta_{ki}\delta_{\ell\ell} + J_{i\ell} J_{\ell k} \big) \exp\left(\frac{1}{2} \sum_{a,b} J_{ab} J_{ba} \right) \Bigg|_{J=0} \\
=& \sum_{i,j,k,\ell} \frac{\partial}{\partial J_{ij}} \big(\delta_{ki}\delta_{\ell\ell} J_{kj} + \delta_{ji}\delta_{k\ell} J_{\ell k} + J_{i\ell}\delta_{j\ell}\delta_{kk} + J_{i\ell} J_{\ell k} J_{kj} \big) \exp\left(\frac{1}{2} \sum_{a,b} J_{ab} J_{ba} \right) \Bigg|_{J=0} \\
=& \sum_{i,j,k,\ell} \big(\delta_{ki}\delta_{\ell\ell}\delta_{ik}\delta_{jj} + \delta_{ji}\delta_{k\ell}\delta_{i\ell}\delta_{jk} + \delta_{ii}\delta_{j\ell}\delta_{j\ell}\delta_{kk} \big) \\
=& \operatorname{trace}(I^3) + \operatorname{trace}(I) + \operatorname{trace}(I^3) \\
=& 2N^3 + N.
\end{aligned}$$

The fibration $\pi : \mathcal{D} \longrightarrow \mathcal{V}$ is defined by mapping a group of j dots to a j-valent vertex. The projection changes a pairing scheme into a **ribbon graph** Γ^R. To visualize the transition, consider the case that $\bullet_{i_1 i_2}$ is paired, or connected, with $\bullet_{a_1 a_2}$. The dot $\bullet_{i_1 i_2}$ is one of the j dots cyclically ordered. So it can be identifies with a half edge of a j-valent vertex V that is placed on an oriented plane. The other dot $\bullet_{i_1 i_2}$ belongs to another set of cyclically ordered dots, so we can identify it with a half edge of another vertex V'. (Of course it is possible that $V = V'$.) The contribution from this pair, $\delta_{i_1a_2}\delta_{i_2a_1}$, can be visualized by connecting the outgoing line labeled by i_1 from V with the incoming line a_2 at V', and i_2 with a_1. On the ribbon graph level, the connection is exactly the same as in Figure 2.8. Thus the quantity $\delta_{i_1a_2}\delta_{i_2a_1}$, called a **propagator** in QFT, is attached to an edge E, and the factor $\delta_{i_1a_2}$ represents one of the oriented boundaries of the ribbon and the other factor, $\delta_{i_2a_1}$, the other oriented boundary.

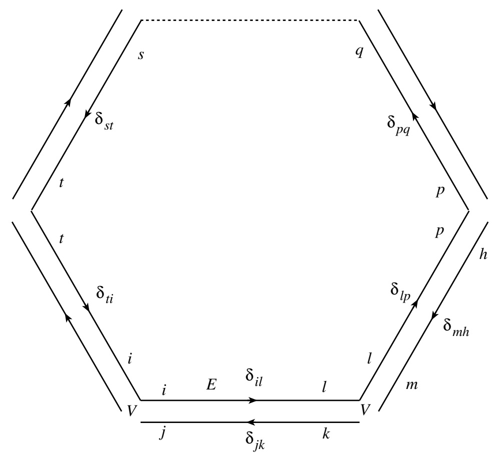

FIGURE 2.10. A propagator around a boundary circuit that is an n-gon.

What happens if we follow a boundary circuit starting with, say $\delta_{i\ell}$? The *next* dot $\bullet_{\ell m}$ represents a half edge incident to vertex V' that follows $\bullet_{k\ell}$ in the cyclic ordering at V'. It is connected to another dot, say $\bullet_{hp}$. Then the factor of the next propagator following $\delta_{i\ell}$ is $\delta_{\ell p}$. In this way, we have a sequence of factors of propagators

$$\delta_{i\ell}\delta_{\ell p}\delta_{pq}\cdots\delta_{st}\delta_{ti}$$

along a boundary circuit of the ribbon graph Γ^R. Note that

$$\sum_{i,\ell,p,q,\cdots,s,t}\delta_{i\ell}\delta_{\ell p}\delta_{pq}\cdots\delta_{st}\delta_{ti} = \mathrm{trace}(I^n) = N,$$

when the boundary circuit is an n-gon (Figure 2.10). Therefore, the product of all propagators for all edges, after taking summation over every index involved, gives $N^{b(\Gamma^R)}$, where $b(\Gamma^R)$ is the number of boundary circuits of Γ^R.

Example 2.4. Example 2.3 produces two ribbon graphs consisting of one vertex and two edges.

We have noted that the fibration $\pi : \mathcal{D} \longrightarrow \mathcal{V}$ has an extra structure for the case of Hermitian matrix integral. For every vertex $V \in \mathcal{V}$, the fiber has a cyclic ordering. Thus the automorphism of the fibration is

$$G = \prod_{j=3}^{2m}\left(\mathbb{Z}/j\mathbb{Z}\right)^{v_j} \ltimes \prod_{j=3}^{2m}\mathfrak{S}_{v_j},$$

whose order, $\prod_j j^{v_j} v_j!$, appears in the coefficient of (2.33). In the same way we proved for the regular graph, the stabilizer subgroup of G of a pairing scheme is identified with the ribbon graph automorphism group. We thus have

$$\prod_{j=3}^{2m} \frac{1}{j^{v_j} v_j!} \left(\operatorname{trace} \left(\frac{\partial}{\partial J} \right)^j \right)^{v_j} \exp\left(\frac{1}{2} \operatorname{trace}(J^2) \right) \Bigg|_{J=0} = \sum_{\substack{\Gamma^R \text{ ribbon graph} \\ v_j(\Gamma^R) = v_j, j=3,\cdots,2m}} \frac{1}{|\operatorname{Aut}(\Gamma^R)|} N^{b(\Gamma^R)}.$$

Plugging it back to (2.33), we obtain

$$\mathcal{A}\big(Z_{\mathcal{H}}(t, N; m)\big) = \sum_{\substack{\Gamma^R \text{ ribbon graph} \\ \text{with valence } j=3,4,\cdots,2m}} \frac{1}{|\operatorname{Aut}(\Gamma^R)|} N^{b(\Gamma^R)} \prod_{j=3}^{2m} t_j^{v_j(\Gamma^R)}. \tag{2.36}$$

The argument of the Krull topology and taking the logarithm for the connected graphs are the same as before. Finally, we have established

$$\lim_{m\to\infty} \log \mathcal{A}\big(Z_{\mathcal{H}}(t, N; m)\big) = \sum_{\Gamma^R \text{ connected ribbon graph}} \frac{1}{|\operatorname{Aut}(\Gamma^R)|} N^{b(\Gamma^R)} \prod_{j\geq 3} t_j^{v_j(\Gamma^R)}. \tag{2.37}$$

2.8. **Möbius Graphs and Non-Orientable Surfaces.** We have observed that a cyclic ordering of half edges at each vertex of a graph is an **intrinsic** condition for the graph to be drawn on an oriented surface. Let us now turn our attention to graphs drawn on a non-orientable surface. What is an intrinsic condition for a graph to be on a non-orientable surface? And how do we find a canonical embedding of a graph into a non-orientable surface? Before answering these questions, let us review some basic facts about non-orientable surfaces.

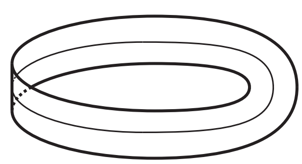

FIGURE 2.11. A Möbius band.

The simplest non-orientable surface is a Möbius band (Figure 2.11). It is created by gluing one pair of parallel edges of a rectangle in a certain manner. We start with an oriented rectangle. Note that the orientation induces a natural orientation of the boundary edges. If we glue a parallel pair of edges in a way preserving the orientation, then we obtain an oriented cylinder. On the other hand, if we glue the same parallel edges in a way inconsistent with the orientation of the rectangle, then we obtain a Möbius band (Figure 2.12 top). The boundary of a Möbius band is a circle. The homotopy type of a Möbius band is that of a circle, and hence, it has Euler characteristic 0.

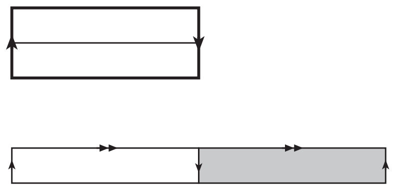

FIGURE 2.12. Making a Möbius band from a rectangle.

It is well-known that when we cut a Möbius band along the middle circle, we obtain an orientable cylinder (Figure 2.12 bottom). Following this cutting process backward, we see that a Möbius band can be constructed by identifying the antipodes of the top circle of a cylinder (Figure 2.13).

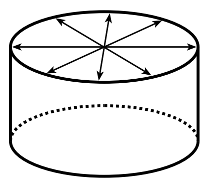

FIGURE 2.13. Möbius band is obtained by identifying the antipodes of the top circle of a cylinder.

Compact non-orientable surfaces without boundary are classified by their Euler characteristic, which takes all integer values less than or equal to 1. A **non-orientable surface of genus** g, denoted by X_g, is constructed as follows. First we remove $g+1$ disjoint disks from a sphere. We then glue a Möbius band to each boundary circle. The surface thus obtained is non-orientable and compact without boundary. Since a Möbius band has Euler characteristic 0, X_g has Euler characteristic $1-g$. We note here that gluing a Möbius band to a boundary circle is topologically the same as identifying the antipodes on the boundary circle. Therefore, X_0 is homeomorphic to a real projective plane $S^2\big/\langle\iota\rangle$, where $\iota : S^2 \longrightarrow S^2$ is the map of a sphere that interchanges the antipodes. X_0 can be also constructed by attaching a disk to the boundary of a Möbius band.

A non-orientable surface X_1 of genus 1 is best known as a **Klein bottle**. It is constructed by gluing the two ends of an oriented cylinder in a way inconsistent with the orientation chosen (Figure 2.14).

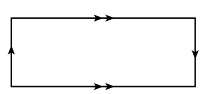

FIGURE 2.14. A Klein bottle.

Of course it is the same as gluing two Möbius bands to a two-punctured sphere. To see that these two different constructions give the same result, let us start with the standard construction, Figure 2.15 top left. Recall that gluing a Möbius band is the same as identifying the antipodes of a boundary circle. First, we cut out a piece $ABEDGF$ from the sphere (Figure 2.15 top right). We then flip the cut-out piece over, and glue it back to the surface (Figure 2.15 bottom). Since we cut the surface along the line segment AB, the segment becomes two arcs a and g. Arc g is at the bottom of the colored piece of Figure 2.15 bottom, because the piece is flipped over. Similarly, the line segment GD becomes two arcs c and e. Arc d represents the same arc AHG, and b is equal to BCD. Originally the arc AFG is glued to AHG by identifying the antipodes of the circle. But since the cut piece is flipped over, arcs d and h are now glued straight, as indicated in Figure 2.15 bottom. The same gluing is done to arcs b and f. At this stage, the surface we have thus constructed is again a sphere with two disks removed. Note that it is homeomorphic to a cylinder. The pair-wise identification of $a = g$ and $c = e$ is indeed the same as gluing two ends of an oriented cylinder in a manner that is inconsistent with the orientation.

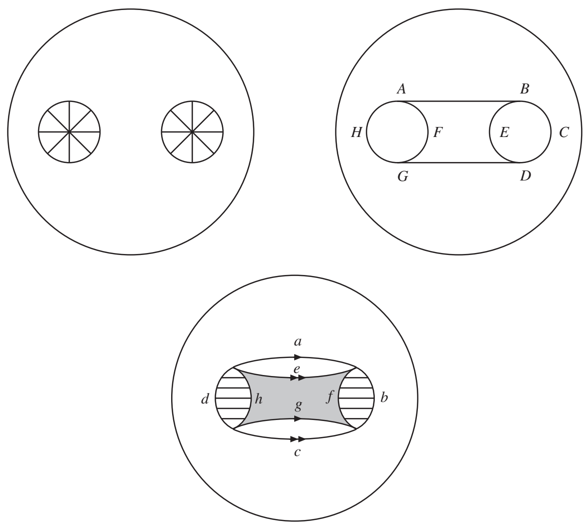

FIGURE 2.15. Two different constructions of a Klein bottle.

The above consideration shows that gluing two Möbius bands to two boundary circles of an oriented punctured sphere is the same as gluing these two circles in an orientation-inconsistent manner. We can now modify our previous construction of X_g in a more visual way. First, let us consider the case when $g = 2k$ is even. We remove $g + 1 = 2k + 1$ disjoint disks from an oriented sphere. To one of the boundary circles, we glue a Möbius band. We note that the surface is already non-orientable. Out of the remaining $g = 2k$ boundary circles, let us form k pairs of two circles. Instead of gluing two Möbius bands to a pair of circles, we simply attach a cylinder. Of course we have to connect the two circles *in an orientation-inconsistent manner*, but since the surface is already non-orientable, we can simply connect the two circles in whichever way we want. In particular, we can just attach a cylinder, or a handle, to a pair of circles. The surface thus obtained looks like one in Figure 2.16.

Now consider an oriented surface Σ_g of even genus $g = 2k$. It has an **orientation-reversing involution**

$$\iota : \Sigma_g \longrightarrow \Sigma_g.$$

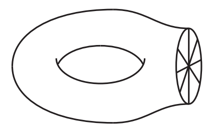

FIGURE 2.16. A non-orientable surface X_2 of genus 2. It is obtained by gluing a Möbius band and a cylinder to a 3-punctured sphere.

An easy way to visualize it is to place the surface Σ_g with half of the handles in one side. An orientation-reversing involution can be given by the antipodal mapping around the center of the surface (Figure 2.17). It is now obvious that

$$X_g \cong \Sigma_g / \langle \iota \rangle.$$

We note that the center of the antipodes is not on the surface. Thus the action of the involution does not have any fixed points on the surface.

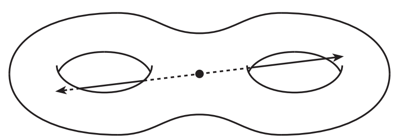

FIGURE 2.17. The antipodal map as an orientation-reversing involution of an oriented surface of even genus. The center of the antipodes is not on the surface. Thus the involution does not have any fixed points on the surface.

The case of an odd genus $g = 2k + 1$ is almost the same. We start with an oriented sphere with $g+1 = 2k+2$ disjoint disks removed. Let us pair all boundary circles into $k + 1$ groups. To the first pair, we attach an oriented cylinder in an orientation-inconsistent manner. It makes sense because both the cylinder and the $(g + 1)$-punctured sphere are oriented. The surface thus obtained is non-orientable and has still $2k$ boundary circles. We then attach a cylinder to each pair of circles to make a compact non-orientable surface without boundary. It looks like one in Figure 2.18.

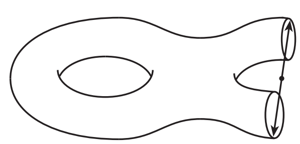

FIGURE 2.18. A non-orientable surface X_3 of genus 3. The two boundary circles have the natural orientation coming from the orientation of the surface. The circles are glued to one another in an orientation-inconsistent manner.

As before, for an odd genus case, we can also find an orientation-reversing involution ι of an oriented surface Σ_g of genus g such that X_g is the quotient of Σ_g by the involution (Figure 2.19).

Thus we have shown the following.

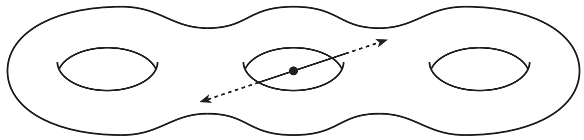

FIGURE 2.19. An orientation-reversing involution of an oriented surface of odd genus.

Proposition 2.23 (Oriented covering of a non-orientable surface)**.** *For every compact non-orientable surface X_g of genus g, there is a compact oriented surface Σ_g and an orientation-reversing involution $\iota : \Sigma_g \longrightarrow \Sigma_g$ such that*

$$X_g \cong \Sigma_g / \langle \iota \rangle.$$

Remark. Since ι does not have any fixed points, the quotient is a topological manifold, and we have

$$\chi(X_g) = \frac{1}{2}\chi(\Sigma_g) = 1 - g.$$

Proposition 2.23 motivates us to introduce the notion of **Möbius graphs**. These are the graphs drawn on an orientable or non-orientable surface.

Definition 2.24 (Möbius graphs)**.** A 2**-color** ribbon graph is a ribbon graph with an element of $\mathbb{Z}/2\mathbb{Z}$ assigned to every edge. An **orientation-color change** at a vertex is an operation on a 2-color ribbon graph that reverses the cyclic order of the vertex and the color of an edge by adding $1 \in \mathbb{Z}/2\mathbb{Z}$ if one of its half edges is incident to the vertex. Thus if an edge is doubly incident to a vertex, then the color of this edge does not change after an orientation-color change at the vertex. Two 2-color ribbon graphs are said to be **equivalent** if one is obtained from the other by a successive application of orientation-color change operations. An equivalence class of a 2-color ribbon graph is called a **Möbius graph**. A **Möbius graph automorphism** is a pair consisting of a permutation of vertices and a permutation of half edges that preserve the incidence relation, color at each edge, and either preserve or reverse the cyclic ordering at each vertex. We can make a ribbon graph a Möbius graph by giving color 0 at each edge. A Möbius graph is said to be **orientable** if it is equivalents to a ribbon graph, and **non-orientable** otherwise.

We can give a **topological realization** of a Möbius graph by indicating the color of an edge with twisting (color 1) or no twisting (color 0). The topological realization is an orientable or non-orientable surface with boundary. Figure 2.20 shows two equivalent Möbius graphs. We note that a Möbius graph has boundary circuits, which are the boundary components of the topological realization, but they are no longer canonically oriented. We can construct a compact connected surface Σ_{Γ^M}, orientable or non-orientable, from a connected Möbius graph Γ^M by attaching a disk to each boundary circuit of Γ^M.

Remark. The notion of Möbius graphs has appeared in the literature in many different names, such as *voltage graphs with a rotation system and the voltage group* $\mathbb{Z}/2\mathbb{Z}$ (cf. [11]). The graphs on a surface, orientable or non-orientable, are studies in the context of map coloring problem for surfaces of genus $g > 1$ in [40]. Since we do not consider any other *voltage groups* than $\mathbb{Z}/2\mathbb{Z}$, we use the more topologically appealing terminology.

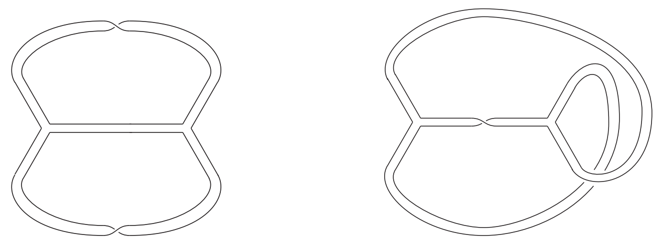

Figure 2.20. Two equivalent Möbius graphs consisting of two vertices, three edges, and one boundary circuit. The graphs are interchanged one another by an orientation-color change operation at the right hand side vertex.

The **opposite** of a ribbon graph is the ribbon graph obtained by reversing the cyclic order at every vertex. Generically the opposite is a different ribbon graph, but they are equivalent as a Möbius graph.

For every connected non-orientable Möbius graph Γ^M, there is a connected orientable Möbius graph Γ^2 and a fixed-point free involution $\iota : \Gamma^2 \longrightarrow \Gamma^2$ such that

$$\Gamma^2/\langle\iota\rangle \cong \Gamma^M. \tag{2.38}$$

We call Γ^2 the **covering graph** of Γ^M, which is unique up to isomorphism. This corresponds to the situation of Proposition 2.23.

The construction of Γ^2 is as follows. First we apply the orientation-color change operation, if necessary, to place all vertices of Γ^M on an oriented plane so that the cyclic ordering at each vertex is consistent with the orientation of the plane. (Of course Γ^M does not have to be planer and its edges may not be placed on the plane.) We then prepare two copies of Γ^M, calling them Γ^M and Γ'^M. Let E be an edge of Γ^M of color 1, incident to vertices V_1 and V_2 (which can be the same vertex), and E', V_1' and V_2' be the corresponding edge and vertices of Γ'^M. Remove E and E' from $\Gamma^M \cup \Gamma'^M$ and connect V_1 and v_2' with an edge $\overline{V_1V_2'}$, and give it color 1. Likewise, connect V_2 and V_1' with an edge $\overline{V_1'V_2}$ and give it color 1. Let us call this procedure **cross-bridge construction** (see Figure 2.21).

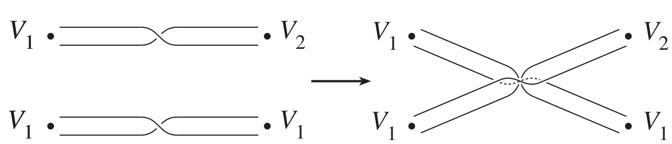

Figure 2.21. The cross-bridge construction.

The covering Γ^2 is obtained by applying the cross-bridge construction to every edge of Γ^M of color 1. The involution ι maps every vertex $V \in \Gamma^M$ to its counterpart $V' \in \Gamma'^M$, every edge of color 0 of Γ^M to its counterpart of Γ'^M preserving its incidence, and a new edge $\overline{V_1V_2'}$ to its cross-bridge partner $\overline{V_1'V_2}$. Figure 2.22 shows the covering graph of the Möbius graph of Figure 2.20.

Let us show that Γ^2 is orientable. First, consider the subset of Γ^M consisting of all vertices and edges of color 0. On this subset we can introduce an orientation consistent with the oriented plane. To the counterpart subset of Γ'^M, we give the

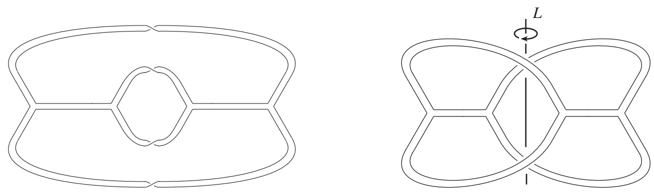

FIGURE 2.22. The covering graph Γ^2 of Figure 2.20. It has 4 vertices, 6 edges, and 2 boundary circuits. The Möbius graph on the left is equivalent to the ribbon graph on the right. The $180°$ rotation ι about the vertical line L is an orientation reversing involution, and the quotient $\Gamma^2/\langle\iota\rangle$ is the original Möbius graph of Figure 2.20.

opposite orientation. These two subsets of Γ^2 are connected only with edges of color 1. Therefore, the orientation of the subsets can be extended consistently to the whole graph Γ^2. By construction, the involution ι is orientation reversing with respect to any orientation we choose on Γ^2. To see that the covering does not depend on the choice of an element of the equivalence class of Γ^M, let us apply an orientation-color change operation at a vertex V of Γ^M, and call it Γ^M_V. The cross-bridge construction is performed on Γ^M_V and its copy Γ'^M_V to make the covering Γ^2_V. Let V' be the copy of V on Γ'^M_V. Apply the orientation-color change operation simultaneously to V and V', and then interchange V and V'. This operation makes Γ^2_V and Γ^2 equivalent.

We note that the covering Γ^2 has twice as many boundary circuits as Γ^M does. It follows from the fact that a boundary circuit of a Möbius graph always passes through even number of twisted (i.e., color 1) edges. To see this, consider an ϵ-neighborhood B_ϵ of a boundary circuit of the topological model of the graph that passes through n twisted edges. The ϵ-neighborhood B_ϵ is orientable since it is a part of the disk attached to create the compact surface Σ_{Γ^M}. We note that B_ϵ consists of n twisted ϵ-bands and other non-twisted bands. Since it is orientable, n is even. Now, from the cross-bridge construction, one sees that the lift of a boundary circuit of Γ^M consists of two boundary circuits of Γ^2 of the same length.

The fixed-point free and orientation-reversing involution

$$\iota : \Gamma^2 \longrightarrow \Gamma^2$$

induces a fixed-point free and orientation-reversing involution

$$\iota : \Sigma_{\Gamma^2} \longrightarrow \Sigma_{\Gamma^2} \tag{2.39}$$

of the compact orientable surface Σ_{Γ^2}. The quotient surface $\Sigma_{\Gamma^2}/\langle\iota\rangle$ is the non-orientable surface Σ_{Γ^M}.

Let Γ^M be a connected non-orientable Möbius graph. The *genus* $g(\Gamma^M)$ of Γ^M is the genus of the compact orientable surface Σ_{Γ^2} associated with the covering Γ^2 of Γ^M. Let $v(\Gamma^M)$, $e(\Gamma^M)$, and $b(\Gamma^M)$ (resp.) denote the number of vertices, edges, and the boundary circuits of Γ^M (resp.). Then we have the genus-Euler characteristic relation

$$v(\Gamma^M) - e(\Gamma^M) + b(\Gamma^M) = 1 - g(\Gamma^M), \tag{2.40}$$

since the Euler characteristic of Σ_{Γ^2} is $2-2g(\Gamma^2)$ and since Γ^2 is a double covering of Γ^M.

2.9. **Symmetric Matrix Integrals.** (This section is under construction.)

Our next target is an integral over the space of real symmetric matrices. Let $\mathcal{S}_N$ denote the space of all real symmetric matrices of size N. The goal of this section is to identify the asymptotic expansion of

$$Z_{\mathcal{S}}(t,N;m) = \frac{1}{C_N}\int_{\mathcal{S}_N} \exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)\exp\left(\sum_{j=3}^{2m}\frac{t_j}{c_j}\mathrm{trace}(X^j)\right)dX, \tag{2.41}$$

where the overall normalization constant C_N and the coefficient $1/c_j$ of the parameter t_j are determined later.

First of all, we have to define the measure of integration dX. Let $X=\big[x_{ij}\big]_{ij}$ be a real symmetric matrix of size N. Since $x_{ij}=x_{ji}$, we define

$$dX = \bigwedge_{i<j} dx_{ij}\wedge\bigwedge_i dx_{ii}.$$

The measure dX is the standard Lebesgue measure of $\mathcal{S}_N$, which is a real vector space of dimension $N(N+1)/2$. We note that

$$\mathrm{trace}(X^2)=\mathrm{trace}(X^tX)=\sum_{i,j}(x_{ij})^2 = 2\sum_{i<j}(x_{ij})^2+\sum_i (x_{ii})^2 \tag{2.42}$$

is a positive definite quadratic form. From (2.42), it is obvious what we should choose as the normalization constant. So we define

$$C_N=\int_{\mathcal{S}_N}\exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)dX = \sqrt{\pi}^{N(N-1)/2}\sqrt{2\pi}^N = 2^{\frac{N}{2}}\pi^{\frac{N(N+1)}{4}}. \tag{2.43}$$

We also note that if we choose positive constants $c_j>0$, then the integral $Z_{\mathcal{S}}(t,N;m)$ is absolutely convergent for $Re(t_{2m})<0$ and arbitrary $t_3,t_4,\cdots,t_{2m-1}$. Therefore, $Z_{\mathcal{S}}(t,N;m)$ is a holomorphic function in

$$t=(t_3,t_4,\cdots,t_{2m-1},t_{2m})\in\mathbb{C}^{2m-3}\times\Omega_\epsilon,$$

where Ω_ϵ is the same domain as in (2.14). In exactly the same way as in the case of the scalar integral (2.15), we obtain the expansion of $Z_{\mathcal{S}}(t,N;m)$ as the Taylor series in $t_3,t_4,\cdots,t_{2m-1}$ and the asymptotic series in t_{2m}:

$$\begin{aligned}&\mathcal{A}\big(Z_{\mathcal{S}}(t,N;m)\big) \\ &= \frac{1}{C_N}\sum_{v_3\geq 0, v_4\geq 0,\cdots,v_{2m}\geq 0}\prod_{j=3}^{2m}\frac{t_j^{v_j}}{c_j^{v_j}v_j!}\int_{\mathcal{S}_N}\exp\left(-\frac{1}{2}\mathrm{trace}(X^2)\right)\prod_{j=3}^{2m}\big(\mathrm{trace}(X^j)\big)^{v_j}dX.\end{aligned} \tag{2.44}$$

We use the following lemmas to calculate this last integral.

Lemma 2.25 (Matrix differentiation)**.** *Let $J=\big[y_{ij}\big]_{1\leq i,j\leq N}$ be a symmetric matrix valued variable, and let*

$$\frac{\partial}{\partial J}=\left[\tfrac{1}{2}\tfrac{\partial}{\partial y_{ij}}\right]_{1\leq i,j\leq N}.$$

Then

$$\left(\mathrm{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n\exp\big(\mathrm{trace}(X^tJ)\big)\bigg|_{J=0}=\big(\mathrm{trace}(X^j)\big)^n.$$

Proof.

$$
\begin{aligned}
&\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^j \exp\big(\operatorname{trace}(X^t J)\big)\Big|_{J=0}\\
=&\left(\left(\frac{1}{2}\right)^j \sum_{a_1,a_2,\cdots,a_j} \frac{\partial}{\partial y_{a_1a_2}}\frac{\partial}{\partial y_{a_2a_3}}\cdots\frac{\partial}{\partial y_{a_ja_1}}\right)\exp\left(\sum_{a,b} x_{ab}y_{ab}\right)\Bigg|_{J=0}\\
=&\left(\frac{1}{2}\right)^j\left(\sum_{a,b} x_{ab}(\delta_{a_1a}\delta_{a_2b}+\delta_{a_1b}\delta_{a_2a})\right)\left(\sum_{a,b} x_{ab}(\delta_{a_2a}\delta_{a_3b}+\delta_{a_2b}\delta_{a_3a})\right)\cdots\\
&\qquad\left(\sum_{a,b} x_{ab}(\delta_{a_ja}\delta_{a_1b}+\delta_{a_jb}\delta_{a_1a})\right)\\
=&\left(\frac{1}{2}\right)^j (x_{a_1a_2}+x_{a_2a_1})(x_{a_2a_3}+x_{a_3a_2})\cdots(x_{a_ja_1}+x_{a_1a_j})\\
=&x_{a_1a_2}x_{a_2a_3}\cdots x_{a_ja_1}\\
=&\operatorname{trace}(X^j).
\end{aligned}
$$

The desired formula follows from repeating the above computation n-times. □

Lemma 2.26 (Source term for symmetric matrix integral)**.** *Let* $J = \big[y_{ij}\big]_{1\le i,j\le N}$ *and* $\frac{\partial}{\partial J} = \left[\frac{1}{2}\frac{\partial}{\partial y_{ij}}\right]_{1\le i,j\le N}$ *be as above. Then*

$$
\begin{aligned}
&\frac{1}{C_N}\int_{\mathcal{S}_N} \exp\left(-\frac{1}{2}\operatorname{trace}(X^2)\right)\big(\operatorname{trace}(X^j)\big)^n dX\\
&\qquad\qquad = \left(\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \exp\left(\frac{1}{2}\operatorname{trace}(J^2)\right)\Bigg|_{J=0}.
\end{aligned}
$$

Proof. Since the integral is absolutely convergent, we can interchange the integration and the differentiation with respect to a parameter, and we obtain

$$
\begin{aligned}
&\frac{1}{C_N}\int_{\mathcal{S}_N} \exp\left(-\frac{1}{2}\operatorname{trace}(X^2)\right)\big(\operatorname{trace}(X^j)\big)^n dX\\
=&\frac{1}{C_N}\int_{\mathcal{S}_N} \exp\left(-\frac{1}{2}\operatorname{trace}(X^2)\right)\left(\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \exp\big(\operatorname{trace}(X^tJ)\big)\Big|_{J=0} dX\\
=&\left(\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \frac{1}{C_N}\int_{\mathcal{S}_N} \exp\left(-\frac{1}{2}\operatorname{trace}(X^2)\right)\exp\big(\operatorname{trace}(X^tJ)\big)\,dX\Bigg|_{J=0}\\
=&\left(\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \frac{1}{C_N}\int_{\mathcal{S}_N} \exp\left(-\frac{1}{2}\operatorname{trace}(X-J)^2\right)\exp\left(\frac{1}{2}\operatorname{trace}(J^2)\right)dX\Bigg|_{J=0}\\
=&\left(\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^j\right)^n \exp\left(\frac{1}{2}\operatorname{trace}(J^2)\right)\Bigg|_{J=0},
\end{aligned}
$$

where we used the translational invariance of the Lebesgue measure dX on $\mathcal{S}_N$. □

As in the case of scalar integral, the quantity

$$\prod_{j=3}^{2m}\left(\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^{j}\right)^{v_j}\exp\left(\frac{1}{2}\operatorname{trace}(J^2)\right)\Bigg|_{J=0}$$

is non-zero only when a pair of $\frac{1}{2}\frac{\partial}{\partial y_{ij}}$'s in the differential operator operates on the exponential function $\exp\left(\frac{1}{2}\operatorname{trace}(J^2)\right)$. Note that

$$\begin{aligned}
&\left(\frac{1}{2}\frac{\partial}{\partial y_{ij}}\right)\left(\frac{1}{2}\frac{\partial}{\partial y_{k\ell}}\right)\exp\left(\frac{1}{2}\operatorname{trace}(J^2)\right)\Bigg|_{J=0}\\
=&\left(\frac{1}{2}\frac{\partial}{\partial y_{ij}}\right)\left(\frac{1}{2}\frac{\partial}{\partial y_{k\ell}}\right)\exp\left(\frac{1}{2}\sum_{a,b}(y_{ab})^2\right)\Bigg|_{J=0}\\
=&\left(\frac{1}{2}\right)^2\sum_{a,b}(\delta_{ia}\delta_{jb}+\delta_{ib}\delta_{ja})(\delta_{ka}\delta_{\ell b}+\delta_{kb}\delta_{\ell a})\\
=&\frac{1}{4}(\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk}+\delta_{i\ell}\delta_{jk}+\delta_{ik}\delta j\ell)\\
=&\frac{1}{2}(\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk}).
\end{aligned}$$

As before, let us introduce v_j sets of j dots for given indices $v_3, v_4, \cdots, v_{2m}$. Since we are dealing with the differentiation by a matrix variable, the j dots are labeled with double indices like

$$\bullet_{a_1a_2}\qquad \bullet_{a_2a_3}\qquad \cdots \qquad \bullet_{a_ja_1}$$

and these labels introduce a **cyclic ordering** of the dots. Then we have

$$\begin{aligned}
&\prod_{j=3}^{2m}\left(\operatorname{trace}\left(\frac{\partial}{\partial J}\right)^{j}\right)^{v_j}\exp\left(\frac{1}{2}\operatorname{trace}(J^2)\right)\Bigg|_{J=0}\\
&=\sum_{\mathcal{P}\text{ pairing scheme}}\ \prod_{\substack{\bullet_{ij}\text{ and }\bullet_{k\ell}\\ \text{are paired in }\mathcal{P}}}\frac{1}{2}(\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk})\\
&=\left(\frac{1}{2}\right)^{\frac{1}{2}\sum_{j=3}^{2m}jv_j}\sum_{\mathcal{P}\text{ pairing scheme}}\ \prod_{\substack{\bullet_{ij}\text{ and }\bullet_{k\ell}\\ \text{are paired in }\mathcal{P}}}(\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk}).
\end{aligned}$$

Now let us choose the constant c_j such that

$$\frac{1}{c_j}=\frac{(\sqrt{2})^j}{j}. \tag{2.45}$$

To be continued...

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