

# ELEMENTARY MODULAR IWASAWA THEORY

HARUZO HIDA

## CONTENTS

1. Curves over a field	3
1.1. Plane curves	3
1.2. Tangent space and local rings	5
1.3. Projective space	8
1.4. Projective plane curve	9
1.5. Divisors	11
1.6. The theorem of Riemann–Roch	12
1.7. Regular maps from a curve into projective space	13
2. Elliptic curves	14
2.1. Abel’s theorem	14
2.2. Weierstrass Equations of Elliptic Curves	15
2.3. Moduli of Weierstrass Type	17
3. Modular forms and functions	20
3.1. Geometric modular forms	20
3.2. Topological Fundamental Groups	21
3.3. Classical Weierstrass $\wp$ -function	23
3.4. Complex Modular Forms	24
3.5. Weierstrass $\zeta$ and $\sigma$ functions	26
3.6. Product $q$ -expansion	28
3.7. Klein forms	29
4. Modular units	32
4.1. Siegel units	33
4.2. Distribution on $p$ -divisible groups	34
4.3. Stickelberger distribution	35
4.4. Rank of distribution	36
4.5. Cusps of $X(N)$	37
4.6. Finiteness of $Cl_{X(N)}$	38
4.7. Siegel units generate $\mathcal{A}_{p^m}^\times$	38
4.8. Fricke–Wohlfahrt theorem	42
4.9. Siegel units and Stickelberger’s ideal	44
4.10. Cuspidal class number formula	46
4.11. Cuspidal class number formula for $X_1(N)$ .	49

We discuss the first three topics of the following list:

- (1) Explicit construction of modular forms and modular functions;
- (2) Determination of units in the elliptic modular function fields (modular units);
- (3) The cuspidal class group of modular curves including a proof of the cuspidal class number formula;
- (4) Construction of units (called elliptic units) of the Hilbert ring class field of imaginary quadratic fields as specialization of modular units;
- (5) Iwasawa theory for imaginary quadratic fields via elliptic units.

If time allows, we further go into the topics (4) and (5). A modular curve is an affine plane curve (i.e., an open Riemann surface) classifying elliptic curves with certain additional structure (called level structure) naturally defined over  $\mathbb{Q}$ . As a Riemann surface, it is a quotient of the upper half complex plane by  $\mathrm{SL}_2(\mathbb{Z})$  (and its subgroups). Adding finite number of points (called cusps), we can complete the curve into a projective curve (i.e., a compact Riemann surface). Most of recent progress in number theory and arithmetic geometry is based on the study of modular curves and modular forms defined on them; e.g., proof of Iwasawa's conjectures (Mazur–Wiles) and Fermat's last theorem (Wiles).

Modular units are the units in the ring of holomorphic functions of the affine modular curve. Divisors supported on cusps modulo principal divisors (divisors of modular units) give the cuspidal class group (which is the torsion subgroup of rational points of the Jacobian of the modular curve). We can determine explicitly the group of modular units via classical results of Weierstrass and Siegel (like the determination of cyclotomic units in the field generated by roots of unity by Dirichlet–Kummer).

Striking points are that the class group is finite and that we have an explicit class number formula of Kubert–Lang in terms of the Dirichlet L-values at  $s = 2$  (while the classical class number formula of Dirichlet/Kummer of the cyclotomic field is in terms of the values at  $s = 1$ ). This is done by generalizing Stickelberger's theory of cyclotomic class groups to the setting of modular curves. This theory gives a base of the proof of the Iwasawa main conjecture by Mazur–Wiles.

For each prime  $p > 2$ ,  $\{\sin(\pi a/p)/\sin(\pi/p)\}_{0 < a < p/2}$  gives independent units in the integer ring of the field of  $p$ -th root of unity for each prime  $p$  (the cyclotomic units). Analogously, the value of modular units at a point on the upper half complex plane belonging to an imaginary quadratic field  $K$  gives independent units in the (certain) Hilbert ring class field over  $K$ , and this is a base of the generalization of the cyclotomic Iwasawa theory to the elliptic Iwasawa theory of imaginary quadratic fields. If time allows, we describe these finer results (possibly including the class number formulas of the ring class field as an index of elliptic units inside the entire units).

We start with a sketch of the theory of affine plane curves and how to compactify it in a projective space. Then we turn analytic and construct rational functions over modular curves in analytic means. This explicit construction helps us to compute cuspidal class groups inside the Jacobian of the projective modular curves.

1. CURVES OVER A FIELD

Any algebraic curve over an algebraically closed field can be embedded into the 3-dimensional projective space  $\mathbf{P}^3$  (e.g., [ALG, IV.3.6]) and any closed curve in  $\mathbf{P}^3$  is birationally isomorphic to a curve inside  $\mathbf{P}^2$  (a plane curve; see [ALG, IV.3.10]), we give some details of the theory of plain curve defined over a field  $k \subset \mathbb{C}$  in this section to accustom with the theory of curves. We also write  $\bar{k}$  for the algebraic closure  $\bar{k}$  of  $k$  inside  $\mathbb{C}$ .

**1.1. Plane curves.** Let  $\mathfrak{a}$  be a principal ideal of the polynomial ring  $k[X, Y]$ . Note that polynomial rings over a field is a unique factorization domain. We thus have prime factorization  $\mathfrak{a} = \prod_{\mathfrak{p}} \mathfrak{p}^{e(\mathfrak{p})}$  with principal primes  $\mathfrak{p}$ . We call  $\mathfrak{a}$  square free if  $0 \leq e(\mathfrak{p}) \leq 1$  for all principal primes  $\mathfrak{p}$ . Fix a square-free  $\mathfrak{a}$ . The set of  $A$ -rational points for any  $k$ -algebra  $A$  of a *plane curve* is given by the zero set

$$V_{\mathfrak{a}}(A) = \{(x, y) \in A^2 \mid f(x, y) = 0 \text{ for all } f(X, Y) \in \mathfrak{a}\}.$$

Obviously, for a generator  $f(X, Y)$  of  $\mathfrak{a}$ , we could have defined

$$V_{\mathfrak{a}}(A) = V_f(A) = \{(x, y) \in A^2 \mid f(\mathbf{x}) = 0\},$$

but this does not depend on the choice of generators and depends only on the ideal  $\mathfrak{a}$ ; so, it is more appropriate to write  $V_{\mathfrak{a}}$ . As an exceptional case, we note  $V_{(0)}(A) = A^2$ . Geometrically, we think of  $V_{\mathfrak{a}}(\mathbb{C})$  as a curve in  $\mathbb{C}^2 = V_{(0)}(\mathbb{C})$  (the 2-dimensional “plane”). This view point is more geometric. In this sense, for any algebraically closed field  $K$  over  $k$ , a point  $x \in V_{\mathfrak{a}}(K)$  is called a geometric point with coefficients in  $K$ , and  $V_{(f)}(K) \subset V_{(0)}(K)$  is called the geometric curve in  $V_{(0)}(K) = K^2$  defined by the equation  $f(X, Y) = 0$ .

By Hilbert’s zero theorem (Nullstellensatz; see [CRT] Theorem 5.4 and [ALG] Theorem I.1.3A), writing  $\bar{\mathfrak{a}}$  the principal ideal of  $\bar{k}[X, Y]$  generated by  $\mathfrak{a}$ , we have

$$(1.1) \quad \bar{\mathfrak{a}} = \{f(X, Y) \in \bar{k}[X, Y] \mid f(x, y) = 0 \text{ for all } (x, y) \in V_{\mathfrak{a}}(\bar{k})\}.$$

Thus we have a bijection

$$\{\text{square-free ideals of } \bar{k}[X, Y]\} \leftrightarrow \{\text{plane curves } V_{\mathfrak{a}}(\bar{k}) \subset V_{(0)}(\bar{k})\}.$$

The association  $V_{\mathfrak{a}} : A \mapsto V_{\mathfrak{a}}(A)$  is a covariant functor from the category of  $k$ -algebras to the category of sets (denoted by *SETS*). Indeed, for any  $k$ -algebra homomorphism  $\sigma : A \rightarrow A'$ ,  $V_{\mathfrak{a}}(A) \ni (x, y) \mapsto (\sigma(x), \sigma(y)) \in V_{\mathfrak{a}}(A')$  as  $0 = \sigma(0) = \sigma(f(x, y)) = f(\sigma(x), \sigma(y))$ . Thus  $\mathfrak{a} = \bar{\mathfrak{a}} \cap k[X, Y]$  is determined uniquely by this functor, but the value  $V_{\mathfrak{a}}(A)$  for an individual  $A$  may not determine  $\mathfrak{a}$ .

From number theoretic view point, studying  $V_{\mathfrak{a}}(A)$  for a small field is important. Thus it would be better regard  $V_{\mathfrak{a}}$  as a functor in some number theoretic setting.

If  $\mathfrak{a} = \prod_{\mathfrak{p}} \mathfrak{p}$  for principal prime ideals  $\mathfrak{p}$ , by definition, we have

$$V_{\mathfrak{a}} = \bigcup_{\mathfrak{p}} V_{\mathfrak{p}}.$$

The plane curve  $V_{\mathfrak{p}}$  (for each prime  $\mathfrak{p} \mid \mathfrak{a}$ ) is called an *irreducible* component of  $V_{\mathfrak{a}}$ . Since  $\mathfrak{p}$  is a principal prime, we cannot further have non-trivial decomposition  $V_{\mathfrak{p}} = V \cup W$  with plane curves  $V$  and  $W$ . A prime ideal  $\mathfrak{p} \subset k[X, Y]$  may decompose into a

product of primes in  $\bar{k}[X, Y]$ . If  $\mathfrak{p}$  remains prime in  $\bar{k}[X, Y]$ , we call  $V_{\mathfrak{p}}$  *geometrically irreducible*.

Suppose that we have a map  $F_A = F(\phi)_A : V_{\mathfrak{a}}(A) \rightarrow V_{\mathfrak{b}}(A)$  given by two polynomials  $\phi_X(X, Y), \phi_Y(X, Y) \in k[X, Y]$  (independent of  $A$ ) such that  $F_A(x, y) = (\phi_X(x), \phi_Y(y))$  for all  $(x, y) \in V_{\mathfrak{a}}(A)$  and all  $k$ -algebras  $A$ . Such a map is called a *regular  $k$ -map* or a  *$k$ -morphism* from a plane  $k$ -curve  $V_{\mathfrak{a}}$  into  $V_{\mathfrak{b}}$ . Here  $V_{\mathfrak{a}}$  and  $V_{\mathfrak{b}}$  are plane curve defined over  $k$ . If  $\mathbb{A}^1 = V_{\mathfrak{b}}$  is the affine line, i.e.,  $V_{\mathfrak{b}}(A) \cong A$  for all  $A$  (taking for example  $\mathfrak{b} = (y)$ ), a regular  $k$ -map  $V_{\mathfrak{a}} \rightarrow \mathbb{A}^1$  is called a *regular  $k$ -function*. Regular  $k$ -functions are just functions induced by the polynomials in  $k[x, y]$  on  $V_{\mathfrak{a}}$ ; so,  $R_{\mathfrak{a}}$  is the ring of regular  $k$ -functions of  $V_{\mathfrak{a}}$  defined over  $k$ .

We write  $\text{Hom}_{k\text{-curves}}(V_{\mathfrak{a}}, V_{\mathfrak{b}})$  for the set of regular  $k$ -maps from  $V_{\mathfrak{a}}$  into  $V_{\mathfrak{b}}$ . Obviously, only  $\phi \pmod{\mathfrak{a}}$  can possibly be unique. We have a commutative diagram for any  $k$ -algebra homomorphism  $\sigma : A \rightarrow A'$ :

$$\begin{array}{ccc} V_{\mathfrak{a}}(A) & \xrightarrow{F_A} & V_{\mathfrak{b}}(A) \\ \sigma \downarrow & & \downarrow \sigma \\ V_{\mathfrak{a}}(A') & \xrightarrow{F_{A'}} & V_{\mathfrak{b}}(A'). \end{array}$$

Indeed,

$$\begin{aligned} \sigma(F_A((x, y))) &= (\sigma(\phi_X(x, y)), \sigma(\phi_Y(x, y))) \\ &= (\phi_X(\sigma(x), \sigma(y)), \phi_Y(\sigma(x), \sigma(y))) = F_{A'}(\sigma(x), \sigma(y)). \end{aligned}$$

Thus the  $k$ -morphism is a *natural transformation of functors* (or a *morphism of functors*) from  $V_{\mathfrak{a}}$  into  $V_{\mathfrak{b}}$ . We write  $\text{Hom}_{COF}(V_{\mathfrak{a}}, V_{\mathfrak{b}})$  for the set of natural transformations (we will see later that  $\text{Hom}_{COF}(V_{\mathfrak{a}}, V_{\mathfrak{b}})$  is a set).

The polynomials  $(\phi_X, \phi_Y)$  induces a  $k$ -algebra homomorphism  $\underline{F} : k[X, Y] \rightarrow k[X, Y]$  by pull-back, that is,  $\underline{F}(\Phi(X, Y)) = \Phi(\phi_X(X, Y), \phi_Y(X, Y))$ . Take a class  $[\Phi]_{\mathfrak{b}} = \Phi + \mathfrak{b}$  in  $B = k[X, Y]/\mathfrak{b}$ . Then look at  $\underline{F}(\Phi) \in k[X, Y]$  for  $\Phi \in \mathfrak{b}$ . Since  $(\phi_X(x), \phi_Y(y)) \in V_{\mathfrak{b}}(\bar{k})$  for all  $(x, y) \in V_{\mathfrak{a}}(\bar{k})$ ,  $\Phi(\phi_X(x, y), \phi_Y(x, y)) = 0$  for all  $(x, y) \in V_{\mathfrak{a}}(\bar{k})$ . By Nullstellensatz,  $\underline{F}(\Phi) \in \bar{\mathfrak{a}} \cap k[X, Y] = \mathfrak{a}$ . Thus  $\underline{F}(\mathfrak{b}) \subset \mathfrak{a}$ , and  $\underline{F}$  induces a (reverse)  $k$ -algebra homomorphism

$$\underline{F} : k[X, Y]/\mathfrak{b} \rightarrow k[X, Y]/\mathfrak{a}$$

making the following diagram commutative:

$$\begin{array}{ccc} k[X, Y] & \xrightarrow{\underline{F}} & k[X, Y] \\ \downarrow & & \downarrow \\ k[X, Y]/\mathfrak{b} & \xrightarrow{\underline{F}} & k[X, Y]/\mathfrak{a}. \end{array}$$

We write  $R_{\mathfrak{a}} = k[X, Y]/\mathfrak{a}$  and call it the affine ring of  $V_{\mathfrak{a}}$ . Here is a useful (but tautological) lemma which is a special case of Yoneda's lemma (in Math 210 series):

**Lemma 1.1.** *We have a canonical isomorphism:*

$$\text{Hom}_{COF}(V_{\mathfrak{a}}, V_{\mathfrak{b}}) \cong \text{Hom}_{k\text{-curves}}(V_{\mathfrak{a}}, V_{\mathfrak{b}}) \cong \text{Hom}_{k\text{-alg}}(R_{\mathfrak{b}}, R_{\mathfrak{a}}).$$

The first association is covariant and the second is contravariant.

Since  $R_{\mathfrak{a}} = \text{Hom}_{k\text{-alg}}(k[X] = R_{(Y)}, R_{\mathfrak{a}}) = \text{Hom}_{k\text{-curves}}(V_{\mathfrak{a}}, \mathbb{A}) = \text{Hom}_{COF}(V_{\mathfrak{a}}, \mathbb{A})$ , we can recover the ring  $R_{\mathfrak{a}}$  as a collection of morphisms from  $V_{\mathfrak{a}}$  into the affine line  $\mathbb{A}$ . Here is a sketch of the proof.

*Proof.* First we note  $V_{\mathfrak{a}}(A) \cong \text{Hom}_{ALG/k}(R_{\mathfrak{a}}, A)$  via  $(a, b) \leftrightarrow (\Phi(X, Y) \mapsto \Phi(a, b))$ . Thus as functors, we have  $V_{\mathfrak{a}}(?) \cong \text{Hom}_{ALG/k}(R_{\mathfrak{a}}, ?)$ . We identify the two functors  $A \mapsto V_{\mathfrak{a}}(A)$  and  $A \mapsto \text{Hom}(R_{\mathfrak{a}}, A)$  in this way, and in this sense, we write the functor corresponding to  $V_{\mathfrak{a}}$  as  $\text{Spec}(R_{\mathfrak{a}})$ . Then the main point of the proof of the lemma is to construct from a given natural transformation  $F \in \text{Hom}_{COF}(V_{\mathfrak{a}}, V_{\mathfrak{b}})$  a  $k$ -algebra homomorphism  $\underline{F} : R_{\mathfrak{b}} \rightarrow R_{\mathfrak{a}}$  giving  $F$  by  $V_{\mathfrak{a}}(A) = \text{Hom}_{ALG/k}(R_{\mathfrak{a}}, A) \ni \phi \xrightarrow{F_A} \phi \circ \underline{F} \in \text{Hom}_{ALG/k}(R_{\mathfrak{b}}, A) = V_{\mathfrak{b}}(A)$ . Then the following exercise finishes the proof, as plainly if we start with  $\underline{F}$ , the above association gives rise to  $F$ .  $\square$

**Exercise 1.2.** Let  $\underline{F} = F_{R_{\mathfrak{a}}}(\text{id}_{R_{\mathfrak{a}}}) \in V_{R_{\mathfrak{b}}}(R_{\mathfrak{a}}) = \text{Hom}_{ALG/k}(R_{\mathfrak{b}}, R_{\mathfrak{a}})$ , where  $\text{id}_{R_{\mathfrak{a}}} \in V_{\mathfrak{a}}(R_{\mathfrak{a}}) = \text{Hom}_{ALG/k}(R_{\mathfrak{a}}, R_{\mathfrak{a}})$  is the identity map. Then prove that  $\underline{F}$  does the job.

We call  $V_{\mathfrak{a}}$  *irreducible* (resp. *geometrically irreducible*) if  $\mathfrak{a}$  is a prime ideal (resp.  $\bar{\mathfrak{a}} = \mathfrak{a}\bar{k}[X, Y]$  is a prime ideal in  $\bar{k}[X, Y]$ ). For a general noetherian  $k$ -algebra  $\mathcal{A}$ , we define  $V_{\mathcal{A}} = \text{Spec}(\mathcal{A})$  to be the functor  $A \mapsto \text{Hom}_{k\text{-alg}}(\mathcal{A}, A) = \text{Spec}(\mathcal{A})(A)$ , and call  $\text{Spec}(\mathcal{A})$  a curve associated to the ring  $\mathcal{A}$ . In the same way as above,  $\mathcal{A} \cong \text{Hom}_{COF}(V_{\mathcal{A}}, \mathbb{A})$ . The curve  $V_{\mathcal{A}}$  is called an affine curve with affine ring  $\mathcal{A}$ . We give many geometric notion for plane curves in two ways, one is an intuitive definition particular to the plane curve and another a ring theoretic interpretation of the notion which is valid for general  $V_{\mathcal{A}}$ . For example,  $\text{Hom}_{k\text{-curves}}(V_{\mathcal{A}}, V_{\mathcal{B}}) = \text{Hom}_{COF}(V_{\mathcal{A}}, V_{\mathcal{B}}) = \text{Hom}_{k\text{-alg}}(\mathcal{B}, \mathcal{A})$ .

An element in the total quotient ring of  $R_{\mathfrak{a}}$  (resp.  $\mathcal{A}$ ) is called a *rational  $k$ -function* on  $V_{\mathfrak{a}}$  (resp.  $V_{\mathcal{A}}$ ). If  $V_{\mathfrak{a}}$  is irreducible, then rational  $k$ -functions form a field. This field is called the *rational function field of  $V_{\mathfrak{a}}$  over  $k$* .

**1.2. Tangent space and local rings.** Suppose  $\mathfrak{a} = (f(X, Y))$ . Write  $V = V_{\mathfrak{a}}$  and  $R = R_{\mathfrak{a}}$ . Let  $P = (a, b) \in V_{\mathfrak{a}}(K)$ . We consider partial derivatives

$$\frac{\partial f}{\partial X}(P) := \frac{\partial f}{\partial X}(a, b) \quad \text{and} \quad \frac{\partial f}{\partial Y}(P) := \frac{\partial f}{\partial Y}(a, b).$$

Then the line tangent to  $V_{\mathfrak{a}}$  at  $(a, b)$  has equation

$$\frac{\partial f}{\partial X}(a, b)(X - a) + \frac{\partial f}{\partial Y}(a, b)(Y - b) = 0.$$

We write corresponding line as  $T_P = V_{\mathfrak{b}}$  for the principal ideal  $\mathfrak{b}$  generated by  $\frac{\partial f}{\partial X}(a, b)(X - a) + \frac{\partial f}{\partial Y}(a, b)(Y - b)$ . We call  $V_{\mathfrak{a}}$  is *non-singular* or *smooth* at  $P = (a, b) \in V_{\mathfrak{a}}(K)$  for a subfield  $K \subset \mathbb{C}$  if this  $T_P$  is really a line; in other word, if  $(\frac{\partial f}{\partial X}(P), \frac{\partial f}{\partial Y}(P)) \neq (0, 0)$ .

*Example 1.1.* Let  $\mathfrak{a} = (f)$  for  $f(X, Y) = Y^2 - X^3$ . Then  $\frac{\partial f}{\partial X}(a, b)(X - a) + \frac{\partial f}{\partial Y}(a, b)(Y - b) = -3a^2(X - a) + 2b(Y - b)$  ( $b^2 = a^3$ ). Thus this curve is singular only at  $(0, 0)$ .

*Example 1.2.* Suppose that  $k$  has characteristic different from 2. Let  $\mathfrak{a} = (Y^2 - g(X))$  for a cubic polynomial  $g(X) = X^3 + aX + b$ . Then the tangent line at  $(x_0, y_0)$  is

given by  $2y_0(X - x_0) - g'(x_0)(Y - y_0)$ . This equation vanishes if  $0 = y_0^2 = g(x_0)$  and  $g'(x_0) = 0$ ; so, singular at only  $(x_0, 0)$  for a multiple root  $x_0$  of  $g(X)$ . Thus  $V_{\mathfrak{a}}$  is a nonsingular curve if and only if  $g(X)$  is separable if and only if its discriminant  $4a^3 - 27b^2 \neq 0$ .

Suppose that  $K/k$  is an algebraic field extension. Then  $K[X, Y]/\mathfrak{a}K[X, Y]$  contains  $R_{\mathfrak{a}}$  as a subring. The maximal ideal  $(X_1 - a_1, \dots, X_n - a_n) \subset K[X, Y]/\mathfrak{a}K[X, Y]$  induces a maximal ideal  $P = (X - a, Y - b) \cap R_{\mathfrak{a}}$  of  $R_{\mathfrak{a}}$ . The local ring  $\mathcal{O}_{V_{\mathfrak{a}}, P}$  at  $P$  is the localization

$$\mathcal{O}_{V_{\mathfrak{a}}, P} = \left\{ \frac{a}{b} \mid b \in R_{\mathfrak{a}}, b \in R_{\mathfrak{a}}P \right\},$$

where  $\frac{a}{b} = \frac{a'}{b'}$  if there exists  $s \in R_{\mathfrak{a}} \setminus P$  such that  $s(ab' - a'b) = 0$ . Write the maximal ideal of  $\mathcal{O}_{V_{\mathfrak{a}}, P}$  as  $\mathfrak{m}_P$ . Then  $\mathfrak{m}_P \cap R = P$ .

**Lemma 1.3.** *The linear vector space  $T_P(K)$  is the dual vector space of  $P/P^2 = \mathfrak{m}_P/\mathfrak{m}_P^2$ .*

In general, for a maximal ideal  $P$  of  $\mathcal{A}$  with residue field  $K$ , we define the tangent space  $T_P := \text{Hom}_K(P/P^2, K)$ .

*Proof.* Write  $\mathfrak{a} = (f)$ . Replacing  $k[X, Y]/(f)$  by  $K[X, Y]/(f)$ , we may assume that  $K = k$ . A  $K$ -derivation  $\partial : \mathcal{O}_{V, P} \rightarrow K$  (at  $P$ ) is a  $K$ -linear map with  $\partial(\phi\varphi) = \varphi(P)\partial(\phi) + \phi(P)\partial(\varphi)$ . Writing  $D_{V, P}$  for the space of  $K$ -derivations at  $P$ , which is a  $K$ -vector space. Plainly for  $\mathbb{A} := V_{(0)}$ ,  $D_{\mathbb{A}, P}$  is a 2-dimensional vector space generated by  $\partial_X : \phi \mapsto \frac{\partial\phi}{\partial X}(P)$  and  $\partial_Y : \phi \mapsto \frac{\partial\phi}{\partial Y}(P)$ . We have a natural injection  $i : D_{V, P} \rightarrow D_{\mathbb{A}, P}$  given by  $i(\partial)(\phi) = \partial(\phi|_V)$ . Note that  $\Omega_{(a, b)} = (X - a, Y - b)/(X - a, Y - b)^2$  is a 2-dimensional vector space over  $K$  generated by  $X - a$  and  $Y - b$ . Thus  $D_{\mathbb{A}, P}$  and  $\Omega_{(a, b)}$  is dual each other under then pairing  $(\alpha(X - a) + \beta(Y - b), \partial) = \partial(\alpha(X - a) + \beta(Y - b))$ . The projection  $k[X, Y] \twoheadrightarrow R$  induces a surjection

$$\Omega_{(a, b)} \rightarrow \Omega_{V, P} = P/P^2,$$

whose kernel is spanned by  $f \bmod (X - a, Y - b)^2 = \frac{\partial f}{\partial X}(a, b)(X - a) + \frac{\partial f}{\partial Y}(a, b)(Y - b)$  if  $\mathfrak{a} = (f)$ , since  $\phi(X, Y) \equiv \frac{\partial\phi}{\partial X}(a, b)(X - a) + \frac{\partial\phi}{\partial Y}(a, b)(Y - b) \bmod (X - a, Y - b)^2$ . Thus the above duality between  $\Omega_{(a, b)}$  and  $D_{\mathbb{A}, (a, b)}$  induces the duality  $\Omega_{V, P} = P/P^2$  and  $T_P(K)$  given by  $(\omega, t) = t(\omega)$ , where we regard  $t$  as a derivation  $\mathcal{O}_{V, P} \rightarrow K$ .  $\square$

We call  $T_P$  the *tangent* space at  $P$  and  $\Omega_P = \Omega_{V, P}$  the *cotangent* space at  $P$  of  $V$ . More generally, a  $k$ -derivation  $\partial : R_{\mathfrak{a}} \rightarrow R_{\mathfrak{a}}$  is a  $k$ -linear map satisfying the Leibniz condition  $\partial(\phi\varphi) = \phi\partial(\varphi) + \varphi\partial(\phi)$  and  $\partial(k) = 0$ . For a  $k$ -derivation as above,  $f\partial : \varphi \mapsto f \cdot \partial(\varphi)$  for  $f \in R_{\mathfrak{a}}$  is again a  $k$ -derivation. The totality of  $k$ -derivation  $\text{Der}_{V_{\mathfrak{a}}/k}$  is therefore an  $R_{\mathfrak{a}}$ -module.

First take  $\mathfrak{a} = (0)$ ; so,  $V_{\mathfrak{a}} = \mathbb{A}^2$ . By the Leibniz relation,  $\partial(X^2) = nX^{n-1}\partial X$ ,  $\partial(Y^m) = mY^{m-1}\partial Y$  and  $\partial(X^2Y^m) = nX^{n-1}Y^m\partial X + mX^2Y^{m-1}\partial Y$  for  $\partial \in \text{Der}_{\mathbb{A}^2/k}$ ; so,  $\partial$  is determined by its value  $\partial(X)$  and  $\partial(Y)$ . Note that  $(\partial X)\frac{\partial}{\partial X} + (\partial Y)\frac{\partial}{\partial Y}$  in  $\text{Der}_{\mathbb{A}^2/k}$  and the original  $\partial$  has the same value at  $X$  and  $Y$ ; so, we have

$$\partial = (\partial X)\frac{\partial}{\partial X} + (\partial Y)\frac{\partial}{\partial Y}.$$

Thus  $\left\{\frac{\partial}{\partial X}, \frac{\partial}{\partial Y}\right\}$  gives a basis of  $Der_{\mathbb{A}^2/k}$ .

Assuming  $V_{\mathfrak{a}}$  nonsingular (including  $\mathbb{A}^2 = V_{(0)}$ ), we write the  $R_{\mathfrak{a}}$ -dual as  $\Omega_{V_{\mathfrak{a}}/k} := \text{Hom}(Der_{V_{\mathfrak{a}}/k}, R_{\mathfrak{a}})$  (the *space of  $k$ -differentials*) with the duality pairing

$$(\cdot, \cdot) : \Omega_{V_{\mathfrak{a}}/k} \times Der_{V_{\mathfrak{a}}/k} \rightarrow R_{\mathfrak{a}}.$$

We have a natural map  $d : R_{\mathfrak{a}} \rightarrow \Omega_{V_{\mathfrak{a}}/k}$  given by  $\phi \mapsto (d\phi : \partial \mapsto \partial(\phi)) \in Der_{V_{\mathfrak{a}}/k}$ . Note

$$(d(\phi\varphi), \partial) = \partial(\phi\varphi) = \phi\partial(\varphi) + \varphi\partial(\phi) = (\phi d\varphi + \varphi d\phi, \partial)$$

for all  $\partial \in Der_{V_{\mathfrak{a}}/k}$ . Thus we have  $d(\phi\varphi) = \phi d\varphi + \varphi d\phi$ , and  $d$  is a  $k$ -linear derivation with values in  $\Omega_{V_{\mathfrak{a}}/k}$ .

Again let us first look into  $\Omega_{\mathbb{A}^2/k}$ . Then by definition  $(dX, \partial) = \partial X$  and  $(dY, \partial) = \partial Y$ ; so,  $\{dX, dY\}$  is the dual basis of  $\left\{\frac{\partial}{\partial X}, \frac{\partial}{\partial Y}\right\}$ . We have  $d\Phi = \frac{\partial\Phi}{\partial X}dX + \frac{\partial\Phi}{\partial Y}dY$  as we can check easily that the left hand side and right hand side as the same value on any  $\partial \in Der_{\mathbb{A}^2/k}$ .

If  $\partial : R_{\mathfrak{a}} = k[X, Y]/(f) \rightarrow R_{\mathfrak{a}}$  is a  $k$ -derivation, we can apply it to any polynomial  $\Phi(X, Y) \in k[X, Y]$  and hence regard it as  $\partial : k[X, Y] \rightarrow R_{\mathfrak{a}}$ . By the above argument,  $Der_k(k[X, Y], R_{\mathfrak{a}})$  has a basis  $\left\{\frac{\partial}{\partial X}, \frac{\partial}{\partial Y}\right\}$  now over  $R_{\mathfrak{a}}$ . Since  $\partial$  factor through the quotient  $k[X, Y]/(f)$ , it satisfies  $\partial(f(X, Y)) = (df, \partial) = 0$ . Thus we have

**Lemma 1.4.** *We have an inclusion  $Der_{V_{\mathfrak{a}}/k} \hookrightarrow (R_{\mathfrak{a}}\frac{\partial}{\partial X} \oplus R_{\mathfrak{a}}\frac{\partial}{\partial Y})$  whose image is given by  $\{\partial \in Der_k(k[X, Y], R_{\mathfrak{a}}) \mid \partial f = 0\}$ . This implies  $\Omega_{V_{\mathfrak{a}}/k} = (R_{\mathfrak{a}}dX \oplus R_{\mathfrak{a}}dY)/R_{\mathfrak{a}}df$  for  $df = \frac{\partial f}{\partial X}dX + \frac{\partial f}{\partial Y}dY$  by duality.*

**Remark 1.5.** *If  $V_{\mathfrak{a}}$  is an irreducible curve; so,  $R_{\mathfrak{a}}$  is an integral domain, for its quotient field  $k(V_{\mathfrak{a}})$ ,  $k(V_{\mathfrak{a}})\Omega_{V_{\mathfrak{a}}/k} = (k(V_{\mathfrak{a}})dX \oplus k(V_{\mathfrak{a}})dY)/k(V_{\mathfrak{a}})df$  is 1 dimensional, as  $df \neq 0$  in  $\Omega_{\mathbb{A}^2/k}$ . In particular, if we pick  $\psi \in R_{\mathfrak{a}}$  with  $d\psi \neq 0$  (i.e., a non-constant), any differential  $\omega \in \Omega_{V_{\mathfrak{a}}/k}$  can be uniquely written as  $\omega = \phi d\psi$  for  $\phi \in k(V_{\mathfrak{a}})$ .*

**Lemma 1.6.** *The following four conditions are equivalent:*

- (1) *A point  $P$  of  $V(\bar{k})$  is a smooth point.*
- (2)  *$\mathcal{O}_{V,P}$  is a local principal ideal domain, not a field.*
- (3)  *$\mathcal{O}_{V,P}$  is a discrete valuation ring with residue field  $\bar{k}$ .*
- (4)  *$\varprojlim_n \mathcal{O}_{V,P}/\mathfrak{m}_P^2 \cong \bar{k}[[T]]$  (a formal power series ring).*

*Proof.* Let  $K = \bar{k}$ . By the above lemma,  $T_P$  is a line if and only if  $\dim T_P(K) = 1$  if and only if  $\dim P/P^2 = 1$ . Thus by Nakayama's lemma,  $P$  is principal. Any prime ideal of  $k[X, Y]$  is either minimal or maximal (i.e, the ring  $k[X, Y]$  has Krull dimension 2). Thus any prime ideal of  $R$  and  $\mathcal{O}_{V,P}$  is maximal. Thus (1) and (2) are equivalent. The equivalence of (2) and (3) follows from general ring theory covered by Math 210 (see [CRT] Theorem 11.2). We leave the equivalence (3)  $\Leftrightarrow$  (4) as an exercise.  $\square$

Write  $x, y$  for the image of  $X, Y \in k[X, Y]$  in  $R_{\mathfrak{a}}$ . Any  $\omega \in \Omega_{V_{\mathfrak{a}}/k}$  can be written as  $\phi dx + \varphi dy$ . Suppose that  $V_{\mathfrak{a}}$  is nonsingular. Since  $\mathcal{O}_{V_{\mathfrak{a}},P} \hookrightarrow k[[T]]$  (for  $P \in V_{\mathfrak{a}}(k)$ ) for a local parameter  $T$  as above,  $\phi, \varphi, x, y$  have the ‘‘Taylor expansion’’ as an element of  $k[[T]]$ , for example,  $x(T) = \sum_{n \geq 0} a_n(x)T^n$  with  $a_n(x) \in k$ . Thus  $dx, dy$  also have a well define expansion, say,  $dx = d(\sum_{n \geq 0} a_n(x)T^n) = \sum_{n \geq 1} a_n(x)T^{n-1}dT$ . Thus

we may expand  $\omega = \phi dx + \varphi dy = \sum_{n \geq 0} a_n(\omega) T^2 dT$  once we choose a parameter  $T$  at  $P$ . This expansion is unique independent of the expression  $\phi dx + \varphi dy$ . Indeed, if we allow meromorphic functions  $\Phi$  as coefficients, as we remarked already, we can uniquely write  $\omega = \Phi dx$  and the above expansion coincides with the Taylor expansion of  $\Phi dx$ . Write  $Max(R_{\mathfrak{a}})$  for the set of maximal ideals of  $R_{\mathfrak{a}}$ . Then plainly, we have a natural inclusion  $V_{\mathfrak{a}}(k) \hookrightarrow Max(R_{\mathfrak{a}})$  sending  $(a, b)$  to  $(x - a, y - b)$  for the image  $x, y$  in  $R_{\mathfrak{a}}$  of  $X, Y \in k[X, Y]$ . For  $P \in Max(R_{\mathfrak{a}})$ , we call  $P$  is smooth on  $V_{\mathfrak{a}}$  if  $\mathcal{O}_{V, P}$  is a discrete valuation ring. By the above exercise, this is consistent with the earlier definition (no more and no less).

For any given affine plane irreducible curve  $V_{\mathfrak{a}}$ , we call  $V_{\mathfrak{a}}$  is normal if  $R_{\mathfrak{a}}$  is integrally closed in its field of fractions.

**Corollary 1.7.** *Any normal irreducible affine plane curve is smooth everywhere.*

*Proof.* By ring theory, any localization of a normal domain is normal. Thus  $\mathcal{O}_{V, P}$  is a normal domain. By the exercise below, we may assume that  $P \cap k[X, Y] \neq (0)$ . Then  $P$  is a maximal ideal, and hence  $K = k[X, Y]/P$  is an algebraic extension of  $k$ . In this case,  $\mathcal{O}_{V, P}$  is a normal local domain with principal maximal ideal, which is a discrete valuation ring (cf. [CRT] Theorem 11.1).  $\square$

**1.3. Projective space.** Let  $A$  be a commutative ring. Write  $A_P$  be the localization at a prime ideal  $P$  of  $A$ . Thus

$$A_P = \left\{ \frac{b}{s} \mid s \in A \setminus P \right\} / \sim,$$

where  $\frac{b}{s} \sim \frac{b'}{s'}$  if there exists  $s'' \in A \setminus P$  such that  $s''(s'b - sb') = 0$ . An  $A$ -module  $M$  is called *locally free* at  $P$  if

$$M_P = \left\{ \frac{m}{s} \mid s \in A \setminus P \right\} / \sim = A_P \otimes_A M$$

is free over  $A_P$ . We call  $M$  locally free if it is free at all prime ideals of  $A$ . If  $\text{rank}_{A_P} M_P$  is constant  $r$  independent of  $P$ , we write  $\text{rank}_A M$  for  $r$ .

Write  $ALG/k$  for the category of  $k$ -algebras; so,  $\text{Hom}_{ALG/k}(A, A')$  is made up of  $k$ -algebra homomorphisms from  $A$  into  $A'$  sending the identity  $1_A$  to the identity  $1_{A'}$ . Here  $k$  is a general base ring, and we write  $ALG$  for  $ALG/\mathbb{Z}$  (as  $ALG$  is the category of all commutative rings with identity). We consider a covariant functor  $\mathbf{P}^2 = \mathbf{P}^2/k : ALG/k \rightarrow SETS$  given by

$$\mathbf{P}^2(A) = \{ L \subset A^{n+1} \mid L \text{ and } A^{n+1}/L \text{ are locally } A\text{-free with } \text{rank } L = 1 \}.$$

This is a covariant functor. Indeed, if  $\sigma : A \rightarrow A'$  is a  $k$ -algebra homomorphism, letting it act on  $A^{n+1}$  component-wise,  $L \mapsto \sigma(L) \otimes_A A'$  induces a map  $\mathbf{P}^2(A) \rightarrow \mathbf{P}^2(A')$ . If  $A$  is a field  $K$ , then  $X$  has to be free of dimension 1 generated by a non-zero vector  $x = (x_0, x_1, \dots, x_n)$ . The vector  $x$  is unique up to multiplication by non-zero elements of  $K$ . Thus we have proven (for a field) of the following

**Lemma 1.8.** *Suppose that  $K$  is a field. Then we have*

$$\mathbf{P}^2(K) \cong \{ \underline{x} = (x_0, x_1, \dots, x_n) \in K^{n+1} \mid x \neq (0, \dots, 0) \} / K^\times.$$



Moreover, writing  $D_i : \text{ALG}/_k \rightarrow \text{SETS}$  for the subfunctor  $D_i(A) \subset \mathbf{P}^2(A)$  made up of the classes  $L$  whose projection to the  $i$ -th coordinate is surjective onto  $A$  (i.e.,  $x_i \neq 0$ ), we have  $\mathbf{P}^2(K) = \bigcup_i D_i(K)$  and  $D_i \cong \mathbb{A}^2$  canonically for all  $k$ -algebras  $A$ . The isomorphism:  $D_i \cong \mathbb{A}^2$  is given by sending  $(x_0, \dots, x_n)$  to  $(\frac{x_0}{x_i}, \dots, \frac{x_n}{x_i}) \in \mathbb{A}^2$  removing the  $i$ -th coordinate.

*Proof.* If  $L \in D_i(A)$ , we have the following commutative diagram

$$\begin{array}{ccc} L & \xrightarrow{\hookrightarrow} & A^{n+1} \\ \parallel \downarrow & & \downarrow \text{\scriptsize } i\text{-th proj} \\ L & \xrightarrow{\sim} & A \end{array}$$

Thus  $L$  is free of rank 1 over  $A$ ; so, it has a generator  $(x_0, \dots, x_n)$  with  $x_i \in A^\times$ . Then  $(x_0, \dots, x_n) \mapsto (\frac{x_0}{x_i}, \dots, \frac{x_n}{x_i}) \in \mathbb{A}^2$  gives rise to a natural transformation of  $D_i$  onto  $\mathbb{A}^2$  (which is an isomorphism of functors).  $\square$

If  $K$  is a field, we write  $(x_0 : x_1 : \dots : x_n)$  for the point of  $\mathbf{P}^2(K)$  represented by  $(x_0, \dots, x_n)$  as only the ratio matters. We assume that  $K$  is a field for a while. When  $n = 1$ , we see  $\mathbf{P}^1(K) = K^\times \sqcup \{\infty\}$  by  $(x : y) \mapsto \frac{x}{y} \in K \sqcup \{\infty\}$ . Thus  $\mathbf{P}^1(\mathbb{R})$  is isomorphic to a circle and  $\mathbf{P}^1(\mathbb{C})$  is a Riemann sphere.

We now assume that  $n = 2$ . Writing  $L = \{(x : y : 0) \in \mathbf{P}^2(K)\}$ . Then  $\mathbf{P}^1 \cong L$  by  $(x : y) \mapsto (x : y : 0)$ ; so,  $L$  is isomorphic to the projective line. We have  $\mathbf{P}^2(K) = D(K) \sqcup L$  for fields  $K$ , where  $D = D_2$ . Thus geometrically (i.e., over fields),  $\mathbf{P}^2$  is the union of the affine plane added  $L$ . We call  $L = L_\infty$  (the line at  $\infty$ ).

**1.4. Projective plane curve.** For a plane curve defined by  $\mathfrak{a} = (f(x, y))$  for  $f(x, y)$  of degree  $m$ ,  $F(X, Y, Z) = Z^m f(\frac{X}{Z}, \frac{Y}{Z})$  is a (square-free) homogeneous polynomial of degree  $m$  in  $k[X, Y, Z]$ . If  $L \in \mathbf{P}^2(A)$ , we can think of  $F(\ell)$  for  $\ell \in L$ . We write  $F(L) = 0$  if  $F(\ell) = 0$  for all  $\ell \in L$ . Thus for any  $k$ -algebra  $A$ , we define the functor  $\overline{V}_\mathfrak{a} : \text{ALG}/_k \rightarrow \text{SETS}$  by

$$\overline{V}_\mathfrak{a}(A) = \{L \in \mathbf{P}^2(A) \mid F(L) = 0\}.$$

If  $A$  is a field  $K$ , we sent  $L \in \mathbf{P}^2(K)$  to its generator  $(a : b : c) \in L$  when we identified  $\mathbf{P}^2(K)$  with the (classical) projective space with homogeneous coordinate. Since  $F(L) = 0$  if and only if  $F(a : b : c) = 0$  in this circumstances, we have

$$\overline{V}_\mathfrak{a}(K) = \{(a : b : c) \in \mathbf{P}^2(K) \mid F(a, b, c) = 0\}$$

which is called a *projective plane  $k$ -curve*. Since  $D_2 \cong \mathbb{A}^2$  canonically via  $(x : y : 1) \mapsto (x, y)$  (and this coordinate is well defined even over  $A$  which is not a field), we have  $\overline{V}_\mathfrak{a}(A) \cap D_2(A) = V_\mathfrak{a}(A)$ . In this sense, we can think of  $\overline{V}_\mathfrak{a}$  as a completion of  $V_\mathfrak{a}$  adding the boundary  $\overline{V}_\mathfrak{a} \cap L_\infty$ . Since in  $D_j \cong \mathbb{A}^2$  ( $j = 0, 1$ ),  $\overline{V}_\mathfrak{a} \cap D_j$  is a plane affine curve (for example,  $\overline{V}_\mathfrak{a} \cap D_0$  is defined by  $F(1, y, z) = 0$ ),  $(L_\infty \cap \overline{V}_\mathfrak{a})(\overline{k})$  is a finite set. Thus  $\overline{V}_\mathfrak{a}$  is a sort of completion/compactification of the (open) affine curve  $V_\mathfrak{a}$  (we sort out this point more rigorously later). Of course, we can start with a homogeneous polynomial  $F(X, Y, Z)$  (or a homogeneous ideal of  $k[X, Y, Z]$  generated by  $F(X, Y, Z)$ ) to define a projective plane curve. Following Lemma 1.1, we define  $\text{Hom}_{\text{proj } k\text{-curves}}(\overline{V}_\mathfrak{a}, \overline{V}_\mathfrak{b}) := \text{Hom}_{\text{COF}}(\overline{V}_\mathfrak{a}, \overline{V}_\mathfrak{b})$ .

*Example 1.3.* Suppose  $\mathfrak{a} = (y^2 - f(x))$  for a cubic  $f(x) = x^3 + ax + b$ . Then  $F(X, Y, Z) = Y^2Z - X^3 - aXZ^2 - bZ^3$ . Since  $L_\infty$  is defined by  $Z = 0$ , we find  $L_\infty \cap \overline{V}_\mathfrak{a} = \{(0 : 1 : 0)\}$  made of a single point (with multiplicity 3). This point we call the origin  $\mathbf{0}$  of  $V_\mathfrak{a}$ .

A projective plane curve  $\overline{V}_\mathfrak{a}$  is non-singular (or smooth) if  $\overline{V}_\mathfrak{a} \cap D_j$  is a non-singular plane curve for all  $j = 0, 1, 2$ . The tangent space at  $P \in \overline{V}_\mathfrak{a}(K)$  is defined as before since  $P$  is in one of  $D_j \cap V_\mathfrak{a}$ . By the above exercise, the tangent space (the dual of  $\mathfrak{m}_P/\mathfrak{m}_P^2$ ) at  $P \in \overline{V}_\mathfrak{a}(K)$  does not depend on the choice of  $j$  with  $P \in \overline{V}_\mathfrak{a} \cap D_j$ . If a projective plane curve  $C$  is irreducible, the rational function field over  $k$  is the field of fraction of  $\mathcal{O}_{C,P}$  for any  $P \in C(\overline{k})$ ; so, independent of  $C \cap D_j$ .

**Lemma 1.9.** *Take a nonzero  $f \in k(C)$ . Then there exist homogeneous polynomials  $G(X, Y, Z), H(X, Y, Z) \in k[X, Y, Z]$  with  $\deg(G) = \deg(H)$  such that  $f(x : y : z) = \frac{H(x,y,z)}{G(x,y,z)}$  for all  $(x : y : z) \in C(\overline{k})$ .*

*Proof.* We may write on  $C \cap D_2$   $f(x, y, 1) = \frac{h(x,y)}{g(x,y)}$ . If  $m = \deg(h) = \deg(g)$ , we just define  $H(X, Y, Z) = h(\frac{X}{Z}, \frac{Y}{Z})Z^m$  and  $G(X, Y, Z) = g(\frac{X}{Z}, \frac{Y}{Z})Z^m$ . If  $\deg(h) > \deg(g)$ , we define  $H(X, Y, Z) = h(\frac{X}{Z}, \frac{Y}{Z})Z^{\deg(h)}$  and  $G(X, Y, Z) = g(\frac{X}{Z}, \frac{Y}{Z})Z^{\deg(h)}$ . If  $\deg(h) < \deg(g)$ , we define  $H(X, Y, Z) = h(\frac{X}{Z}, \frac{Y}{Z})Z^{\deg(g)}$  and  $G(X, Y, Z) = g(\frac{X}{Z}, \frac{Y}{Z})Z^{\deg(g)}$ . Multiplying  $h$  or  $g$  by a power of  $Z$  does not change the above identity  $f(x, y, 1) = \frac{h(x,y)}{g(x,y)}$ , because  $Z = 1$  on  $C \cap D_2$ . Thus by adjusting in this way, we get  $G$  and  $H$ .  $\square$

*Example 1.4.* Consider the function  $\phi = cx + dy$  in  $k(C)$  for  $C = \overline{V}_\mathfrak{a}$  with  $\mathfrak{a} = (y^2 - x^3 - ax - b)$ . Then  $C$  is defined by  $Y^2Z - X^3 - aXZ^2 - bZ^3 = 0$ , and

$$\phi(X : Y : Z) = c\frac{X}{Z} + d\frac{Y}{Z} = \frac{cX + dY}{Z}.$$

So  $\phi$  has pole of order 3 at  $Z = 0$  (as the infinity on  $C$  has multiplicity 3) and three zeros at the intersection of  $L := \{cx + dy = 0\}$  and  $C \cap D_2 \cap L$ .

Take a projective nonsingular plane  $k$ -curve  $C/k$ . Put  $C_i = C \cap D_i$  which is an affine nonsingular plane curve. Then we have well defined global differentials  $Der_{C_i/k}$ . Since  $\partial : Der_{C_i/k}$  induces  $\partial_P : \mathcal{O}_{C_i,P} \rightarrow K$  for any  $P \in C_i(K)$  by  $f \mapsto \partial(f)(P)$ , we have  $\partial_P \in T_P$ . If  $\partial_i \in Der_{C_i/k}$  given for each  $i = 0, 1, 2$  satisfies  $\partial_{i,P} = \partial_{j,P}$  for all  $(i, j)$  and all  $P \in (D_i \cap D_j)(\overline{k})$ , we call  $\partial = \{\partial_i\}_i$  a global tangent vector defined on  $C$ . Plainly the totality  $T_{C/k}$  of global tangent vectors are  $k$ -vector space. The  $k$ -dual of  $T_{C/k}$  is called the space of  $k$ -differentials over  $k$  and written as  $\Omega_{C/k}$ . It is known that  $\Omega_{C/k}$  is finite dimensional over  $k$ .

**Corollary 1.10.** *Suppose that  $C$  is non-singular. Each  $\phi \in k(C)$  induces  $\phi \in \text{Hom}_{\text{proj } k\text{-curves}}(C, \mathbf{P}^1)$ . Indeed, we have  $k(C) \sqcup \{\infty\} \cong \text{Hom}_{\text{proj } k\text{-curves}}(C, \mathbf{P}^1)$ , where  $\infty$  stands for the constant function sending all  $P \in C(A)$  to the image of  $\infty \in \mathbf{P}^1(k)$  in  $\mathbf{P}^1(A)$ .*

*Proof.* We prove only the first assertion. Suppose  $k = \overline{k}$ . Write  $\phi(x : y : z) = \frac{h(x,y,z)}{g(x,y,z)}$  as a reduced fraction by the above lemma. For  $L \in C(A) \subset \mathbf{P}^2(A)$ , we consider the sub  $A$ -module  $\phi(L)$  of  $A^2$  generated by  $\{(h(\ell), g(\ell)) \in A^2 | \ell \in L\}$ . We now show

that  $\phi(L) \in \mathbf{P}^1(A)$ ; so, we will show that the map  $C(A) \ni L \mapsto \phi(L) \in \mathbf{P}^1(A)$  induces the natural transformation of  $C$  into  $\mathbf{P}^1$ . If  $A$  is local, by Lemma 1.8,  $L$  is generated by  $(a, b, c)$  with at least one unit coordinate. Then any  $\ell \in L$  is of the form  $\lambda(a, b, c)$  and therefore  $\phi(\ell) = \lambda^{\deg(h)} \phi(a, b, c)$ . Thus  $\phi(L) = A \cdot \phi(a, b, c)$ . Since  $A$  is a  $k$ -algebra,  $k$  is naturally a subalgebra of the residue field  $A/\mathfrak{m}$  of  $A$ . Since  $\phi(P)$  for all  $P \in C(k)$  is either a constant in  $k$  or  $\infty$ , we may assume that  $(h(P), g(P)) \neq (0, 0)$  for all  $P \in C(k)$ . Since  $(a, b, c) \not\equiv 0 \pmod{\mathfrak{m}}$  as  $(a, b, c)$  generates a direct summand of  $A^3$ . Thus  $(h(a, b, c), g(a, b, c)) \not\equiv (0, 0) \pmod{\mathfrak{m}}$ . After tensoring  $A/\mathfrak{m}$  over  $A$ ,  $(A/\mathfrak{m})^2/(\phi(L)/\mathfrak{m}\phi(L))$  is one dimensional. Thus by Nakayama's lemma (e.g., [CRT] Theorem 2.2-3),  $A/\phi(L)$  is generated by a single element and has to be a free module of rank 1 as  $\phi(L)$  is a free  $A$ -module of rank 1. Thus  $\phi(L) \in \mathbf{P}^1(A)$ . If  $k$  is not algebraically closed, replacing  $A$  by  $\bar{A} = A \otimes_k \bar{k}$ , we find  $\phi(L) \otimes_k \bar{k} \in \mathbf{P}^2(\bar{k})$  and hence  $\phi(L) \otimes_A A/\mathfrak{m} \in \mathbf{P}^2(k)$ , which implies  $\phi(L) \in \mathbf{P}^2(A)$ .

If  $A$  is not necessarily local, applying the above argument to the local ring  $A_P$  for any prime ideal  $P$  of  $A$ , we find that  $\phi(L)_P = \phi(L_P)$  and  $A_P^2/\phi(L_P)$  are free of rank 1; so,  $\phi(L)$  and  $A^2/\phi(L)$  are locally free of rank 1; therefore,  $\phi(L) \in \mathbf{P}^2(A)$ .

Now it is plain that  $L \mapsto \phi(L)$  induces a natural transformation of functors.  $\square$

**1.5. Divisors.** The divisor group  $\text{Div}(C)$  of a non-singular projective geometrically irreducible plane curve  $C$  is a formal free  $\mathbb{Z}$ -module generated by points  $P \in C(\bar{k})$ . When we consider a point  $P$  as a divisor, we write it as  $[P]$ . For each divisor  $D = \sum_P m_P [P]$ , we define  $\deg(D) = \sum_P m_P$ . Since  $C$  is nonsingular, for any point  $P \in C(\bar{k})$ ,  $\mathcal{O}_{C,P}$  is a DVR, and the rational function field  $\bar{k}(C)$  is the quotient field of  $\mathcal{O}_{C,P}$  (regarding  $C$  as defined over  $\bar{k}$ ). Thus if we write the valuation  $v_P : \bar{k}(C) \rightarrow \mathbb{Z} \cup \{\infty\}$  for the additive valuation of  $\mathcal{O}_{C,P}$ , we have a well defined  $v_P(f) \in \mathbb{Z}$  for any non-zero rational  $\bar{k}$ -function  $f \in \bar{k}(C)$ . Since  $\mathfrak{m}_P = (t_P)$  and  $t_P^{v_P(f)} \parallel f$  in  $\mathcal{O}_{C,P}$ ,  $f$  has a zero of order  $v_P(f)$  at  $P$  if  $v_P(f) > 0$  and a pole of order  $|v_P(f)|$  if  $v_P(f) < 0$ . In other words, the Taylor expansion of  $f$  at  $P$  is given by  $\sum_n a_n(f) t_P^n$  and  $v_P(f) = \min(n : a_n(f) \neq 0)$ . For a global doifferential  $\omega \in \Omega_{C/\bar{k}}$ , we have its Taylor expansion  $\sum_n a_n(\omega) t_P^n dt_P$  at each  $P \in C(\bar{k})$ ; so, we may also define  $v_P(\omega) := \min(n : a_n(\omega) \neq 0)$ . We extend this definition for meromorphic differentials  $k(C) \cdot \Omega_{C/k} = \{f \cdot \omega \mid f \in k(C), \omega \in \Omega_{C/k}\}$ . Here we quote Bézout's theorem:

**Theorem 1.11.** *Let  $C$  and  $C'$  be two plane projective  $k$ -curves inside  $\mathbf{P}^2$  defined by relatively prime homogeneous equations  $F(X, Y, Z) = 0$  and  $G(X, Y, Z) = 0$  of degree  $m$  and  $n$  respectively. Then counting with multiplicity,  $|C(\bar{k}) \cap C'(\bar{k})| = m \cdot n$ .*

If  $C$  is smooth at  $P \in C \cap C'$  in  $C \cap D_2$ ,  $\phi = \frac{G(X,Y,Z)}{Z^2}$  is a function vanishing at  $P$ . The multiplicity of  $P$  in  $C \cap C'$  is just  $v_P(\phi)$ . More generally, if  $P = (a, b)$  is not necessarily a smooth point, writing  $C \cap D_2 = V_{\mathfrak{a}}$  and  $C' \cap D_2 = V_{\mathfrak{b}}$  for principal ideals  $\mathfrak{a}, \mathfrak{b}$  in  $\bar{k}[X, Y]$  and regarding  $P$  as an ideal  $(X - a, Y - b) \subset \bar{k}[X, Y]$ , the multiplicity is given by the dimension of the localization  $(\bar{k}[x, y]/\mathfrak{a} + \mathfrak{b})_P$  over  $\bar{k}$ . The same definition works well for any points in  $C \cap D_0$  and  $C \cap D_1$ . One can find the proof of this theorem with (possibly more sophisticated) definition of multiplicity in a text of algebraic geometry (e.g. [ALG] Theorem I.7.7).

Since there are only finitely many poles and zeros of  $f$ , we can define the divisors  $\operatorname{div}(f) = \sum_{P \in C(k)} v_P(f)[P]$ ,  $\operatorname{div}_0(f) = \sum_{P \in C(k), v_P(f) > 0} v_P(f)[P]$  and  $\operatorname{div}_\infty(f) = \sum_{P \in C(k), v_P(f) < 0} v_P(f)[P]$  of  $f$ . Similarly, for meromorphic differential  $\omega$ , we define again  $\operatorname{div}(\omega) = \sum_P v_P(\omega)[P]$ . By Lemma 1.9,  $f(x : y : z) = \frac{h(x:y:z)}{g(x:y:z)}$  for a homogeneous polynomial  $h, g$  in  $\bar{k}[x, y, z]$  of the same degree. If the degree of equation defining  $C$  is  $m$  and  $C'$  is defined by  $h(X, Y, Z) = 0$ ,  $\deg_0(\operatorname{div}(f)) = |C(\bar{k}) \cap C'(\bar{k})| = m \deg(h) = m \deg(g) = \deg_\infty(\operatorname{div}(f))$ . This shows  $\deg(\operatorname{div}(f)) = 0$  as  $\sum_{P, v_P(f) > 0} m_P = m \deg(h)$  and  $-\sum_{P, v_P(f) < 0} m_P = m \deg(g)$ .

**Lemma 1.12.** *Let  $C$  be a nonsingular projective plane curve. For any  $f \in \bar{k}(C)$ ,  $\deg(\operatorname{div}(f)) = 0$ , and if  $f \in \bar{k}(C)$  is regular at every  $P \in C$ ,  $f$  is a constant in  $\bar{k}$ .*

**Lemma 1.13.** *If  $f \in k(C)$  satisfies  $\deg(\operatorname{div}_0(f)) = \deg(\operatorname{div}_\infty(f)) = 1$ ,  $f : C \rightarrow \mathbf{P}^1$  induces an isomorphism of projective plane curve over  $k$ .*

*Proof.* By the proof of Corollary 1.10,  $\deg(\operatorname{div}_0(f))$  is the number of points over 0 (counting with multiplicity) of the regular map  $f : C \rightarrow \mathbf{P}^1$ . By taking off a constant  $\alpha \in k \subset \mathbf{P}^1$  to  $f$ ,  $\deg(\operatorname{div}_0(f - \alpha)) = 1 = \deg(\operatorname{div}_\infty(f - \alpha))$ , and  $|f^{-1}(\alpha)| = \deg(\operatorname{div}_0(f - \alpha)) = 1$ ; so, we find that  $f$  is 1-1 onto. Thus  $f$  is an isomorphism.  $\square$

Write  $\operatorname{Div}^0(C) = \{D \in \operatorname{Div}(C/\bar{k}) \mid \deg(D) = 0\}$ . Inside  $\operatorname{Div}^0(C)$ , we have the subgroup  $\{\operatorname{div}(f) \mid f \in \bar{k}(C)^\times\}$ . We call two divisors  $D, D'$  *linearly equivalent* if  $D = \operatorname{div}(f) + D'$  for  $f \in \bar{k}(C)$ . We call that  $D$  and  $D'$  are *algebraically equivalent* if  $\deg(D) = \deg(D')$ . The quotient groups  $J(C) = \operatorname{Div}^0(C) / \{\operatorname{div}(f) \mid f \in k(C)^\times\}$  and  $\operatorname{Pic}(C) = \operatorname{Div}(C) / \{\operatorname{div}(f) \mid f \in k(C)^\times\}$  are called the jacobian and the Picard group of  $C$ , respectively. Sometimes,  $J(C)$  is written as  $\operatorname{Pic}^0(C)$  (the degree 0 Picard group).

**1.6. The theorem of Riemann–Roch.** We write  $D = \sum_P m_P[P] \geq 0$  (resp.  $D > 0$ ) for a divisor  $D$  on  $C$  if  $m_P \geq 0$  for all  $P$  (resp.  $D \geq 0$  and  $D \neq 0$ ). For a divisor  $D$  on  $C_{\bar{k}}$

$$L(D) = \{f \in \bar{k}(C) \mid \operatorname{div}(f) + D \geq 0\} \cup \{0\}.$$

Plainly,  $L(D)$  is a vector space over  $\bar{k}$ . It is known that  $\ell(D) = \dim_{\bar{k}} L(D) < \infty$ . For  $\phi \in k(C)^\times$ ,  $L(D) \ni f \mapsto f\phi \in L(D - \operatorname{div}(\phi))$  is an isomorphism. Thus  $\ell(D)$  only depends on the class of  $D$  in  $\operatorname{Pic}(C)$ .

*Example 1.5.* Let  $C = \mathbf{P}^1$ . For a positive divisor  $D = \sum_{a \in \bar{k}} m_a[a]$  with  $m_a \geq 0$  and  $m_a > 0$  for some  $a$ , regarding  $a \in \bar{k}$  as a point  $[a] \in \mathbf{P}^1(\bar{k}) = \bar{k} \sqcup \{\infty\}$ . On  $\mathbb{A}^1(\bar{k}) = \bar{k}$ , forgetting about the infinity,  $\operatorname{div}(f) + D \geq 0$  if  $f = \frac{g(x)}{\prod_a (x-a)^{m_a}}$  for a polynomial  $g(x)$ . If  $\deg(D) \geq \deg(g(x))$ , the function  $f$  does not have pole at  $\infty$ . Thus  $L(D) = \{g(x) \mid \deg(g(x)) \leq \deg(D)\}$  and we have  $\ell(D) = 1 + \deg(D)$  if  $D > 0$ . If  $C$  is a plane projective curve, we can write  $f = \frac{h(X,Y,Z)}{g(X,Y,Z)}$  as a reduced fraction by Lemma 1.9. Write  $D = \sum_P m_P[P]$ , and put  $|D| = \{P \mid D = \sum_P m_P[P] \text{ with } m_P \neq 0\}$ . If  $|D|$  is inside  $D_2 \cap C \subset \mathbb{A}^2$  and  $D > 0$ , we may assume that  $V_{(g(X,Y,1))} \cap C$  contains  $|D|$ . Then not to have pole at  $C \setminus D_2$ ,  $\deg(h)$  has to be bounded; so,  $\ell(D) < \infty$ . Since  $L(D) \subset L(D_+)$  in general, writing  $D = D_+ + D_-$  so that  $D_+ \geq 0$  and  $-D_- \geq 0$ , this shows  $\ell(D) < \infty$ .

**Theorem 1.14** (Riemann-Roch). *Let  $C = \overline{V}_a$  be a non-singular projective curve defined over a field  $k$ . Then for  $g = \dim_{\overline{k}} \Omega_{C/\overline{k}}$  and a divisor  $K$  of degree  $2g - 2$  of the form  $\text{div}(\omega)$  for a meromorphic differential  $\omega$  on  $C$  such that  $\ell(D) = 1 - g + \text{deg}(D) + \ell(K - D)$  for all divisor  $D$  on  $C(\overline{k})$  and the equality holds for sufficiently positive divisor  $D$ . If  $g = 1$ ,  $K = 0$ .*

The divisor  $K$  is called a *canonical divisor*  $K$  (whose linear equivalence class is unique). Note that

$$L(K) = \{f \in \overline{k}(C) \mid \text{div}(f\omega) = \text{div}(f) + \text{div}(\omega) \geq 0\} \cong \Omega_{C/\overline{k}}$$

by  $f \mapsto f\omega \in \Omega_{C/\overline{k}}$ . Then by the above theorem,

$$g(C) = \dim \Omega_{C/\overline{k}} = \ell(K) = 1 - g + \text{deg}(K) + \ell(0) = 2 + \text{deg}(K) - g(C),$$

and from this, we conclude  $\text{deg}(K) = 2g(C) - 2$ . One can find a proof of this theorem in any algebraic geometry book (e.g., [ALG] IV.1 or [GME] Theorem 2.1.3).

**Corollary 1.15.** *If  $g(C) = 1$ , then  $\ell(D) = \text{deg}(D)$  if  $\text{deg}(D) > 0$ .*

*Proof.* For a non-constant  $f \in \overline{k}(E)$ ,  $\text{deg}(\text{div}(f)) = 0$  implies that  $f$  has a pole somewhere. If  $D > 0$ ,  $f \in L(-D)$  does not have pole; so, constant. Since  $D > 0$ ,  $f$  vanishes at  $P \subset D$ . Thus  $f = 0$ . More generally, if  $\text{deg}(D) > 0$  and  $\phi \in L(-D)$ , then  $0 > \text{deg}(-D) = \text{deg}(\phi) - \text{deg}(D) \geq 0$ ; so,  $\phi = 0$ . Thus if  $\text{deg}(D) > 0$ , then  $\ell(-D) = 0$ . Since  $K = 0$ , we have by the Riemann-Roch theorem that  $\ell(D) = \text{deg}(D) + \ell(0 - D) = \text{deg}(D)$  if  $\text{deg}(D) > 0$ .  $\square$

Because of  $\text{deg}(\text{div}(f)) = 0$ , if  $D \gg 0$ ,  $\ell(-D) = 0$ . Thus in particular  $\ell(K - D) = 0$  if  $D \gg 0$ . Thus the above theorem implies what Riemann originally proved:

**Corollary 1.16** (Riemann). *Let  $C = \overline{V}_a$  be a non-singular projective curve defined over a field  $k$ . Then there exists a non-negative integer  $g = g(C)$  such that  $\ell(D) \geq 1 - g + \text{deg}(D)$  for all divisor  $D$  on  $C(\overline{k})$  and the equality holds for sufficiently positive divisor  $D$ .*

By the above example, we conclude  $g(\mathbf{P}^1) = 0$  from the corollary.

**Exercise 1.17.** *Prove  $\Omega_{\mathbf{P}^1/\overline{k}} = 0$ .*

**1.7. Regular maps from a curve into projective space.** Tak a divisor  $D$  on a nonsingular projective plane curve  $C$ . Suppose  $\ell(D) = n > 0$ . Take a basis  $(f_1, f_2, \dots, f_n)$  of  $L(D)$ . Thus we can write  $f_j = \frac{h_j}{g_j}$  with homogeneous polynomials  $g_j, h_j$  having  $\text{deg}(g_j) = \text{deg}(h_j)$ . Replacing  $(g_j, h_j)$  by  $(g'_0 := g_1 g_2 \cdots g_n, h'_j := h_j g^{(j)})$  for  $g^{(j)} = \prod_{i \neq j} g_i$ , we may assume  $\text{deg}(g'_j) = \text{deg}(h'_j)$  for all  $j$ , and further dividing them by the GCD of  $(h'_1, \dots, h'_n, g'_0)$ , we may assume that  $f_j = \frac{h_j}{g_0}$  with  $\text{deg}(h_j) = \text{deg}(g_0)$  for all  $j$  and  $(g_0, h_1, \dots, h_n)$  do not have nontrivial common divisor.

**Lemma 1.18.** *Let the assumptions on  $(g_0, h_1, \dots, h_n)$  be as above. Suppose that  $(g_0(P), h_1(P), \dots, h_n(P)) \neq (0, 0, \dots, 0)$  for all  $P \in C(\overline{k})$ . Define  $L \in C(A) \subset \mathbf{P}^2(A)$ ,  $\phi_A(L)$  for an  $A$ -submodule of  $A^3$  generated by  $\phi(\ell) = (g_0(\ell), h_1(\ell), \dots, h_n(\ell)) \in A^{n+1}$  for all  $\ell \in L$ . Then  $\phi = \{\phi_A\}_A : C \rightarrow \mathbf{P}^2$  is a  $k$ -morphism of the projective plane  $k$ -curve  $C$  into  $\mathbf{P}^2_{/k}$ .*

The proof of the above lemma is the same as that of Corollary 1.10; so, we leave it to the reader.

## 2. ELLIPTIC CURVES

An *elliptic curve*  $E/k$  is a non-singular projective geometrically irreducible plane curve with point  $\mathbf{0}_E$  specified having  $g(E) = 1$ . Here we define  $g(E)$ , regarding  $E$  is defined over  $\bar{k}$ . We study elliptic curves in more details.

**2.1. Abel's theorem.** When we regard  $P \in E(k)$  as a divisor, we just write  $[P]$ . So  $3[P]$  is a divisor supported on  $P$  with multiplicity 3. We prove

**Theorem 2.1** (Abel). *Let  $E/k$  be an elliptic curve with origin  $\mathbf{0}_E$ . The correspondence  $P \mapsto [P] - [\mathbf{0}_E]$  induces a bijection  $E(\bar{k}) \cong J(E)$ . In particular,  $E(\bar{k})$  is an abelian group.*

*Proof.* Injectivity: if  $[P] - [Q] = [P] - [\mathbf{0}_E] - ([Q] - [\mathbf{0}_E]) = \text{div}(f)$  with  $P \neq Q$  in  $E(\bar{k})$ , by Lemma 1.13,  $f$  is an isomorphism. This is wrong as  $g(\mathbf{P}^1) = 0$  while  $g(E) = 1$ . Thus  $P = Q$ .

Surjectivity: Pick  $D \in \text{Div}^0(E)$ . Then  $D + [\mathbf{0}_E]$  has degree 1; so,  $\ell(D + [\mathbf{0}_E]) = 1$  by Corollary 1.15, and we have  $\phi \in L(D + [\mathbf{0}_E])$ . Then  $\text{div}(\phi) + D + [\mathbf{0}_E] \geq 0$  and has degree 1. Any non-negative divisor with degree 1 is a single point  $[P]$ . Thus  $D + [\mathbf{0}_E]$  is linearly equivalent to  $[P]$ ; so, the map is surjective.  $\square$

**Corollary 2.2.** *If  $0 \neq \omega \in \Omega_{E/\bar{k}}$ , then  $\text{div}(\omega) = 0$ .*

*Proof.* Since  $E(\bar{k})$  is a group, for each  $P \in E(\bar{k})$ ,  $\mathcal{T}_P : Q \mapsto Q + P$  gives an automorphism of  $E$ . Thus  $\omega \circ \mathcal{T}_P$  is another element in  $\Omega_{E/\bar{k}}$ . Since  $\dim \Omega_{E/\bar{k}} = 1$ , we find  $\omega \circ \mathcal{T}_P = \lambda(P)\omega$  for  $\lambda \in \bar{k}$ . Since  $\omega \neq 0$ , at some point  $P \in E(\bar{k})$ ,  $v_P(\omega) = 0$ . Since  $v_Q(\omega \circ \mathcal{T}_P) = v_{P+Q}(\omega)$  and we can bring any point to  $P$  by translation, we have  $v_P(\omega) = 0$  everywhere. Thus  $\text{div}(\omega) = 0$ .  $\square$

We can show easily  $\lambda(P) = 1$  for all  $P$  (see [GME] §2.2.3). The nonzero differentials  $\omega$  in  $\Omega_{E/k}$  are called *nowhere vanishing differentials* as  $\text{div}(\omega) = 0$ . They are unique up to constant multiple.

**Exercise 2.3.** *Take a line  $L$  defined by  $aX + bY + cZ$  on  $\mathbf{P}^2$  and suppose its intersection with an elliptic curve  $E \subset \mathbf{P}^2$  to be  $\{P, Q, R\}$ . Prove that  $[P] + [Q] + [R] \sim 3[\mathbf{0}_E]$ .*

A field  $k$  is called a *perfect field* if any finite field extension of  $k$  is separable (i.e., generated by  $\theta$  over  $k$  whose minimal equation over  $k$  does not have multiple roots). Fields of characteristic 0 and finite fields are perfect.

**Remark 2.4.** *If  $k$  is perfect,  $\bar{k}/k$  is possibly an infinite Galois extension; so, by Galois theory, we have a bijection between open subgroups  $G$  of  $\text{Gal}(\bar{k}/k)$  and finite extensions  $K/k$  inside  $\bar{k}$  by  $G \mapsto \bar{k}^G = \{x \in \bar{k} \mid \sigma(x) = x \text{ for all } \sigma \in G\}$  and  $K \mapsto \text{Gal}(\bar{k}/K)$ . Since the isomorphism  $E(\bar{k}) \cong J(E)$  is Galois equivariant, we have*

$$E(K) \cong J(E)^{\text{Gal}(\bar{k}/K)} = \{D \in J(E) \mid \sigma(D) = D \text{ for all } \sigma \in G\},$$

where  $\sigma \in \text{Gal}(\bar{k}/k)$  acts on  $D = \sum_P m_P [P]$  by  $\sigma(D) = \sum_P m_P [\sigma(P)]$ . Basically by definition, we have

$$J(E)(K) := J(E)^{\text{Gal}(\bar{k}/K)} = \frac{\{D \in \text{Pic}^0(E) \mid \sigma(D) = D\}}{\{\text{div}(f) \mid f \in K(E)^\times\}}.$$

Since any subfield  $K \subset \bar{k}$  is a union of finite extensions, the identity  $E(K) \cong J(E)(K)$  is also true for an infinite extension  $K/k$  inside  $\bar{K}$ . Actually we have a good definition of  $\text{Pic}(E)(A)$  for any  $k$ -algebra  $A$ , and we can generalize the identity  $E(K) \cong J(E)(K)$  to all  $k$ -algebras  $A$  in place of fields  $K$  inside  $\bar{k}$ .

**2.2. Weierstrass Equations of Elliptic Curves.** We now embed  $E/k$  into the two-dimensional projective space  $\mathbf{P}_{/k}^2$  using a basis of  $L(3[\mathbf{0}])$  and determine the equation of the image in  $\mathbf{P}_{/k}^2$ . Choose a parameter  $T = t_{\mathbf{0}}$  at the origin  $\mathbf{0} = \mathbf{0}_E$ . We first consider  $L(n[\mathbf{0}])$  which has dimension  $n$  if  $n > 0$ . We have  $L([\mathbf{0}]) = k$  and  $L(2[\mathbf{0}]) = k1 + kx$ . Since  $x$  has to have a pole of order 2 at  $\mathbf{0}$ , we may normalize  $x$  so that  $x = T^{-2}(1 + \text{higher terms})$  in  $k[[T]]$ . Here  $x$  is unique up to translation:  $x \mapsto x + a$  with  $a \in k$ . Then  $L(3[\mathbf{0}]) = k1 + kx + ky$ . We may then normalize  $y$  so that  $y = -T^{-3}(1 + \text{higher terms})$  (following the tradition, we later rewrite  $y$  for  $2y$ ; thus, the normalization will be  $y = -2T^{-3}(1 + \text{higher terms})$  at the end). Then  $y$  is unique up to the affine transformation:  $y \mapsto y + ax + b$  ( $a, b \in k$ ).

**Proposition 2.5.** *Suppose that the characteristic of the base field  $k$  is different from 2 and 3. Then for a given pair  $(E, \omega)$  of an elliptic curve  $E$  and a nowhere-vanishing differential  $\omega$  both defined over  $k$ , we can find a unique base  $(1, x, y)$  of  $L(3[\mathbf{0}])$  such that  $E$  is embedded into  $\mathbf{P}_{/k}^2$  by  $(1, x, y)$  whose image is defined by the affine equation*

$$(2.1) \quad y^2 = 4x^3 - g_2x - g_3 \quad \text{with } g_2, g_3 \in k,$$

and  $\omega$  on the image is given by  $\frac{dx}{y}$ . Conversely, a projective algebraic curve defined by the above equation is an elliptic curve with a specific nowhere-vanishing differential  $\frac{dx}{y}$  if and only if the discriminant  $\Delta(E, \omega) = g_2^3 - 27g_3^2$  of  $4X^3 - g_2X - g_3$  does not vanish.

An equation of an elliptic curve  $E$  as in (2.1) is called a *Weierstrass equation* of  $E$ , which is determined by the pair  $(E, \omega)$ .

*Proof.* By the dimension formulas, counting the order of poles at  $\mathbf{0}$  of monomials of  $x$  and  $y$ , we have

$$\begin{aligned} L(4[\mathbf{0}]) &= k + kx + ky + kx^2, \\ L(5[\mathbf{0}]) &= k + kx + ky + kx^2 + kxy \quad \text{and} \\ L(6[\mathbf{0}]) &= k + kx + ky + kx^2 + kxy + kx^3 \\ &= k + kx + ky + kx^2 + kxy + ky^2, \end{aligned}$$

from which the following relation results,

$$(2.2) \quad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \quad \text{with } a_j \in k,$$

because the poles of order 6 of  $y^2$  and  $x^3$  have to be canceled. We homogenize the equation (2.2) by putting  $x = \frac{X}{Z}$  and  $y = \frac{Y}{Z}$  (and multiplying by  $Z^3$ ). Write  $C$  for

the projective plane  $k$ -curve in  $\mathbf{P}^2$  defined by the (homogenized) equation. Thus we have a  $k$ -regular map:  $\phi : E \rightarrow C \subset \mathbf{P}^2$  given by  $P \mapsto (x(P) : y(P) : 1)$ . Thus the function field  $k(E)$  contains the function field  $k(C)$  by the pull back of  $\phi$ . By definition,  $k(C) = k(x, y)$ . Since  $\text{div}_\infty(x) = 2[\mathbf{0}_E]$  for  $x = \frac{X}{Z} : E \rightarrow \mathbf{P}^1$ , this gives a covering of degree 2; so,  $[k(E) : k(x)] = 2$ . Similarly  $[k(E) : k(y)] = 3$ . Since  $[k(E) : k(C)]$  is a common factor of  $[k(E) : k(x)] = 2$  and  $[k(E) : k(y)] = 3$ , we get  $k(E) = k(C)$ . Thus if  $C$  is smooth,  $E \cong C$  by  $\phi$  as a smooth geometrically irreducible curve is determined by its function field. Therefore, assuming  $C$  is smooth,  $E/k$  can be embedded into  $\mathbf{P}^2/k$  via  $P \mapsto (x(P), y(P))$ . The image is defined by the equation (2.2).

Let  $T$  be a local parameter at  $\mathbf{0}_E$  normalized so that

$$\omega = (1 + \text{higher degree terms})dT.$$

Anyway  $\omega = (a + \text{higher degree terms})dT$  for  $a \in k^\times$ , and by replacing  $T$  by  $aT$ , we achieve this normalization. The parameter  $T$  normalized as above is called a parameter adapted to  $\omega$ . Then we may normalize  $x$  so that  $x = T^{-2} + \text{higher degree terms}$ . We now suppose that 2 is invertible in  $k$ . Then we may further normalize  $y$  so that  $y = -2T^{-3} + \text{higher degree terms}$  (which we will do soon but not yet; so, for the moment, we still assume  $y = T^{-3} + \text{higher degree terms}$ ).

The above normalization is not affected by variable change of the form  $y \mapsto y + ax + b$  and  $x \mapsto x + a'$ . Now we make a variable change  $y \mapsto y + ax + b$  in order to remove the terms of  $xy$  and  $y$  (i.e., we are going to make  $a_1 = a_3 = 0$ ):

$$\begin{aligned} (y + ax + b)^2 + a_1x(y + ax + b) + a_3(y + ax + b) \\ = y^2 + (2a + a_1)xy + (2b + a_3)y + \text{polynomial in } x. \end{aligned}$$

Assuming that 2 is invertible in  $k$ , we take  $a = -\frac{a_1}{2}$  and  $b = -\frac{a_3}{2}$ . The resulting equation is of the form  $y^2 = x^3 + b_2x^2 + b_4x + b_6$ . We now make the change of variable  $x \mapsto x + a'$  to make  $b_2 = 0$ :

$$y^2 = (x + a')^3 + b_2(x + a')^2 + b_4(x + a') + b_6 = x^3 + (3a' + b_2)x^2 + \dots$$

Assuming that 3 is invertible in  $k$ , we take  $a' = -\frac{b_2}{3}$ . We can rewrite the equation as in (2.1) (making a variable change  $-2y \mapsto y$ ). By the variable change as above, we have  $y = -2T^{-3}(1 + \text{higher terms})$ , and from this, we conclude  $\omega = \frac{dx}{y}$ . The numbers  $g_2$  and  $g_3$  are determined by  $T$  adapted to a given nowhere-vanishing differential form  $\omega$ .

If the discriminant  $\Delta(E, \omega)$  of  $g(x) = 4x^3 - g_2x - g_3$  vanishes,  $C$  has only singularity at  $(x_0 : 0 : 1)$  for a multiple root  $x_0$  of  $g(x) = 0$ . If  $g(x)$  has a double zero,  $C$  is isomorphic over  $\bar{k}$  to the curve defined by  $y^2 = x^2(x - a)$  for  $a \neq 0$ . Let  $t = \frac{x}{y}$ . Then for  $P \in E(\bar{k})$  mapping to  $(0, 0)$ ,  $v_P(y) = v_P(x)$ ; so,  $P$  is neither a zero nor a pole of  $t$ . The function  $t$  never vanishes outside  $\mathbf{0}_E$  (having a pole at  $(a, 0)$ ). It has a simple zero at  $\mathbf{0}_E$  by the normalization of  $x$  and  $y$ . Thus  $\text{deg}(\text{div}_0(t)) = 1$ , and  $\bar{k}(C) = \bar{k}(t)$ , which is impossible as  $k(C) = k(E)$  and  $g(E) = 1$ . The case of triple zero can be excluded similarly. Thus we conclude  $\Delta(E, \omega) \neq 0$  ( $\Leftrightarrow C$  is smooth: Example 1.3), and we have  $E \cong C$  by  $\phi$ .



Conversely, we have seen that any curve defined by equation (2.1) is smooth in Example 1.3 if the cubic polynomial  $F(X) = 4X^3 - g_2X - g_3$  has three distinct roots in  $k$ . In other words, if the discriminant  $\Delta(E, \omega)$  of  $F(X)$  does not vanish,  $E$  is smooth.

For a given equation,  $Y^2 = F(X)$ , the algebraic curve  $E$  defined by the homogeneous equation  $Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$  in  $\mathbf{P}_{/k}^2$  has a rational point  $\mathbf{0} = (0, 1, 0) \in E(k)$ , which is  $\infty$  in  $\mathbf{P}^2$ . Thus  $E$  is smooth over  $k$  if and only if  $\Delta(E, \omega) \neq 0$  (an exercise following this proof).

We show that there is a canonical nowhere-vanishing differential  $\omega \in \Omega_{E/k}$  if  $E$  is defined by (2.1). If such an  $\omega$  exists, all other holomorphic differentials  $\omega'$  are of the form  $f\omega$  with  $\text{div}(f) \geq 0$ , which implies  $f \in k$ ; so,  $g = \dim_k \Omega_{E/k} = 1$ , and  $E/k$  is an elliptic curve. It is an easy exercise to show that  $y^{-1}dx$  does not vanish on  $E$  (an exercise following this proof).

We summarize what we have seen. Returning to the starting elliptic curve  $E/k$ , for the parameter  $T$  at the origin, we see by definition

$$x = T^{-2}(1 + \text{higher degree terms}) \quad \text{and} \quad y = -2T^{-3}(1 + \text{higher degree terms}).$$

This shows

$$\frac{dx}{y} = \frac{-2T^{-3}(1 + \dots)}{-2T^{-3}(1 + \dots)}dT = (1 + \text{higher degree terms})dT = \omega.$$

Thus the nowhere-vanishing differential form  $\omega$  to which  $T$  is adapted is given by  $\frac{dx}{y}$ . Conversely, if  $\Delta \neq 0$ , the curve defined by  $y^2 = 4x^3 - g_2x - g_3$  is an elliptic curve over  $k$  with origin  $\mathbf{0} = \infty$  and a standard nowhere-vanishing differential form  $\omega = \frac{dx}{y}$ . This finishes the proof.  $\square$

- Exercise 2.6.** (1) If  $C$  is defined by  $y^2 = x^3$ , prove  $k(C) = k(t)$  for  $t = \frac{x}{y}$ .  
 (2) Compute  $v_P(dx/y)$  explicitly at any point  $P$  on  $E(\bar{k})$ .  
 (3) Show that if  $\Delta \neq 0$ , the curve defined by  $y^2 = 4x^3 - g_2x - g_3$  is also smooth at  $\mathbf{0} = \infty$ .

**2.3. Moduli of Weierstrass Type.** We continue to assume that the characteristic of  $k$  is different from 2 and 3. Suppose that we are given two elliptic curves  $(E, \omega)_{/k}$  and  $(E', \omega')_{/k}$  with nowhere-vanishing differential forms  $\omega$  and  $\omega'$ . We call two pairs  $(E, \omega)$  and  $(E', \omega')$  isomorphic if we have an isomorphism  $\varphi : E \rightarrow E'$  with  $\varphi^*\omega' = \omega$ . Here for  $\omega' = fdg$ ,  $\varphi^*\omega' = (f \circ \varphi)d(g \circ \varphi)$ ; in other words, if  $\sigma : k(E') \rightarrow k(E)$  is the isomorphism of the function fields associated with  $\varphi$ ,  $\varphi^*\omega' = \sigma(f)d(\sigma(g))$ . Let  $T'$  be the parameter at the origin  $\mathbf{0}$  of  $E'$  adapted to  $\omega'$ . If  $\varphi : (E, \omega) \cong (E', \omega')$ , then the parameter  $T = \varphi^*T' \bmod T^2$  is adapted to  $\omega$  (because  $\varphi^*\omega' = \omega$ ). We choose coordinates  $(x, y)$  for  $E$  and  $(x', y')$  for  $E'$  relative to  $T$  and  $T'$  as above. By the uniqueness of the choice of  $(x, y)$  and  $(x', y')$ , we know  $\varphi^*x' = x$  and  $\varphi^*y' = y$ . Thus the Weierstrass equations of  $(E, \omega)$  and  $(E', \omega')$  coincide. We write  $g_2(E, \omega)$  and  $g_3(E, \omega)$  for the  $g_2$  and  $g_3$  of the coefficients of the Weierstrass equation of  $(E, \omega)$ . If a field  $K$  has characteristic different from 2 and 3, we have

$$(2.3) \quad \mathcal{P}(K) := [(E, \omega)_{/K}] \cong \{(g_2, g_3) \in K^2 \mid \Delta(E, \omega) \neq 0\} \cong \text{Hom}_{\text{ALG}}(\mathcal{R}, K),$$

where  $\mathcal{R} := \mathbb{Q}[g_2, g_3, \frac{1}{g_2^3 - 27g_3^2}]$  (the polynomial ring of variables  $g_j$  with  $g_2^3 - 27g_3^2$  inverted) and  $[\cdot]$  indicates the set of isomorphism classes of the objects inside the bracket and  $\text{Spec}(R)(K)$  for a ring  $R$  is the set of all algebra homomorphisms:  $R \rightarrow K$ . The last isomorphism sends  $(g_2, g_3)$  to the algebra homomorphism  $\phi$  with  $\phi(X) = g_2$  and  $\phi(Y) = g_3$ .

There is an elliptic curve  $\mathbb{E}$  defined by  $Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$  over  $\mathcal{R}$ . This is a universal curve in the sense that for any pair  $(E, \omega)_{/A}$  defined by  $Y^2Z = 4X^3 - a_2XZ^2 - a_3Z^3$  with  $\omega = \frac{dX}{Y}$  over  $A$ , we have a unique morphism  $\mathcal{R} \xrightarrow{\varphi} A$  such that  $\varphi(g_j) = a_j$  induces the pair  $(E, \omega)$ . In other words, (2.3) means that for each  $\phi \in \text{Hom}_{ALG}(\mathcal{R}, K)$ , there is a unique object  $(E, \omega)_{/K}$  defined by the equation  $Y^2Z = 4X^3 - \phi(g_2)XZ^2 - \phi(g_3)Z^3$  with  $\omega = \frac{dX}{Y}$  (not just an isomorphism class of  $(E, \omega)_{/K}$ ; so, for such representability, it is absolutely necessary that  $\text{Aut}(E, \omega)_{/K} = \{\text{id}\}$  as otherwise, we would have several choices in the isomorphism class of  $(E, \omega)_{/K}$ ).

We now classify elliptic curves  $E$  eliminating the contribution of the differential from the pair  $(E, \omega)$ . If  $\varphi : E \cong E'$  for  $(E, \omega)$  and  $(E', \omega')$ , we have  $\varphi^*\omega' = \lambda\omega$  with  $\lambda \in K^\times$ , because  $\varphi^*\omega'$  is another nowhere-vanishing differential. Therefore we study  $K^\times$ -orbit:  $(E, \omega) \bmod K^\times$  under the action of  $\lambda \in K^\times$  given by  $(E, \omega)_{/K} \mapsto (E, \lambda\omega)_{/K}$ , computing the dependence of  $g_j(E, \lambda\omega)$  ( $j = 2, 3$ ) on  $\lambda$  for a given pair  $(E, \omega)_{/K}$ . Let  $T$  be the parameter adapted to  $\omega$ . Then  $\lambda T$  is adapted to  $\lambda\omega$ . We see

$$\begin{aligned} x(E, \omega) &= \frac{(1 + T\phi(T))}{T^2} \Rightarrow x(E, \lambda\omega) = \frac{(1 + \text{higher terms})}{(\lambda T)^2} = \lambda^{-2}x(E, \omega), \\ y(E, \omega) &= \frac{(-2 + T\psi(T))}{T^3} \Rightarrow y(E, \lambda\omega) = \frac{(-2 + \text{higher terms})}{(\lambda T)^3} = \lambda^{-3}y(E, \omega). \end{aligned}$$

Since  $y^2 = 4x^3 - g_2(E, \omega)x - g_3(E, \omega)$ , we have

$$\begin{aligned} (\lambda^{-3}y)^2 &= 4\lambda^{-6}x^3 - g_2(E, \omega)\lambda^{-6}x - \lambda^{-6}g_3(E, \omega) \\ &= 4(\lambda^{-2}x)^3 - \lambda^{-4}g_2(E, \omega)(\lambda^{-2}x) - \lambda^{-6}g_3(E, \omega). \end{aligned}$$

This shows

$$(2.4) \quad g_2(E, \lambda\omega) = \lambda^{-4}g_2(E, \omega) \quad \text{and} \quad g_3(E, \lambda\omega) = \lambda^{-6}g_3(E, \omega).$$

Thus we have

**Theorem 2.7.** *If two elliptic curves  $E_{/K}$  and  $E'_{/K}$  are isomorphic, then choosing nowhere-vanishing differentials  $\omega_{/E}$  and  $\omega'_{/E'}$ , we have  $g_j(E', \omega') = \lambda^{-2j}g_j(E, \omega)$  for  $\lambda \in K^\times$ . The constant  $\lambda$  is given by  $\varphi^*\omega' = \lambda\omega$ .*

We define the  $J$ -invariant of  $E$  by  $J(E) = \frac{(12g_2(E, \omega))^3}{\Delta(E, \omega)}$ . Then  $J$  only depends on  $E$  (not the chosen differential  $\omega$ ). If  $J(E) = J(E')$ , then we have

$$\frac{(12g_2(E, \omega))^3}{\Delta(E, \omega)} = \frac{(12g_2(E', \omega'))^3}{\Delta(E', \omega')} \iff g_j(E', \omega') = \lambda^{-2j}g_j(E, \omega)$$

for a twelfth root  $\lambda$  of  $\Delta(E, \omega)/\Delta(E', \omega')$ . Note that the twelfth root  $\lambda$  may not be in  $K$  if  $K$  is not algebraically closed.

Conversely, for a given  $j \notin \{0, 1\}$ , the elliptic curve defined by  $y^2 = 4x^3 - gx - g$  for  $g = \frac{27j}{j-1}$  has  $J$ -invariant  $12^3 j$ . If  $j = 0$  or  $1$ , we can take the following elliptic curve with  $J = 0$  or  $12^3$ . If  $J = 0$ , then  $y^2 = 4x^3 - 1$  and if  $J = 12^3$ , then  $y^2 = 4x^3 - 4x$ . Thus we have

**Corollary 2.8.** *If  $K$  is algebraically closed, then  $J(E) = J(E') \Leftrightarrow E \cong E'$  for two elliptic curves over  $K$ . Moreover, for any field  $K$ , there exists an elliptic curve  $E$  with a given  $J(E) \in K$ .*

**Exercise 2.9.** (1) *Prove that  $g_j(E', \omega') = \lambda^{-2j} g_j(E, \omega)$  for suitable  $\omega$  and  $\omega'$  and a suitable twelfth root  $\lambda$  of  $\Delta(E, \omega)/\Delta(E', \omega')$  if  $J(E) = J(E')$ .*  
 (2) *Explain what happens if  $J(E) = J(E')$  but  $E \not\cong E'$  over a field  $K$  not necessarily algebraically closed.*

Note that  $\mathcal{R}$  is a graded ring such that  $g_2$  is of degree 4 and  $g_3$  is of degree 6. Then the degree 0 subring  $\mathcal{R}_0 = \mathbb{Q}[J]$ . Note that  $\text{Spec}(\mathbb{Q}[J])$  is the 1 dimensional affine space  $\mathbb{A}^1$ , as  $\text{Spec}(\mathbb{Q}[J])(A) = \text{Hom}_{\text{alg}}(\mathbb{Q}[J], A) \cong A$  by  $\phi \mapsto \phi(J) \in A$ .

Consider functors

$$(2.5) \quad \begin{aligned} \mathcal{P}_{1,N}(A) &:= \left[ (E, \omega, \phi)_{/A} \mid \phi : N^{-1}\mathbb{Z}/\mathbb{Z} \hookrightarrow E[N] := \text{Ker}(E \xrightarrow{N} E) \right], \\ \mathcal{E}_{1,N}(A) &:= \left[ (E, \phi)_{/A} \mid \phi : N^{-1}\mathbb{Z}/\mathbb{Z} \hookrightarrow E[N] := \text{Ker}(E \xrightarrow{N} E) \right] \end{aligned}$$

for a positive integer  $N$ . The functor has natural transformation  $\mathcal{P}_{1,N} \rightarrow \mathcal{P}$  sending  $(E, \omega, \phi)_{/A}$  to  $(E, \omega)_{/A}$ . This is represented by  $\mathcal{Y}_1(N) := \mathbb{E}[N] - \bigcup_{0 < d \mid N, d \neq N} \mathbb{E}[d]$  which is affine in the sense that  $\mathcal{Y}_1(N) = \text{Spec}(\mathcal{R}_{1,N})$  for the ring  $\mathcal{R}_{1,N} := \text{Hom}_{\text{COF}}(\mathcal{Y}_1(N), \mathbb{A})$ , and  $\mathcal{R}_{1,N}$  is finite locally free over  $\mathcal{R}$ . By the action  $\omega \mapsto \lambda\omega$  on  $\mathcal{P}_{1,N}$ , the multiplicative group  $\mathbb{G}_m$  given by  $A \mapsto A^\times$  acts on  $\mathcal{R}_{1,N}$ . The subring  $\mathcal{A}_{1,N} := H^0(\mathbb{G}_m, \mathcal{R}_{1,N})$  fixed by this action represents  $\mathcal{E}_{1,N}$ , i.e.,

$$\mathcal{E}_{1,N}(A) = \text{Hom}_{\mathbb{Q}\text{-alg}}(\mathcal{A}_{1,N}, A).$$

The corresponding plane curve  $Y_1(N)_{/\mathbb{Q}} := \text{Spec}(\mathcal{A}_{1,N})$  is the modular curve of level  $\Gamma_1(N)$ . In other words, for a triple  $(E, \omega, \phi : \mathbb{Z}/N\mathbb{Z} \hookrightarrow E[N])_{/A}$ , the value  $\phi(1)$  give rise to a unique point of  $\mathcal{Y}_1(N)(A)$  over the point corresponding  $(E, \omega)_{/A} \in \mathcal{P}_{1,N}(A)$ .

Similarly,

$$(2.6) \quad \begin{aligned} \mathcal{P}_{1,N}(A) &:= \left[ (E, \omega, \phi)_{/A} \mid \phi : (N^{-1}\mathbb{Z}/\mathbb{Z})^2 \cong E[N] := \text{Ker}(E \xrightarrow{N} E) \right], \\ \mathcal{E}_{1,N}(A) &:= \left[ (E, \phi)_{/A} \mid \phi : (N^{-1}\mathbb{Z}/\mathbb{Z})^2 \cong E[N] := \text{Ker}(E \xrightarrow{N} E) \right] \end{aligned}$$

is represented by

$$\mathcal{Y}(N) = \{(x, y) \in \mathbb{E}[N] \times_{\mathcal{Y}} \mathbb{E}[N] \mid x \wedge y \in \mathbb{E}[N] \wedge \mathbb{E}[N] - \bigcup_{0 < d \mid N, d \neq N} \mathbb{E}[d] \wedge \mathbb{E}[d]\}$$

and  $Y(N) = \text{Spec}(\mathcal{A}_N)$  with  $\mathcal{A}_N = H^0(\mathbb{G}_m, \mathcal{R}_N)$ , respectively, where  $\mathcal{Y}(N) = \text{Spec}(\mathcal{R}_N)$  for  $\mathcal{R}_N := \text{Hom}_{\text{COF}}(\mathcal{Y}(N), \mathbb{A})$

**Remark 2.10.** *We will see later that the curve  $Y(N)$  is irreducible over  $\mathbb{Q}$  but becomes reducible over  $\mathbb{Q}[\mu_N]$  (the cyclotomic field of  $N$ -th root of unity), looking into the  $q$ -expansion of Weierstrass  $\wp$ -functions. The function field  $\mathbb{Q}(Y(N))$  therefore contains  $\mathbb{Q}[\mu_N]$  as algebraic closure of  $\mathbb{Q}$  in it.*

### 3. MODULAR FORMS AND FUNCTIONS

We give an algebraic definition of modular forms and then relate it to classical theory (due to Weierstrass, Klein, Fricke).

**3.1. Geometric modular forms.** Let  $A$  be an algebra over  $\mathbb{Q}$ . We restrict the functor  $\mathcal{P}$  to  $ALG_{/A}$  and write the restriction  $\mathcal{P}_{/A}$ . Then by (2.3), for  $\mathcal{R}_A := A[g_2, g_3, \frac{1}{\Delta}]$ ,

$$\mathcal{P}_{/A}(?) = \text{Hom}_{ALG_{/A}}(\mathcal{R}_A, ?).$$

A morphism of functors  $\phi : \mathcal{P}_{/A} \rightarrow \mathbb{A}_{/A}^1$  is by definition given by maps  $\phi_R : \mathcal{P}_{/A}(R) \rightarrow \mathbb{A}^1(R) = R$  indexed by  $R \in ALG_{/A}$  such that for any  $\sigma : R \rightarrow R'$  in  $\text{Hom}_{ALG_{/A}}(R, R')$ ,  $\phi_{R'}((E, \omega) \otimes_R R') = \sigma(f((E, \omega)_{/R}))$ . Note that  $\mathbb{A}_{/A}^1(?) = \text{Hom}_{ALG_{/A}}(A[X], ?)$  by  $R \ni a \leftrightarrow (\varphi : A[X] \rightarrow R) \in \text{Hom}_{ALG_{/A}}(A[X], ?)$  with  $\varphi(X) = a$ . Thus in particular,

$$\phi_{\mathcal{R}_A} : \mathcal{P}(\mathcal{R}_A) = \text{Hom}_{ALG_{/A}}(\mathcal{R}_A, \mathcal{R}_A) \rightarrow \mathbb{A}^1(A[X], \mathcal{R}_A) = \mathcal{R}_A.$$

Thus  $\phi_{\mathcal{R}_A}(\text{id}_{\mathcal{R}_A}) \in \mathcal{R}_A$ ; so, write  $\phi_{\mathcal{R}_A}(\text{id}_{\mathcal{R}_A}) = \Phi(g_2, g_3)$  for a two variable rational function  $\Phi(x, y) \in A[x, y, \frac{1}{x^3 - 27y^2}]$ . Let  $\mathbf{E}_{/\mathcal{R}_A}$  be the universal elliptic curve over  $\mathcal{R}_A$  defined by  $Y^2Z = 4X^3 - g_2XZ^2 - g_3Z^3$  with the universal differential  $\omega = \frac{dX}{Y}$ . If we have  $(E, \omega)_{/R}$ , we have a unique  $A$ -algebra homomorphism  $\sigma : \mathcal{R}_A \rightarrow R$  given by  $\sigma(g_j) = g_j(E, \omega)$ ; in other words,  $(E, \omega)_{/R} \cong (\mathbf{E}, \omega)_{\mathcal{R}_A} \otimes_{\mathcal{R}_A} R$ . Thus

$$\begin{aligned} \phi_R(E, \omega) &= \phi_R((\mathbf{E}, \omega) \otimes_{\mathcal{R}_A} R) = \sigma(\phi_{\mathcal{R}_A}(\mathbf{E}, \omega)) \\ &= \sigma(\phi_{\mathcal{R}_A}(\text{id}_{\mathcal{R}_A})) = \Phi(\sigma(g_2), \sigma(g_3)) = \Phi(g_2(E, \omega), g_3(E, \omega)). \end{aligned}$$

**Theorem 3.1.** *Any functor morphism  $\phi : \mathcal{P}_{/A} \rightarrow \mathbb{A}_{/A}^1$  is given by a rational function  $\Phi \in \mathcal{R}_A$  of  $g_2$  and  $g_3$  so that  $\phi(E, \omega) = \Phi(g_2(E, \omega), g_3(E, \omega))$  for every elliptic curve  $(E, \omega)$  over an  $A$ -algebra.*

Define a weight function  $w : A[g_2, g_3] \rightarrow \mathbb{Z}$  by  $w(g_2^a g_3^b) = 4a + 6b$ , and for general polynomials  $\Phi = \sum_{a,b} c_{a,b} g_2^a g_3^b$ , we put  $w(\Phi) = \max(w(g_2^a g_3^b) | c_{a,b} \neq 0)$ . A polynomial  $\Phi = \sum_{a,b} c_{a,b} g_2^a g_3^b$  of  $g_2$  and  $g_3$  is called *isobaric* if  $c_{a,b} \neq 0 \Rightarrow 4a + 6b = w$ .

A weight  $w$  modular form defined over  $A$  is a morphism of functors  $\mathcal{P}_{/A} \rightarrow \mathbb{A}_{/A}^1$  given by an isobaric polynomial of  $g_2$  and  $g_3$  of weight  $w$  with coefficients in  $A$ . Write  $G_w(A)$  for the  $A$ -module of modular forms of weight  $w$ . Then  $f \in G_w(A)$  is a functorial rule assigning each isomorphism class of  $(E, \omega)_{/R}$  for an  $A$ -algebra  $R$  an element  $f(E, \omega) \in R$  satisfying the following properties:

- (G0)  $f \in A[g_2, g_3]$ ,
- (G1) If  $(E, \omega)$  is defined over an  $A$ -algebra  $R$ , we have  $f(E, \omega) \in R$ , which depends only on the isomorphism class of  $(E, \omega)$  over  $R$ ,
- (G2)  $f((E, \omega) \otimes_R R') = \sigma(f(E, \omega))$  for  $A$ -algebra homomorphism  $\sigma : R \rightarrow R'$ ,
- (G3)  $f((E, \lambda\omega)_{/R}) = \lambda^{-w} f(E, \omega)$  for any  $\lambda \in R^\times$ .

**Exercise 3.2.** For a field  $K \supset \mathbb{Q}$ , prove for  $0 < w \in 2\mathbb{Z}$ ,

$$\dim_K G_w(K) = \begin{cases} \left[ \frac{w}{12} \right] & \text{if } w \equiv 2 \pmod{12}, \\ \left[ \frac{w}{12} \right] + 1 & \text{otherwise.} \end{cases}$$

We can define modular form of level  $N$  rational over  $A$  replacing by the functor  $\mathcal{P}$  in the previous section by  $\mathcal{P}_{1,N}$  or  $\mathcal{P}_N$ . We list the functorial properties:

- (G<sub>N</sub>0)  $f$  is integral over  $A[g_2, g_3]$ ,
- (G<sub>N</sub>1) If  $(E, \omega, \phi)$  is defined over an  $A$ -algebra  $R$ , we have  $f(E, \omega, \phi) \in R$ , which depends only on the isomorphism class of  $(E, \omega)$  over  $R$ ,
- (G<sub>N</sub>2)  $f((E, \omega, \phi) \otimes_R R') = \sigma(f(E, \omega, \phi))$  for  $A$ -algebra homomorphism  $\sigma : R \rightarrow R'$ ,
- (G<sub>N</sub>3)  $f((E, \lambda\omega, \phi)_{/R}) = \lambda^{-w} f(E, \omega, \phi)$  for any  $\lambda \in R^\times$ .

Here  $\phi$  is the level structure depending on our choice of  $\mathcal{P}_{1,N}$  and  $\mathcal{P}_N$ .

We then define  $G_k(\Gamma_1(N); A)$  (resp.  $G_k(\Gamma(N); A)$ ) for the space of modular forms of weight  $k$  defined over  $A$  for  $\mathcal{P}_{1,N}$  (resp.  $\mathcal{P}_N$ ).

**3.2. Topological Fundamental Groups.** In the following three sections, we would like to give a sketch of Weierstrass' theory of elliptic curves defined over the complex field  $\mathbb{C}$ . By means of Weierstrass  $\mathcal{P}$ -functions, we can identify  $E(\mathbb{C})$  (for each elliptic curve  $E_{/\mathbb{C}}$ ) with a quotient of  $\mathbb{C}$  by a lattice  $L$ . In this way, we can identify  $[(E, \omega)_{/\mathbb{C}}]$  with the space of lattices in  $\mathbb{C}$ . This method is analytic.

We can deduce from the analytic parameterization (combining with geometric technique of Weil-Shimura) many results on the moduli space of elliptic curves, like, the exact field of definition of the moduli, determination of the field of moduli (of each member), and so on (e.g., [IAT] Chapter 6). We have come here in a reverse way: starting algebraically, mainly by the Riemann-Roch theorem, we have determined a unique Weierstrass equation over  $A$  for a given pair  $(E, \omega)_{/A}$ , and therefore, we know the exact shape of the moduli space before setting out in studying analytic method. After studying analytic theory over  $\mathbb{C}$ , combining these techniques, we start studying modular units.

Let  $(E, \omega)_{/\mathbb{C}}$  be an elliptic curve over  $\mathbb{C}$ . Then

$$E(\mathbb{C}) = E(g_2, g_3)(\mathbb{C}) = \{(x : y : z) \in \mathbf{P}^2(\mathbb{C}) \mid y^2z - 4x^3 + g_2z^2x + g_3z^3 = 0\},$$

and  $E(\mathbb{C})$  is a compact Riemann surface of genus 1. A path  $\gamma : y \rightarrow x$  on  $E(\mathbb{C})$  is a piecewise smooth continuous map  $\gamma$  from the interval  $[0, 1]$  into  $E(\mathbb{C})$  (under the Euclidean topology on  $E(\mathbb{C})$ ) such that  $\gamma(0) = y$  and  $\gamma(1) = x$ . Two paths  $\gamma, \gamma' : x \rightarrow x$  are homotopy equivalent (for which we write  $\gamma \approx \gamma'$ ) if there is a bi-continuous map  $\varphi : [0, 1] \times [0, 1] \rightarrow E(\mathbb{C})$  such that  $\varphi(0, t) = \gamma(t)$  and  $\varphi(1, t) = \gamma'(t)$ . Let  $\mathcal{Z}$  be the set of all equivalence classes of paths emanating from  $\mathbf{0}$ .

More generally, for each complex manifold  $M$ , we can think of the space  $\mathcal{Z} = \mathcal{Z}(M)$  of homotopy classes of paths emanating from a fixed point  $x \in M$ . An open neighborhood  $U$  of  $x$  is called *simply connected* if  $\mathcal{Z}(U) \cong U$  by projecting  $(\gamma : x \rightarrow y)$  down to  $y$ . For example, if  $U$  is diffeomorphic to an open disk with center  $x$ , it is simply connected (that is, every loop is equivalent to  $x$ ). If  $\gamma : x \rightarrow y$  and  $\gamma' : y \rightarrow z$

are two paths, we define their product path  $\gamma\gamma' : x \rightarrow z$  by

$$\gamma\gamma'(t) = \begin{cases} \gamma(2t) & \text{if } 0 \leq t \leq 1/2 \\ \gamma'(2t-1) & \text{if } 1/2 \leq t \leq 1. \end{cases}$$

By this multiplication,  $\pi_M = \pi(M, x) = \{\gamma \in \mathcal{Z}(M) | \gamma : x \rightarrow x\} / \approx$  becomes a group called *the topological fundamental group* of  $M$ . Taking a fundamental system of neighborhoods  $\mathcal{U}_y$  of  $y \in M$  made of simply connected open neighborhoods of  $y$ , we define a topology on  $\mathcal{Z}(M)$  so that a fundamental system of neighborhoods of  $\gamma : x \rightarrow y$  is given by  $\{\gamma U | U \in \mathcal{U}_x\}$ . Then  $\pi_M$  acts on  $\mathcal{Z}(M)$  freely without fixed points. By definition, we have a continuous map  $\pi : \pi_M \backslash \mathcal{Z}(M) \rightarrow M$  given by  $\pi(\gamma : x \rightarrow y) = y$ , which is a local isomorphism. Since  $\pi^{-1}(x) = \{x\}$ ,  $\pi : \pi_M \backslash \mathcal{Z}(M) \cong M$  is a homeomorphism. Since  $\pi : \mathcal{Z}(M) \rightarrow M$  is local isomorphism, we can regard  $\mathcal{Z}(M)$  as a complex manifold. This space  $\mathcal{Z}(M)$  is called a universal covering space of  $M$ .

We now return to the original setting:  $\mathcal{Z} = \mathcal{Z}(E(\mathbb{C}))$ , and write  $\Pi = \pi(E, \mathbf{0})$ . Since  $E(\mathbb{C})$  is a commutative group, writing its group multiplication additively, we define the sum  $\gamma + \gamma'$  on  $\mathcal{Z}$  by, noting that  $\gamma$  and  $\gamma'$  originate at the origin  $\mathbf{0}$ ,

$$(\gamma + \gamma')(t) = \begin{cases} \gamma(2t) & \text{if } 0 \leq t \leq 1/2 \\ \gamma(1) + \gamma'(2t-1) & \text{if } 1/2 \leq t \leq 1. \end{cases}$$

Then  $(\gamma + \gamma')(1) = \gamma(1) + \gamma'(1)$ , and we claim that  $\gamma + \gamma' \approx \gamma' + \gamma$ . In fact, on the square  $[0, 1] \times [0, 1]$ , we consider the path  $\alpha$  on the boundary connecting the origin  $(0, 0)$  and  $(1, 1)$  passing  $(0, 1)$ , and write  $\beta$  the opposite path from  $(0, 0)$  to  $(1, 1)$  passing  $(1, 0)$ . They are visibly homotopy equivalent. Thus we have a continuous map  $\phi : [0, 1] \times [0, 1] \rightarrow [0, 1] \times [0, 1]$  such that  $\phi(0, t) = \alpha(t)$  and  $\phi(1, t) = \beta(t)$ . Define

$$f : [0, 1] \times [0, 1] \rightarrow E(\mathbb{C}) \text{ by } f(t, t') = \gamma(t) + \gamma'(t').$$

Then it is easy to see  $f \circ \phi(0, t) = (\gamma' + \gamma)(t)$  and  $f \circ \phi(1, t) = (\gamma + \gamma')(t)$ .

By the above addition,  $\mathcal{Z}$  is an additive complex Lie group. Since  $\gamma + \gamma' = \gamma\gamma'$  if  $\gamma \in \Pi$  and  $\gamma' \in \mathcal{Z}$  by definition,  $\Pi$  is an additive subgroup of  $\mathcal{Z}$  and  $\Pi \backslash \mathcal{Z} \cong E(\mathbb{C})$ , where the quotient is made through the group action.

Now we define, choosing a  $C^\infty$ -path  $[\gamma]$  in each class of  $\gamma \in \mathcal{Z}$  and a nowhere vanishing differential form  $\omega$  on  $E$ , a map  $I : \mathcal{Z} \rightarrow \mathbb{C}$  by  $\gamma \mapsto \int_{[\gamma]} \omega \in \mathbb{C}$ . Since  $\omega$  is holomorphic on  $\mathcal{Z}$ , the value of  $I$  is independent of the choice of the representative  $[\gamma]$  by Cauchy's integration theorem. Since  $\omega$  is translation invariant on  $E(\mathbb{C})$ , it is translation invariant on  $\mathcal{Z}$  and  $I(\gamma + \gamma') = I(\gamma) + I(\gamma')$ . In particular,  $I$  is a local homeomorphism because  $E(\mathbb{C})$  is one dimensional and for simply connected  $U$ ,  $\mathcal{Z}(U) \cong I(U)$ . The pair  $(E(\mathbb{C}), \omega)$  is isomorphic locally to the pair of the additive group  $\mathbb{C}$  and  $du$  for the coordinate  $u$  on  $\mathbb{C}$ , because  $du$  is the unique translation invariant differential (up to constant multiple). Since  $I^{-1}([\mathbf{0}]) = \{\mathbf{0}\}$ ,  $I$  is a linear isomorphism into  $\mathbb{C}$ . For an open neighborhood  $U$  of  $\mathbf{0}$  with  $U \cong \mathcal{Z}(U) \ni \gamma \mapsto I(\gamma) = \int_\gamma \omega \in \mathbb{C}$  giving an isomorphism onto a small open disk  $D$  in  $\mathbb{C}$  centered at 0, we have two  $\gamma_1, \gamma_2 \in U$  giving rise to a two  $\mathbb{R}$ -linearly independent  $I(\gamma_j)$  ( $j = 1, 2$ ). Then  $I(m\gamma_1 + n\gamma_2) = mI(\gamma_1) + nI(\gamma_2)$  for all  $m, n \in \mathbb{Z}$ . Replacing  $\gamma_j$  by  $\frac{1}{a}\gamma_j \in \mathcal{Z}(U)$

such that  $I(\frac{1}{a}\gamma_j) = \frac{I(\gamma_j)}{a}$  for any positive integer  $a$ , by the same argument, we find  $I(m\gamma_1 + n\gamma_2) = mI(\gamma_1) + nI(\gamma_2)$  for all  $m, n \in \mathbb{Q}$ ; so,  $I$  is a surjective isomorphism.

This also shows that if  $\alpha : E \rightarrow E$  is an endomorphism of  $E$  with  $\alpha(\mathbf{0}_E) = \mathbf{0}_E$ ,  $\alpha$  lifts an endomorphism of  $\mathcal{Z}$  sending a path  $\gamma$  from  $\mathbf{0}_E$  to  $z \in \mathbb{C}$  to a path  $\alpha(\gamma)$  from  $\alpha(\mathbf{0}_E) = \mathbf{0}_E$  to  $\alpha(z)$ . In particular,  $\alpha(\gamma + \gamma') = \alpha(\gamma) + \alpha(\gamma')$ . Thus  $\alpha$  induces a linear map from  $\mathbb{C} = \mathcal{Z}$  to  $\mathbb{C}$ . Since  $\alpha$  is holomorphic (as it is a polynomial map of the coordinates of  $\mathbf{P}_{/\mathbb{C}}^2$ ),  $\alpha$  is a  $\mathbb{C}$ -linear map. We thus get a natural inclusion:

$$(3.1) \quad \text{End}(E_{/\mathbb{C}}) \hookrightarrow \mathbb{C}.$$

Writing  $L = L_E$  for  $I(\Pi)$ , we can find a base  $w_1, w_2$  of  $L$  over  $\mathbb{Z}$ . Thus we have a map

$$\mathcal{P}(\mathbb{C}) \ni (E, \omega) \longmapsto L_E \in \{L|L : \text{lattice in } \mathbb{C}\} = \text{Lat},$$

and we have  $(E(\mathbb{C}), \omega) \cong (\mathbb{C}/L_E, du)$ . Therefore the map:  $\mathcal{P}(\mathbb{C}) \rightarrow \text{Lat}$  is injective. We show its surjectivity in the next subsection.

By the above fact combined with (3.1), we get

**Proposition 3.3.** *We have a ring embedding  $\text{End}(E_{/\mathbb{C}}) \hookrightarrow \{u \in \mathbb{C} | u \cdot L_E \subset L_E\}$ , and hence  $\text{End}(E_{/\mathbb{C}})$  is either  $\mathbb{Z}$  or an order of an imaginary quadratic field.*

*Proof.* The first assertion follows from (3.1). Pick  $\alpha \in \text{End}(E_{/\mathbb{C}})$  corresponding  $u \in \mathbb{C}$  as above. Note that  $L_E = \mathbb{Z}w_1 + \mathbb{Z}w_2$ . Then  $uw_1 = aw_1 + bw_2$  and  $uw_2 = cw_1 + dw_2$  for integers  $a, b, c, d$ . In short, writing  $w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$  and  $\rho(\alpha) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ , we get  $uw = \rho(\alpha)w$ ; so,  $\rho : \text{End}(E_{/\mathbb{C}}) \rightarrow M_2(\mathbb{Z})$  is a ring homomorphism. By the first assertion, the image has to be an order of imaginary quadratic field or just  $\mathbb{Z}$ .  $\square$

When  $\text{End}(E_{/\mathbb{C}}) \neq \mathbb{Z}$ ,  $E$  is said to have *complex multiplication*.

**3.3. Classical Weierstrass  $\wp$ -function.** For a given  $L \in \text{Lat}$ , we define the Weierstrass  $\wp$ -functions by

$$\begin{aligned} x_L(u) = \wp(u) &= \frac{1}{u^2} + \sum_{\ell \in L - \{0\}} \left\{ \frac{1}{(u - \ell)^2} - \frac{1}{\ell^2} \right\} = \frac{1}{u^2} + \frac{g_2}{20}u^2 + \frac{g_3}{28}u^4 + \dots \\ y_L(u) = \wp'(u) &= -\frac{2}{u^3} - 2 \sum_{\ell \in L - \{0\}} \frac{1}{(u - \ell)^3} = -2u^{-3} + \dots, \end{aligned}$$

where

$$g_2 = g_2(L) = 60 \sum_{\ell \in L - \{0\}} \frac{1}{\ell^4} \quad \text{and} \quad g_3 = g_3(L) = 140 \sum_{\ell \in L - \{0\}} \frac{1}{\ell^6}.$$

Then  $\varphi = y_L^2 - 4x_L^3 + g_2x_L + g_3$  is holomorphic everywhere. Since these functions factors through the compact space  $\mathbb{C}/L$ ,  $\varphi$  has to be constant, because any non-constant holomorphic function is an open map (the existence of power series expansion and the implicit function theorem). Since  $x_L$  and  $y_L$  do not have constant terms, we conclude  $\varphi = 0$ .

We have obtained a holomorphic map  $(x_L, y_L) : \mathbb{C}/L - \{0\} \rightarrow \mathbf{A}_{/\mathbb{C}}^2$ . Looking at the order of poles at  $\mathbf{0}$ , we know the above map is of degree 1, that is, an isomorphism onto its image and extends to

$$\Phi = (x_L : y_L : 1) = (u^3x_L : u^3y_L : u^3) : \mathbb{C}/L \rightarrow \mathbf{P}_{/\mathbb{C}}^2.$$

Thus we have an elliptic curve  $E_L = \Phi(\mathbb{C}/L) = E(g_2(L), g_3(L))$ . We then have

$$\omega_L = \frac{dx_L}{y_L} = du.$$

This shows

**Theorem 3.4.** (Weierstrass) *We have  $[(E, \omega)_{/\mathbb{C}}] \cong Lat$ .*

We would like to make the space  $Lat$  a little more explicit. We see easily that  $w_1, w_2 \in (\mathbb{C}^\times)^2$  span a lattice if and only if  $\text{Im}(w_1/w_2) \neq 0$ . Let  $\mathfrak{H} = \{z \in \mathbb{C} \mid \text{Im}(z) > 0\}$ . By changing the order of  $w_1$  and  $w_2$  without affecting their lattice, we may assume that  $\text{Im}(w_1/w_2) > 0$ . Thus we have a natural isomorphism of complex manifolds:

$$\mathcal{B} = \left\{ v = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \in (\mathbb{C}^\times)^2 \mid \text{Im}(w_1/w_2) > 0 \right\} \cong \mathbb{C}^\times \times \mathfrak{H} \quad \text{via} \quad \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \mapsto (w_2, w_1/w_2).$$

Since  $v$  and  $v'$  span the same lattice  $L$  if and only if  $v' = \alpha v$  for  $\alpha \in SL_2(\mathbb{Z})$ ,

$$Lat \cong SL_2(\mathbb{Z}) \backslash \mathcal{B}.$$

This action of  $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$  on  $\mathcal{B}$  can be interpreted on  $\mathbb{C}^\times \times \mathfrak{H}$  as follows:

$$\alpha(u, z) = (cu + d, \alpha(z)) \quad \text{for} \quad \alpha(z) = \frac{az + b}{cz + d}.$$

By definition,  $\wp(u)$  is an even function. Let  $L = \mathbb{Z}w_1 + \mathbb{Z}w_2$  and put  $w_3 = w_1 + w_2$ . Put  $e_j := \wp(\frac{w_j}{2})$ . Then  $\wp(u) - e_j$  has zero at  $\frac{w_j}{2}$ . Since  $\wp$  is even, the order of zero is even; so,  $\wp'(u)$  has zero at  $\frac{w_j}{2}$ . Therefore,  $e_j$  are roots of  $4x^3 - g_2x - g_3$ . Comparing the zeros and poles of  $\wp'^2$  and  $4(\wp - e_1)(\wp - e_2)(\wp - e_3)$ , we get

$$(3.2) \quad \wp'^2 = 4(\wp - e_1)(\wp - e_2)(\wp - e_3) \quad \text{and} \quad \Delta = 16[(e_1 - e_2)(e_2 - e_3)(e_3 - e_1)]^2.$$

Since  $E(\mathbb{C})$  is smooth  $\Leftrightarrow \Delta \neq 0$ ,  $\Delta$  never vanishes over  $\mathfrak{H}$ .

**3.4. Complex Modular Forms.** We want to write down definitions of modular forms over  $\mathbb{C}$ . We consider  $f \in G_w(\mathbb{C})$ . Writing  $L(v) = L(w_1, w_2)$  for the lattice spanned by  $v \in \mathcal{B}$ , we can regard  $f$  as a holomorphic function on  $\mathcal{B}$  by  $f(v) = f(E_{L(v)}, \omega_{L(v)})$ . Then the conditions (G0–3) can be interpreted as

$$\begin{aligned} \text{(G0)} & \quad f \in \mathbb{C}[g_2(v), g_3(v)]; \\ \text{(G1)} & \quad f(\alpha v) = f(v) \text{ for all } \alpha \in SL_2(\mathbb{Z}); \\ \text{(G2)} & \quad f \in \mathbb{C}[g_2(v), g_3(v), \Delta(v)^{-1}]; \\ \text{(G3)} & \quad f(\lambda v) = \lambda^{-w} f(v) \quad (\lambda \in \mathbb{C}^\times). \end{aligned}$$

We may also regard  $f \in G_w(\mathbb{C})$  as a function on  $\mathfrak{H}$  by  $f(z) = f(v(z))$  for  $v(z) = 2\pi i \begin{pmatrix} z \\ 1 \end{pmatrix}$  ( $z \in \mathfrak{H}$ ). Here multiplying  $\begin{pmatrix} z \\ 1 \end{pmatrix}$  by  $2\pi i$  is to adjust the rationality coming from  $q$ -expansion to the rationality coming from the universal ring  $\mathbb{Q}[g_2, g_3]$ , as we will see later  $(2\pi i)^{-j} g(\begin{pmatrix} z \\ 1 \end{pmatrix})$  has Fourier expansion in  $\mathbb{Q}[[q]]$  for  $q = \exp(2\pi iz)$ . Then we have the following interpretation:

$$\begin{aligned} \text{(G0)} & \quad f \in \mathbb{C}[g_2(z), g_3(z)]; \\ \text{(G1,3)} & \quad f(\alpha(z)) = f(z)(cz + d)^w \text{ for all } \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}); \\ \text{(G2)} & \quad f \in \mathbb{C}[g_2(z), g_3(z), \Delta(z)^{-1}]. \end{aligned}$$



Since  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}(z) = z + 1$ , any  $f \in \mathbb{C}[g_2(z), g_3(z), \Delta^{-1}(z)]$  is translation invariant. Defining  $\mathbf{e}(z) = \exp(2\pi iz)$  for  $i = \sqrt{-1}$ , the function  $\mathbf{e} : \mathbb{C} \rightarrow \mathbb{C}^\times$  induces an analytic isomorphism:  $\mathbb{C}/\mathbb{Z} \cong \mathbb{C}^\times$ . Let  $q = \mathbf{e}(z)$  be the variable on  $\mathbb{C}^\times$ . Since  $f$  is translation invariant,  $f$  can be considered as a function of  $q$ . Thus it has a Laurent expansion  $f(q) = \sum_{n \gg -\infty} a(n, f)q^n$ . We have the following examples (see the following section and [LFE] Chapter 5):

$$(3.3) \quad \begin{aligned} 12g_2 &= 1 + 240 \sum_{n=1}^{\infty} \left\{ \sum_{0 < d|n} d^3 \right\} q^2 \in \mathbb{Z}[[q]]^\times, \\ -6^3 g_3 &= 1 - 504 \sum_{n=1}^{\infty} \left\{ \sum_{0 < d|n} d^5 \right\} q^2 \in \mathbb{Z}[[q]]^\times, \\ \Delta &= q \prod_{n=1}^{\infty} (1 - q^n)^{24} \in q(\mathbb{Z}[[q]]^\times). \end{aligned}$$

This shows that

$$J = \frac{(12g_2)^3}{\Delta} = q^{-1} + \cdots \in q^{-1}(1 + \mathbb{Z}[[q]]).$$

In particular, we may regard  $g_2$  and  $g_3$  as elements of  $\mathbb{Q}[[q]]$ .

We consider a projective plane curve  $E_{\infty/\mathbb{Z}[[q]]}$  called the Tate curve defined over the power series ring  $\mathbb{Z}[[q]]$  by the equation  $Y^2Z = 4X^3 - g_2(q)XZ^2 - g_3(q)Z^3$  and define  $\omega_\infty = \frac{dX}{Y}$ . Since  $\Delta$  is a unit in  $\mathbb{Z}[1/6]((q)) := \mathbb{Z}[1/6][[q]][\frac{1}{q}]$ , we see that  $(E_\infty, \omega_\infty)$  gives an elliptic curve over  $\mathbb{Z}[1/6]((q))$  with nowhere vanishing differential  $\omega_\infty$ . For any  $f \in G_w(A)$ ,  $f(q) = f((E_\infty, \omega_\infty) \otimes_{\mathbb{Q}((q))} A((q))) \in A[[q]]$  is called the  $q$ -expansion of  $f$ . In particular, if  $f \in G_w(\mathbb{C})$ , the  $q$ -expansion  $f(q)$  coincides with the analytic Fourier expansion via  $q = \mathbf{e}(z)$ , because  $f$  is an isobaric polynomial in  $g_2$  and  $g_3$  and by definition  $g_2(q)$  and  $g_3(q)$  are their analytic expansions.

Write  $\mathbf{P}^1(J)_{/\mathbb{Q}}$  for the projective line over  $\mathbb{Z}[\frac{1}{6}]$  whose coordinate is given by  $J$  (in other words,  $\mathbf{P}^1(J) = D_0 \cup D_1$  over local rings with  $D_1 = \mathbb{A}^1$  defined by the affine ring  $\mathbb{Z}[\frac{1}{6}][J]$ ). Since the coordinate at  $\infty$  of  $\mathbf{P}^1(J)$  can be given by  $J^{-1}$  ( $J^{-1} \in q(1 + q\mathbb{Z}[[q]])$ ), we know that  $\mathbb{Z}[[q]] = \mathbb{Z}[[J^{-1}]]$  and

$$(3.4) \quad \widehat{\mathcal{O}}_{\mathbf{P}^1(J), \infty} \cong \mathbb{Z}[1/6][[q]] \text{ via } q\text{-expansion,}$$

where  $\widehat{\mathcal{O}}_{\mathbf{P}^1(J), \infty}$  is the  $(q)$ -adic completion of the local ring  $\mathcal{O}_{\mathbf{P}^1(J), \infty}$  at  $\infty$ .

Since we have

$$M_1(\mathbb{C}) = \text{Lat}/\mathbb{C}^\times = \mathfrak{H} \times \mathbb{C}^\times / (SL_2(\mathbb{Z}) \times \mathbb{C}^\times) \cong SL_2(\mathbb{Z}) \backslash \mathfrak{H},$$

which is isomorphic to  $\mathbf{P}^1(J) - \{\infty\}$  by  $J$ . Thus we see that (G0) over  $\mathbb{C}$  is equivalent to

(G0')  $f$  is a holomorphic function on  $\mathfrak{H}$  satisfying the automorphic property (G1,3), and its analytic  $q$ -expansion  $f(q)$  is contained in  $\mathbb{C}[[q]]$ .

More generally, for modular forms  $f \in G_w(A)$ , we can interpret (G0) as

(G0'')  $f : \mathcal{P}/A \rightarrow \mathbb{A}_{/A}^1$  is a morphism of functors satisfying the automorphic property (G3) in §3.1, and its algebraic  $q$ -expansion  $f(E_\infty, \omega_\infty)$  is contained in  $A[[q]]$ .

**3.5. Weierstrass  $\zeta$  and  $\sigma$  functions.** Pick a lattice  $L$  of  $\mathbb{C}$ ; so, for  $E(\mathbb{C}) = \mathbb{C}/L$ ,  $L = \pi_1(E(\mathbb{C}))$ , and put  $L' := \{\ell \in L \mid \ell \neq 0\}$ . We define Weierstrass  $\sigma$ -function by the following infinite product:

$$(\sigma) \quad \sigma(u) = \sigma(u; L) = u \prod_{\ell \in L'} \left(1 - \frac{u}{\ell}\right) \exp\left(\frac{u}{\ell} + \frac{1}{2}\left(\frac{u}{\ell}\right)^2\right).$$

Plainly  $\sigma(\lambda u; \lambda L) = \lambda \sigma(u; L)$  (homogeneous of degree 1). Taking the logarithmic derivative of  $\sigma$  formally, we get Weierstrass  $\zeta$ -function given by the following infinite sum:

$$(\zeta) \quad \zeta(u) = \zeta_W(u) = \frac{1}{u} + \sum_{\ell \in L'} \left[ \frac{1}{u - \ell} + \frac{1}{\ell} + \frac{u}{\ell^2} \right],$$

which converges absolutely and locally uniformly outside  $L$  as the denominator of the term is of degree 2 in  $\ell$ . Therefore  $\sigma$  also converges. By definition,

$$\zeta'(u) = \frac{d\zeta}{du}(u) = -\wp(u) \quad \text{and} \quad \zeta(\lambda u; \lambda L) = \frac{1}{\lambda} \zeta(u; L) \quad (\text{homogeneous of degree } -1).$$

Since  $\wp$ -function is periodic, we find  $\zeta(u + \ell) = \zeta(u) + \eta(\ell)$  for a linear map  $\eta = \eta_W : L \rightarrow \mathbb{C}$ . If  $L = \mathbb{Z}w_1 + \mathbb{Z}w_2$  with  $\text{Im}(w_1/w_2) > 0$ , we define  $\eta_j := \eta(w_j)$  and extend  $\eta$  to a  $\mathbb{R}$ -linear map from  $\mathbb{C}$  into  $\mathbb{C}$  by  $\eta(a_1w_1 + a_2w_2) := a_1\eta_1 + a_2\eta_2$ . Also we can easily check

$$\sigma(-u) = -\sigma(u) \quad \text{and} \quad \zeta(-u) = -\zeta(u).$$

**Proposition 3.5** (Legendre relation). *We have  $\eta_2w_1 - \eta_1w_2 = 2\pi i$ .*

*Proof.* Take the fundamental parallelogram  $P$  with 4 vertices  $\alpha$ ,  $\alpha + w_j$  and  $\alpha + w_1 + w_2$  so that  $L \cap P = \{0\}$ . Write the path connecting  $\alpha, \alpha + w_1$  as  $\gamma$ . Then  $\gamma + w_2$  connect  $\alpha + w_2$  and  $\alpha + w_1 + w_2$ . Similarly, writing  $\delta$  for the path connecting  $\alpha$  to  $\alpha + w_2$ , then  $\delta + w_1$  connects  $\alpha + w_1$  and  $\alpha + w_1 + w_2$ . On the one hand, we have

$$\begin{aligned} \int_{\partial P} \zeta(u) du &= \int_{\gamma} \zeta(u) du - \int_{\gamma} \zeta(u + w_2) du + \int_{\delta} \zeta(u + w_1) du - \int_{\delta} \zeta(u) du \\ &= \int_{\gamma} \zeta(u) du - \int_{\gamma} \zeta(u) du + \eta_2 \int_{\gamma} du + \int_{\delta} \zeta(u) du + \eta_1 \int_{\delta} du - \int_{\delta} \zeta(u) du \\ &= \eta_2 w_1 - \eta_1 w_2 \end{aligned}$$

On the other hand,  $\zeta(u)$  has pole of residue 1 at 0 and no other pole in  $P$ ; so,

$$\int_{\partial P} \zeta(u) du = 2\pi i \cdot \text{Res}_{u=0} \zeta = 2\pi i$$

as desired. □

**Theorem 3.6.** *For  $a \in \mathbb{C}$  not in  $L$ , we have*

$$\wp(u) - \wp(a) = -\frac{\sigma(u+a)\sigma(u-a)}{\sigma^2(u)\sigma^2(a)}.$$

*Proof.* We may assume that  $a$  in the parallelogram  $P$ . The function  $\wp(u) - \wp(a)$  has zeros at  $a$  and  $-a$  (as  $\wp$  is an even function) and has a double pole at 0. The product expansion of  $\sigma$  tells us that the same holds for  $\frac{\sigma(u+a)\sigma(u-a)}{\sigma^2(u)}$ ; so,

$$\wp(u) - \wp(a) = C \frac{\sigma(u+a)\sigma(u-a)}{\sigma^2(u)}$$

for a constant  $C$ . We see easily

$$\lim_{u \rightarrow 0} \sigma^2(u)/u^2 = 1 \quad \text{and} \quad \lim_{u \rightarrow 0} u^2 \wp(u) = 1.$$

So  $C = -1/\sigma^2(a)$ . □

A theta function for  $E$  is an entire function  $\theta$  satisfying the following functional equation”

$$(\theta) \quad \theta(u + \ell) = \theta(z) \exp(2\pi i[l(u, \ell) + c(\ell)]),$$

where  $l$  is  $\mathbb{C}$ -linear in  $u$  and  $\mathbb{R}$ -linear in  $\ell$  and a function  $c : L \rightarrow \mathbb{C}$ . There is a trivial theta function  $\phi(u) = \exp(au^2 + bu)$  as

$$\phi(u + \ell) = \exp(a(u + \ell)^2 + b(u + \ell)) = \phi(u) \exp(2a u \ell + c(\ell))$$

for  $c(\ell) := a\ell^2 + b\ell$ .

**Theorem 3.7.** *The  $\sigma$ -function is a theta function; i.e.,*

$$\sigma(u + \ell) = \psi(\ell) \exp(\eta(\ell)(u + \ell/2)) \sigma(u),$$

$$\text{where } \psi(\ell) = \begin{cases} 1 & \text{if } \ell \in 2L, \\ -1 & \text{if } \ell \notin 2L. \end{cases}$$

*Proof.* We have

$$\frac{d}{du} \log \frac{\sigma(u + \ell)}{\sigma(u)} = \eta(\ell)$$

by definition. By the fundamental theorem of Calculus, we get

$$\log \frac{\sigma(u + \ell)}{\sigma(u)} = \eta(\ell)u + c(\ell).$$

Exponentiating, we get

$$\sigma(u + \ell) = \sigma(u) \exp(\eta(\ell)u + c(\ell)).$$

Here  $c \pmod{2\pi i}$  is well defined.

Define  $\psi(\ell) := \exp(\eta(\ell)u + c(\ell)) / \exp(\eta(\ell)(u + \ell/2))$ , and we compute  $\psi$ . Suppose  $\ell \notin 2L$ . Set  $u = -\ell/2$ . Then  $\sigma(\ell/2) = \psi(\ell)\sigma(-\ell/2)$ ; so,  $\psi(\ell) = -1$  as  $\sigma$  is an odd function and  $\sigma(\ell/2) \neq 0$  by  $\ell \notin 2L$ .

We see

$$\begin{aligned} \psi(2\ell) \exp(\eta(2\ell)(u + \ell)) &= \frac{\sigma(u + 2\ell)}{\sigma(u)} = \frac{\sigma(u + 2\ell)}{\sigma(u + \ell)} \frac{\sigma(u + \ell)}{\sigma(u)} \\ &= \psi(\ell)^2 \exp(\eta(\ell)(u + \frac{3}{2}\ell) + \eta(\ell)(u + \frac{1}{2}\ell)). \end{aligned}$$

In other words, we have  $\psi(\ell) = \psi(\ell/2)^2$ . Iterating this to reach  $\ell/2^n \notin L$  first time, we get  $\psi(\ell) = (-1)^{2n} = 1$ .  $\square$

**3.6. Product  $q$ -expansion.** Hereafter we take  $L = \mathbb{Z} + \mathbb{Z}z$  for  $z \in \mathfrak{H}$ . We write  $(u; L)$  as  $(u, z)$ . Define  $\varphi(z) = \varphi(u, z) := \exp(-\frac{1}{2}\eta_2 u^2) q_u^{1/2} \sigma(u; z)$ , where  $\eta_2 = \eta(1)$  for  $\eta = \eta_W : L \rightarrow \mathbb{C}$  given by  $\eta(\ell) = \zeta(u + \ell) - \zeta(u)$  and  $q_u = \exp(2\pi i u)$ . Then we see

$$(3.5) \quad \varphi(u+1) = \varphi(u) \quad \text{and} \quad \varphi(u+z) = -\frac{1}{q_u} \varphi(u).$$

By Theorem 3.7, we find

$$\begin{aligned} \varphi(u+1) &= \exp\left(-\frac{1}{2}\eta_2(u+1)^2 + \pi i(u+1)\right) \sigma(u+1) \\ &= -\exp\left(-\frac{1}{2}\eta_2(u+1)^2 + \pi i(u+1) + \eta_2\left(u + \frac{1}{2}\right)\right) \sigma(u) = \varphi(u). \end{aligned}$$

Similarly for  $\eta_1 = \eta(z)$

$$\begin{aligned} \varphi(u+z) &= \exp\left(-\frac{1}{2}\eta_2(u+z)^2 + \pi i(u+z)\right) \sigma(u+\tau) \\ &= -\exp\left(-\frac{1}{2}\eta_2(u+z)^2 + \pi i(u+z) + \eta_1(u+z/2)\right) \sigma(u). \end{aligned}$$

By Legendre's relation:  $\eta_2 z - \eta_1 = 2\pi i$ , we can eliminate  $\eta_1$  and get (3.5). Indeed, we have

$$\begin{aligned} &-\frac{1}{2}\eta_2(u+z)^2 + \pi i(u+z) + \eta_1(u+z/2) \\ &= -\frac{1}{2}\eta_2 u^2 - \eta_2 u z - \frac{1}{2}\eta_2 z^2 + \pi i(u+z) + (\eta_2 z - 2\pi i)u + \frac{(\eta_2 z - 2\pi i)z}{2} \\ &= -\frac{1}{2}\eta_2 u^2 + \pi i(u+z) - 2\pi i u - \pi i z = -\frac{1}{2}\eta_2 u^2 + \pi i u - 2\pi i u \end{aligned}$$

as desired.

**Theorem 3.8.** *Let  $q_w = \exp(2\pi i w)$  for  $w = u, z$ . Then we have*

$$\sigma(u, z) = (2\pi i)^{-1} \exp\left(\frac{1}{2}\eta_2 z^2\right) (q_u^{1/2} - q_u^{-1/2}) \prod_{n=1}^{\infty} \frac{(1 - q_z^n q_u)(1 - q_z^n / q_u)}{(1 - q_z^n)^2}.$$

*Proof.* We prove the equivalent form of the product formula of  $\varphi(u, z) = g(u)$ :

$$g(u) := (2\pi i)^{-1} (q_u - 1) \prod_{n=1}^{\infty} \frac{(1 - q_z^n q_u)(1 - q_z^n / q_u)}{(1 - q_z^n)^2}.$$

By definition, it is easy to see  $g(u+1) = g(u)$  and  $g(u+z) = -\frac{1}{q_u} g(u)$  and that  $g$  has exactly the same zeros of order 1 at  $u = 0$  as  $\sigma$  (and hence as  $\varphi$ ). Thus  $\varphi/g$  is an entire function on  $E(\mathbb{C}) = \mathbb{C}/L$ , and hence  $\varphi = C \cdot g$  for a constant  $C$ . Then we compute  $C$  by  $C = \lim_{u \rightarrow 0} \varphi(u)/g(u) = 1$ .  $\square$

**Corollary 3.9.** *We have*

$$\Delta(z) = (2\pi i)^{12} q_z \prod_{n=1}^{\infty} (1 - q_z^n)^{24}.$$

This can be proven by the above theorem because  $\Delta = 16[(e_1 - e_2)(e_2 - e_3)(e_3 - e_1)]^2$  by (3.2) with  $e_j = \wp(\frac{w_j}{2})$  and invoking Theorem 3.6

$$e_i - e_j = \wp(w_i/2) - \wp(w_j/2) = -\frac{\sigma((w_i + w_j)/2)\sigma((w_i - w_j)/2)}{\sigma^2(w_i/2)\sigma^2(w_j/2)}.$$

Define the Dedekind  $\eta$ -function (as a 24-th root of  $\Delta$  by

$$(7) \quad \eta(z) = \eta_D(z) = q_z^{1/24} \prod_{n=1}^{\infty} (1 - q_z^n).$$

**Theorem 3.10** (R. Dedekind). *We have  $\eta_D(z+1) = \eta_D(z)$  and  $\eta_D(-1/z) = \sqrt{-iz}\eta_D(z)$ , where the square root  $\sqrt{-iz}$  has positive value on the imaginary axis in  $\mathfrak{H}$ .*

*Proof.* The first formula follows from the definition. Since  $\Delta$  is on  $\text{SL}_2(\mathbb{Z})$  of weight 12,  $\frac{\eta_D(-1/z)}{\sqrt{z}\eta_D(z)}$  is holomorphic on  $\mathfrak{H}$  with  $\left| \frac{\eta_D(-1/z)}{\sqrt{z}\eta_D(z)} \right| = 1$ . By maximum principle, it must be a constant  $C$ . Putting  $z = i$ , we have  $1 = C\sqrt{i}$ ; so,  $C = \sqrt{-i}$ .  $\square$

**3.7. Klein forms.** We modify the Weierstrass  $\sigma$ -function into the so-called Klein form which is a modular form of weight 1. Let  $L = L(v) = \mathbb{Z}w_1 + \mathbb{Z}w_2$  for  $v = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$  and  $\eta = \eta_W$  be Weierstrass eta function given by  $\eta_W(\ell; L) = \zeta(u + \ell; L) - \zeta(\ell; L)$  for  $\ell \in L \in \text{Lat}$  (not the Dedekind eta function). Extend  $\eta_W : L \rightarrow \mathbb{C}$  linearly to the  $\mathbb{Q}$  span  $\mathbb{Q} \cdot L$ .

**Definition 3.1.** For  $a = (a_1, a_2) \in \mathbb{Q}^2 - \mathbb{Z}^2$  (a row vector), define  $\mathfrak{k}_a(v) = \mathfrak{k}_a(L(v)) := \exp(-\eta_W(a \cdot v; L(v))a \cdot v/2)\sigma(a \cdot v; L(v))$ , where  $a \cdot v = a_1w_1 + a_2w_2$  (matrix product).

Since  $\sigma(\lambda u, \lambda L) = \lambda\sigma(u, L)$  (homogeneous of weight  $-1$ ) and  $\zeta(\lambda u, \lambda L) = \lambda^{-1}\zeta(u, L)$  (homogeneous of weight 1), we have

$$(3.6) \quad \mathfrak{k}_a(\lambda v) = \lambda \mathfrak{k}_a(v).$$

**Theorem 3.11.** *Let  $a \in \frac{1}{N}\mathbb{Z}^2$  but  $a \notin \mathbb{Z}^2$ . Then  $\mathfrak{k}_a$  is a meromorphic modular form of weight  $-1$  on  $\Gamma(2N^2)$  (whose poles concentrated at cusps), and  $\mathfrak{k}_a^{2N}$  is on  $\Gamma(N)$  and if  $N$  is odd,  $\mathfrak{k}_a^N$  is on  $\Gamma(N)$ .*

*Proof.* Since  $\mathfrak{k}_a(v) = \mathfrak{k}_a(L(v)) := \exp(-\eta_W(a \cdot v; L(v))a \cdot v/2)\sigma(a \cdot v; L(v))$ , for  $\alpha \in \text{SL}_2(\mathbb{Z})$ , we have

$$\begin{aligned} \mathfrak{k}_{a\alpha}(v) &= \exp(-\eta_W(a\alpha \cdot v; L(v))a\alpha \cdot v/2)\sigma(a\alpha \cdot v; L(v)) \\ &= \exp(-\eta_W(a \cdot \alpha v; L(v))a \cdot \alpha v/2)\sigma(a \cdot \alpha v; L(v)). \end{aligned}$$

In short

$$(3.7) \quad \mathfrak{k}_{a\alpha}(v) = \mathfrak{k}_a(\alpha v).$$

We are going to show for  $b = (b_1, b_2) \in \mathbb{Z}^2$

$$(3.8) \quad \mathfrak{k}_{a+b}(v) = \varepsilon(a, b)\mathfrak{k}_a(v) \quad \text{for } \varepsilon(a, b) = (-1)^{b_1b_2+b_1+b_2} \mathbf{e}((a_1b_2 - a_2b_1)/2),$$

where  $\mathbf{e}(z) = \exp(2\pi iz)$ . By (3.7) combined with (3.8), as long as  $\varepsilon(a, m\mathbb{Z}^2)^j = 1$  for  $0 < m \in \mathbb{Z}$ , we find that  $\mathfrak{k}_a^j$  is on  $\Gamma(mN)$  as  $\Gamma(mN)$  induces an identity on  $(N^{-1}\mathbb{Z}/m\mathbb{Z})^2$ .

By Theorem 3.7, we know

$$\sigma(u + \ell) = \psi(\ell) \overline{\exp(\eta(\ell)(u + \ell/2))} \sigma(u)$$

with  $\psi(\ell) = \begin{cases} 1 & \text{if } \ell \in 2L, \\ -1 & \text{if } \ell \notin 2L. \end{cases}$  We can then write  $\psi(b_1 w_1 + b_2 w_2) = (-1)^{b_1 b_2 + b_1 + b_2}$  for  $b = (b_1, b_2) \in \mathbb{Z}^2$ . Take  $\ell = b \cdot v \in L(v)$ . Then

$$\begin{aligned} \mathfrak{k}_{a+\ell}(v) &= \exp(-\eta((a+b) \cdot v)(a+b) \cdot v/2) \sigma((a+\ell) \cdot v) \\ &= \psi(\ell) \exp(-\eta((a+b) \cdot v)(a+b) \cdot v/2) \exp(\eta(b \cdot v)(a \cdot v + b \cdot v/2)) \sigma(a \cdot v) \\ &= \psi(\ell) \exp(-\eta((a+b) \cdot v)(a+b) \cdot v/2 + \eta(b \cdot v)(2a+b) \cdot v/2 + \eta(a \cdot v)a \cdot v/2) \mathfrak{k}_a(v) \end{aligned}$$

The inside of the exponential function is

$$\begin{aligned} &-\eta((a+b) \cdot v)(a+b) \cdot v/2 + \eta(b \cdot v)(2a+b) \cdot v/2 + \eta(a \cdot v)a \cdot v/2 \\ &= \frac{1}{2}(-a_2 b_1(\eta_2 w_1 - \eta_1 w_2) + b_2 a_1(\eta_2 w_1 - \eta_1 w_2)) \stackrel{(*)}{=} \pi i(a_1 b_2 - a_2 b_1). \end{aligned}$$

Here the identity at (\*) is by Legendre's relation  $\eta_2 w_1 - \eta_1 w_2 = 2\pi i$ . Thus (3.8) follows.

Thus if  $m = N^2$ , we have  $\mathfrak{k}_{a+\ell}(v) = \psi(\ell) \mathfrak{k}_a(v)$  and hence, if  $m = 2N^2$ ,  $\mathfrak{k}_a$  depends only on  $a \in N^{-1}\mathbb{Z}^2/2N\mathbb{Z}^2$  as desired. Since  $\varepsilon(a, b)$  is a  $2N$ -th root of unity,  $\mathfrak{k}_a^{2N}$  only depends on  $a \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$ ; so, it is on  $\Gamma(N)$  in the stabilizer of  $a \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$ .

To deal with  $\mathfrak{k}_{a'}^N$  for odd  $N$ , we need to be more careful. Let  $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(N)$ . Write  $a' = (r, s)/N$ . Then  $a'\alpha = (\frac{r}{N} + (\frac{a-1}{N}r + \frac{c}{N}s), \frac{s}{N} + (\frac{b}{N}b + \frac{d-1}{N}s))$ . Therefore

$$(3.9) \quad \mathfrak{k}_{a'}(\alpha v) = \mathfrak{k}_{a'\alpha}(v) = \mathfrak{k}_{a'+b'}(v) = \varepsilon_{a'}(\alpha) \mathfrak{k}_{a'}(v),$$

where from (3.8), for  $Na' = (r, s)$ ,

$$\varepsilon_{a'}(\alpha) = -(-1)^{(\frac{a-1}{N}r + \frac{c}{N}s + 1)(\frac{b}{N}r + \frac{d-1}{N}s + 1)} \mathbf{e}((br^2 + (d-a)rs - cs^2)/2N^2)$$

because  $b' = (\frac{a-1}{N}r + \frac{c}{N}s, \frac{b}{N}b + \frac{d-1}{N}s)$ .

We now need to show

$$(3.10) \quad \varepsilon_{a'}(\alpha)^N = 1 \quad \text{if } N \text{ is odd.}$$

Equivalently, as  $N \equiv 1 \pmod{2}$ ,

$$(3.11) \quad ((a-1)r + cs + 1)(br + (d-1)s + 1) + (br^2 + (d-a)rs - cs^2) \equiv 1 \pmod{2}.$$

If  $r, s \in 2\mathbb{Z}$ , it is plain. Suppose  $(r, s) \equiv (1, 0) \pmod{2}$ . Then (3.11) is equivalent to

$$a(b+1) + b \equiv 1 \pmod{2}.$$

If  $b$  is odd, this is plain. By  $ad - bc = 1$ , if  $b$  is even, then  $a$  is odd; so, the result follows. In the same way, we can settle the case where  $(r, s) \equiv (0, 1) \pmod{2}$ .

Now suppose  $(r, s) \equiv (1, 1) \pmod{2}$ . Then (3.11) is equivalent to

$$(a+c)(b+d) + a + b + c + d \equiv 1 \pmod{2}.$$

Again if  $b$  is even, then  $a$  and  $d$  are odd, and hence

$$(a+c)(b+d) + a + b + c + d \equiv (1+c) + (1+c) + 1 \equiv 1 \pmod{2}.$$

We can settle the case where  $c$  even in the same way. If  $bc \equiv 1 \pmod{2}$ , then  $a$  or  $d$  is even. Supposing that  $a$  is even, again we have

$$(a+c)(b+d) + a + b + c + d \equiv 1 + d + 1 + 1 + d \equiv 1 \pmod{2}.$$

We settle the case where  $d$  even in the same way. This finishes the proof.  $\square$

Let  $p \geq 5$  be a prime and put  $N = p^m$  for  $m > 0$ . For  $m : (N^{-1}\mathbb{Z}/\mathbb{Z})^2 - \{(0,0)\} \rightarrow \mathbb{Z}$ , we say  $m$  satisfies  $(Q_N)$  if

$$(Q_N) \quad \sum_{a \neq 0} m(a)a_1^2 \equiv \sum_{a \neq 0} m(a)a_2^2 \equiv \sum_{a \neq 0} m(a)a_1a_2 \equiv 0 \pmod{N^{-1}\mathbb{Z}},$$

which is equivalent to, writing  $a_1 = \frac{r}{N}$ ,  $a_2 = \frac{s}{N}$  and  $m(a) = m(r, s)$ ,

$$(Q'_N) \quad \sum_{(r,s) \neq 0} m(r, s)r^2 \equiv \sum_{(r,s) \neq 0} m(r, s)s^2 \equiv \sum_{(r,s) \neq 0} m(r, s)rs \equiv 0 \pmod{N},$$

**Remark 3.12.** If  $m(a) = 1$  for all  $a$ , then

$$\sum_a m(a)a_1^2 = N^{-2} \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} r^2 = N^{-1} \sum_{r=0}^{N-1} r^2 = \frac{(N-1)(2N-1)}{6} \in \mathbb{Z}.$$

Similarly

$$\sum_a m(a)a_1a_2 = N^{-2} \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} rs = N^{-2} \left( \sum_{r=0}^{N-1} r \right)^2 = N^{-2} \frac{N^2(N-1)^2}{4} \in \mathbb{Z}.$$

Thus a constant function  $m$  satisfies  $(Q_N)$ .

We like to prove

**Theorem 3.13** (D. Kubert). *Let  $m$  is as above. Then  $\mathfrak{k}^m = \prod_{a \neq 0} \mathfrak{k}_a^{m(a)}$  is on  $\Gamma(N)$  if and only if  $m$  satisfies  $(Q_N)$ .*

We give a sketch of the proof done by D. Kubert.

*Proof.* Let  $\mathbf{e}(z) = \exp(2\pi iz)$  and define  $\varepsilon_a(\alpha)$  by  $\mathfrak{k}_{a\alpha}(v) = \varepsilon_a(\alpha)\mathfrak{k}_a$  for  $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(N)$ . By (3.8), we have  $\varepsilon(a, b) = (-1)^{b_1b_2+b_1+b_2} \exp(\pi i(a_1b_2 - a_2b_1))$ . Replacing  $b$  by  $(\frac{a-1}{N}r + \frac{c}{N}s, \frac{b}{N}b + \frac{d-1}{N}s)$  as in (3.9) and writing  $a = (\frac{r}{N}, \frac{s}{N})$ ,

$$\varepsilon_a(\alpha) = -(-1)^{(\frac{a-1}{N}r + \frac{c}{N}s+1)(\frac{b}{N}r + \frac{d-1}{N}s+1)} \mathbf{e}((br^2 + (d-a)rs - cs^2)/2N^2).$$

Then we need to compute  $\varepsilon^m(\alpha) = \prod_a \varepsilon_a(\alpha)^{m(a)}$ . Define, writing  $m(r, s) = m(\frac{r}{N}, \frac{s}{N})$ ,

$$(3.12) \quad E(r, s; a, b, c, d) := m(r, s) \left[ \frac{ab}{N^2} r^2 + \frac{c(d-1) - c}{N^2} s^2 + \left( \frac{b}{N} + \frac{a-1}{N} \right) r \right. \\ \left. + \left( \frac{d-1}{N} + \frac{c}{N} \right) s + \left( \frac{bc}{N^2} + \frac{(a-1)(d-1)}{N^2} \right) rs + \frac{d-a}{N^2} rs \right].$$

In the definition of this formula in [MUN, page 69], the term  $\frac{d-a}{N^2}rs$  is wrongly written as  $\frac{d-a}{N}rs$ . Then we have, for  $Z := \{(r, s) \in \mathbb{Z}^2 \cap [0, N]^2 \mid (r, s) \neq (0, 0)\}$ ,

$$\varepsilon^m(\alpha) = \exp(\pi i \sum_{(r,s) \in Z} E(r, s; a, b, c, d)).$$

From this, taking  $\alpha = \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix}$ , we have  $E(r, s; 1, N, 0, 1) = m(r, s)(\frac{1}{N}r^2 + r)$  and hence

$$\sum_{(r,s)} m(r, s)(rN + r^2) \equiv 0 \pmod{2N}.$$

Similarly, taking  $\alpha = \begin{pmatrix} 1 & 0 \\ N & 1 \end{pmatrix}$ , we conclude

$$\sum_{(r,s)} m(r, s)(sN + s^2) \equiv 0 \pmod{2N}.$$

However from (3.10), we can replace the modulus  $2N$  by  $N$ . This implies the first two identities of  $(Q_N)$ .

As for the third term involving  $rs$ , again we only need to compute modulo  $N$  in place of  $2N$ . We take  $\alpha = \begin{pmatrix} 1-N & N \\ -N & 1+N \end{pmatrix}$ . Since we know the first two identities of  $(Q_N)$ , we can ignore the square terms in  $r, s$  and also terms only having denominator  $N$  of (3.12). Then

$$\begin{aligned} E(r, s; 1-N, N, -N, 1+N) &\equiv m(r, s) \left( \frac{bc}{N^2} + \frac{(a-1)(d-1)}{N^2} + \frac{d-a}{N^2} \right) rs \\ &\equiv m(r, s) \frac{2}{N} rs \pmod{\mathbb{Z}}. \end{aligned}$$

Thus the last identity of  $(Q_N)$  holds if  $\mathfrak{k}^m$  is on  $\Gamma(N)$ .

Conversely, we already know from (3.10) that  $\mathfrak{k}_a$  is on  $\Gamma(N^2)$  if  $N = p^m$  with  $p \geq 5$ . Writing  $\gamma \in \Gamma(N)$  as  $\gamma = 1 + N\gamma_1$  and sending  $\gamma_1$  to  $(\gamma_1 \bmod N) \in M_2(\mathbb{Z}/N\mathbb{Z})$ , we can easily verify that  $i : \Gamma(N)/\Gamma(N^2) \cong \mathfrak{sl}_2(\mathbb{Z}/N\mathbb{Z}) = \{\gamma_1 \in M_2(\mathbb{Z}/N\mathbb{Z}) \mid \text{Tr}(\gamma_1) = 0\}$ . This is an isomorphism of groups regarding  $\mathfrak{sl}_2(\mathbb{Z}/N\mathbb{Z})$  as an additive group. Then it is easy to see that  $i \begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix}$ ,  $i \begin{pmatrix} 1 & 0 \\ N & 1 \end{pmatrix}$  and  $i \begin{pmatrix} 1-N & N \\ -N & 1+N \end{pmatrix}$  generate the right-hand-side. Therefore the relation  $(Q_N)$  is sufficient for  $\mathfrak{k}^m$  is on  $\Gamma(N)$ .  $\square$

#### 4. MODULAR UNITS

The units  $\mathcal{A}_N^\times$  and  $\mathcal{A}_{1,N}^\times$  of affine ring of the modular curves  $Y_1(N)$  and  $Y(N)$  are called modular units. They are holomorphic and does not vanishes on  $\mathfrak{H}$ , and indeed often they have product expansion similar to the  $\Delta$ -function. The compactification  $\mathbf{P}^1(J)$  of  $\mathcal{A}_1 = \mathbb{Q}[J]$  has one cusp  $\infty$ . Since  $\mathcal{A} = \mathcal{A}_{1,N}$  and  $\mathcal{A}_N$  are finite flat over  $Y(1) = \text{Spec}(\mathcal{A}_1)$ , the normalization of  $\mathbf{P}^1(J)$  in  $\mathcal{A}$  gives a projective curve  $X$  finite flat over  $\mathbf{P}^1(J)$ . The curve  $X$  for  $\mathcal{A}_{1,N}$  (resp.  $\mathcal{A}_N$ ) is denoted by  $X_1(N)$  (resp.  $X(N)$ ). Points above  $\infty$  is called cusps of  $X$ . The cuspidal divisor group of  $X$  is defined to be

$$Cl_X := \frac{\{D \in \bigoplus_{P:\text{cusps}} \mathbb{Z}[P] \mid \deg(D) = 0\}}{\{\text{div}(u) \mid u: \text{modular units}\}}.$$



We study this cuspidal divisor group and prove its finiteness in this section and a formula for its order in the following section. Moreover we see it is a cyclic module over a suitable group algebra.

**4.1. Siegel units.** We now introduce a typical modular unit called Siegel units whose order at the infinity cusp is essentially given by the rational values of the second Bernoulli polynomial  $B_2(x) = x^2 - x + \frac{1}{6}$  (so, the values of Hurwitz zeta functions at  $s = -1$ ).

**Definition 4.1.** We define the Siegel unit  $g_a(z)$  for  $a = (a_1, a_2) \in \mathbb{Q}^2 \cap [0, 1]^2$  by

$$g_a(z) = \mathfrak{k}_a(z)\Delta(z)^{1/12},$$

where  $\Delta^{1/12}(z) = 2\pi i q^{1/12} \prod_{n=1}^{\infty} (1 - q^n)^2$  for  $q = \exp(2\pi iz)$ . For a general  $a \in \mathbb{Q}^2$ , taking the fraction part  $\langle a \rangle := (\langle a_1 \rangle, \langle a_2 \rangle)$  we define  $g_a := g_{\langle a \rangle}$ . Here  $x - \langle x \rangle \in \mathbb{Z}$  with  $\langle x \rangle \in [0, 1)$  for  $x \in \mathbb{R}$ .

Note that  $\Delta^{1/12}(z)$  is the square of the Dedekind  $\eta$ -function  $\eta_D$  up to constant.

**Proposition 4.1.** *We have*

$$(4.1) \quad g_a(z) = -q^{B_2(a_1)/2} \exp(2\pi i a_2(a_1 - 1)/2) (1 - \exp(2\pi i a_2) q^{a_1}) \\ \times \prod_{n=1}^{\infty} [(1 - \exp(2\pi i a_2) q^{n+a_1}) (1 - \exp(-2\pi i a_2) q^{n-a_1})],$$

where  $q = \exp(2\pi iz)$ .

*Proof.* This is just a computation out of the product expansion of the  $\sigma$ -function (combined with definition of  $\mathfrak{k}_a$ ) and  $\Delta^{1/12}$ . An important point later is the leading term  $q^{B_2(a_1)/2}$ ; so, we describe the computation for the leading term. Here is the contribution of each of  $\sigma$ ,  $\mathfrak{k}_a$  and  $\Delta^{-1/12}$  to the leading term:

$$\begin{aligned} (\sigma) & \exp\left(\frac{1}{2}\eta_2(a_1 z + a_2)^2\right); \\ (\mathfrak{k}) & \exp\left(-(\eta_1 a_1 + \eta_2 a_2)(a_1 z + a_2)/2\right); \\ (\Delta) & q^{1/12} = \exp\left(\frac{2\pi i z}{12}\right). \end{aligned}$$

The product of the terms  $(\sigma)$ ,  $(\mathfrak{k})$  and  $(\Delta)$  has inside exp the following:

$$\begin{aligned} & \frac{1}{2}\eta_2(a_1 z + a_2)^2 - (\eta_1 a_1 + \eta_2 a_2)(a_1 z + a_2)/2 + \frac{2\pi i z}{12} \\ & = \frac{1}{2}(\eta_1 a_1^2 z - \eta_1 a_1 a_2 - \eta_2 a_2 a_1 z - \eta_2 a_2^2 + \eta_2 a_1^2 z^2 + 2\eta_2 a_1 a_2 z + \eta_2 a_2^2) + \frac{2\pi i z}{12} \\ & = \frac{1}{2}a_1^2(-\eta_1 + \eta_2 z) + \frac{a_1 a_2}{2}(-\eta_1 + \eta_2 z) + \frac{2\pi i z}{12} \stackrel{(*)}{=} 2\pi i \left(\frac{a_1^2}{2} + \frac{a_1 a_2}{2}\right) + \frac{2\pi i z}{12} \\ & = 2\pi i (B_2(a_1)/2 + a_2(a_1 - 1)/2). \end{aligned}$$

The identity  $(*)$  follows from Legendre's relation:  $\eta_2 z - \eta_1 = 2\pi i$ .  $\square$

By the analysis of Klein forms and the fact that  $\Delta \in S_{12}(\mathrm{SL}_2(\mathbb{Z}))$ , we get

**Theorem 4.2.** *Suppose that  $0 \neq a \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2$ . Then  $g_a^{12N}$  is a modular function on  $\Gamma(N)$ , and it does not have poles and zeros on  $\mathfrak{H}$ . Further for  $\alpha \in \mathrm{SL}_2(\mathbb{Z})$ , we have  $g_{a\alpha}^{12N}(z) = g_a^{12N}(\alpha(z))$ ; therefore,  $g_{a\alpha}(z) = \zeta_{a,\alpha} g_a(\alpha(z))$  for  $\zeta_{a,\alpha} \in \mu_{12N}(\overline{\mathbb{Q}})$ .*

Note that  $g_{a\alpha}/g \circ \alpha$  is holomorphic everywhere over  $X(2N^2)(\mathbb{C})$  by Theorem 3.11; so, it is a non-zero constant. Thus the ratio is a  $12N$ -th root of unity.

**Remark 4.3.** *We see that  $g_a^{12N}$  generate the function field  $\mathbb{Q}(Y(N))$  over  $\mathbb{Q}$  as it is generated by  $\Delta^{-1/12}\wp(a \cdot v; L(v))$  for a running over  $(N^{-1}\mathbb{Z}/\mathbb{Z})^2$  (essentially logarithmic derivative of  $g_a$ ) basically by definition. Since its  $q$ -expansion involves  $N$ -th roots of unity,  $\mathbb{Q}(Y(N))$  becomes reducible over  $\mathbb{Q}[\mu_N]$  (see [EFN, Chapter 7] to see Shimura's proof of the fact that each geometrically irreducible component of  $Y(N)$  is defined over  $\mathbb{Q}[\mu_N]$ ).*

**Proposition 4.4.** *We have  $\prod_{Na=0, a \neq 0} g_a^{12N} \in \mathbb{Q}^\times$ , where  $a$  runs over  $N^{-1}\mathbb{Z}^2 \cap [0, 1)^2$ .*

*Proof.* Since the product  $\varepsilon := \prod_{Na=0, a \neq 0} g_a^{12N}$  is invariant under  $\text{Gal}(X(N)/\mathbf{P}^1(J)) = \text{Gal}(\mathbb{Q}(Y(N))/\mathbb{Q}(J))$ , we find that  $\varepsilon \in \mathbb{Q}[J]^\times$ ; so,  $\varepsilon$  is a constant in  $\mathbb{Q}^\times$ .  $\square$

**Remark 4.5.** *It is known that  $\prod_{Na=0, a \neq 0} g_a^{12N} = N^{12N}$  (see [MUN, Chapter 2, §4]).*

**Theorem 4.6.** *Let  $Z_N := \{a \in N^{-1}\mathbb{Z}^2 \cap [0, 1)^2 \mid a \neq (0, 0)\}$  for  $N = p^m$  (with a prime  $p \geq 5$ ), and put  $g^m := \prod_{a \in Z_N} g_a^{m(a)}$  for a function  $m : Z_N \rightarrow \mathbb{Z}$ . Then  $g^m \in \mathcal{A}_N^\times$  if and only if  $\sum_{a \in Z_N} m(a) \equiv 0 \pmod{12}$  and  $(Q_N)$  is satisfied.*

*Proof.* By Theorem 3.13,  $\mathfrak{k}^m$  is on  $\Gamma(N)$  if and only if  $(Q_N)$  is satisfied. Let  $M := \sum_{a \in Z_N} m(a)$ . Thus we need to show  $\Delta^{M/12}$  is on  $\Gamma(N)$  if and only if  $M \equiv 0 \pmod{12}$ . Since  $p \geq 5$ , we need to prove that  $f := \Delta^{1/2}$  is on  $\Gamma(6)$  but not on  $\text{SL}_2(\mathbb{Z})$ . For  $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{SL}_2(\mathbb{Z})$ , we have  $f|\alpha(z) = f(\alpha(z))(cz + d)^{-6} = \chi(\alpha)f(z)$ . Indeed, by  $q$ -expansion, we have  $\chi\left(\begin{smallmatrix} 1 & \\ 0 & 1 \end{smallmatrix}\right) = -1$  and by Theorem 3.10,  $\chi\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = -1$ . Since  $\Delta$  is on  $\text{SL}_2(\mathbb{Z})$  and the above two elements generate  $\text{SL}_2(\mathbb{Z})$ ,  $\chi$  is a character of  $\text{SL}_2(\mathbb{Z})$  of order 2. By the strong approximation theorem (e.g., [LFE, §6.1]), the abelian quotient of  $\text{SL}_2(\mathbb{Z})$  is isomorphic to the abelian quotient of  $\text{SL}_2(\mathbb{Z}/6\mathbb{Z})$ . Therefore  $f$  is on  $\Gamma(6)$  as desired.  $\square$

Actually, for  $\eta = \eta_D$ , we know  $\eta^2 = \Delta^{1/12}$  is on  $\Gamma(12)$ ,  $\eta^4$  is on  $\Gamma(3)$ ,  $\Delta^{1/2}$  is on  $\Gamma(2)$  (see [MUN, §3.5]).

**4.2. Distribution on  $p$ -divisible groups.** We seek to prove that  $\mathcal{A}_N^\times$  for the affine ring  $\mathcal{A}_N$  of  $X = Y(N)$  is generated by Siegel units. This is to show

$$Cl_X := \frac{\{D \in \bigoplus_{P:\text{cusps}} \mathbb{Z}P \mid \deg(D) = 0\}}{\langle \text{div}(g_a) \mid a \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2 \rangle},$$

where the denominator is the  $\mathbb{Z}$ -span of  $\text{div}(g_a)$ . Hereafter we assume that  $N = p^m$  for a prime  $p \geq 5$  for simplicity. See [MUN] for a general theory. To show linear independence of  $\{\text{div}(g_a)\}_{a \in (N^{-1}\mathbb{Z}/\mathbb{Z})^2}$ , we recall a theory of distributions.

Let  $W$  be the unique discrete valuation ring unramified over  $\mathbb{Z}_p$  of rank 2 (the Witt vector ring with coefficients in  $\mathbb{F}_{p^2}$ ). Then the field  $K$  of fractions of  $W$  is the unique unramified quadratic extension of  $\mathbb{Q}_p$  generated by  $(p^2 - 1)$ -st roots of unity. Let  $V = \mathbb{Z}_p$  or  $W$  and put  $Q = V \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$  (the field of fractions of  $V$ ). We say  $\mu : Q/V = \bigcup_n p^{-n}V/V \rightarrow A$  for an abelian group  $A$  with identity is a *distribution* of

weight  $k \geq 0$  if  $\mu : Q/V \rightarrow A$  is a function satisfying

$$p^{k(m-n)} \sum_{p^{m-n}y=x} \mu(y) = \mu(x)$$

for all  $m > n \geq 0$ . If we modify  $\mu$  into  $\mu_{p^n}(a) = p^{kn} \mu(\frac{a}{p^n})$  for  $a \in V/p^n V$ , this is equivalent to the usual distribution relation  $\sum_{a \equiv b \pmod{p^n}} \mu_{p^m}(a) = \mu_{p^n}(b)$  for all  $a \in V/p^m V$  and  $b \in V/p^n V$ .

We restrict  $\mu$  to  $p^{-m}V/V$  and write it as  $\mu|_m$  if we need to indicate its order  $p^m$ . For a distribution  $\mu$ ,  $\text{rank}(\mu|_m) = \dim_{\mathbb{Q}} \text{Im}(\mu|_m) \otimes_{\mathbb{Z}} \mathbb{Q}$ . A distribution  $\mathcal{M} : p^{-m}V/V \rightarrow \mathcal{A}$  is called universal if for any distribution  $\mu : p^{-m}V/V \rightarrow A$ , there exists a unique homomorphism  $\phi : \mathcal{A} \rightarrow A$  such that  $\mu = \phi \circ \mathcal{M}$ . Universal distribution is unique up to isomorphism if exists. Indeed  $\mathcal{A} = \bigoplus_{a \in p^{-m}V/V} \mathbb{Z}(a)/R$  (the free abelian group generated by symbol  $(a)$  indexed by  $a \in p^{-m}V/V$  and  $R$  is submodule generated by  $(x) - p^{k(m-n)} \sum_{p^{m-n}y=x} (y)$  for all possible  $x \in p^{-m}V/V$ ). Plainly  $\mathcal{M}(x) = (x)$  is the universal distribution.

**4.3. Stickelberger distribution.** Let  $C_m := (V/p^m V)^\times$  and consider its group ring  $A[C_m]$  (assuming  $A$  is a ring). For  $a \in C_m$ , we write  $\sigma_a$  for the group element in  $\mathbb{Q}[C_m]$  corresponding to  $a \in C_m$ . We define  $St_\mu = St_{\mu|_m} : p^{-m}V/V \rightarrow A[C_m]$  by  $St_\mu(x) := \sum_{a \in C_m} \mu(ax) \sigma_a^{-1}$ , which is obviously a distribution. Let  $B_n(x)$  be  $n$ -th Bernoulli polynomial; so,

$$\frac{te^{tx}}{e^t - 1} = \sum_{n=1}^{\infty} B_n(x) \frac{x^n}{n!}.$$

By definition,  $B_n(1-x) = (-1)^n B_n(x)$ . For example,  $B_1(x) = x - \frac{1}{2}$  and  $B_2(x) = x^2 - x + \frac{1}{6}, \dots$ . Fix  $0 < n \in \mathbb{Z}$ . Let  $V = \mathbb{Z}_p$  and define  $\beta(x) = p^{n-1} B_n(\langle x \rangle)$  for  $x \in p^{-m}\mathbb{Z}/\mathbb{Z}$  with representative  $0 \leq \langle x \rangle < 1$ . Here is an obvious way of creating a distribution:

**Lemma 4.7.** *Let  $V = \mathbb{Z}_p$ , and fix  $0 < n \in \mathbb{Z}$ . Then  $\beta$  is a distribution with values in  $\mathbb{Q}$ .*

*Proof.* Define Hurwitz zeta function by

$$\zeta(s, x) = \sum_{n=0}^{\infty} \frac{1}{(x+n)^s} \quad (\text{Re}(s) > 1, 0 < x \leq 1).$$

This function can be analytically continued to  $s \in \mathbb{C}$  and holomorphic outside  $s = 1$  and known that  $\zeta(1-n, x) = -\frac{B_n(x)}{n}$  ( $0 < n \in \mathbb{Z}$ ) for the Bernoulli polynomial  $B_n(x)$  (cf. [LFE, §2.3]). By definition,

$$f^{-s} \zeta(s, \frac{a}{f}) = \sum_{\substack{n \equiv a \\ \text{mod } f, n > 0}}^{\infty} \frac{1}{n^s}.$$

Thus

$$\sum_{a \in (p^{-m'}\mathbb{Z}/\mathbb{Z})^\times, p^{m'} - m a = b} p^{-m's} \zeta(s, \frac{a}{p^{m'}}) = p^{-ms} \zeta(s, \frac{b}{p^m}).$$

Thus  $\beta$  is a distribution of weight  $n - 1$ . □

For a function  $f : \mathbb{Z}/p^m\mathbb{Z} \rightarrow \mathbb{C}$ , we define

$$\zeta(s, f) = \sum_{n=1}^{\infty} f(n)n^{-s} = \sum_{a \in p^{-m}\mathbb{Z}/\mathbb{Z}} f(p^m a)p^{-ms}\zeta(s, \langle a \rangle).$$

We have the following functional equation of Hurwitz zeta function (e.g. [LFE, §2.3]):

$$(4.2) \quad \zeta(s, f) = \frac{1}{2\pi i} \left( \frac{2\pi}{p^m} \right)^s \Gamma(1-s) [\zeta(1-s, \widehat{f}^-) e^{\pi i s/2} - \zeta(1-s, \widehat{f}) e^{-\pi i s/2}],$$

where  $\widehat{f}(x) = \widehat{f}(-x)$  and  $\widehat{f}$  is the Fourier transform over  $\mathbb{Z}/p^m\mathbb{Z}$  given by

$$\widehat{f}(a) = \sum_{b \in \mathbb{Z}/p^m\mathbb{Z}} f(b) \exp(-2\pi i \langle p^{-m} ab \rangle).$$

We consider the trace map  $\text{Tr} : W \rightarrow \mathbb{Z}_p$ , which induces a surjection  $\text{Tr} : K/W \rightarrow \mathbb{Q}_p/\mathbb{Z}_p$ , where  $K = W \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$ . Then  $\beta \circ \text{Tr}$  is a distribution defined on  $K/W$ . We then have Stickelberger distribution  $St_{\beta|_m}$  and  $St_{\beta \circ \text{Tr}|_m}$  with values in  $\mathbb{Q}[C_m]$ .

**4.4. Rank of distribution.** For a distribution  $\mu|_m : C_m \rightarrow A$ , we write  $\langle \mu|_m \rangle$  be the submodule of  $A$  generated by the values of  $\mu|_m$ . Then we define  $\text{rank } \mu|_m = \dim_{\mathbb{Q}} \langle \mu|_m \rangle \otimes_{\mathbb{Z}} \mathbb{Q}$ . By distribution relation, the value at an element  $x$  of order  $p^n \neq 0$  is the sum of values at  $y$  with  $py = x$ . Thus we can modify the value at 0 to be zero taking off  $\mu(0) \sum_a \sigma_a$  from  $St_{\mu}$ . This does not affect distribution relation and the new modified distribution will be denoted as  $St_{\mu}^0$ . This modified distribution satisfies  $\deg(St_{\mu|_m}^0) = 0$  for the degree map of the group algebra  $A[C_m]$ . Since By definition,  $B_n(1-x) = (-1)^n B_n(x)$ ; so,  $B_n(\langle x \rangle) = (-1)^n B_n(\langle x \rangle)$ . Thus  $St_{\beta \circ \text{Tr}}^0$  factors through  $C_m/\{\pm 1\}$  if  $n$  is even.

**Theorem 4.8.** *Let  $k$  be the weight of  $\beta$ . For a prime  $p$ , if  $p^m > 3$ , we have*

$$\text{rank } St_{\beta \circ \text{Tr}|_m}^0 = \begin{cases} |(W/p^m W)^\times / \{\pm 1\}| - 1 & \text{if } (-1)^{k+1} = 1, \\ |(W/p^m W)^\times / \{\pm 1\}| & \text{if } (-1)^{k+1} = -1. \end{cases}$$

*Proof.* Write  $St_m = St_{\beta \circ \text{Tr}|_m}$  simply. Since

$$St_m(x) = \sum_{a \in C_m} \beta(\text{Tr}(ax)) \sigma_a^{-1},$$

we have  $St(x) \cdot \sigma_b = \sum_{a \in C_m} \beta(\text{Tr}(ax)) \sigma_{ab}^{-1} = \sum_{a \in C_m} \beta(\text{Tr}(abx)) \sigma_a^{-1} = St_m(bx)$ . Thus the module  $\langle St_m \rangle$  is stable under the multiplication by  $\sigma_b$ ; so, to prove that  $\text{rank } St_m = |C_m| - 1$  or  $|C_m|$ , we need to show the  $\chi$ -eigenspace of  $\langle St_m \rangle \otimes_{\mathbb{Z}} \overline{\mathbb{Q}}$  is non-trivial for all characters  $\chi \neq 1$  of  $C_m$  with  $\chi(-1) = (-1)^n$ . We therefore compute

$$\sum_{a \in p^{-m}W/W} \beta(\langle \text{Tr}(a) \rangle) \chi(p^m a) = \sum_{a \in C_m} \beta(\langle \text{Tr}(a/p^m) \rangle) \chi(a).$$

By distribution relation,  $\langle St_m \rangle \supset \langle St_{m'} \rangle$  for all  $m' < m$ , we only need to do this for primitive characters modulo  $p^m$ . Then

$$\sum_{a \in C_m} \beta(\langle \text{Tr}(a/p^m) \rangle) \chi(\text{Tr}(a)) = \sum_{a \in C_m} p^{-ms} \chi(\text{Tr}(a)) \zeta(s, \langle \text{Tr}(a/p^m) \rangle) |_{s=1-n} = \zeta(1-n, f_{\chi}),$$

where  $f_\chi(a) = \sum_{x \in C_m, \text{Tr}(x)=a} \chi(x)$  for  $a \in \mathbb{Z}/p^m\mathbb{Z}$ . We compute the Fourier transform of  $f_\chi$ :

$$\begin{aligned} \widehat{f}_\chi(a) &= \sum_{b=1}^N f_\chi(b) e^{-2\pi i b \langle a/p^m \rangle} = \sum_{b=1}^N \sum_{x \in C_m, \text{Tr}(x)=b} \chi(x) e^{-2\pi i \langle ba/p^m \rangle} \\ &= \sum_{x \in C_m} \chi(x) e^{-2\pi i \langle \text{Tr}(xa/p^m) \rangle} = G(\chi) \overline{\chi}(-a) \end{aligned}$$

for the Gauss sum  $G(\chi) = \sum_{x \in C_m} \chi(x) e^{2\pi i \langle \text{Tr}(xp^{-m}) \rangle} \neq 0$  (as  $\chi$  is primitive). Thus by the functional equation (4.2), we get from  $-nN^{n-1} \zeta(1-n, \langle a \rangle) = B_n(\langle a \rangle)$

$$\begin{aligned} (4.3) \quad & \sum_{a \in C_m} \beta(\langle \text{Tr}(a/p^m) \rangle) \chi(\text{Tr}(a)) \\ &= \frac{-nN^{n-1}}{2\pi i} \left( \frac{2\pi}{p^m} \right)^s \Gamma(1-s) [\zeta(1-s, \widehat{f}_\chi^-) e^{\pi i s/2} - \zeta(1-s, \widehat{f}_\chi) e^{-\pi i s/2}]_{s=1-n} \\ &= \frac{-nN^{n-1}}{2\pi i} \left( \frac{2\pi}{p^m} \right)^{1-n} \Gamma(n) [e^{\pi i(1-n)/2} - \chi(-1) e^{-\pi i(1-n)/2}] G(\chi) L(n, \overline{\chi}|_Z) \\ &= \frac{-nN^{n-1}}{2\pi i} \left( \frac{2\pi}{p^m} \right)^{1-n} \Gamma(n) [i^{1-n} - \chi(-1)(-i)^{1-n}] G(\chi) L(n, \overline{\chi}|_Z) \\ &= \frac{-nN^{n-1}}{2\pi i} \left( \frac{2\pi}{p^m} \right)^{1-n} \Gamma(n) i^{n-1} [(-1)^{n-1} - \chi(-1)] G(\chi) L(n, \overline{\chi}|_Z) \end{aligned}$$

which does not vanish if and only if  $\chi(-1) = (-1)^n$ . This finishes the proof.  $\square$

4.5. **Cusps of  $X(N)$ .** We prove

**Theorem 4.9.** *Let  $C_m = (W/p^m W)^\times$  and embed  $C_m$  into  $\text{GL}_2(\mathbb{Z}/p^m\mathbb{Z})$  by its action on  $W/p^m W \cong (\mathbb{Z}/p^m\mathbb{Z})^2$ . Then  $C_m/\{\pm 1\}$  acts on the cusp of  $X(p^m)$  freely and transitively.*

*Proof.* Let  $N = p^m$ . The set  $S_N$  of cusps is one to one onto correspondence with  $\Gamma(N) \backslash \mathbf{P}^1(\mathbb{Q})/\{\pm 1\}$ . Since  $Y(N)$  classifies elliptic curves  $E$  with level structure  $\phi : (\mathbb{Z}/N\mathbb{Z})^2 \cong E[N]$ ,  $\text{GL}_2(\mathbb{Z}/N\mathbb{Z})$  acts on  $Y(N)$  and  $X(N)$  and  $X(N)/\text{GL}_2(\mathbb{Z}/N\mathbb{Z}) = \mathbf{P}^1(J)$ . Since  $\{\pm 1\} \subset \text{Aut}(E)$ , the action factors through  $\text{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$ . The action sends the Siegel unit  $g_a^{12N}$  to  $g_{a\alpha}^{12N}$  as in Theorem 4.2, and hence as is shown by Shimura,  $\overline{\alpha} \in \text{GL}_2(\mathbb{Z}/N\mathbb{Z})$  in the image  $\overline{\alpha}$  of  $\alpha \in \text{SL}_2(\mathbb{Z})$  acts on  $\mathbb{Q}(X(N))$  by  $f \mapsto f \circ \alpha$  and on  $\mathbb{Q}(X(N)) \cap \overline{\mathbb{Q}} = \mathbb{Q}(\mu_N)$ ,  $g \in \text{GL}_2(\mathbb{Z}/N\mathbb{Z})$  acts by  $\zeta_N \mapsto \zeta_N^{\det(g)}$ . Also  $\begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}$  acts on  $q$ -expansion of  $\mathbb{Q}(X(N))$  through its coefficients by  $\sigma_a : \zeta_N \mapsto \zeta_N^d$ . This follows from the fact that  $\mathbb{Q}(X(N)) = \mathbb{Q}(g_a^{12N})$  for  $g_a$  in Theorem 4.2.

Writing one geometrically irreducible component of  $X(N)$  as  $X^\circ(N)$ , we therefore have  $\mathbb{Q}(X(N)) = \mathbb{Q}(\mu_N)(X^\circ(N))$  and  $X^\circ(N)$  is geometrically irreducible over  $\mathbb{Q}(\mu_N)$ . Thus  $\text{Gal}(X(N)/\mathbf{P}^1(J)) \cong \text{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$  canonically. Since  $\mathbf{P}^1(J)$  has one cusp  $\infty$ , the cusp of  $\Gamma(N) \backslash \mathfrak{H}$  is one to one onto  $\text{SL}_2(\mathbb{Z})(\infty) \cong \text{SL}_2(\mathbb{Z}) \backslash \mathbf{P}^1(\mathbb{Q}) \cong \Gamma(N) \backslash \text{SL}_2(\mathbb{Z})/\Gamma_\infty$  for

$$\Gamma_\infty = \left\{ \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \mid m \in \mathbb{Z} \right\}.$$

Since  $\Gamma(N)\backslash\mathfrak{H}$  gives rise to a geometrically irreducible component of  $X(N)$  indexed by  $\text{Gal}(\mathbb{Q}(\mu_N)/\mathbb{Q})$ . Thus the stabilizer of  $\infty$  in  $\text{Gal}(X(N)/\mathbf{P}^1(J)) = \text{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$  is given by the image  $M(\mathbb{Z}/N\mathbb{Z})$  of

$$\left\{ \begin{pmatrix} 1 & u \\ 0 & d \end{pmatrix} \mid u \in \mathbb{Z}/N\mathbb{Z}, d \in (\mathbb{Z}/N\mathbb{Z})^\times \right\}.$$

We have

$$S_N = \text{GL}_2(\mathbb{Z}/N\mathbb{Z})/M(\mathbb{Z}/N\mathbb{Z}) \cong \left\{ v := \begin{pmatrix} a \\ c \end{pmatrix} \in (\mathbb{Z}/N\mathbb{Z})^2 \mid \text{order of } v = N \right\} / \{\pm 1\}.$$

Taking a basis  $\tau, 1$  of  $W$  over  $\mathbb{Z}_p$ , we have for  $w = a + b\tau$  with  $a, b \in \mathbb{Z}_p$ ,

$$w(1, \tau) = (w, w\tau) = (1, \tau) \begin{pmatrix} a & * \\ b & * \end{pmatrix}.$$

This shows that  $C_m M(\mathbb{Z}/N\mathbb{Z}) = \text{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\}$  (the Iwasawa decomposition) and  $S_N \cong C_m/\{\pm 1\}$ , and the action of  $C_m$  on the left is transitive, as desired.  $\square$

**4.6. Finiteness of  $Cl_{X(N)}$ .** Note that the parameter at the cusp  $\infty \in X(N)$  is  $q^{1/N}$  (as the parameter at the cusp of  $\mathbf{P}^1(J)$  is  $q$ ). We take a basis of  $W$  of the form  $(\frac{1}{2}, \tau)$  with  $\text{Tr}(\tau) = 0$ . Then  $\text{Tr}(a_1 \frac{1}{2} + a_2 \tau) = a_1$ . In this way, we identify  $W \cong \mathbb{Z}_p^2$  and  $p^{-m}W/W$  with  $(p^{-m}\mathbb{Z}/\mathbb{Z})^2$ . Then we consider Siegel units  $g_a$  for  $a \in p^{-m}W/W$  with  $\text{Tr}(a) = a_1$ . Identify  $S_{p^m}$  with  $C_m/\{\pm 1\}$ . Thus by Proposition 4.1, noting that the parameter at  $\infty$  of  $X(N)$  is given by  $q^{1/N}$ , we find

$$(4.4) \quad \text{div}(g_a) = N \sum_{b \in C_m/\{\pm 1\}} \frac{1}{2} B_2(\langle \text{Tr}(ab) \rangle) \sigma_b^{-1}(\infty),$$

Though this formula is defined for  $0 \neq a \in p^{-m}W/W$ , we can put  $g_0 := 1$  and then the  $a \mapsto \text{div}(g_a) \in \text{Div}^0(X(N))$  is a distribution proportional to  $St := St_{\beta \circ \text{Tr}|_m}^0$ , where  $\text{Div}^0(X(N))$  is the degree 0 divisor group of  $X(N)$ . Write  $\text{Div}^0(S_N) \subset \text{Div}^0(X(N))$  for the subgroup generated by cusps of  $X(N)$ . Thus  $Cl_{X(N)}$  is a surjective image of  $\text{Div}^0(S_N)/\langle \text{div}(g_a^{12p^m})_{a \in p^{-m}W/W} \rangle$ . By Theorem 4.8,  $\text{Div}^0(S_N)/\langle \text{div}(g_a^{12p^m})_{a \in p^{-m}W/W} \rangle$  is a finite group. Therefore we obtain

**Theorem 4.10.** *If  $N = p^m > 3$  for a prime  $p$ , we have  $Cl_{X(N)}$  is finite, and  $\mathcal{A}_N^\times/\mathbb{Q}[\mu_N]^\times$  is non-trivial.*

If  $N \leq 3$ ,  $X(N)$  has genus 0; so,  $\mathcal{A}_N = \mathbb{Q}[\mu_N][\lambda]$  for a modular function  $\lambda$ . This implies  $\mathcal{A}_N = \mathbb{Q}[\mu_N]^\times$  and  $Cl_{X(N)}$  with  $1 \leq N \leq 3$  is trivial.

**4.7. Siegel units generate  $\mathcal{A}_{p^m}^\times$ .** We start with a theorem due to Shimura but we have given a sketch of a proof by using a solution of the moduli problem for the functor  $\wp$ :

**Theorem 4.11.** *If  $f \in \mathbb{Q}(X)$  for  $X = X_0(N)$ ,  $X_1(N)$  and  $X(N)$ , for every  $\sigma \in \text{GL}_2(\mathbb{Z}/N\mathbb{Z}) = \text{Gal}(X(N)/\mathbf{P}^1(J))$ , the  $q$ -expansion of  $f^\sigma$  has bounded denominator; in other word, for some  $0 < M \in \mathbb{Z}$ ,  $Mf^\sigma(q)$  has integral  $q$ -expansion coefficients.*

For a function  $m : \mathbb{Q}^2/\mathbb{Z}^2 \rightarrow \mathbb{Z}$  with finite support, we define  $g(m) := \prod_a g_{\langle a \rangle}^{m(a)}$ . Here  $\langle a \rangle = (\langle a_1 \rangle, \langle a_2 \rangle)$ . Since  $g_{a+b} = cg_a$  with  $c \in \mu_{2N}$  for  $b \in \mathbb{Z}^2$  as long as  $Na \in \mathbb{Z}^2$  (see (3.8)), up to constant, for any choice of representatives  $[a]$  of  $a \in \mathbb{Q}^2/\mathbb{Z}^2$ ,

$g(m)/\prod_a g_{[a]}^{m(a)} \in \mu_{2N}(\mathbb{C})$ . Note that  $\text{End}(\mathbb{Q}^2/\mathbb{Z}^2) = M_2(\widehat{\mathbb{Z}})$  for  $\widehat{\mathbb{Z}} = \varprojlim_N \mathbb{Z}/N\mathbb{Z} = \prod_l \mathbb{Z}_l$ . By Theorems 3.11 and 4.2,  $g_a \bmod$  scalars only depends on  $a \bmod \mathbb{Z}^2$ , we can think of  $m\sigma(a) := m(a\sigma^{-1})$  (modulo scalars). Then  $\sigma \in \text{GL}_2(\widehat{\mathbb{Z}}) = \varprojlim_N \text{GL}_2(\mathbb{Z}/N\mathbb{Z})$  acts (modulo scalars) on  $g(m)$  by  $g(m) \mapsto \sigma g(m) := g(m\sigma)$ .

We write  $g(m) = \sum_{n=n_0}^{\infty} a_n q^{n/N}$  with  $a_{n_0} \neq 0$  and put  $g^*(m) := a_{n_0}^{-1} q^{-n_0} g(m)$  which is determined uniquely for any scalar multiples of  $g(m)$ . Since

$$g_a = q^{(1/2)B_2(\langle a_1 \rangle)} (1 - \exp(2\pi i a_2) q^{a_1}) \prod_{n=1}^{\infty} (1 - \exp(2\pi i a_2) q^{n+a_1}) (1 - \exp(-2\pi i a_2) q^{n-a_1}),$$

we see

$$(4.5) \quad g_a^* = \begin{cases} (1 - e^{2\pi i a_2} q^{a_1}) \prod_{n=1}^{\infty} (1 - e^{2\pi i a_2} q^{n+a_1}) (1 - e^{-2\pi i a_2} q^{n-a_1}) & \text{if } \langle a_1 \rangle \neq 0, \\ \prod_{n=1}^{\infty} (1 - e^{2\pi i a_2} q^{n+a_1}) (1 - e^{-2\pi i a_2} q^{n-a_1}) & \text{if } \langle a_1 \rangle = 0. \end{cases}$$

An immediate feature from this definition is

**Lemma 4.12.** *We have  $g_a^* \circ \alpha = g_{a\alpha}^*$  for  $\alpha \in \text{GL}_2(\widehat{\mathbb{Z}})$  and  $g_a^* \in \mathbb{Z}[\mu_N][[q_N]]$ .*

Since  $\mathfrak{k}_a$  only depends on  $a \in N^{-1}\mathbb{Z}^2/2N\mathbb{Z}^2 \cong \mathbb{Z}/2N^2\mathbb{Z}$ ,  $g_a$  only depends on  $a \in N^{-1}\mathbb{Z}/12N\mathbb{Z}^2$  (basically by (3.8)),  $\text{GL}_2(\widehat{\mathbb{Z}})$  ( $\widehat{\mathbb{Z}} = \prod_l \mathbb{Z}_l$ ) acts on  $g_a$  via its quotient  $\text{GL}_2(\mathbb{Z}/12N^2)$  as  $\text{Aut}(N^{-1}\mathbb{Z}/12N\mathbb{Z}^2) = \text{GL}_2(\mathbb{Z}/12N^2\mathbb{Z})$ .

*Proof.* By (3.8),  $\mathfrak{k}_{a+b} = \varepsilon \mathfrak{k}_a$  and  $\Delta^{1/12} \circ \alpha = \epsilon \Delta^{1/12}$  for roots of unity  $\varepsilon, \epsilon$ . Therefore  $g_a = \mathfrak{k}_a \Delta^{1/12}$  also satisfies  $g_{a\alpha} = \zeta(\alpha) g_a$  for a root of unity  $\zeta(\alpha)$  (dependent on  $\alpha$ ). Thus  $(g_a \circ \alpha)^* = (\zeta(\alpha) g_{a\alpha})^* = g_{a\alpha}^*$ . The integrality of  $g_a^*$  is obvious from the product expansion.  $\square$

Here by (3.8),  $g_a \circ \alpha = \varepsilon g_a$  for a root of unity  $\varepsilon$ ; so, Therefore, for  $q_N = q^{1/N}$  if  $Na \in \mathbb{Z}^2$ , we have  $g_a^* = 1 + q_N f(q_N)$  for  $f(q_N) \in \mathbb{Q}[\mu_N][[q_N]]$ . Thus as a formal power series, for any  $1 < M \in \mathbb{Z}$ , we have a well defined

$$(g_a^*)^{1/M} = \sum_{n=0}^{\infty} \binom{1/M}{n} q_N^n f(q_N)^n \in \mathbb{Q}[\mu_N][[q_N]]$$

with

$$\begin{aligned} \binom{1/M}{n} &= \frac{(1/M)((1/M) - 1)((1/M) - 2) \cdots ((1/M) - n + 1)}{n!} \\ &= \frac{(1 - M)(1 - 2M) \cdots (1 - (n - 1)M)}{M^n n!}. \end{aligned}$$

However the coefficients of  $g_a^{*1/M}$  and hence of  $g^*(m)^{1/M}$  would have growing denominator caused by the denominator of  $1/M$ ; in other words, if  $g^*(m)^{1/M}$  remains integral, we expect that  $m$  is divisible by  $M$ . We explore this expectation. By the above formula, we remark that for a prime  $l$ ,

$$(4.6) \quad g^*(m)^{1/l} \text{ remains in } \mathbb{Q}[\mu_N][[q^{1/N}]] \text{ if } g^*(m) \in \mathbb{Q}[\mu_N][[q^{1/N}]].$$

Fix a prime  $p$  and consider a primitive root of unity  $\zeta_{p^n} := \exp(\frac{2\pi i}{p^n})$ . We need the following lemma:

**Lemma 4.13.** *The ring  $\mathbb{Z}[\zeta_{p^n}]$  is the integer ring of  $\mathbb{Q}[\zeta_{p^n}]$  and the roots of  $\Phi_n(X)$  (i.e., all primitive  $p^n$ -th roots gives rise to a basis of  $\mathbb{Z}[\zeta_{p^n}]$  over  $\mathbb{Z}$ ,  $p$  fully ramifies in  $\mathbb{Z}[\zeta_{p^n}]$ , and  $\zeta_{p^n} - 1$  generates a unique prime ideal of  $\mathbb{Z}[\zeta_{p^n}]$  over  $p$ . Moreover, taking a decomposition  $\mu_{p^n} = \bigsqcup_{j=1}^{p^{n-1}} \zeta_j \mu_p$ ,  $S = \mu_{p^n} - \{\zeta_j\}_j$  is a basis of  $\mathbb{Z}[\zeta_{p^n}]$  over  $\mathbb{Z}$ .*

*Proof.* Note that  $\zeta_p$  satisfies the equation  $\Phi_1(X) = \frac{X^p-1}{X-1} = 1 + X + X^2 + \cdots + X^{p-1} = \prod_{j=1}^{p-1} (X - \zeta_p^j)$ . Since  $\Phi_1(X+1) = \frac{(X+1)^p-1}{X} = \sum_{j=1}^p \binom{p}{j} X^{j-1} = X^{p-1} + pX^{p-2} + \cdots + p$  is an Eisenstein polynomial,  $\Phi_1(X+1)$  is irreducible, and hence  $\Phi_1(X)$  is irreducible. In the same manner,  $\zeta_{p^n}$  is a root of  $\Phi_n(X) = \frac{X^{p^n}-1}{X^{p^{n-1}}-1} = \Phi_1(X^{p^{n-1}})$  is irreducible. Therefore  $\mathbb{Q}(\zeta_{p^n})/\mathbb{Q}$  is a field extension of degree equal to  $\deg(\Phi_n) = p^{n-1}(p-1)$ . In particular, in  $\mathbb{Q}(\zeta_{p^n})/\mathbb{Q}$ ,  $p$  fully ramifies with  $\zeta_{p^n} - 1$  giving the unique prime ideal  $(\zeta_{p^n} - 1)$  over  $p$ ; so,  $(\zeta_{p^n} - 1)^{p^{n-1}(p-1)} = (p)$  in  $\mathbb{Q}[\zeta_{p^n}]$ , and  $\mathbb{Z}_p[\zeta_{p^n} - 1]$  is the  $p$ -adic integer ring of  $\mathbb{Q}_p[\zeta_{p^n}]$ . For all other prime  $l \neq p$ , taking a prime  $\mathfrak{l}$  of  $\mathbb{Q}[\zeta_{p^n}]$  above  $l$ ,  $\bar{\zeta}_{p^n} := (\zeta_{p^n} \bmod \mathfrak{l})$  is a primitive  $p^n$ -th root in a finite field of characteristic  $l$ ; so,  $\bar{\zeta}_{p^n}^j$  are all distinct for  $j = 1, \dots, p^{n-1}(p-1)$ . Thus  $\mathbb{Q}_l[\zeta_{p^n}]$  is unramified at  $l$ , and  $\mathbb{Z}_l[\zeta_{p^n}]$  is an unramified valuation ring over  $\mathbb{Z}_l$ ; so, it is the  $l$ -adic integer ring of  $\mathbb{Q}_l[\zeta_{p^n}]$ . This shows that  $\mathbb{Z}[\zeta_{p^n}] \cong \mathbb{Z}[X]/(\Phi_n(X))$  is the integer ring of  $\mathbb{Q}[\mu_{p^n}]$  (cf. [CRT, §9]).

Since  $\sum_{\zeta \in \mu_p} \zeta = 0$ , therefore  $\sum_{\zeta \in \mu_p} \zeta_j \zeta = 0$ . Since  $\mathbb{Z}[\mu_{p^n}] = \sum_{\zeta \in \mu_{p^n}} \mathbb{Z}\zeta$ , by  $\sum_{\zeta \in \mu_p} \zeta_j \zeta = 0$ , we find  $\mathbb{Z}[\mu_{p^n}] = \sum_{\zeta \in S} \mathbb{Z}\zeta$ . Since  $|S| = \deg \Phi_n(X) = [\mathbb{Q}[\mu_{p^n}] : \mathbb{Q}]$ , the set  $S$  must be a basis of  $\mathbb{Z}[\mu_{p^n}]$ .  $\square$

**Corollary 4.14.** *For a prime  $l$ , if  $l \mid \sum_{\zeta \in \mu_{p^n}} a_\zeta \zeta$  in  $\mathbb{Z}[\mu_{p^n}]$ , for any pair  $\zeta, \zeta'$  with  $\zeta' \zeta^{-1} \in \mu_p$ ,  $l \mid (a_\zeta - a_{\zeta'})$ .*

*Proof.* Decompose  $\mu_{p^n} = \bigsqcup_{j=1}^{p^{n-1}} \zeta_j \mu_p$ . By Lemma 4.13,  $\sum_{j=1}^{p^{n-1}} \sum_{\zeta \in \mu_p \zeta_j - \{\zeta_j\}} b_\zeta \zeta$  is divisible by  $l$  if and only if  $l \mid b_\zeta$  for all  $\zeta \in \mu_{p^n} - \{\zeta_j\}_j$ . Since  $-\zeta_j = \sum_{\zeta \in \mu_p \zeta_j - \{\zeta_j\}} \zeta$ , we have  $\sum_{\zeta \in \mu_{p^n}} a_\zeta \zeta = \sum_j \sum_{\mu_p \zeta_j - \{\zeta_j\}} (a_\zeta - a_{\zeta_j}) \zeta$ , which shows the result.  $\square$

We say

- an element  $m : \mathbb{Q}^2/\mathbb{Z}^2 \rightarrow \mathbb{Z}$  has prime period if for any  $a$  with  $m(a) \neq 0$ , there exists a prime  $p$  such that  $pa \in \mathbb{Z}^2$ ,
- $0 < n \in \mathbb{Z}$  occurs as a denominator of  $m$  if there exists  $a \in \mathbb{Q}^2/\mathbb{Z}^2$  with order  $n$  such that  $m(a) \neq 0$ .

Note that  $g^*(m\sigma/l)$  has  $l$ -integral coefficients for all  $\sigma \in \mathrm{GL}_2(\widehat{\mathbb{Z}})$  by Lemma 4.12

**Lemma 4.15.** *Let  $l$  be a prime number. Assume that  $m$  has prime period. Then there exists an expression  $g^*(m) = cg^*(m')$  with a constant  $c$  such that  $m'$  has values in  $l\mathbb{Z}$  and every denominator occurring in  $m'$  also occur in  $m$ .*

*Proof.* For each prime  $p$ , we put  $g_p^*(m/l) = \prod_{pa=0} (g_a^*)^{m(a)/l}$ . Then by (4.6),  $g_p^*(m/l)$  lives in  $\mathbb{Q}[\mu_p][[q^{1/p}]]$  with the identity  $g_p^*(m/l)(0) = 1$  of its leading term. Therefore, the coefficient in  $q^{1/p}$  of  $g_p^*(m/l)$  and  $g^*(m/l)$  are equal. In the same way, the coefficient in  $q^{1/p}$  of  $g_p^*(m\sigma/l)$  and  $g^*(m\sigma/l)$  are equal for any  $\sigma \in \mathrm{GL}_2(\widehat{\mathbb{Z}})$ .

We claim

If  $a, b$  in  $\mathbb{Q}^2/\mathbb{Z}^2$  has order  $p$ , then  $l \mid m(a) - m(b)$ .



Writing  $\langle a \rangle$  for a subgroup generated by  $a$ , to prove the claim we may assume  $\langle a \rangle \cap \langle b \rangle = \{0\}$  as we can write  $m(a) - m(b) = m(a) - m(a') + m(a') - m(b)$  for a choice  $a' \in (p^{-1}\mathbb{Z}/\mathbb{Z})^2$  with  $\langle a' \rangle \cap \langle a \rangle = \langle a' \rangle \cap \langle b \rangle = \{0\}$ . Thus  $a, b$  is a basis over  $\mathbb{F}_p$  of  $(p^{-1}\mathbb{Z}/\mathbb{Z})^2$ , and we can find  $\sigma \in \mathrm{GL}_2(\mathbb{F}_p)$  so that  $a\sigma = (\frac{1}{p}, 0)$  and  $b\sigma = (\frac{1}{p}, \frac{1}{p})$ . Then the coefficient in  $q^{1/p}$  of  $g_p^*(m\sigma/l)$  is equal to the coefficient in  $q^{1/p}$  of the following product:

$$\prod_c \left[ (1 - \exp(2\pi i c_2) q^{c_1}) \prod_{n=1}^{\infty} (1 - \exp(2\pi i c_2) q^{c_1} q^n) (1 - \exp(-2\pi i c_2) q^{-c_1} q^n) \right]^{m(c)/l},$$

where  $c$  runs over elements of the following form:

$$c = a\sigma + j(b\sigma - a\sigma) \quad (j = 0, \dots, p-1).$$

This is because the coefficient of  $q^{1/p}$  comes from  $a$  with  $a_1 = \frac{1}{p}$ . Therefore the coefficient is equal to

$$-\frac{1}{l} \sum_c m(c) \zeta_c \quad \text{for } \zeta_c = \exp\left(\frac{2\pi i j}{p}\right).$$

Thus by Corollary 4.14,  $l$ -integrality of this sum implies  $l|m(c) - m(c')$  for any two  $c \neq c'$ . We take  $j = 0, 1$  (i.e.,  $c = a\sigma$  and  $c' = a\sigma + b\sigma - a\sigma = b\sigma$ ), and we get  $l|m(a\sigma) - m(b\sigma)$  as claimed.

We now finish the proof. Take  $a_0$  of order  $p$ . Then we have

$$\prod_{pa=0, a \neq 0} = \prod_{pa=0, a \neq 0} g_a^{m(a)-m(a_0)} \prod_{pa=0, a \neq 0} g_a^{m(a_0)}.$$

By Proposition 4.4, we know  $c := \prod_{pa=0, a \neq 0} g_a^{m(a_0)} \in \mathbb{Q}^\times$ , and by the above claim  $m' = m - m(a_0)$  has values in  $l\mathbb{Z}$ .  $\square$

We now prove

**Theorem 4.16.** *Let  $l, p$  be prime numbers. Assume that  $g^*(m\sigma/l) := \prod_a (g_{a\sigma^{-1}}^*)^{m(a)/l}$  has  $l$ -integral coefficients for all  $\sigma \in \mathrm{GL}_2(\widehat{\mathbb{Z}})$  and that  $m$  is supported on  $(p^{-r}\mathbb{Z}/\mathbb{Z})^2$  for some  $r > 0$ . Then there exists an even function  $m' : \mathbb{Q}^2/\mathbb{Z}^2 \rightarrow l\mathbb{Z}$  supported on  $(p^{-r'}\mathbb{Z}/\mathbb{Z})^2$  for some  $r' > 0$  such that*

$$\prod_a g_a^{m(a)} = c \prod_a g_a^{m'(a)}$$

for a constant  $c \in \mathbb{Q}[\mu_{p^r}]^\times$ .

The left-hand-side resides in  $\mathcal{A}_N^\times$  and the right-hand-side is stable under  $\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ ; so,  $c \in \mathbb{Q}[\mu_{p^r}]^\times$ .

*Proof.* The case of  $r = 1$  follows from Lemma 4.15. We proceed by induction on  $r$ ; so, we may assume that  $m(a) \neq 0$  for some  $a$  of order  $p^r$ . We claim

$$\text{If } p^r a = p^r b = 0 \text{ in } \mathbb{Q}^2/\mathbb{Z}^2 \text{ with } p(a-b) = 0, \text{ then } l|m(a) - m(b).$$

If  $\langle a \rangle = \langle b \rangle$ , choose  $a'$  of order  $p^r$  so that  $\langle a' \rangle \neq \langle a \rangle$ . Then again  $m(a) - m(b) = m(a) - m(a') + m(a') - m(b)$ , and hence, we may assume that  $\langle a \rangle \neq \langle b \rangle$  (so, in particular,  $a - b \neq 0$ ). Choose  $c$  so that  $p^{r-1}c = b - a$ . As  $p(b - a) = 0$  and (we

may assume)  $b - a \neq 0$ ,  $c$  has order  $p^r$ . Since  $a$  and  $a - b$  form a basis of  $(p^{-1}\mathbb{Z}/\mathbb{Z})^2$ , by Nakayama's lemma,  $a$  and  $c$  form a basis of  $(p^{-r}\mathbb{Z}/\mathbb{Z})^2$ . Through base-change via multiplication by  $\sigma$ , we may assume that  $a = \left(\frac{1}{p^r}, 0\right)$  and  $c = \left(0, \frac{1}{p^r}\right)$ . Define  $g_{p^r}^*(m/l) := \prod_{p^r a=0, p^{r-1}a \neq 0} (g_a^*)^{m(a)/l}$ . As in the proof of Lemma 4.15, we look into the coefficient in  $q^{1/p^r}$  of  $g_{p^r}^*(m/l)$  and  $g^*(m/l)$  which are equal. It is given by

$$-\frac{1}{l} \sum_{j=0}^{p^r-1} m \left(\frac{1}{p^r}, \frac{j}{p^r}\right) \zeta_{p^r}^j \quad (\zeta_{p^r} = \exp(\frac{2\pi i}{p^r})).$$

Again by Corollary 4.14, we find  $l|m(a) - m(b)$  as claimed.

Given a coset  $C$  of  $(p^{-1}\mathbb{Z}/\mathbb{Z})^2$  in  $(p^{-r}\mathbb{Z}/\mathbb{Z})^2$ , we see  $p \cdot C = \{c\}$  for a single  $c \in (p^{1-r}\mathbb{Z}/\mathbb{Z})^2$ , and the distribution relation Proposition 4.4 tells us

$$\prod_{x \in C} g_x = \lambda g_c$$

for a constant  $\lambda \in \mathbb{Q}[\mu_{12p^r}]^\times$ . Taking  $c_0 \in C$ , we see

$$\prod_{x \in C} g_x^{m(x)} = \prod_{x \in C} g_x^{m(x)-m(c_0)} \prod_{x \in C} g_x^{m(c_0)} = \lambda^{m(c_0)} g_c^{m(c_0)} \prod_{x \in C} g_x^{m(x)-m(c_0)}.$$

Since  $c$  has order  $\leq p^{r-1}$ , for  $\prod_c g_c^{m(c_0)}$ , induction assumption applies. As for the remaining  $\prod_{x \in C} g_x^{m(x)-m(c_0)}$ , we have proven  $l|m(x) - m(c_0)$ . This finishes the proof.  $\square$

Since we already know that Siegel unit  $\{g_a^{12N}\}_{a \in p^{-m}W/W}$  (for  $N = p^m$ ) generate subgroup of finite index in  $\mathcal{A}_N^\times$ . So for any  $g \in \mathcal{A}_N^\times$ , we find  $0 < M \in \mathbb{Z}$  such that  $g^M = g(m)$  for some  $m : (p^{-m}\mathbb{Z}/\mathbb{Z})^2 - \{(0,0)\} \rightarrow \mathbb{Z}$ . Then by Theorem 4.16, we can remove prime-factor by prime-factor from  $M$  replacing the exponent  $m$ ; so, we obtain

**Corollary 4.17.** *Suppose  $p \geq 5$  and that  $N$  is a  $p$ -power. Then  $g \in \mathcal{A}_N^\times$ , there exists  $m : (N^{-1}\mathbb{Z}/\mathbb{Z})^2 - \{(0,0)\} \rightarrow \mathbb{Z}$  such that  $g = g(m)$  up to constants.*

Thus the next step is to make explicit the set

$$\{m : (N^{-1}\mathbb{Z}/\mathbb{Z})^2 - \{(0,0)\} \rightarrow \mathbb{Z} | g(m) \in \mathcal{A}_N^\times\}.$$

**4.8. Fricke–Wohlfahrt theorem.** We need a theorem due to Fricke–Klein (1980) and its generalization by Wohlfahrt (1963). Let  $\Delta \subset \mathrm{SL}_2(\mathbb{Z})$  be a subgroup of finite index. Such a group  $\Delta$  is called a congruence subgroup if  $\Delta$  contains a principal congruence subgroup

$$\Gamma(N) = \{\alpha \in \mathrm{SL}_2(\mathbb{Z}) | \alpha \equiv 1 \pmod{N \cdot M_2(\mathbb{Z})}\}$$

for a positive integer  $N$ . If  $\Delta$  is a congruence subgroup, the smallest  $N$  such that  $\Delta \supset \Gamma(N)$  is called the arithmetic level of  $\Delta$ . For each cusp  $s$  of general  $\Delta$ , let  $\Delta_s = \{\alpha \in \Delta | \alpha(s) = s\}$ . Taking  $\gamma = \gamma_s \in \mathrm{SL}_2(\mathbb{Z})$  with  $\gamma(s) = \infty$ , we find

$$\gamma\{\pm 1\}\Delta_s\gamma^{-1} \subset \left\{ \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \mid m \in \mathbb{Z} \right\} =: \Gamma_\infty$$

as a subgroup of finite index  $N_s$ . Since the coset  $\Gamma_\infty\gamma_s$  is determined uniquely by  $s$ ,  $N_s$  depends only on  $s$ . We define the geometric level of  $\Delta$  to be the least common multiple of  $N_s$  for  $s$  running over all cusps of  $\Delta$ .

**Theorem 4.18.** *If  $\Delta$  is a congruence subgroup of  $\mathrm{SL}_2(\mathbb{Z})$ , its geometric level and arithmetic level coincide. In other words, writing  $N$  for the geometric level of  $\Delta$ , if  $\Delta \supset \Gamma(N')$  and  $E(N)$  is the minimal normal subgroup of  $\mathrm{SL}_2(\mathbb{Z})$  generated by  $\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix}$ , then  $\Gamma(N) = \Gamma(N')E(N)$ .*

We give a proof due to Wohlfarth. Let  $\Gamma := \Gamma(N')E(N)$ . We define  $M$  to be the order of  $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  in  $\mathrm{SL}_2(\mathbb{Z})/\Gamma$ . Since  $U = jTj^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$  also has order  $M$ . Plainly  $M|N$ . Thus we need to prove that  $\Gamma(M) \subset \Gamma(N)$ , which shows  $M = N$ .

*Proof.* We will define subsets  $E_1$ ,  $E_2$ , and  $E_3$  of  $\Gamma(M)$  whose union equals  $\Gamma(M)$  and show that each  $E_i$  is a subset of  $\Gamma(N)$ . To this end, let  $A \in \Gamma(M)$  and write  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ . First, let  $E_1$  be the collection of elements whose upper right and lower left entries are divisible by  $N$  and suppose  $A \in E_1$ . If we let

$$A_0 = \begin{pmatrix} a & ad-1 \\ 1-ad & d(2-ad) \end{pmatrix}$$

then  $\det(A_0) = 1$ , and since  $ad - bc = 1$ , we have  $ad \equiv 1 \pmod{N}$  and so  $A \equiv A_0 \pmod{N}$ . Equivalently,  $AA_0^{-1} - 1 \in \Gamma(N)$ . It suffices to show  $A_0 \in \Gamma(N)$ . This follows from writing  $V = TU^{-1}T^3U^{-1}T$  and computing  $A_0 = (ST^{d-1}S)^{-1}(VT^{a-1}V^{-1})T^{d-1}$  with  $S = TU^{-1}T$  is an element of  $E(N)$  since  $a \equiv d \equiv 1 \pmod{M}$ .

Now let  $E_2$  be the set of matrices whose diagonal entries are relatively prime to  $N$ . If  $A \in E_2$  then  $\gcd(a, N) = 1$  so  $a$  is a unit modulo  $N$  say with inverse  $a'$ . We can write  $c = M\tilde{c}$  and let  $k = -a'\tilde{c}$ . Then

$$(ST^m S)^{-k} A = \begin{pmatrix} 1 & 0 \\ mk & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c+amk & d+bm k \end{pmatrix}$$

and  $c + amk = m(\tilde{c} - a'a\tilde{c})$  is congruent to 0 modulo  $N$ . Let  $\tilde{d} = d + bm k$ . Taking the determinant of  $(ST^m S)^{-k} A$ , we see that  $\tilde{d}$  is relatively prime to  $N$ . Therefore we can also choose  $\ell$  so that  $b + \tilde{d}m\ell \equiv 0 \pmod{N}$ . In this case,

$$T^{m\ell} (ST^m S)^{-k} A = \begin{pmatrix} 1 & m\ell \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c+amk & \tilde{d} \end{pmatrix} = \begin{pmatrix} * & b+\tilde{d}m\ell \\ c+amk & \tilde{d} \end{pmatrix}$$

and this is an element of  $E_1$ . This says  $A$  is an element of  $E(M)E_1$  and is therefore in  $\Gamma(N)$ . Finally, define  $E_3 = \Gamma(M) \setminus (E_1 \cup E_2)$ . If  $A \in E_3$ , then  $\gcd(a, n) \neq 1$  or  $\gcd(d, n) \neq 1$ . Suppose first that  $\gcd(a, n) \neq 1$ . Since  $ad - bc = 1$ , we know  $\gcd(a, b) = 1$ . Then since  $a \equiv 1 \pmod{M}$ , we also have  $\gcd(a, bM) = 1$ . Recall Dirichlet's theorem on primes in arithmetic progressions: Suppose  $r$  and  $s$  are relatively prime integers. Then  $\{r + st | t \in \mathbb{Z}\}$  contains infinitely many prime numbers. In particular, taking  $r = a$  and  $s = bm$ , we can choose an integer  $k$  so that  $a + bm k$  is a prime number larger than  $N$ . Then

$$A(ST^m S)^{-k} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ mk & 1 \end{pmatrix} = \begin{pmatrix} a+bm k & b \\ c+dm k & d \end{pmatrix}$$

has upper left entry relatively prime to  $N$ . If  $\gcd(d, n) = 1$ , then  $A(ST^m S)^{-k} \in E_2$  and we are done. Otherwise, let  $\tilde{a} = a + bm k$ , let  $\tilde{c} = c + dm k$  and note that  $\tilde{a}d - b\tilde{c} = 1$  so  $\gcd(\tilde{c}, d) = 1$ . As  $\gcd(d, m) = 1$ , just as above we can choose  $\ell$  so that  $d + \tilde{c}m\ell$  is a prime number larger than  $N$ . Then

$$A(ST^m S)^{-k} T^{m\ell} = \begin{pmatrix} \tilde{a} & b \\ \tilde{c} & d \end{pmatrix} \begin{pmatrix} 1 & m\ell \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \tilde{a} & * \\ \tilde{c} & d+\tilde{c}m\ell \end{pmatrix}$$

and  $A(ST^m S)^{-k} T^{m\ell}$  must be in  $E_2$ . Thus  $A$  is an element of  $E_2 E(M)$  and the theorem follows.  $\square$

**4.9. Siegel units and Stickelberger's ideal.** We put

$$\begin{aligned}
 \mathfrak{D} &= \mathfrak{D}_N = \bigoplus_{P:\text{cusp of } X(N)} \mathbb{Z}[P], \\
 \mathfrak{D}^0 &= \mathfrak{D}_N^0 = \{D \in \bigoplus_{P:\text{cusp of } X(N)} \mathbb{Z}[P] \mid \deg(D) = 0\}, \\
 \mathfrak{F} &= \mathfrak{F}_N = \{\text{div}(u) \in \bigoplus_{P:\text{cusp of } X(N)} \mathbb{Z}[P] \mid u \in \mathcal{A}_N^\times\}, \\
 \mathfrak{S}_N &= \{\text{div}(g^m) \mid g^m = \prod_{a \in N^{-1}\mathbb{Z}^2/\mathbb{Z}^2 - \{0\}} g_a^{m(a)} \in \mathcal{A}_N\}, \\
 \mathfrak{S} &= \{\text{div}(g^m) \mid g^m = \prod_{a \in \mathbb{Q}^2/\mathbb{Z}^2 - \{0\}} g_a^{m(a)} \in \mathcal{A}_N\}
 \end{aligned}
 \tag{4.7}$$

Then we have  $\mathfrak{D}^0 \supset \mathfrak{F} \supset \mathfrak{S}$  and  $Cl_{X(N)} = \mathfrak{D}^0/\mathfrak{F}$ , which is a finite abelian group by Theorem 4.10.

**Theorem 4.19** (Kubert–Lang). *We have  $\mathfrak{S} = \mathfrak{S}_N = \mathfrak{F}_N$  if  $N = p^r$  for a prime  $p \geq 5$ .*

When  $N$  is not a prime power as in the theorem, then  $[\mathfrak{F} : \mathfrak{S}]$  is a 2-power (see [MUN, Theorem 5.1.1]).

*Proof.* We know that the  $q$ -expansion coefficients of  $g^m$  lies in  $\mathbb{Q}[\mu_N]$  if  $m$  is supported on  $N^{-1}\mathbb{Z}^2 - \mathbb{Z}^2$ . Since  $g_a \circ \alpha = g_{a\alpha}$  up to non-zero constant by Theorem 4.2, if an integer which is not a factor of  $N$  occurs as a divisor of  $m$ ,  $g^m \circ \alpha \notin \mathbb{C}[[q^{1/N}]]$  by looking into the leading term of  $g^m$ . Therefore to have  $g^m \in \mathcal{A}_N$ , it is necessary to have  $m$  supported on  $N^{-1}\mathbb{Z}^2 - \mathbb{Z}^2$ . Thus  $\mathfrak{S} \subset X_N := \{\text{div}(g^m) \mid m : N^{-1}\mathbb{Z}^2/\mathbb{Z}^2 - \{0\} \rightarrow \mathbb{Z}\}$ . Since  $X_N \subset \mathbb{Q}[\mu_N][[q_{N^2}]]$ , we need to show that  $X_N \cap \mathbb{C}(X(N)) = \mathfrak{S}$ . This follows from Theorem 4.18.  $\square$

By Theorem 4.9, we identify  $R = R_N = \mathbb{Z}[C_m/\{\pm 1\}]$  with  $\mathfrak{D}$  by sending  $\alpha \in C_m/\{\pm 1\}$  to  $[\alpha(\infty)]$ . Recall (4.4):

$$\text{div}(g_a) = N \sum_{b \in C_m/\{\pm 1\}} \frac{1}{2} B_2(\langle \text{Tr}(ab) \rangle) \sigma_b^{-1}.$$

We then put

$$(4.8) \quad \theta = N \sum_{b \in C_m/\{\pm 1\}} \frac{1}{2} B_2(\langle \text{Tr}(b/N) \rangle) \sigma_b^{-1} \in \mathbb{Q} \cdot R_N = \mathbb{Q}[C_m/\{\pm 1\}].$$

Taking  $a = \text{Tr}(1/N)$ , we find  $\text{div}(g_a) = \theta$ ; so,

$$(4.9) \quad \deg(\theta) = \deg(\text{div}(g_a)) = 0.$$

Since  $\text{div}(g_a \circ \alpha) = \text{div}(g_{a\alpha})$  for  $\alpha \in \text{GL}_2(\mathbb{Z})$ , we find  $\mathfrak{S} = R \cap R\theta$  by Fricke–Wohlfahrt theorem, where  $R\theta$  is the fractional ideal inside  $\mathbb{Q}[C_m/\{\pm 1\}]$  generated by  $\theta$ . As in

§4.6, we identify  $W_m = W/p^m W$  with  $(\mathbb{Z}/p^m \mathbb{Z})^2$  by sending  $\frac{1}{2}a_1 + a_2\tau$  to  $a = (a_1, a_2)$ . Note that  $C_m = W_m - pW_m$ ; so, if  $a, b \in W_m$  are generators of  $\mathbb{Z}/p^m \mathbb{Z}$ -module  $W_m$ ,  $a, b, a+b$  and  $a-b$  belong to  $C_m$ , where the sum  $a+b$  and difference  $a-b$  is computed in the additive group  $W_m$ . Let  $I$  be the ideal of  $R_N$  generated by parallelograms:

$$\pi(a, b) = (a + b) + (a - b) - 2(a) - 2(b) \in R_N,$$

where we regard  $a, b \in C_m \subset (\mathbb{Z}/p^m \mathbb{Z})^2$  and  $(a + b), (a - b), (a)$  and  $(b)$  is the image of  $a + b, a - b, a$  and  $b$  in  $C_m/\{\pm 1\}$ .

**Lemma 4.20.** *Let  $N = p^m$  with  $m \geq 1$  and  $B_N$  be the set of pairs  $(a, b)$  in  $\mathbb{Z}^2/N\mathbb{Z}^2$  such that  $a, b, a + b$  and  $a - b$  all have order  $N$ . Define  $I$  by the ideal of  $R_N$  generated by  $\{\pi(a, b)\}_{(a,b) \in B_N}$ . Then  $I$  contains  $\sum_{a \in C_m} m(a/N)[a]$  if and only if  $m : N^{-1}\mathbb{Z}^2/\mathbb{Z}^2 - \{(0, 0)\} \rightarrow \mathbb{Z}$  satisfies  $(Q_N)$ .*

If  $m(a) \neq 0$  for  $a$  of order  $d$  less than  $N$  with  $N = dM$ , by the distribution relation:  $g_a = c \prod_{Mb=a} g_b$  for  $b$ 's of order  $N$ , replacing  $m$  by  $m'$  such that  $m'(b) = m(a)$  for all  $b$  with  $Mb = a$  and  $m'(a) = 0$  without changing any other values, we may assume that  $m(a) \neq 0$  implies that  $a$  has order  $N$  (i.e.,  $m$  is supported on  $C_m$ ).

*Proof.* Multiplying  $N$ , we prove the equivalence between

$$(Q'_N) \quad \sum_{(r,s) \neq 0} m(r/N, s/N)r^2 \equiv \sum_{(r,s) \neq 0} m(r/N, s/N)s^2 \equiv \sum_{(r,s) \neq 0} m(r/N, s/N)rs \equiv 0 \pmod{N},$$

and  $\sum_{(r,s)} m(\frac{r}{N}, \frac{s}{N})[(r, s)] \in I$ . Without taking congruence modulo  $N$  of  $r$  and  $s$ , it is a plain computation to check that  $\pi(a, b)$  for  $a = (r, s)$  and  $b = (r', s')$  satisfies the following condition:

$$(Q) \quad \sum_{(r,s) \neq 0} m(r/N, s/N)r^2 = \sum_{(r,s) \neq 0} m(r/N, s/N)s^2 = \sum_{(r,s) \neq 0} m(r/N, s/N)rs = 0.$$

For example,

$$\text{Left-hand-side of } (Q) \text{ for } \pi(a, b) = (r + r')^2 + (r - r')^2 - 2r^2 - 2r'^2 = 0.$$

Since  $r$  and  $s$  are actually classes modulo  $N$ , the equality in  $(Q)$  becomes the identity  $(Q'_N)$  modulo  $N$ .

Let  $\deg(m) = \sum_{(r,s)} m(r, s)$ . To forget about congruence classes, we argue in the group ring  $\mathbb{Z}[\mathbb{Z}^2]$  regarding  $m = \sum_{(r,s)} m(r/N, s/N)(r, s) \in \mathbb{Z}[\mathbb{Z}^2]$ . In other words, assuming  $(Q)$  for  $m \in \mathbb{Z}[\mathbb{Z}^2]$  and prove that  $m$  is a linear combination of  $\pi(a, b)$  for finitely many  $(a, b)$  as long as  $\deg(m)$  is even. Since we take modulo  $N$  at the end and  $N$  is odd, by multiplying 2 (which is a unit in  $\mathbb{Z}/N\mathbb{Z}$ ), we may assume that  $\deg(m)$  is even. Let  $\mathcal{I}$  be the  $\mathbb{Z}$ -linear span in  $\mathbb{Z}[\mathbb{Z}^2]$  generated by  $\pi(a, b)$  with  $a + b, a - b, a, b$  all having order  $N$  in  $(\mathbb{Z}/N\mathbb{Z})^2$ . Since the image of  $\mathcal{I}$  in  $R_N$  and  $NR_N$  span the ideal of elements satisfying  $(Q'_N)$ , we are going to prove that any  $m : \mathbb{Z}^2 \rightarrow \mathbb{Z}$  satisfying  $(Q)$  is in  $\mathcal{I}$ .

Put  $h(r, s) = |r| + |s|$  and define  $h(m) = \max_{m(r,s) \neq 0} h(r, s)$ . Suppose  $m(r, s) \neq 0$  with maximal  $h(r, s) \geq 3$ . As we remarked,  $a := (r, s)$  has order  $N$  in  $(\mathbb{Z}/N\mathbb{Z})^2$ . Since the argument is the same for each quadrant by rotation of  $\pm 90$  and  $180$  degrees we may assume  $r \geq 0$  and  $s > 0$  with  $r + s \geq 3$ . We now show that we are able to

choose  $(0, 0) \neq (i, j) \in [0, r/2] \times [0, s/2]$  such that  $x = (r - i, s - j)$  and  $y = (i, j)$  have both order  $N$  in  $(\mathbb{Z}/N\mathbb{Z})^2$ . Then  $\pi(x, y) = (x + y) + (x - y) - 2(x) - 2(y) = (a) + (r - 2i, r - 2j) - 2(x) - 2(y)$ , and  $m' := m - m(a)\pi(x, y)$  removes  $a$  from its support and  $h(m') \leq h(m)$ , and repeating this process, eventually we can bring  $m$  modulo  $\mathcal{I}$  to another function supported in  $P := \{(c, d) : |\pm c \pm d| \leq 3\}$ .

We separate our argument in the following four cases:

- (1)  $(r, s) \bmod p \in (\mathbb{Z}/p\mathbb{Z})^2 - \{0, 1\}^2$ ,
- (2)  $(r, s) \equiv (0, 1) \pmod{p}$ ,
- (3)  $(r, s) \equiv (1, 0) \pmod{p}$ ,
- (4)  $(r, s) \equiv (1, 1) \pmod{p}$ .

First suppose  $r \geq 2$  and  $s \geq 2$ , we may choose in Case (1) and (2),  $y = (1, 0)$ ; Case (3)  $y = (0, 1)$  and Case (4)  $y = (2, 0)$  (as  $r > 2$  and  $s > 2$  in Case (4) by  $N \geq 5$ ). If  $r = 1$  (then  $s \geq 2$ ), we are in Case (3), and  $y = (0, 1)$  still works. In the same way, if  $s = 1$ ,  $y = (1, 0)$  works. If  $r = 0$ ,  $s \geq 3$ , we can choose  $y = (0, j)$  with  $j \in \{1, 2\}$  so that  $s - 2j \not\equiv 0 \pmod{p}$ . This is possible as  $p \geq 5$ . If  $s = 0$ , we take  $y = (j, 0)$  with  $j \in \{1, 2\}$  so that  $r - 2j \not\equiv 0 \pmod{p}$ . Thus we are able to choose  $x, y$  as desired, and by induction, we may now assume that  $m$  is supported on  $P_0 := \{(\pm 2, 0), (0, \pm 2), (\pm 1, 0), (0, \pm 1), (\pm 1, \pm 1)\}$ . We can explicitly check that  $m : P_0 \rightarrow \mathbb{Z}$  satisfying  $(Q'_N)$  implies that  $m$  is a linear combination of  $\pi(a, b)$ 's.

To pin down, we follow [MUN, §3.1]. Note that  $(1, 1) + (1, -1) \equiv 2(1, 0) + 2(0, 1) \pmod{\mathcal{I}}$  and  $(2, 0) + (0, 2) \equiv 2(1, 1) + 2(1, -1) \pmod{\mathcal{I}}$ . By this, we are reduced to a linear combination of  $P_1 := \{(1, 0), (0, 1), (1, 1), (2, 0)\}$ . Note that  $8(1, 0) - 2(2, 0) = \alpha - \beta - 2\gamma \in \mathcal{I}$  with  $\alpha = \pi((2, 0), (3, 0)) - \pi((1, 0), (4, 0))$ ,  $\beta = \pi((1, 0), (2, 0))$  and  $\gamma = \pi((1, 0), (3, 0))$ . Suppose  $m : P_1 \rightarrow \mathbb{Z}$  satisfies  $(Q'_N)$ , and write  $x = m(1, 0)$ ,  $y = m(0, 1)$ ,  $z = m(2, 0)$  and  $u = m(1, 1)$ . Then, we have

$$x + 4z + u = 0, y + u = 0, u = 0 \Rightarrow y = u = 0$$

Since  $\deg(m) = x + z$  is even,  $z$  is even. Writing  $z = 2a$  and  $x = -8a$ . Since  $8(1, 0) - 2(2, 0) \in \mathcal{I}$ , we are done.  $\square$

This lemma shows

**Theorem 4.21.** *Let  $I_{12} := \{\alpha \in I \mid \deg(\alpha) \equiv 0 \pmod{12}\}$ . Then we have  $\mathfrak{S}_N = R_N \theta \cap R_N = I_{12} \theta$ .*

4.10. **Cuspidal class number formula.** Let  $R^0 = R_N^0 = \{D \in R_N \mid \deg(D) = 0\}$ . Then we have

**Theorem 4.22.** *Suppose  $N = p^m$  ( $m > 0$ ) with a prime  $p \geq 5$ . Then we have*

$$|Cl_{X(N)}| = [R_N^0 : \mathfrak{S}_N] = \frac{6N^3}{|C_m|} \prod_{\chi \neq 1} \frac{N}{2} B_{2, \chi}$$

for the generalized Bernoulli number  $B_{2, \chi} = \sum_{a \in C_m / \{\pm 1\}} B_2(\langle \frac{\text{Tr}(a)}{N} \rangle)$ , where  $\chi$  runs over all even character of  $C_m$ .

The above  $B_{2, \chi}$  is a non-zero multiple of  $L(-1, \chi)$  (see (4.3)).

We prepares a series of lemmas to prove the formula. Let  $s := \sum_{\sigma \in C_m / \{\pm 1\}} \sigma \in R_N$  and put  $\theta' := \theta - \frac{N}{12}s$ . Then for  $D \in R_N$ ,  $D \cdot s = \deg(D)s$ , and  $\text{div}(\Delta) = N \cdot s$ .

**Lemma 4.23.** *We have  $R_N^0 \cap (I\theta' + R \operatorname{div}(\Delta)) = I_{12}\theta$ .*

*Proof.* Since  $\xi s = \deg(\xi)s$  and  $\operatorname{div}(\Delta) = Ns$ , we find  $R \operatorname{div}(\Delta) = \mathbb{Z} \operatorname{div}(\Delta)$ . For simplicity, write  $G := C_m/\{\pm 1\}$ . Recall  $\deg(\theta) = 0$  by (4.9); so,

$$\deg(\theta') = -\frac{N}{12}|G| = -\frac{N}{12}\deg(s).$$

Thus, for  $\xi \in I$ ,  $\xi\theta = \xi\theta' + \frac{1}{12}\deg(\xi)Ns = \xi\theta' + \frac{1}{12}\deg(\xi)\operatorname{div}(\Delta) \in R^0$ . This implies

$$R_N^0 \cap (I\theta' + R \operatorname{div}(\Delta)) = R_N^0 \cap (I\theta' + \mathbb{Z} \operatorname{div}(\Delta)) \subset I_{12}\theta.$$

The reverse inclusion follows from Theorem 4.21.  $\square$

**Lemma 4.24.** *Let  $R^d := \{D \in R \mid \deg(D) \in d\mathbb{Z}\}$ . Then we have*

$$\deg(R^0 + I\theta' + RNs) = \deg(I\theta' + RNs) = \frac{N|G|}{6}\mathbb{Z};$$

so,  $R^d = R^0 + I\theta' + RNs$  for  $d = \frac{N|G|}{6}$ .

*Proof.* Since  $\theta$  has degree 0 and  $\mathbb{Q} \cdot R^0$  is the augmentation ideal of  $\mathbb{Q}[G]$ ,  $\deg(I\theta) = 0$ . Thus we conclude

$$\deg(I\theta') = \deg(I(\frac{N}{12}s)) = 2\frac{N|G|}{12}\mathbb{Z} = \frac{N|G|}{6}\mathbb{Z}.$$

Here we used the fact that  $\deg(\pi(a, b)) = -2$  and hence  $\deg(I) = 2\mathbb{Z}$ . Since  $|G| = (p^2 - 1)/2 \equiv 0 \pmod{6}$ ,  $\deg(I\theta') \subset \mathbb{Z}$ . Therefore

$$\deg(\deg(I\theta' + RNs)) = \frac{N|G|}{6}\mathbb{Z} + N|G|\mathbb{Z} = \frac{N|G|}{6}\mathbb{Z}$$

as  $p \geq 5$ .  $\square$

Then for  $d = \frac{N|G|}{6}$ , we have inclusions:

$$R \supset R^d \supset I\theta' + RN \cdot s \supset I\theta'.$$

**Lemma 4.25.** *We have  $[R : R^d] = d$  and  $[I\theta' + RN \cdot s : I\theta'] = \frac{|G|}{12}$ .*

*Proof.* By the surjectivity of  $\deg : R \rightarrow \mathbb{Z}$ , we see

$$[R : R^d] = [\deg(R) : \deg(R^d)] = [\mathbb{Z} : d\mathbb{Z}] = d.$$

By the isomorphism theorem, we have  $[I\theta' + RN \cdot s : I\theta'] = [RN \cdot s : RN \cdot s \cap I\theta']$ . We argue via modular forms. The module  $I\theta'$  is the module generated by divisors of  $\mathfrak{k}^m$  for  $m : N^{-1}\mathbb{Z}^2/\mathbb{Z}^2 - \{0\} \rightarrow \mathbb{Z}$  satisfying  $(Q_N)$ . Thus  $D \in I\theta' \cap RN \cdot s = I\theta' \cap R(\operatorname{div}(\Delta))$  means that  $D = \operatorname{div}(\mathfrak{k}^m) = \nu \operatorname{div}(\Delta)$  with  $\nu \in \mathbb{Z}$ . Thus  $\mathfrak{k}^m = C \cdot \Delta^\nu$  with  $0 \neq C \in \mathbb{C}$ . Comparing the weight, we get

$$\nu = -\frac{1}{12} \sum_a m(a).$$

Since  $g_a = \mathfrak{k}_a \Delta^{1/12}$ , we find  $g^m = C$ . This implies that  $m$  is a constant function as  $g^m \circ \alpha = g^{m \circ \alpha}$  for  $\alpha = \operatorname{GL}_2(\mathbb{Z}/N\mathbb{Z})/\{\pm 1\} = \operatorname{Gal}(X(N)/\mathbf{P}^1(J))$  (and  $g_a$  for  $a \in N^{-1}\mathbb{Z}^2/\mathbb{Z}^2 - \{0\}$  are independent; see Proposition 4.4). Since a constant function  $m$  satisfies  $(Q_N)$  (see Remark 3.12),  $m$  is arbitrary, and  $\xi Ns \mapsto \xi \operatorname{div}(\Delta) = Ns \deg(\xi) \in$

$\mathbb{Z}Ns$  sends  $RNs$  (resp.  $RNs \cap I\theta'$ ) isomorphic to  $\mathbb{Z}Ns$  (resp.  $\frac{|G|}{12}\mathbb{Z}Ns$ ). This we conclude  $[RN \cdot s : RN \cdot s \cap I\theta'] = \frac{|G|}{12}$ .  $\square$

For any two lattices  $L$  and  $L'$  of a finite dimensional  $\mathbb{Q}$ -vector space, we define  $[L : L'] = \frac{[L:L \cap L']}{[L':L \cap L']} \in \mathbb{Q}$ , which behaves just like index. For example, we have

$$[L : L'][L' : L''] = [L : L'']$$

and the value equals the index  $[L : L']$  if  $L \supset L'$ . Moreover for a linear transformation  $T$  taking  $L$  onto  $L'$ , we have  $[L : L'] = |\det(T)|$ . Since  $[R^0 : R^0\theta] \neq 0$  (by Theorem 4.10) and  $\deg(\theta') = -\frac{N|G|}{12} \neq 0$ , we find  $[R : R\theta'] \neq 0$ .

**Lemma 4.26.** *We have  $[R_N : R_N\theta'] = \deg(\theta') \prod_{\chi \neq 1} \chi(\theta)$ .*

*Proof.* The index is just the determinant of the multiplication by  $\theta'$  on  $R$ . The determinant can be calculated in  $\mathbb{C}[G]$  which is the product of the eigenvalues on each eigenspaces. The  $\chi$  eigenspace is of diemnsion 1, and hence we get  $\chi(\theta)$  as its eigenvalue if  $\chi \neq 1$  (since  $\theta \in R^0$ ). For the trivial eigenspace,  $R/R^0$ , the eigenvalue is  $\deg(\theta') = \frac{N|G|}{12}$ .  $\square$

**Lemma 4.27.** *We have  $[R : I] = [R\theta' : I\theta'] = N^3$ .*

*Proof.* We may identify  $C_m$  with  $P := \{(r, s) \in (\mathbb{Z}/N\mathbb{Z})^2 \mid (r) + (s) = \mathbb{Z}/N\mathbb{Z}\}$  (having order  $N$ ). Then we claim that  $R/I$  is free of rank 3 with basis  $(0, 1), (1, 0), (1, 1)$ . Indeed, in the proof of Lemma 4.20, we have shown that  $R/I$  is generated by  $P_1 := \{(1, 0), (0, 1), (1, 1), (2, 0)\}$  and  $8(1, 1) \equiv 2(2, 0) \pmod{I}$ ; so, we can remove  $(2, 0)$ .

We need to show that for any given  $(r, s) \in P$ , we have a unique triple  $(x, y, z) \in (\mathbb{Z}/N\mathbb{Z})^3$  such that  $(r, s) + x(1, 0) + y(0, 1) + z(1, 1)$  such that this sum satisfies  $(Q'_N)$ . The condition  $(Q'_N)$  is equivalent to

$$r^2 + x + z = 0, s^2 + y + z = 0 \quad \text{and} \quad rs + z = 0,$$

which can be solved:

$$x = -rs - r^2, y = -rs - s^2 \quad \text{and} \quad z = -rs.$$

$\square$

**Proof of Theorem 4.22:** Recall  $d = \frac{N|G|}{6} = [R : R^d]$ . We have

$$\begin{aligned} |Cl_{X(N)}| &= [R^0 : I_{12}\theta] \stackrel{\text{Lemma 4.23}}{=} [R^0 : R^0 \cap (I\theta' + RN \cdot s)] \\ &\stackrel{\text{Lemma 4.24}}{=} [R^d : (I\theta' + RN \cdot s)] = \frac{[R : I\theta']}{[(I\theta' + RN \cdot s) : I\theta'][R : R^d]} \\ &\stackrel{\text{Lemma 4.25}}{=} \frac{12 \cdot 6}{N|G|^2} [R : R\theta'] [R\theta' : I\theta'] \stackrel{\text{Lemma 4.27}}{=} \frac{12 \cdot 6N^2}{|G|^2} [R : R\theta'] \\ &\stackrel{\text{Lemma 4.26}}{=} \frac{12 \cdot 6N^2}{|G|^2} \deg(\theta') \prod_{\chi \neq 1} \chi(\theta) = \frac{6N^3}{|G|} \prod_{\chi \neq 1} \chi(\theta), \end{aligned}$$

where the last identity follows from  $\deg(\theta') = \frac{N|G|}{12}$ . This finishes the proof as  $\chi(\theta) = \frac{N}{2} B_{2,\chi}$  by definition.



4.11. **Cuspidal class number formula for  $X_1(N)$ .** We give a description of the cuspidal class number formula for  $X_1(N)$  without proof. Let  $N = p^r$  for a prime  $p$ . Let  $X_0(p) = \Gamma_0(p) \backslash (\mathfrak{H} \sqcup \mathbf{P}^1(\mathbb{Q}))$ , where

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid c \equiv 0 \pmod{N} \right\}.$$

It is easy to see

$$\mathrm{SL}_2(\mathbb{Z}) = \Gamma_0(p) \sqcup \bigsqcup_{j=1}^p \Gamma_0(p) \delta \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix}$$

for  $\delta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ . Thus the set of cusps  $S_0(N)$  of  $X_0(N)$  can be explicitly given by

$$S_0(p) = \Gamma_0(p) \backslash \mathrm{SL}_2(\mathbb{Z}) / \Gamma_\infty = \{0, \infty\} = \{\delta(\infty), \infty\},$$

where

$$\Gamma_\infty := \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) \mid \gamma(\infty) = \infty\} = \left\{ \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \mid m \in \mathbb{Z} \right\}.$$

Thus one can classify the set  $S_1(p^r)$  cusps of  $X_1(p^r)$  into three classes:

$$S_1(p^r) = S^\infty \sqcup S_m \sqcup S^0.$$

The classification goes as follows: Since  $\begin{pmatrix} x & b \\ y & d \end{pmatrix} \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} x & xm+b \\ y & ym+d \end{pmatrix}$  for  $\begin{pmatrix} x & b \\ y & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ ,

$$\Gamma_1(p^r) \backslash \mathrm{SL}_2(\mathbb{Z}) / \Gamma_\infty \cong \left\{ \mathbf{x} := \begin{pmatrix} x \\ y \end{pmatrix} \in (\mathbb{Z}/N\mathbb{Z})^2 \mid \mathbf{x} \text{ has order } N \right\} / \sim,$$

where  $\begin{pmatrix} x \\ y \end{pmatrix} \sim \begin{pmatrix} x+my \\ y \end{pmatrix}$  for  $m \in \mathbb{Z}$ . Thus

$$S^0 \cong \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \mid y \in (\mathbb{Z}/N\mathbb{Z})^\times / \{\pm 1\} \right\} \quad \text{and} \quad S^\infty \cong \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in (\mathbb{Z}/N\mathbb{Z})^\times / \{\pm 1\} \right\}.$$

The set  $S^\infty$  is made up of cusps unramified over  $\infty \in X_0(p)$ . The action of  $W := \begin{pmatrix} 0 & -1 \\ p^r & 0 \end{pmatrix}$  induces  $S^0 \cong S^\infty$ . On  $S^?$ , the group  $G_r = (\mathbb{Z}/p^r\mathbb{Z})^\times / \{\pm 1\}$  acts transitively and freely, and hence, writing  $\mathfrak{D}^? := \bigoplus_{s \in S^?} \mathbb{Z}[s]$  and  $\mathfrak{D}_0^? := \{D \in \mathfrak{D}^? \mid \deg(D) = 0\}$ , we have  $\mathbb{Z}[G_r] \cong \mathfrak{D}^?$  by sending  $\sum_g a_g g$  to  $\sum_g a_g g(?)$  for  $? = 0, \infty$ .

Take a model  $X_1(p^r)$  of  $\Gamma_1(p^r) \backslash (\mathfrak{H} \sqcup \mathbf{P}^1(\mathbb{Q}))$  classifying a pair  $(E, \mu_{p^r} \hookrightarrow E)$ , and its dual model  $X_1^*(p^r)$  classifying  $(E, \mathbb{Z}/p^r \hookrightarrow E)$ . Write  $X_1(p^r) - \{\text{cusps}\} = \mathrm{Spec}(\mathcal{A}_{1,N})$  and  $X_1^*(p^r) - \{\text{cusps}\} = \mathrm{Spec}(\mathcal{A}_{1,N}^*)$ . Define

$$(4.10) \quad \begin{aligned} \mathfrak{F}^\infty &:= \{\mathrm{div}(u) \mid u \in \mathcal{A}_{1,N}^?\} \quad \text{with} \quad \mathcal{A}_{1,N}^\infty := \{u \in \mathcal{A}_{1,N}^\times \mid \mathrm{Supp}(\mathrm{div}(u)) \subset S^\infty\} \\ \mathfrak{F}^0 &:= \{\mathrm{div}(u) \mid u \in \mathcal{A}_{1,N}^{*,0}\} \quad \text{with} \quad \mathcal{A}_{1,N}^{*,0} := \{u \in (\mathcal{A}_{1,N}^*)^\times \mid \mathrm{Supp}(\mathrm{div}(u)) \subset S^0\}. \end{aligned}$$

The partial cuspidal group is defined as follows:

$$Cl_{X_1(N)}^\infty := \mathfrak{D}_0^\infty / \mathfrak{F}^? \quad \text{and} \quad Cl_{X_1^*(N)}^0 := \mathfrak{D}_0^0 / \mathfrak{F}^0.$$

By the action of the involution  $W$ , we have  $Cl_{X_1(N)}^\infty \cong Cl_{X_1^*(N)}^0$ . In this case, the Stickelberger element is the traditional one:

$$\theta = N \sum_{b \in G} \frac{1}{2} B_2(\langle b/N \rangle) \sigma_b^{-1},$$

where  $\sigma_b \in \mathrm{Gal}(X_1(p^r)/X_0(p^r))$  associated to  $b \in G$ . Let  $I$  be the ideal of  $R = \mathbb{Z}[G]$  generated by  $\sigma_c - c^2$  and  $I_0 := \{\xi \in I \mid \deg(\xi) = 0\}$ . Then the following lemma is easier than Theorem 4.21 (see [MUN, Lemma 6.1.2]):

**Lemma 4.28.** For  $\theta' = \theta - \frac{p^r}{12}s$  with  $s = \sum_{\sigma \in G} \sigma$ ,  $R\theta' \cap R = I\theta'$  and  $R_0\theta' \cap R = I_0\theta' = I_0\theta$ .

Consider functions  $m : (p^{-r}\mathbb{Z}/\mathbb{Z} \rightarrow p^{1-r}\mathbb{Z}/\mathbb{Z})/\{\pm 1\} \rightarrow \mathbb{Z}$  satisfying

- (1) For every coset  $C$  of  $p^{-1}\mathbb{Z}/\mathbb{Z}$ ,  $\sum_{a \in C} m(a) = 0$ ,
- (2)  $\sum_a m(a)(p^r a)^2 \equiv 0 \pmod{p^r}$ .

It is shown in [MUN, Theorem 3.1] that the units satisfying the above two conditions generates  $\mathcal{A}_{1,N}^{*,0}$  up to scalars. This enable us to get

**Theorem 4.29** (Kubert–Lang). *We have*

$$|Cl_{X_1(N)}^\infty| = |Cl_{X_1^*(N)}^0| = \frac{p^{p^{r-1}}}{p^{2r-2}} \prod_{\chi: G \rightarrow \overline{\mathbb{Q}}^\times, \chi \neq 1} \frac{1}{2} B_{2,\chi}.$$

But we do not know the cyclicity of  $Cl_{X_1(N)}$  over  $\mathbb{Z}[G]$  except for the case where  $N = p$ .

#### REFERENCES

- [ALG] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977.
- [CRT] H. Matsumura, *Commutative Ring Theory*, Cambridge Studies in Advanced Mathematics **8**, Cambridge Univ. Press, New York, 1986.
- [EEK] A. Weil, *Elliptic Functions according to Eisenstein and Kronecker*, Springer, 1976
- [EFN] S. Lang, *Elliptic functions*, second edition, GTM **112**, 1987, Springer.
- [GME] H. Hida, *Geometric Modular Forms and Elliptic Curves*, World Scientific, Singapore, 2000.
- [IAT] G. Shimura, *Introduction to the Arithmetic Theory of Automorphic Functions*, Princeton University Press, Princeton, NJ, and Iwanami Shoten, Tokyo, 1971.
- [LFE] H. Hida, *Elementary Theory of L-Functions and Eisenstein Series*, LMSST **26**, Cambridge University Press, Cambridge, England, 1993.
- [MUN] D. Kubert and S. Lang, *Modular units*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Science], **244**, 1981, Berlin, New York: Springer.