LECTURE NOTES ON ELLIPTIC CURVES

HAN YU

MA426, University of Warwick, 2022-2023 Term 2.
Textbook:
Washington, Elliptic Curves: Number Theory and Cryptography, 2nd Edition.
Suggested knowledge base:
1. complex analysis: Freitag and Busam, Complex Analysis
2. algebraic curves: Fulton, Algebraic curves
3. for those who need proper algebra knowledge: Knapp, Basic Algebra and a second volume named Advanced Algebra.
Warning: Do not trust anything in the notes without thinking! Do your own math work!

Contents

2020 *Mathematics Subject Classification.*

1. Introduction

Elliptic curves are irreducible algebraic curves of degree 3. They are really the third simplest algebraic curves (of course, linear curves or quadratic curves are simpler). We do not assume any knowledge of algebraic geometry in this module. However, it is highly recommended to have some (Fulton). Now, we briefly recall some fundamental notions.

1.1. **Affine curves.** Given a field $\mathbb{K}$. We consider the ring of polynomials $R = \mathbb{K}[X, Y]$. Given an ideal $I \subset \mathbb{K}[X, Y]$. We consider the ring R/I. The Krull dimension of R/I (or the transcendence degree of the fraction field of R/I when I is prime) can be $0, 1, 2$.

Consider the set

$$V(I) = \{(x, y) \in \mathbb{K}^2 : f(x, y) = 0, \forall f \in I\}.$$

If $\mathbb{K}$ is algebraically closed, then by Hilbert's Nullstellensatz we see that $V(I)$ is not empty if and only if I is a proper ideal. Moreover, f vanishes on $V(I)$ if and only if $f \in \sqrt{I}$ (of course, if I is prime then $I = \sqrt{I}$).

Now consider an ideal I (not necessarily prime) with R/I of Krull dimension one. We say that $V(I)$ is an algebraic curve and name it with C. In our situation,

$I = (f)$. The generator f is a polynomial in $\mathbb{K}[X, Y]$ whose degree is the degree of C.

1.2. **Projective curves.** The projective plane over $\mathbb{K}$ is defined to be

$$P^2(\mathbb{K}) = \{\mathbb{K}^3 \setminus \{(0, 0, 0)\}/ \sim,$$

where the equivalence condition is

$$(x, y, z) \sim (x', y', z') \iff \exists \lambda \neq 0, (x, y, z) = \lambda(x', y', z').$$

A polynomial $f \in \mathbb{K}[X, Y, Z]$ is homogeneous if all the monomials of f has the same degree. If f is homogeneous, then zeros of f can be viewed as a well-defined subset of $P^2(\mathbb{K})$.

Given a polynomial $f(X, Y)$ in $\mathbb{K}[X, Y]$, we can always insert Z into each monomials of f and make it a homogeneous polynomial $\tilde{f}$ in $\mathbb{K}[X, Y, Z]$. For example, $y^2 - x^3 - 4x$ can be turned into $y^2z - x^3 - 4xz^2$. Of course, there are more than one way to make this happen. We always want to use as less z as possible. Observe that $\tilde{f}(x, y, 1) = f(x, y)$.

Given an affine curve C, one can perform the above procedure to find a projective curve which is going to be called the projective closure of C.

1.3. **Towards algebraic geometry.** There is nothing to stop us from considering $\mathbb{K}[X_1, X_2, \ldots, X_n]/I$ where I is an ideal in $\mathbb{K}[X_1, X_2, \ldots, X_n]$. One can consider $V(I)$ in $\mathbb{K}^n$ or $P^n(\mathbb{K})$ after we make I homogeneous. Once there are more than two variables and the polynomials in I have a high degree, $V(I)$ can be rather complicated. Many of the methods and results for elliptic curves can be generalized for more complicated algebraic varieties.

1.4. **Reducibility.** Consider $\mathbb{K}[X, Y]/I$ which provides us with a curve C. We have not required that I is a prime ideal. For example, if

$$f(X, Y) = L_1(X, Y)L_2(X, Y)L_3(X, Y)$$

for linear forms L_1, L_2, L_3, then the curve C is actually a union of three (possibly multiple) lines. Those three lines are components of C. If I is prime, then C has only one component. In this case, we say that C is irreducible.

This discussion can be easily extended for general algebraic varieties.

1.5. **Singularity.** Let C be an irreducible curve. Let $f(X,Y)$ be the defining irreducible polynomial. We can form the (symbolic) gradient ∇f (since f is polynomial, we can just perform the derivatives as if we are dealing with functions with real-valued variables).

A point P on C is singular if $\nabla f(P) = 0$. Otherwise, we say that P is a regular point on C. If all points on C are regular, we say that C is regular.

1.6. **What can be $\mathbb{K}$?** For many reasons, $\mathbb{C}$ is our first choice for $\mathbb{K}$. In this case, we can study elliptic curves as Riemann surfaces on which one has rich analytic structures.

For number theory, one is often interested in $\mathbb{Q}$ or finite fields $\mathbb{F}_q$. One can as well consider finite extensions of $\mathbb{Q}$ or F_q. In those cases, the base field is not algebraically closed. This creates many difficulties. Overcoming those difficulties generates huge rewards.

The study of elliptic curves over finite fields receives benefits from advances in computer technology. It finds its way into cryptography as well as division and prime test algorithms.

1.7. **An example of elliptic curve.** Given a field $\mathbb{K}$. Consider the polynomial

$$f = y^2 - (x^3 - ax^2 - b), a, b \in \mathbb{K}.$$

Then $V(I)$ in $\mathbb{K}^2$ is an elliptic curve with Weierstrass form. We often use the symbol $E(\mathbb{K})$ for an elliptic curve over $\mathbb{K}$.

In this course, we shall only consider elliptic curves in this form. There are more general situations, e.g.

$$y^2 - (x^3 - ax^2 - bx - c).$$

However, if $\mathbb{K}$ is algebraically closed, then one can always perform algebraic transformations (bijective maps defined via rational functions) to reduce to Weierstrass form.

We need to use Bézout's theorem.

Definition 1.1. *Let $\mathbb{K}$ be algebraically closed. Consider two affine plane curves C_f, C_g over $\mathbb{K}$ given by f, g. Suppose that f, g do not have common factor. Consider the point $(0,0)$ and the ring*

$$\mathcal{O} = K[X,Y]_{(0,0)}/(f,g).$$

The intersection multiplicity $I_{f,g}(0,0)$ at $(0,0)$ between C_f, C_g at $(0,0)$ is defined to be the $\mathbb{K}$-dimension of $\mathcal{O}$.

Remark 1.2. *If C_f, C_g do not vanish at $(0,0)$ at the same time, then (f,g) contains a unit of $\mathbb{K}[X,Y]_{(0,0)}$ and thus $(f,g) = \mathbb{K}[X,Y]_{(0,0)}$. We see that $\mathcal{O}$ is trivial and has dimension zero. Conversely, if $I_{f,g}(0,0) = 0$ then (f,g) must not be proper and thus either f or g does not vanish at $(0,0)$.*

Remark 1.3. *Why $(0,0)$? In fact, given any point (x,y). One can find an affine map ϕ sending (x,y) to $(0,0)$. In this way, one can study $I_{f,g}(x,y)$. Of course, one has to check that this $I_{f,g}(x,y)$ does not depend on the choice of this affine map ϕ.*

Theorem 1.4 (Bézout's theorem)**.** *Given two projective plane curves C_f, C_g without common factors. Suppose that the base field is algebraically closed. Then we have*

$$\sum_{P \in C_f \cap C_g} I_{f,g}(P) = \deg g \deg f.$$

We do not prove this theorem here. See Fulton chapter 5. In particular, we see that as long as $\deg f \deg g > 0$, $C_f \cap C_g \neq \emptyset$.

Now, we are all set for the following result.

Theorem 1.5. *Let $\mathbb{K}$ be algebraically closed and $2, 3 \in \mathbb{K}^*$. Given C_f with $\deg f = 3$. Then there is an affine map ϕ sending f into Weierstrass form.*

Remark 1.6. *The condition that $\mathbb{K}$ is algebraically closed can be dropped. The general proof requires Riemann-Roch.*

Proof. We shall consider $f \in \mathbb{K}[X,Y,Z]$ as a homogeneous polynomial of degree three. The key step is to find an inflexion point P on C_f and uses an affine map to translate P into $(0,1,0)$. An inflexion point is a point where the tangent intersects C_f with a multiplicity at least three. In our situation, by Bézout's theorem, this tangent line intersects C_f at P only.

Now, consider the Hessian polynomial

$$H_f = \det(f_{x_i x_j})_{1 \leq i,j \leq 3},$$

$x_1 = X, x_2 = Y, x_3 = Z$. Then one can show that $V(H_f) \cap C_f$ at inflection points.

Using Bézout, we see that $V(H_f) \cap C_f$ is not empty so that an inflexion P can be chosen. Denote L_P be the tangent line at P.

We can now choose an affine transformation sending L_P to $Z = 0$ and P to $(0, 1, 0)$. This will change f into another homogeneous polynomial g. Now since $Z = 0$ intersects C_g at $(0, 1, 0)$ only we see that if $g(X, Y, 0) = 0$ then $(x, y) = (0, 1)$. Thus the polynomial g can be written as

$$g(X, Y, Z) = a_0Y^2Z + a_1XYZ + a_3YZ^2 + \text{homogeneous degree 3 polynomial in } X, Z.$$

Up to now, we have not used the condition that $2, 3$ are invertible in $\mathbb{K}$. For further reduction, this condition is crucial. The rest of the proof is left to the reader. □

Warning: In case 2 or 3 is the characteristic of $\mathbb{K}$, one can not obtain Weierstrass form. In those cases, one has to deal with the more general cubics. Fortunately, developing the theory of elliptic curves does not really rely on the fact that our cubic is of Weierstrass form.

1.8. **singularity of elliptic curves.** Now we have a strong result to work on elliptic curves with Weierstrass equations:

$$C : y^2 = x^3 + ax + b.$$

We see that

$$\nabla f(x, y) = (3x^2 + a, 2y).$$

This in order that ∇f is never zero, we shall have that whenever $y = 0$, $3x^2 + a \neq 0$. As $y = 0$ if and only if $y^2 = 0$ we see that we should have that two polynomials

$$3x^2 + a, x^3 + ax + b$$

do not have common roots (or in other words, common factors). This happens if and only the resultant

$$Res(3x^2 + a, x^3 + ax + b) = 4a^3 + 27b^2 \neq 0.$$

We proved the following result.

Theorem 1.7. *Give $C : y^2 = x^3 + ax + b$, we define $\Delta(C) = 4a^3 + 27b^2$ as the discriminant of C. Then C is regular if and only if $\Delta \neq 0$. This result holds for any fields.*

Remark 1.8. *This result only holds for elliptic curves in Weierstrass form. This restriction is not at all strong if the base field is algebraically closed. However, for some special fields, we have to work with general elliptic curves.*

If $char(\mathbb{K}) = 2$*, we have a further problem as* $2y$ *is always* 0*. Thus as long as* $3x^2 + a$ *has a zero, our curve is singular.*

2. Motivations

Having seen the definitions and basic properties of elliptic curves, we now see some motivations.

2.1. **Fermat's Last Theorem.**

One of the most famous applications of elliptic curves is perhaps Wiles' proof of Fermat's last theorem. The proof itself is too deep to be covered in this course. We will only briefly mention some of the ideas.

Theorem 2.1 (Fermat's Last Theorem proved by A. Wiles+some helps)**.** *For* $n \geq 3$*, if* $x^n + y^n = z^n$ *and* $x, y, z \in \mathbb{Z}$ *then* $xyz = 0$*.*

We can reduce the general Fermat's equation for all $n \geq 3$ to prime n and we can also assume that x, y, z are coprime. Earlier approaches for this problem were around algebraic number theory. In this direction, we know that if the number field $Q(\zeta_p)$ (ζ_p is a primitive root of unity) has class number $h_p = 1$, then we can carry out an argument of factorisation which gives us a proof. More generally, the same argument can be used with some twists if $(h_p, p) = 1$. This method works for many primes p, e.g. all primes less or equal to 19 (class number one). In fact, the only exceptions under 100 are $37, 59, 67$.

For small values of p, e.g. $p = 3, 4, 5$ there is another method, probably developed by Fermat, that is called the descend. To start, we assume the existence of a solution which is minimal in some sense (e.g. with the smallest possible x). Then we can perform some lengthy algebra (coming from a group structure of points on certain elliptic curves) to show the existence of a strictly smaller solution. Such a contradiction leads us to the conclusion of FLT for those specific $p's$.

A crucial turnaround happened when Frey introduced the following special elliptic curve

$$C_{Frey} : Y^2 = X(X - x^p)(X + y^p)$$

where x, y, z is a hypothetical non-trivial solution to $x^p + y^p = z^p$ with x being odd and y being even. It can be shown that under mod q for each prime q,

the X part of C_{Frey} does not have a triple root (it can sometimes have a double root). Such a curve is called semistable.

Given any elliptic curve over $\mathbb{Z}$, we can associated a ζ function which reflects its reduction to $\mod q$ for each q:

$$L_C(s) = \prod_{q\text{: bad primes}} \frac{1}{1-a_q q^{-s}} \prod_{q\text{: good primes}} \frac{1}{1-a_q q^{-s}+q^{1-2s}}.$$

The good primes are so that under $\mod q$ the curve C is regular. It can be proved that there are only finitely many primes that are not good. (This is of a similar taste to Dedekind's theorem on the ramification of primes in number fields.) Thus we do not have to consider bad primes. The number a_q is defined so that

$$\#C(F_q) = q + 1 - a_q$$

where $C(F_q)$ is the elliptic curve considered in F_q. Later on, we will provide more details on elliptic curves over finite fields. We can then write our L_C as

$$L_C(s) = \sum_n \frac{a_n}{n^s}$$

with suitable integers a_n. This is then a Dirichlet series. We can then associate a Fourier series

$$f_C(\tau) = \sum_{n\geq 1} a_n e^{2\pi i \tau}.$$

A result of Hecke tells us that if a Dirichlet series is a meromorphic function with a certain functional equation (just like the Riemann zeta function), then the associated Fourier series is a modular form with a certain boundary condition. (We will learn some basics of modular forms in the next section.) In this case, we say that the Dirichlet series is modular.

Now we come back to C_{Frey}, its discriminant is $(xyz)^{2p}$. This is a rather special discriminant. Under this condition, we expect that $C_{Frey}(E_q)$ has a point of order p (in terms of a group structure on elliptic curves) at all but finitely q (those good primes not equal to p). This will give us a strong restriction of a_q, i.e. $a_q - q - 1$ divides p. This expectation is rather strong however Ribet managed to walk around with some similar, more involved but more realistic condition that holds for C_{Frey} and such a condition also gives strong relations among $a_q's$ and those relations prevent $L_{C_{Frey}}$ from being modular.

Theorem 2.2 (Ribet)**.** *C_{Frey} is not modular. More precisely, $L_{C_{Frey}}$ is not modular.*

The final bit of the idea comes from A. Wiles who proved the following result.

Theorem 2.3 (conjectured by Taniyama and Shimura)**.** *If C is semistable, then C (or L_C) is modular.*

From the above two results, we deduce the Fermat's Last Theorem.

2.2. **Birch and Swinnerton-Dyer conjecture.** We learned from the previous subsection that for each elliptic curve C, we can construct a zeta function L_C. This zeta series should carry a great amount of information for C. In fact, the following result was conjectured by Birth and Swinnerton-Dyer. This conjecture is still open and it is one of problems that the CMI offers one million dollars for the solution. (I'm not sure if the CMI takes inflation into account. The longer the problem is open, the less appealing the prize is. Of course, even today, there are easier ways to earn one million dollars.)

Conjecture 2.4. *Let C be an elliptic curve over $\mathbb{Q}$. Then the order of zero of $L_C(s)$ at $s = 1$ is the rank of the group of points on C.*

Remark 2.5. *We will learn that points on C forms an abelian group. In the case of this conjecture, such a group can be written as $\mathbb{Z}^r \bigoplus T$ where T is a finite group and $r \geq 0$ is the rank. We will prove this as the Mordell-Weil Theorem. In particular, if $L_C(1) \neq 0$, then C contains only finitely many rational points.*

2.3. **Rational mappings.** We learned that under linear mappings, we can transform any elliptic curve into Weierstrass form with some conditions on the base field. If we are allowed to use a more general class of transforms, then we can map more curves into elliptic curves with Weierstrass form.

First, we introduce rational maps. We say that a map (coordinate change) is birational if

$$(x, y) \to (s, t)$$

is such that s, t are rational functions in terms of x, y. In addition, the inverse functions $(s, t) \to (x, y)$ are also defined with rational functions. This defines rational maps in the affine setting. For the projective setting, the situation is similar, however, we need to be sure that the rational functions are homogeneous.

There is a hard conjecture about inverses of polynomial maps. See the Jacobian conjecture.

Here is an example from Washington's book (Theorem 2.17).

Example 2.6 (Transform a quartic equation into elliptic curve)**.** *Let $\mathbb{K}$ be a field whose characteristic is not two. Consider the equation*

$$v^2 = au^4 + bu^3 + cu^2 + du + q^2,$$

$a, b, c, d, q \in \mathbb{K}$. *Let*

$$x = \frac{2q(v+q) + du}{u^2}, y = \frac{4q^2(v+q) + 2q(du + cu^2) - (d^2u^2/2q)}{u^3}.$$

This is a birational map sending v, u to x, y. The inverse map is

$$u = \frac{2q(x+c) - (d^2/2q)}{y}, v = -q + \frac{u(ux-d)}{2q}.$$

Then x, y satisfies the following equation

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,$$

where

$$a_1 = d/q, a_2 = c - (d^2/4q^2), a_3 = 2qb, a_4 = -4q^2a, a_6 = a_2a_4.$$

Since rational maps are defined via rational function. We know that if C is a curve defined over $\mathbb{Q}$, then after a rational map, the transformed curve $\tilde{C}$ is still defined over $\mathbb{Q}$. There is a one-one correspondence between rational points on C and $\tilde{Q}$.

Often, as in the example, we have a high degree curve (also known as high genus). We can try to use a rational map to reduce the degree to be as small as possible (can be intuitively thought of as unwarping the curve). The lower the degree is, the simpler the curve is.

3. Elliptic curves over $\mathbb{C}$

We now work with base field $\mathbb{C}$. Before we start, let us make some non-mathematical discussions.

One of the advantages of considering algebraic varieties over $\mathbb{C}$ is that we can make use of many nice properties of the field of complex numbers. In our case,

we consider elliptic curves over $\mathbb{C}$. Then we have a rich selection of analytic methods in studying those elliptic curves. If the defining polynomials are inside a smaller field, e.g. $\mathbb{Q}$ then it is possible to transfer our knowledge in complex numbers to $\mathbb{Q}$. The analytic part of the argument often uses calculus of residues and obtains some counting properties of zeros and poles of certain functions or (differential) forms. Those properties of zeros/poles are actually what we need for developing the theory of elliptic curves. A more algebraic approach in this direction is contained in Riemann-Roch theory which can be found in Fulton.

One big disadvantage of overly considering $\mathbb{C}$ is that we often rely too much on analytic methods and overlook many underlying algebraic structures. If we want to consider elliptic curves over fields of positive characteristics(e.g. finite fields), then almost surely our knowledge in $\mathbb{C}$ cannot be transferred. The theory of Elliptic curves over finite fields is indeed very interesting, not only for mathematicians (there are now industrial cryptography algorithms based on elliptic curves over finite fields). Actually, it is possible to perform tricks that transfer knowledge from $\mathbb{C}$ to fields with positive characteristics, e.g. we can try to mod p. However, a subtler issue is that we might have to work with a field extension that is inseparable which is a significant difference between fields with 0 characteristic and positive characteristics.

3.1. **Lattice in $\mathbb{C}$.** A lattice $\Lambda \subset \mathbb{C}$ is a discrete (additive) subgroup of $\mathbb{C}$ of rank 2. Let us see some examples.

Example 3.1. $\mathbb{Z}+\mathbb{Z}i$;$\mathbb{Z}+\mathbb{Z}\rho$ *($\rho^3 = 1$ is a non-real root)*;$\mathbb{Z}+\mathbb{Z}\tau$ *($\tau \in \mathcal{H}$ the upper semi plane);*

The following are not examples of lattices.

Example 3.2 (non-example)**.** $\mathbb{Z}$ *(rank 1, discrete);* $\mathbb{Z}+\mathbb{Z}\sqrt{2}$ *(rank 2, not discrete);*

not examinable: The notion of lattice is more general than what we have here. In fact, given a Lie group G and a discrete subgroup $\Lambda \subset G$, one can form the quotient G/Λ (or $\Lambda\backslash G$ as G may not be abelian). This quotient may not be a group (unless Λ is normal). Often, we require that, under the Haar measure, G/Λ should have a finite measure or even compact. More generally, we consider a group G acts on a space X and consider X/Λ. We will meet such an example shortly after.

3.2. **Fundamental domain (Torus).** Given a lattice Λ. We can form the quotient $T_\Lambda = \mathbb{C}/\Lambda$. We also have the natural quotient map $Q : \mathbb{C} \to \mathbb{C}/\Lambda$. Via this quotient map, we can add (group, topology, analytic) structures on T_Λ.

A fundamental domain is any collection of points $F \subset \mathbb{C}$ so that the quotient map Q is bijective. In practice, we do not really want to deal with general sets. In our case, F can be chosen to be a half-open parallelogram. Then $\mathbb{C}/\Lambda$ can be viewed as glueing the two pairs of sides of this parallelogram (like a doughnut). This is a torus.

Of course, there is more than one way to choose a fundamental domain.

3.3. **Elliptic functions (not to be confused with elliptic curves).**

Definition 3.3. *Given a meromorphic function f. We say that f is elliptic if there is a lattice Λ and f is invariant under the action of Λ, i.e. $\forall \lambda \in \Lambda, z \in \mathbb{C}$, $f(z+\lambda) = f(z)$.*

We see that for f to be elliptic, it is enough to check $f(.+\lambda_1) = f(.), f(.+\lambda_2) = f(.)$ for any basis $\{\lambda_1, \lambda_2\}$ of Λ. Trivial examples of elliptic functions include constants (as functions).

3.4. **Liouville Theorems and Abel's Theorem.** We have not yet seen any non-constant elliptic functions. Before we construct one, let us first examine some basic properties of elliptic functions.

Theorem 3.4 (Liouville Theorems)**.** *Let Λ be a lattice and F be its fundamental parallelogram. Let f be an elliptic function with respect to Λ.*

1. *If f is analytic, then f is a contant.*
2. *We have*
$$\sum_{z \in F, f(z)=\infty} Res(f; z) = 0.$$
Here, $Res(f; z)$ is the residue of f at z.
3. *f has as many zeros as poles.*

Proof. 1. Since F has a compact closure, we see that $f(F)$ is bounded. As f is an elliptic function w.r.t Λ, we see that $f(\mathbb{C}) = f(F)$. Thus f is a bounded analytic function and it is then constant (by another theorem of Liouville).

2. We can translate F if necessary to achieve that no pole of f is contained in the boundary of F. We can now perform the following contour integral along ∂F,

$$\int_{\partial F} f = 2\pi i \sum_{z\in F, f(z)=\infty} Res(f;z).$$

Because f is periodic w.r.t Λ, we see that integrals along opposite sides of ∂F cancel each other and the result concludes.

3. Consider the elliptic function $g = f'/f$. This is the logarithmic derivative of f. We want to perform a contour integral along ∂F. Translate F is necessary so that no poles nor zeros are on ∂F. In this way, $\int_{\partial F} g$ is well defined.

If $f(w) = 0$ for some w, then we see that

$$f(z) = (z-w)^k h(z-w)$$

for some meromorphic function h with $h(0) \notin \{0, \infty\}$ and some integer $k > 0$. Then we see that

$$f'(z)/f(z) = \frac{k(z-w)^{k-1}h(z-w) + (z-w)^k h'(z-w)}{(z-w)^k h(z-w)} = k\frac{1}{z-w} + \frac{h'(z-w)}{h(z-w)}.$$

The second part above is analytic around $z = w$. This implies that

$$Res(g;w) = Res(k/(z-w);w) = k.$$

If $f(w) = \infty$ for some w, then we see that

$$f(z) = (z-w)^{-k} h(z-w)$$

for some meromorphic function h with $h(0) \notin \{0, \infty\}$ and some integer $k > 0$. Then a similar argument as above tells us that

$$Res(g;w) = -k.$$

Since g is periodic w.r.t Λ as well, as in part (2), we see that

$$0 = \int_{\partial F} g = \sum Res(g;z) = \sum \text{order of zeros} - \sum \text{order of poles}.$$

This proves the result. □

We see that elliptic functions should have the same amount of zeros and poles. If we prescribe a set of zeros and a set of poles, can we find an elliptic function with these zeros and poles? The following result gives us an answer.

Theorem 3.5 (Abel)**.** *Let Λ be a lattice with F.D. F. Then let $n > 0$ be an integer, $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ be two disjoint collections of points in F. We allow $a's$ not to be different from each other and we all $b's$ not to be different from each other. Then there is an elliptic function f with zeros at $a_1, \ldots, a_n$ and poles at $b_1, \ldots, b_n$ iff*

$$\sum_i a_i - \sum_i b_i \in \Lambda.$$

This result implies in particular that there are many elliptic functions. We will prove this result after we construct our first elliptic function–the Weierstrass function.

Remark 3.6. *As mentioned before, a more algebraic way of stating the above results is via the notion of divisors (Riemann-Roch). If time permits, we have a chance to have a look at it. One advantage of this divisor approach is that we no longer need to work in $\mathbb{C}$. More towards this direction is Grothendieck's theory (Grothendieck-Riemann-Roch).*

3.5. **Weierstrass function.** We now explicitly construct an elliptic function and then find its roots and poles.

Now we have a lattice Λ. We want to construct a function f which is periodic w.r.t Λ. From Liouville's theorem 2, we know that such a function can not have only one pole of order one (simple pole). The next simplest guess would be a pole of order two. We could try something like

$$f_2(z) = \sum_{\lambda \in \Lambda} \frac{1}{(z-\lambda)^2}.$$

From the algebraic structure of the sum, $f_2(z)$ is formally periodic w.r.t Λ. However, there is a huge problem. The series does not converge absolutely and may not define a meromorphic function.

A simple comparison test finds us the following result.

Lemma 3.7. *For integer > 2 the series*

$$f_k(z) = \sum_{\lambda \in \Lambda} \frac{1}{(z-\lambda)^k}$$

converges absolutely and defines a meromorphic function. f_k has poles at Λ. Those poles have order k.

From above we obtain f_3, an elliptic function with one pole (mod Λ) of order three. We can now integrate this function and obtain a function with one pole (mod Λ) of order two. For functions with poles, it is complicated to deal with integrals because integration depends on paths. Luckily, we can construct the integral in a straightforward way.

Definition 3.8. *The following series*

$$\wp(z;\Lambda) = \frac{1}{z^2} + \sum_{\lambda\in\Lambda\setminus\{0\}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right)$$

is the Weierstrass function w.r.t Λ.

We have to check the convergence of the defining series.

Theorem 3.9. *The series defining $\wp$ is absolutely convergent for $z \notin \Lambda$ and defines an analytic function there. We extend the domain by setting $\wp(\lambda) = \infty$ for $\lambda \in \Lambda$. Thus $\wp() = \wp(;\Lambda)$ is a meromorphic function. It is elliptic and has poles of order two at Λ. There are no other poles.*

Proof. Consider the sum

$$\sum_{\lambda\in\Lambda\notin\{0\}} \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}.$$

The bracket in the sum reads

$$\frac{2z\lambda - z^2}{\lambda^2(z-\lambda)2}.$$

For large λ, it has norm of size

$$\sim \frac{1}{|\lambda|^3}.$$

From here we see the absolute convergence for each fixed $z \notin \Lambda$.

It is easy to see that the only poles of $\wp$ are at Λ and they are of order two. We now show the periodicity. Take derivative we see that

$$\wp'(z) = -\frac{2}{z^3} - \sum_{\lambda\in\Lambda\setminus\{0\}} \frac{2}{(z-\lambda)^3} = -2f_3(z).$$

Thus $\wp'$ is periodic w.r.t Λ. From here we see that for each $\lambda \in \Lambda$ there is a constant C_λ such that

$$\wp(z+\lambda) - \wp(z) = C_\lambda.$$

Therefore we have for $\lambda_1, \lambda_2 \in \Lambda$

$$\wp(z+\lambda_1+\lambda_2) - \wp(z) = C_{\lambda_1+\lambda_2}$$

and

$$\wp(z+\lambda_1+\lambda_2) - \wp(z) = \wp(z+\lambda_1+\lambda_2) - \wp(z+\lambda_1) + \wp(z+\lambda_1) - \wp(z) = C_{\lambda_1} + C_{\lambda_2}.$$

So we have

$$C_{\lambda_1+\lambda_2} = C_{\lambda_1} + C_{\lambda_2}.$$

We can find a basis of Λ formed by $\{w_1, w_2\}$. Then we have

$$\wp(z+w_1) - \wp(z) = C_{w_1}.$$

Setting $z = -w_1/2$ (observe that $-w_1/2 \notin \Lambda$) we see that

$$\wp(w_1/2) - \wp(-w_1/2) = C_{w_1}.$$

Observe that Λ is invariant under $z \to -z$ and this shows that

$$\wp(z) = \wp(-z)$$

for $z \notin \Lambda$. This implies that

$$C_{w_1} = 0.$$

Similarly, $C_{w_2} = 0$. As w_1, w_2 for a basis of Λ we see that

$$C_\lambda = 0$$

for all $\lambda \in \Lambda$. Thus we have for all $z \notin \Lambda$ and $\lambda \in \Lambda$,

$$\wp(z+\lambda) = \wp(z).$$

This shows that $\wp$ is periodic w.r.t Λ. □

From the proof, we see that $\wp$ is even and $\wp'$ is odd. Thus $\wp$ has two zeros mod Λ (counting multiplicities) and $\wp'$ has three zeros mod Λ (counting multiplicities). It turns out that we can explicitly determine the zeros of $\wp'$.

Lemma 3.10 (Zeros of $\wp'$)**.** $\wp'(z) = 0$ *if and only if* $z \in (\Lambda/2) \setminus \Lambda$.

Remark 3.11. *Given a basis w_1, w_2 of Λ we have three half lattice points*

$$w_1/2, w_2/2, (w_1+w_2)/2.$$

They are the only three zeros (mod Λ) of $\wp'$.

Proof. As $\wp'$ is odd and periodic w.r.t Λ and $w_1/2, -w_1/2$ has difference $w_1 \in \Lambda$, we see that

$$\wp'(w_1/2) = \wp'(-w_1/2) = -\wp'(w_1/2).$$

This implies that $\wp'(w_1/2) = 0$. Similarly, $\wp'(w_2/2) = 0$ and $\wp'((w_1+w_2)/2) = 0$. Since $\wp'$ has only one order three pole mod Λ, Liouville's theorem part 3 shows that we have already identified all zeros of $\wp'$. □

The map $\wp : \mathbb{C}/\Lambda \to \mathbb{C} \cup \{\infty\}$ is mostly two-one with four exceptions at $\wp^{-1}(w_1/2), \wp^{-1}(w_2/2), \wp^{-1}(w_1/2+w_2/2), \wp^{-1}(\infty)$ where ramifications happen.

Now we prove Abel's theorem.

Proof of Abel's Theorem: necessity. For the necessity, let f be periodic w.r.t Λ. Consider the function

$$h(z) = z\frac{f'(z)}{f(z)}.$$

This h is not elliptic. Instead, we have

$$h(z+\lambda) - h(z) = \lambda\frac{f'(z)}{f(z)}.$$

Next, consider the integral

$$\int_{\partial F} h = 2\pi i\Big(\sum_{z\in F, f(z)=0} z - \sum_{z\in F, f(z)=\infty} z\Big).$$

We perform the integral $\int_{\partial F} h$. Now, integrals of opposite sides do not cancel each other. For example, consider two sides $[a, a+w_1], [a+w_2, a+w_2+w_1]$. We have

$$\int_a^{a+w_1} h + \int_{a+w_2+w_1}^{a+w_2} h = \int_a^{a+w_1} w_2\frac{f'(z)}{f(z)}.$$

On $[a, a+w_1]$, $f(z)$ has no poles nor zeros (this can be achieved by translating F if necessary). Thus we see that there is an analytic function g with $e^g = f$. We have

$$\int_a^{a+w_1} \frac{f'(z)}{f(z)} = h(a+w_1) - h(a).$$

Since $f(a+w_1) = f(a)$, we see that $e^{g(a)} = e^{g(a+w_1)}$. This happens if $g(a) - g(a+w_1) \in 2\pi i\mathbb{Z}$. Similar argument can be performed on the other pair of sides and we see that

$$\int_{\partial F} h \in 2\pi i\mathbb{Z}w_1 + 2\pi i\mathbb{Z}w_2 = 2\pi i\Lambda.$$

This implies that

$$\sum_{z\in F, f(z)=0} z - \sum_{z\in F, f(z)=\infty} z \in \Lambda.$$

□

Proof of sufficiency. For sufficiency, we need a help of a special function. In order to motivate the situation, we recall the theory of meromorphic functions on $\mathbb{C}\cup\{\infty\}$. The space $\mathbb{C}\cup\{\infty\}$ is a compact Riemann surface of genus zero. Our torus $\mathbb{C}/\Lambda$ can be shown to be a compact Riemann surface of genus one. We do not require knowledge of the theory of Riemann surfaces. A uniform treatment will be carried out in what is known as the Riemann-Roch theory (which can be found in Fulton).

Now, for $\mathbb{C}\cup\{\infty\}$, if we prescribe a set of zeros and a set of poles, how can we find a function with this set of zeros and poles. A necessary condition is the number of poles and zeros should be the same. This turns out to be sufficient as well. Indeed, let $a_1, \ldots, a_n$ be a collection of zeros and $b_1, \ldots, b_n$ be a collection of poles. A naive idea is the function

$$f(z) = \frac{\prod_{i=1}^n (z-a_i)}{\prod_{i=1}^n (z-b_i)}.$$

However, we allow $a's$ or $b's$ be ∞. This is easily overcome by considering

$$f(z) = \frac{\prod_{a_i\neq\infty} (z-a_i)}{\prod_{b_i\neq\infty} (z-b_i)}.$$

This function also has the prescribed pole or zero at ∞.

For $\mathbb{C}/\Lambda$, the idea is to find a function σ, preferably elliptic, with only one zero (z_0). Then we can find

$$f(z) = \frac{\prod_{i=1}^n \sigma(z - a_i - z_0)}{\prod_{i=1}^n \sigma(z - b_i - z_0)}.$$

Unfortunately, such a function σ does not exist according to Liouville's theorem part (2). So some twists need to be done.

We will instead construct an analytic function $\sigma : \mathbb{C} \to \mathbb{C}$ such that

1. For each $\lambda \in \Lambda$, there are some a_λ, b_λ with $\sigma(z + \lambda) = e^{a_\lambda z + b_\lambda}\sigma(z)$ for all $z \in \mathbb{C}$.
2. σ has only one zero (of order 1) $\mod \Lambda$.

Once we constructed σ, we can find

$$f(z) = \frac{\prod_{i=1}^n \sigma(z - a_i - z_0)}{\prod_{i=1}^n \sigma(z - b_i - z_0)}.$$

Since we allow $a_1, \ldots, a_n$, $b_1, \ldots, b_n$ to be any representative $\mod \Lambda$ and we know that $\sum_i (a_i - b_i) \in \Lambda$, we can choose specific a_i, b_i in $\mathbb{C}$ with

$$\sum_i (a_i - b_i) = 0.$$

We still have to show that f is elliptic. This is easy:

$$f(z+\lambda) = \frac{\prod_{i=1}^n \sigma(\lambda + z - a_i - z_0)}{\prod_{i=1}^n \sigma(\lambda + z - b_i - z_0)} = f(z)\frac{e^{\sum_i a_\lambda(z - a_i - z_0)}}{e^{\sum_i a_\lambda(z - b_i - z_0)}} = f(z)e^{\sum_i a_\lambda(b_i - a_i)} = f(z).$$

This proves the sufficient part of Abel's theorem.

Now all that is left is the construction of σ. A straightforward way is to define

$$\sigma(z) = z \prod_{\lambda \in \Lambda \setminus \{0\}} \left(1 - \frac{z}{\lambda}\right) e^{(z/\lambda) + (1/2)(z/\lambda)^2}.$$

The infinite product converges because

$$\left|\left(1 - \frac{z}{\lambda}\right) e^{(z/\lambda) + (1/2)(z/\lambda)^2}\right| \lesssim 1/|\lambda|^3.$$

With $x = z/\lambda$, the factor $e^{x + (1/2)x^2}$ was introduced to make the infinite product converge. The particular form of this factor is motivated by $1/(1-x) = e^{-\log(1-x)}$

and the Taylor expansion for $\log(1-x)$ is

$$\log(1-x) = -x - x^2/2 - x^3/3 \dots.$$

Thus if we cut-off at the N-th term we obtain for $|x<1|$ that

$$\log(1-x) = -x - x^2/2 - x^3/3 - \cdots - x^N/N + o(x^N).$$

Then $(1-x)e^{-(-x-x^2/2-x^3/3-\cdots-x^N/N)} = (1-x)(1-x)^{-1}e^{o(x^N)} = 1 + o(x^N)$. This idea was introduced by Weierstrass we can write any analytic function as an infinite product called the Weierstrass product formula. Of course, the infinite product formula is also subject to convergence.

Consider the function $h(z) = \sigma(z+\lambda)$. This function has exactly the same set of zeros as $\sigma(z)$. This $h(z)/\sigma(z)$ is analytic and non-vanishing on $\mathbb{C}$. Thus we can write

$$\sigma(z+\lambda) = \sigma(z)e^{g(z)}$$

for some analytic function g. To show that $g(z)$ is linear, it is enough to show that $g'' = 0$. Direct computing shows that

$$g'(z) = \frac{\sigma'(z+\lambda)}{\sigma(z+\lambda)} - \frac{\sigma'(z)}{\sigma(z)}.$$

To show that $g'' = 0$ is to show that

$$u = (\sigma'/\sigma)'$$

is elliptic. Again, direct computation shows that

$$\frac{\sigma'}{\sigma}(z) = \frac{1}{z} + \sum_{\lambda\in\Lambda\setminus\{0\}} \left(\frac{1}{\lambda} + \frac{z}{\lambda^2} - \frac{1}{\lambda - z}\right).$$

Differentiate, we obtain

$$u(z) = -\frac{1}{z^2} + \sum_{\lambda\in\Lambda\setminus\{0\}} \left(\frac{1}{\lambda^2} - \frac{1}{(z-\lambda)^2}\right) = -\wp(z).$$

So u is indeed elliptic and thus $g'' = 0$ and thus g is linear. To conclude, we found a function σ that fits our needs. The zeros of σ are exactly Λ. The magical z_0 appeared when finding f can be any point in λ...which can be 0.

This finishes the proof of Abel's theorem. □

Remark 3.12. *The function σ has much more impressive properties. We will return to this once we introduced modular forms.*

3.6. **From elliptic functions to elliptic curves.** We now bridge the connection between elliptic functions and elliptic curves (over $\mathbb{C}$). Our gadgets are $\wp, \wp'$. We know that $\wp$ is even and $\wp'$ is odd.

We first want to have a result that tells the structure of the set of elliptic functions. First, we observe that the set of elliptic functions is a field under the usual addition and multiplication.

(Recall that our lattice is always Λ) Let f be an elliptic function. Then we see that

$$\frac{f(z)+f(-z)}{2}$$

is even and elliptic.

For now, we assume that f is itself even and elliptic. Suppose that the poles of f are in Λ. Then around 0 we have the Laurent series

$$f(z)=\sum_{n=-2k, n \text{ even}}^{\infty} a_n z^n.$$

We also know that

$$\wp=\frac{1}{z^2}+\sum_{n=0, n \text{ even}} b_n z^n.$$

Later we will find the coefficients b_n. In particular, $b_0=0$.

Now $f-a_{-2k}\wp^k$ is even and elliptic with poles in Λ of smaller order than that of f. Continue this way we eventually obtain a polynomial P such that

$$f-P(\wp)$$

is elliptic and analytic and therefore a constant. So we see that f is actually a polynomial in $\wp$.

If f has poles not in Λ, say a. The for some $N>0$, $(\wp(z)-\wp(a))^N f(z)$ does not have a pole at a. In this way, we can find a polynomial R so that $R(\wp)f$ has poles only in Λ which implies that $R(\wp)f$ is a polynomial in $\wp$ which implies that f is a rational function in $\wp$. Thus we proved the following result.

Theorem 3.13. *If f is even and elliptic, then $f \in \mathbb{C}(\wp)$. If f is odd, then $f/\wp'$ is even and in $\mathbb{C}(\wp)$. In general, the field of elliptic functions is equal to the field*

$$\mathbb{C}(\wp) + \wp'\mathbb{C}(\wp).$$

Now, we already have a rather satisfactory result for the field of elliptic functions. Remember that we are into elliptic curves, not elliptic functions!

Again, we write

$$\wp(z) = \frac{1}{z^2} + \sum_{n>0, n \text{ even}} b_n z^n,$$

$$\wp'(z) = -\frac{2}{z^3} - \sum_{n>0, n \text{ even}} n b_n z^{n-1}.$$

To cancel the pole, we first compute

$$(\wp'(z))^2 - 4(\wp(z))^3 = -20b_2 z^{-2} - 28b_4 + \text{analytic part vanishing at } 0.$$

We still have a pole in the expression, to cancel it we write

$$(\wp'(z))^2 - 4(\wp(z))^3 + 20b_2\wp(z) + 28b_4 = \text{analytic part vanishing at } 0.$$

The LHS is elliptic and the RHS is analytic and vanishing at zero therefore the whole expression is constantly zero. Thus we see that

$$(\wp'(z))^2 - 4(\wp(z))^3 - 20b_2\wp(z) - 28b_4 = 0.$$

If we write $X = \wp, Y = \wp'$ we see that

$$Y^2 = 4X^2 - 20b_2X - 28b_4.$$

We know that $\wp$ is surjective to $\mathbb{C} \cup \{\infty\}$ (consider $\wp(z) - w$). Once we know $\wp$ we know $(\wp')^2$ and because $\wp'$ is odd, we can achieve either of the square roots. Thus we proved the following result.

Theorem 3.14. *The function $z \to (\wp(z), \wp'(z))$ (meromorphically) maps $\mathbb{C}/\Lambda$ in mostly one-one onto the elliptic curve $\mathbb{C}[X,Y]/(Y^2 - (4X^2 - 20b_2X - 28b_4))$. The point $[0] \in \mathbb{C}/\Lambda$ maps to the point at ∞ of the curve.*

Remark 3.15. *For each (x, y) on the curve, we have exactly two points z, z' with $\wp(z) = \wp(z') = x$. Note that z, z' might be the same point. As $\wp$ is even and $\wp'$ is odd, we see that $z = -z'$ and we see that z, z' are mapped to $(x, y), (x, -y)$. So we obtain a one-one map!*

Don't forget that we will compute b_n. Also, we will show that any regular elliptic curve over $\mathbb{C}$ can be obtained via such a map for a suitable lattice Λ.

We put the computation results here for convenience ($G_\cdot$ are Eisenstein series)

$$20b_2 = 60G_4, 28b_4 = 140G_6.$$

The equation of the elliptic curve is

$$Y^2 = 4X^2 - 60G_4 - 140G_6.$$

3.7. **Computation of coefficients.** We have

$$\wp(z) = \frac{1}{z^2} + \sum_{n\geq 0, n \text{ even}} b_n z^n.$$

We want to find b_n. The clue in our hands is

$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda\in\Lambda\setminus\{0\}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right).$$

Cauchy's theorem tells us that

$$b_n = \frac{f^{(n)}(0)}{(n)!}$$

where $f = \wp - z^{-2}$ which is analytic. Evaluating the above series at 0 tells us that

$$f(0) = 0$$

and therefore $b_0 = 0$.

Direct computing shows that for $n > 0$,

$$f^{(n)}(z) = \sum_{\lambda\in\Lambda\setminus\{0\}} \frac{(-1)^{n+1}(n+1)!}{(z-\lambda)^{n+2}}.$$

From here we see that

$$b_n = (-1)^{n+1}\frac{(n+1)!}{n!} \sum_{\lambda\in\Lambda\setminus\{0\}} \frac{1}{\lambda^{n+2}}.$$

Of course, we only need to consider n being even.

Definition 3.16. *For $n \geq 3$ being integer,*

$$G_n(\Lambda) = \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{\lambda^n}$$

is called Eisenstein series of Λ.

Thus we have

$$b_n = (n+1)G_{n+2}.$$

We specially remark that $G_n(\Lambda)$ is a function of Λ. Such a function is a modular form. The theory of modular form is very important for elliptic curves. This fact cannot be overstated.

3.8. **A group structure.** We established a map

$$E : \mathbb{C}/\Lambda \to V(Y^2 - (4X^4 - 60G_4X - 140G_6)).$$

There is a small technical point. The point $E[0]$ is not quite defined as it is a pole of $\wp$ and $\wp'$. We simply set $E[0] = \infty$. This ∞ is the point at infinity of the curve. Or more formally, we consider $V(Y^2 - (4X^4 - 60G_4X - 140G_6))$ as a projective curve. Then everything is well defined.

On $\mathbb{C}/\Lambda$ there is a very natural additive structure coming from $\mathbb{C}$. We can write as

$$[z_1] + [z_2] = [z_1 + z_2].$$

This group structure defines a group structure on the corresponding elliptic curve $(Y^2 = 4X^4 - 60G_4X - 140G_6)$. The point is to describe this group structure on the curve.

Let P, Q two different points on $E(\mathbb{C})$. For now, we assume that none of P, Q is ∞. Consider the projective line passing through P, Q:

$$l_{P,Q} = t_1P + t_2Q : (t_1, t_2) \in \mathbb{C}^2 \setminus \{(0,0)\}.$$

We see that $l_{P,Q} \cap E(\mathbb{C})$ already contains two points P, Q. From Bézout's theorem, we know that $l_{P,Q} \cap E(\mathbb{C})$ should contain three points (counting multiplicities). Thus there is a third point R. Let P, Q, R has preimages

$$p, q, r = E^{-1}(P), E^{-1}(Q), E^{-1}(R).$$

We claim that $[p] + [q] + [r] = [0]$. To finish the picture, we check what is $[-p]$. We know that $E(p) = (\wp(p), \wp'(p))$ and $E(-p) = (\wp(-p), \wp'(-p), 1) =$

$(\wp(p), -\wp'(p), 1)$. From here we have a complete picture of the group structure on $E(\mathbb{C})$.

Theorem 3.17. *Let P, Q be points on $E(\mathbb{C})$. The projective line $l_{P,Q}$ intersects $E(\mathbb{C})$ at a third point $R = (x_R, y_R, z_R)$. We define*

$$P + Q = (x_R, -y_R, z_R).$$

This '+' defines a group on points $E(\mathbb{C})$. The point ∞ is the group unit.

Proof. Recall the discussion before this theorem. We should now prove our claim that $[p] + [q] + [r] = [0]$. In other words,

$$[p] + [q] + [r] = [0] \iff P, Q, R \text{ are colinear.}$$

Therefore we can first choose three points $p, q, r = -(p + q)$ and show that the points $P = E(p), Q = E(q), R = E(r)$ are colinear. In more detail, we have three points

$$P = (\wp(p), \wp'(p), 1), Q = (\wp(q), \wp'(q), 1), R = (\wp(p+q), -\wp'(p+q), 1).$$

In order to show that P, Q, R are colinear we need to show that

$$\det \begin{bmatrix} \wp(p) & \wp'(p) & 1 \\ \wp(q) & \wp'(q) & 1 \\ \wp(p+q) & -\wp'(p+q) & 1 \end{bmatrix} = 0.$$

To show this, we define the function,

$$f(z) = \det \begin{bmatrix} \wp(p) & \wp'(p) & 1 \\ \wp(q) & \wp'(q) & 1 \\ \wp(z) & -\wp'(z) & 1 \end{bmatrix}.$$

This function is a polynomial in $\wp, \wp'$ so it is an elliptic function. We can see that f has three poles (counting multiplicities, $\mod \Lambda$). As there are already two obvious roots p, q. Since f should have three roots and Abel's theorem tells us that the third root r should satisfy

$$p + q + r = 0 \mod \Lambda.$$

This is exactly $[p] + [q] + [r] = [0]$.

We are almost done. It remains to check the special situations when $P = Q$ and when one or both of P, Q are ∞. All those cases follow from the continuity

of the map E. Notice that when $P = Q$, $l_{P,Q}$ should be the tangent line of $E(\mathbb{C})$ at P. □

We have a rather satisfactory description of the group structure on $E(\mathbb{C})$. This indicates that there should be a rational formula for $\wp(z+w)$. We now try to find it. This is not difficult as we almost have done it in the above proof. Consider the vanishing determinant

$$\det \begin{bmatrix} \wp(p) & \wp'(p) & 1 \\ \wp(q) & \wp'(q) & 1 \\ \wp(p+q) & -\wp'(p+q) & 1 \end{bmatrix} = 0.$$

If nothing is ∞ we obtain three colinear points

$$(\wp(p), \wp'(p)), (\wp(q), \wp'(q)), (\wp(p+q), -\wp'(p+q)).$$

Let us write the line as

$$Y = mX + b.$$

Then the slope m is

$$m = \frac{\wp'(p) - \wp'(q)}{\wp(p) - \wp(q)}.$$

Recall the equation of $E(\mathbb{C})$:

$$(mX + b)^2 = 4X^3 - 60G_4X - 140G_6.$$

From this equation, we see that its three roots satisfy

$$\wp(p) + \wp(q) + \wp(p+q) = \frac{m^2}{4}.$$

Thus we see that

$$\wp(p+q) = \frac{1}{4}\left(\frac{\wp'(p) - \wp'(q)}{\wp(p) - \wp(q)}\right)^2 - \wp(z) - \wp(w).$$

Taking limit $p \to q$ we can also obtain the doubling formula

$$\wp(2q) = \frac{1}{4}\left(\frac{\wp''(q)}{\wp'(q)}\right)^2 - 2\wp(q).$$

As $\wp'^2 = 4\wp^3 - 60G_4\wp - 140G_6$ we see that

$$2\wp''\wp' = 12\wp^2\wp' - 60G_4\wp'$$

so that

$$2\wp'' = 12\wp^2 - 60G_4.$$

Thus we have

$$\wp(2q) = \frac{1}{16}\frac{(12\wp(q)^2 - 60G_4)^2}{4\wp(q)^3 - 60G_4\wp(q) - 140G_6} - 2\wp(q)$$

is a rational relation. In principle, it is possible to deduce $\wp(kq)$ as a rational function of $\wp(q)$ for each integer $k \geq 2$.

3.9. **An algebraic proof of the group structure for elliptic curves not only over $\mathbb{C}$.** We have defined and proved the existence of a group structure over the points on elliptic curves over $\mathbb{C}$. This automatically gives us the group structures for elliptic curves over any field with 0 characteristic with algebraic closure contained in $\mathbb{C}$, e.g. any number fields. For many reasons, we will provide another proof that allows base fields other than $\mathbb{C}$. Among other things, we will need this for our later study on elliptic curves over finite fields, say.

For each field $\mathbb{K}$, we consider $E(\mathbb{K})$ as the set of points on an elliptic curve over $\mathbb{K}$. For the proof of the group structure, we will actually work with $E(\overline{\mathbb{K}})$ and argue that $E(\mathbb{K})$ forms a subgroup. After our study of $E(\mathbb{C})$ we have a hint of how the group structure can be described.

Theorem 3.18. *Assume only that $char(\mathbb{K}) \notin \{2, 3\}$. Consider $E = E(\overline{\mathbb{K}})$ in the projective plane. For two points $P, Q \in E$, we construct the line $l_{P,Q}$ passing through P, Q. This line intersects E at a third point which we denote as R. We define the group structure to be*

$$P + Q + R = 0$$

where 0 denotes the ∞ point which acts as the group identity.

Moreover, $E(\mathbb{K}) \subset E(\overline{\mathbb{K}})$ is a subgroup.

Proof. The commutativity can be checked easily. The fact that ∞ acts as the identity can also be checked easily. The difficult part is associativity, i.e. we need to check that

$$(P + Q) + R = P + (Q + R).$$

Unlike the case over $\mathbb{C}$, where we proved a 'generic' version of the above identity, i.e. this identity holds for most of the points in $E(\mathbb{C})$. Then a continuity argument

will help us extend our result to all points in $E(\mathbb{C})$. We do not have this argument here as the continuity argument is out-of-table. We can try to construct field metrics (or Adeles) to embed our field into some nice metric space. Then we have the continuity argument. However, this approach is much more complicated than how we will proceed now.

We consider the 'generic case', i.e. none of P, Q, R is ∞, none of the two points is the same, P, Q, R are not collinear. The 'special' cases can be checked with direct arguments similar to the generic case, but simpler.

Let us consider lines l_1, l_2, l_3, m_1, m_2, m_3 so that

$$l_1 = l_{P,Q}, l_2 = l_{\infty,Q+R}, l_3 = l_{R,P+Q}$$

and

$$m_1 = l_{P,Q+R}, m_2 = l_{Q,R}, m_3 = l_{\infty,P+Q}.$$

See the illustrative [Picture] which does not reflect the actual positions of the points.

We know that eight points $P, Q, R, P+Q, -(P+Q), Q+R, -(Q+R), \infty$ are on $E(\overline{\mathbb{K}})$. We now show that the ninth point on the intersections of $l's, m's$ is also contained in $E(\overline{\mathbb{K}})$ and this point is both $(P+Q)+R$ and $P+(Q+R)$. This will finish the proof.

If we write $l_1 : aX + bY + cZ = 0$ then we can replace X with a linear form in X, Y. We want to study the field of rational functions over $\mathbb{K}[X, Y, Z]/(l_1)$. Reduction modulo (l_1) is well-defined for homogeneous polynomials. For example, let our curve be $C : c(X, Y, Z)$. Then inserting the expression of X as $X(Y, Z)$ we obtain a homogeneous polynomial $c(X(Y, Z), Y, Z)$ of degree three. By construction, $c(X(Y, Z), Y, Z) = c(X, Y, Z) \mod (l_1)$. So it is enough to study $\overline{c}(Y, Z) = c(X(Y, Z), Y, Z)$. Consider the degree three 'curve' $m_1 \cup m_2 \cup m_3$ whose equation is the multiplication of the linear forms defining m_1, m_2, m_3. For convenience, we write $m_i(X, Y, Z)$ for them. We can also write $m_1 m_2 m_3$ as a degree three homogeneous polynomial $\overline{m} = \overline{m_1 m_2 m_3}(Y, Z)$. Next, it is clear that $\overline{m}$ and $\overline{c}$ has three common zeros coming from $P, Q, -(P+Q)$, i.e. the (Y, Z) parts of $P, Q, -(P+Q)$. For any point T not equal to $P, Q, -(P+Q)$ in the (Y, Z) part, we can define $\alpha \in \overline{\mathbb{K}}$ such that

$$\overline{c}(T) = \alpha \overline{m}(T).$$

Then $\overline{c}-\alpha\overline{m}$ is a homogeneous polynomial of degree 3 that vanishes at least four different points coming from $P, Q, -(P+Q), T$. This implies that $\overline{c}-\alpha\overline{m}=0$. Thus we see that

$$c-\alpha m_1m_2m_3$$

is a multiple of l_1. Similarly,

$$c-\beta l_1l_2l_3$$

is a multiple of m_3 for some $\beta\in\overline{\mathbb{K}}$. This implies that the polynomial

$$f=c-\alpha m_1m_2m_3-\beta l_1l_2l_3$$

is a multiple of l_1 and m_3. As l_1, m_3 are coprime, we see that $l_1m_3|f$. Since f is of degree three, we see that f must be

$$f=l_1m_3h$$

for some linear form h. Consider the non-collinear points $(Q+R), -(Q+R), R$. We see that f vanishes at all of them. Also, l_1m_3 does not vanish at either of them otherwise either l_1 or m_3 would intersect C with four different points which is not possible. Then h vanishes at all of them and this implies that h must be the zero form for otherwise, $(Q+R), -(Q+R), R$ are collinear. This implies that f is zero. Consider $l_1l_2l_3$ and $m_1m_2m_3$. Nine intersection points occur. We already knew that eight of them are contained in the curve C. For the ninth of them P_9, we have

$$0=c(P_9)-\beta l_1l_2l_3(P_9)-\alpha m_1m_2m_3(P_9)$$

and so

$$c(P_9)=0.$$

Thus P_9 is also contained in the curve C. This is what we wanted to prove. □

3.10. **Modular forms/lattice.** Although we do not plan to cover modular forms in this lecture, we nonetheless need to at least have a few results at hand.

Earlier, we always fixed a lattice and then form the theory of elliptic curves. Here, we change our point of view by considering the space of all lattices.

Given two lattices Λ_1, Λ_2, we consider them as equivalent if there is a complex number $a\neq 0$ with

$$\Lambda_1=a\Lambda_2.$$

This says that we can rotate and scale Λ_2 to obtain Λ_1. Observe that Λ_1, Λ_2 are equivalent, then for each elliptic function f_{λ_1} w.r.t. Λ_1 we see that $f_{\lambda_2} : z \to f_{\lambda_1}(az)$ is elliptic w.r.t. Λ_2. In this way, the fields of elliptic functions w.r.t. Λ_1, Λ_2 are isomorphic. We emphasise that the key point here is really the special form of isomorphism. Observe that any two fields of elliptic functions are isomorphic by linking their $\wp$ and $\wp'$. In general, such a correspondence is not made by a change of variable.

Now for any lattice Λ we can rotate and scale so that $1 \in \Lambda$ as one of the members forming a basis. We can choose the other basis member τ such that $\operatorname{Im} \tau > 0$. In other words, $\tau \in \mathbb{H}$. There is more than one possibility for τ.

Let $\tau, \tau' \in \mathbb{H}$, when are $\mathbb{Z} + \mathbb{Z}\tau$, $\mathbb{Z} + \mathbb{Z}\tau'$ equivalent? This happens iff there is some $a \neq 0$ with

$$\mathbb{Z} + \mathbb{Z}\tau = a(\mathbb{Z} + \mathbb{Z}\tau).$$

Then we should have integers $\alpha, \beta, \gamma, \delta$,

$$\tau' = a(\alpha\tau + \beta), 1 = a(\gamma\tau + \delta).$$

Therefore we see that

$$\tau' = \frac{\alpha\tau + \beta}{\gamma\tau + \delta}.$$

We can change a to a^{-1} and exchange τ, τ' to obtain that for some integers $\alpha', \beta', \gamma', \delta'$ with

$$\tau = \frac{\alpha'\tau' + \beta'}{\gamma'\tau' + \delta'}.$$

Let $A = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$. The map

$$M_A : z \to \frac{\alpha z + \beta}{\gamma z + \delta}$$

is a Möbius transformation on $\mathbb{H}$. In this way, we can define an action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$ by requiring $A \in SL_2(\mathbb{Z})$ acting on $\mathbb{H}$ by M_A. This says that for each $A \in SL_2(\mathbb{Z})$,

$$M_A = (M_{A^{-1}})^{-1}$$

and for $A, B \in SL_2(\mathbb{Z})$,

$$M_{AB} = M_A \circ M_B$$

check this!

Observe that

$$\operatorname{Im}\left(\frac{\alpha\tau+\beta}{\gamma\tau+\delta}\right)=\frac{D\operatorname{Im}\tau}{|\gamma\tau+\delta|^2}$$

where

$$D=\det\begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}.$$

Thus as long as $D>0$ we see that the Möbius map sends $\mathbb{H}$ to $\mathbb{H}$. This is the case for matrices in $SL_2(\mathbb{Z})$.

Recall our relations between τ,τ' with

$$A=\begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}, A'=\begin{bmatrix}\alpha' & \beta'\\ \gamma' & \delta'\end{bmatrix}.$$

We see that

$$M_{AA'}=Id$$

and this implies that

$$AA'=Id.$$

Thus we see that A,A' are invertible matrices with integer entries. In order that $M_A, M_{A'}$ take values in $\mathbb{H}$, A,A' should have positive determinants. This implies that A,A' are in $SL_2(\mathbb{Z})$.

Now we have a very nice way of describing the space of lattices. Recall that each lattice can be associated a $\tau\in\mathbb{H}$ and two lattices with τ,τ' are equivalent precisely when $\tau=M_A(\tau')$ for some $A\in SL_2(\mathbb{Z})$. Consider the quotient space $\mathbb{H}/SL_2(\mathbb{Z})$. Recall that we have an action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$. This quotient space is the space of orbits under this action.

Unlike the Λ action on $\mathbb{C}$, this $SL_2(\mathbb{Z})$ action on $\mathbb{H}$ is not abelian (since it is not an abelian group). We will find a fundamental domain for this action. This time, this domain is not a parallelogram. In fact, it is not even bounded!

Theorem 3.19. *Consider the set*

$$F=\{\tau\in\mathbb{H}:|\tau|>1,|\operatorname{Re}\tau|<1/2\}$$
$$\cup\{z:\operatorname{Re}z=-1/2,|z|\geq 1\}\cup\{z:\operatorname{Re}z\in[-1/2,0],|z|=1\}.$$

This is a fundamental domain of the action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$. This means that

1. *For each $\tau\in\mathbb{H}$, there is a matrix $A\in SL_2(\mathbb{Z})$ with $M_A(\tau)\in F$.*

2. For each $\tau \neq \tau'$ in the interior of F, no A is such that $M_A(\tau) = \tau'$. In fact, no A is such that $M_A(\tau) \in F$.

Proof. 1. Let $A =\in SL_2(\mathbb{Z})$ and

$$M_A(\tau) = \frac{a\tau + b}{c\tau + d}.$$

From here we see that

$$\operatorname{Im} M_A(\tau) = \frac{\operatorname{Im} \tau}{|c\tau + d|^2}.$$

Since c, d are integers, we see that there is at least one minimum of $|c\tau + d|$. We assume from the beginning that our A is taken with such a property. In particular, for any other $B \in SL_2(\mathbb{Z})$ we have

$$\operatorname{Im} M_B(\tau) \leq \operatorname{Im} M_A(\tau).$$

We define (high because it is the highest point in the orbit of τ under the action of $SL_2(\mathbb{Z})$)

$$\tau_{high} = M_A(\tau).$$

Next, observe that we can change the real part of τ_{high} by actions of $T(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$ for integers t. With a suitable t, we can translate τ_{high} into the band $\{\operatorname{Re} z \in [-1/2, 1/2]\}$. Thus, we assume from the beginning that $\operatorname{Re} \tau_{high} \in [-1/2, 1/2]$. We can now perform the reflection by the unit circle $R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ after which we obtain a point with imaginary part $\operatorname{Im} \tau_{high}/|\tau_{high}|^2$. Thus we should have

$$\operatorname{Im} \tau_{high} \geq \operatorname{Im} \tau_{high}/|\tau_{high}|^2$$

and this says $|\tau_{high}| \geq 1$. Thus we see that for each τ there is an $A \in SL_2(\mathbb{Z})$ so that $M_A(\tau) \in \overline{F}$. For the boundary case, one can achieve it by acting $T(\pm 1)$ or R. Moreover, $T(n), R$ are what we need for this. In fact, the group generated by $T(n), R$ is a subgroup of $SL_2(\mathbb{Z})$. We could have used this subgroup instead of $SL_2(\mathbb{Z})$.

2. We first have $|c\tau + d| \geq 1$ for c, d being integers (not all zeros). Next, if $M_A(\tau) \in \overline{F}$ then

$$1 \leq | - cM_A(\tau) + a| = \frac{1}{c\tau + d}.$$

Thus $|c\tau+d| = 1$. From here and the fact that $\tau \in \overline{F}$ we conclude that c, d can only be $0, \pm 1$. Replacing the roles of $\tau, M_A(\tau)$, i.e. we consider $M_{A^{-1}}M_A(\tau), M_A(\tau)$, we see that a, d, c, d must be all be 0 or ± 1. From here we see that τ must be one of the 'corner' point $(\pm 1 + \sqrt{3}i)/2$ or the point i.

Thus, F is a fundamental domain of the $SL_2(\mathbb{Z})$ action on $\mathbb{H}$. Moreover, we see that $T(n), R$ generate $SL_2(\mathbb{Z})$. □

3.11. j-invariant and other examples of modular forms.

Definition 3.20 (meromorphic modular form)**.** *Let $k \in \mathbb{Z}$. A modular form of weight k is a meromorphic function*

$$f : \mathbb{H} \cup \infty \to \mathbb{C} \cup \{\infty\}$$

such that for all $A \in SL_2(\mathbb{Z})$,

$$f(M_A(\tau)) = (c\tau + d)^k f(\tau)$$

and f has only finitely many poles in $\mathbb{H}/SL_2(\mathbb{Z})$.

In particular, f does not have an essential singularity at ∞. Next, we remark that $\mathbb{Z}/SL_2(\mathbb{Z})$ is not compact although it has finite Lebesgue measure. For this reason, it is not automatic that f should have finitely many poles in the fundamental domain as we do not have the compactness argument as in the torus case.

If $k = 0$, then f is invariant under the action of $SL_2(\mathbb{Z})$. That is to say, f is well defined as a function over $\mathbb{H}/SL_2(\mathbb{Z})$. In this case, we say that f is a modular function for $SL_2(\mathbb{Z})$ (like our elliptic functions for Λ).

We now provide examples.

Example 3.21. *Let $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$ be a lattice. The Eisenstein series of weight $k \geq 3$ is*

$$G_k(\tau) = \sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{\lambda^k}.$$

It is easy to show that for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z})$,

$$G_k(M_A(\tau)) = (c\tau + d)^k G_k(\tau).$$

Recall that for each Λ (or τ) we have an elliptic curve

$$y^2 = 4x^3 - 60G_4x - 140G_6$$

for simplicity, we define

$$g_2(\tau) = 60G_k(\tau), g_3 = 140G_6(\tau).$$

Then the discriminant of proportional (up to some powers of 2) to $\Delta(\tau) = g_2(\tau)^3 - 27g_3(\tau)^2$. Thus Δ is a modular form of weight 12. Then $j(\tau) = g_3(\tau)^3/\Delta(\tau)$ is a modular function.

Definition 3.22. *For lattice $\Lambda = \mathbb{Z}+\mathbb{Z}\tau$. The elliptic curve $y^2 = 4x^3 - g_2(\tau)x - g_3$ has j-invariant*

$$j(\tau) = g_2(\tau)^3/\Delta(\tau)$$

where $\Delta(\tau) = g_2(\tau)^3 - 27g_3(\tau)^2$. Since we know that $4\wp^3 - g_2\wp - g_3 = (\wp - e_1)(\wp - e_2)(\wp - e_3)$ with different e_1, e_2, e_3, we see that $\Delta(\tau) \neq 0$.

Suppose that for any chosen $j \in \mathbb{C}$, we have $j(\tau) = j$ for some τ. Then for the scaled lattice $\mu\Lambda$ we have

$$g_3(\mu\Lambda) = \mu^{-3}g_3(\Lambda).$$

Thus, we can choose μ to make $g_3(\mu\Lambda) = b$ for any chosen b. What we can do here is that for each fixed value j, a, b such that $j = a^3/(a^3 - 27b^2)$, we can find a lattice Λ such that $g_3(\Lambda) = b$ and $j(\Lambda) = j$. Then we must have $g_2(\lambda)^3 = a^3$. In order to have $g_2(\Lambda) = a$ we can replace Λ by some $\mu\Lambda$ with μ is a root of unity of order 3. Such a change will not change the value of Δ and b. However, g_2 is then changed to $\mu^{-2}g_2$ and as μ ranges over roots of unity of order three, μ^{-2} ranges over the same set. This implies that we can achieve that $g_2(\Lambda) = a$. We have proved the following result.

Lemma 3.23. *If $j : \mathbb{H} \to \mathbb{C} \cup \{\infty\}$ is surjective, then for each $a, b \in \mathbb{C}$ with $a^3 - 27b^2 \neq 0$ there is a lattice Λ such that the elliptic curve w.r.t Λ is deifned via the equation*

$$y^2 = 4x^2 - ax - b.$$

Remark 3.24. *$a^3 - 27b^2 \neq 0$ is crucial for the previous argument as otherwise $j = sth/0$ which is not nicely defined.*

Observe that $\lim_{\operatorname{Im}\tau\to\infty} G_k(\tau) = 2\zeta(k)$ where $\zeta(k) = \sum_{n=1}^{\infty} n^{-k}$. We only consider k being even integers. In those cases, $\zeta(k)$ is known. In particular, we have $(\zeta(4) = \pi^4/90, \zeta(6) = \pi^6/945)$

$$\lim_{\operatorname{Im}\tau\to\infty} \Delta(\tau) = (120\zeta(4))^3 - 27(280\zeta(6))^2 = 0.$$

Modular forms with vanishing $\lim_{\operatorname{Im}\tau\to\infty}$ are cusp forms.

Since $g_3(\tau)$ has finite value for $\operatorname{Im}\tau \to \infty$, we have $\lim_{\operatorname{Im}\tau\to\infty} j(\tau) = \infty$, i.e. ∞ is a pole of j.

Theorem 3.25. *$j : \mathbb{H} \to \mathbb{C} \cup \{\infty\}$ is surjective.*

Proof. j is a modular function which can be seen as a meromorphic function on $\mathbb{H}/SL_2(\mathbb{Z})$, attaching ∞, we obtain the Riemann sphere. The fact that $\lim_{\operatorname{Im}\tau\to\infty} j(\tau) = \infty$ shows that j is well-defined at ∞. Thus j induces a meromorphic function between Riemann spheres. In more details, we have a meromorphic bijection $q : \overline{C} \to \mathbb{H}/SL_2(\mathbb{Z}) \cup \{\infty\}$, then $\tilde{j} = j \circ q$ is meromorphic $\overline{C} \to \overline{\mathbb{C}}$. If $\tilde{j}$ is not surjective, then we have find a point $w \in \mathbb{C}$ not in the image of $\tilde{j}$ (∞ is attained) and we look at

$$g(z) = \frac{1}{\tilde{j}(z) - w}.$$

If a neighbourhood of w is disjoint from the image of $\tilde{j}$ then g is bounded, analytic and therefore a constant. Thus $\tilde{j}$ must be constant. This can be easily checked to be not the case, e.g. we know that ∞ is in the image of $\tilde{j}$ so $\tilde{j}$ and also j must be constantly ∞ and this is clearly not the case.

Thus we see that the image of $\tilde{j}$ must be dense in $\overline{C}$. Let $w \in \mathbb{C}$, we see that there is a sequence $z_1, z_2, \cdots \in \overline{C}$ with $\tilde{j}(z_i) \to w$. With out loss of generality we assume that $\lim_i z_i = z_\infty$ and we see that $\tilde{j}(z_\infty) = w$. This shows that $\tilde{j}$ is surjective and thus j is surjective.

In fact, any meromorphic map from any compact Riemann surface to the Riemann sphere is either constant or surjective. □

The mysterious function q can be chosen to be j itself! For this, we need to show that j is also injective. For this, we need the following result which can be seen as the Liouville Theorem for modular forms (a more precise statement should be the Riemann-Roch for modular forms).

Theorem 3.26 ($k/12$-formula)**.** *Let f be a modular form of weight k. Suppose that f has no zeros or poles at roots of unity of order 6 modulo the action of $SL_2(\mathbb{Z})$. Then we have*

$$\sum_{zeros} 1 - \sum_{poles} 1 = \frac{k}{12}.$$

Remark 3.27. *Here, zeros, poles are considered in $\mathbb{H}/SL_2(\mathbb{Z}) \cup \{\infty\}$. Do not forget about zeros or poles at the ∞! What if we want to consider zeros and poles at the roots of unity of order 6? We need to consider three points, $P_1 = e^{\pi i/3}, P_2 = i, P_3 = e^{2\pi i/3}$. For P_1, P_3 we associate a weight $1/3$, for P_2 we put weight $1/2$. So the general formula should be*

$$\sum_{zeros} weights - \sum_{poles} weights = \frac{k}{12}$$

for all points other than P_1, P_2, P_3, their weights are all one.

Proof. We do not prove this result in detail. As for Liouville Theorems of torus, we perform line integrals along the boundary of (any) fundamental domain. In our case, we choose the fundamental domain to be as in Theorem 3.19. We consider

$$\int_{\partial F} g(d)dz$$

where f is a modular form of weight k and $g = f'/f$. The integrals along the two infinite lines cancel as a translation by $+1$ is represented by $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Therefore the $(c\tau + d)^k$ part of the definition of modular form is simply one. Thus f is $\mathbb{Z}$-periodic and so is g.

We only need to consider the integral along the arc. This arc is symmetric by the action $\tau \to -1/\tau$. For g we simply have

$$g(-1/\tau) = \tau^2 g(\tau) + k\tau.$$

This is because $z \to -1/z$ is represented by $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Because of this addition $k\tau$ term, the integral on the arc is not zero. The arc can be parameterised as

$Arc(t) = e^{\pi it}, t \in [1/3, 2/3]$. Thus we see that

$$\begin{aligned}
-\int_{Arc} g(z)dz &= \int_{1/3}^{2/3} g(Arc(t))e^{\pi it}(\pi i)dt \\
&= \int_{1/3}^{1/2} g(Arc(t))e^{\pi it}(\pi i)dt + \int_{1/2}^{2/3} g(Arc(t))e^{\pi it}(\pi i)dt \\
&= \int_{1/3}^{1/2} g(Arc(t))e^{\pi it}(\pi i)dt + \int_{1/2}^{2/3} e^{-2\pi it} g(Arc(1-t))e^{\pi it}(\pi i)dt \\
&+ \int_{1/2}^{2/3} ke^{-\pi it}e^{\pi it}(\pi i)dt \\
&= \int_{1/3}^{1/2} g(Arc(t))e^{\pi it}(\pi i)dt - \int_{1/2}^{1/3} g(Arc(t))e^{-\pi i(1-t)}(\pi i)dt \\
&+ \int_{1/2}^{2/3} k(\pi i)dt \\
&= \pi i \frac{k}{6}.
\end{aligned}$$

This gives

$$\sum_{zeros} 1 - \sum_{poles} 1 = \frac{1}{2\pi i}\frac{k\pi i}{6} = \frac{k}{12}.$$

There is one technical point. We cannot use Cauchy's theorem on an unbounded domain. Instead, we need to enclose our domain with horizontal lines with larger and larger imaginary parts and take the limit. One can check that the integrals on those horizontal lines contribute to the zero/pole of f at ∞.

For roots of unity of order 6, we can enclose them by arcs. For P_1, P_3 those are 1/6-arcs and for P_2, it is 1/2-arc. Since P_1, P_3 are congruent modulo the $SL_2(\mathbb{Z})$ action those two points will join the force. A careful analysis of the integrals on those arcs will give us the weights mentioned in the remark. We omit the full detail. □

Theorem 3.28. *j is injective.*

Proof. j is a modular function (form of weight 0). Since $|j(\tau)| < \infty$ for $\tau \in \mathbb{H}$ we see that j has one pole at ∞. As $j = g_3^2/\Delta$, we see that the order of this pole is the order of zero of Δ at ∞. However, Δ is a modular form of weight 12 without any poles. Thus for Δ we see that

$$\sum_{zeros} 1 = 1$$

so that ∞ is the only simple zero of Δ. This implies that ∞ is the only simple pole of j. For each $w \in \mathbb{C}$, we consider

$$f(\tau) = j(\tau) - w.$$

Then f has one simple pole. Also, for f,

$$\sum_{zeros} 1 - \sum_{poles} 1 = 0.$$

Thus, we see that f has only one simple zero or a multiple zero at a root of unity of order 6 since there is no way to split one into non-trivial multiples of $1/2$ and $1/3$. This says that $j(\tau) = w$ has exactly one solution for $\tau \in \mathbb{H}/SL_2(\mathbb{Z})$. This is the injectivity we are looking for. □

Up to now, we showed that for each regular elliptic curve over $\mathbb{C}$, we can always find a lattice that is associated with this elliptic curve. If we replace this lattice with an equivalent lattice, then we do not change the j-invariant of the elliptic curve. However, we do change the coefficients g_2, g_3. Conversely, if two lattices are not equivalent, then the j-invariant of their elliptic curves are not the same. Later, we shall have a closer look at this correspondence.

3.12. **Complex multiplication and isogeny: the tale of the 'almost integer' $e^{\pi\sqrt{163}}$.** Some lattices are more symmetric than others. For example, the Gaussian lattice ($\tau = i$) is symmetric under the multiplication of i. Such a symmetry does not hold for the lattice with, say, $\tau = e + \pi i$. Thus we expect the elliptic curve assigned with the Gaussian lattice also has some symmetry. This is what we explore now.

Definition 3.29. *Let Λ be a lattice. The symmetries of Λ are defined to be*

$$Sym(\Lambda) = \{\alpha \in \mathbb{C} : \alpha\Lambda \subset \Lambda\}.$$

The set $Sym(\Lambda)$ is clearly an additive subgroup of $\mathbb{C}$. In fact, more is true.

Theorem 3.30. *For each lattice Λ, $Sym(\Lambda)$ is a ring. In fact, this ring contains $\mathbb{Z}$ as a subring and is contained in a ring of integers of a quadratic number field.*

Proof. Let $\alpha, \beta, \gamma \in Sym(\Lambda)$. We see that $\alpha(\beta+\gamma)\Lambda$ is the same as $(\alpha\beta+\alpha\gamma)\Lambda$. For each $\lambda \in \Lambda$, we see that

$$(\alpha\beta+\alpha\gamma)\lambda \in \Lambda$$

as both $\alpha\beta\lambda$ and $\alpha\gamma\lambda$ are in Λ. This shows that $Sym(\Lambda)$ is in fact a subring of $\mathbb{C}$.

It is simple to see that $\mathbb{Z}$ is a subring of $Sym(\Lambda)$. To show that $Sym(\Lambda)$ is contained in a quadratic integer ring, we let w_1, w_2 be a basis of Λ. For $\alpha \in Sym(\Lambda)$, we have

$$\alpha w_1 = aw_1 + bw_2,$$
$$\alpha w_2 = cw_1 + dw_2$$

for some integers a, b, c, d. We see that

$$\begin{bmatrix} a-\alpha & b \\ c & d-\alpha \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0. \end{bmatrix}$$

This implies that

$$\det \begin{bmatrix} a-\alpha & b \\ c & d-\alpha \end{bmatrix} = 0$$

and from here we see that α is a quadratic integer. Moreover, $\mathrm{Tr}(\alpha) = a+d$ and $\mathrm{Norm}(\alpha) = ad-bc$. Next, for a different $\alpha' \in Sym(\Lambda)$ we see that $\alpha, \alpha', \alpha+\alpha'$ are all in possibly different quadratic number fields. Suppose that $\alpha \in \mathbb{Q}(\sqrt{d})$ and $\alpha \in \mathbb{Q}(\sqrt{d'})$ for squarefree numbers d, d'. Then if $d \neq d'$, the only chance that $\alpha+\alpha'$ is quadratic is that at least one of α, α' should be in $\mathbb{Z}$. This implies that each pair α, α', must be in the same quadratic field. Thus we proved that $Sym(\Lambda)$ is a subring of a ring of integers of a quadratic number field. □

Definition 3.31. *For a number field $\mathbb{K}$, an order is a subring R of $\mathcal{O}_{\mathbb{K}}$ (the ring of integers in $\mathbb{K}$) which is finitely generated as a $\mathbb{Z}$-module and $R \otimes_{\mathbb{Z}} \mathbb{Q} = \mathbb{K}$.*

In this section, the tensor product $\otimes$ will be almost always $\otimes_{\mathbb{Z}}$. Without further notice, we just write $\otimes$ instead of $\otimes_{\mathbb{Z}}$.

Theorem 3.32. *Either $Sym(\Lambda) = \mathbb{Z}$ or else $Sym(\Lambda)$ is an order of a quadratic number field.*

Proof. We have to show that if $\mathbb{Z}$ is strictly contained in $R = Sym(\Lambda) \subset \mathcal{O}_{\mathbb{K}}$ for a quadratic number field, then $R \otimes \mathbb{Q}$ is $\mathbb{K}$. We know that $\mathbb{R} \cap Sym(\Lambda) = \mathbb{Z}$ because rational integers are in $\mathbb{Z}$. Thus we see that $Sym(\Lambda)$ must contain some non-real elements. We know that $\mathbb{K}$, as a quadratic number field, has dimension two as a $\mathbb{Q}$-linear space. Thus we have

$$[\mathbb{K} : \mathbb{Q}] = [R \otimes \mathbb{Q} : \mathbb{Q}] = 2$$

this implies that

$$[\mathbb{K} : R \otimes \mathbb{Q}] = 1$$

and this implies that

$$\mathbb{K} = R \otimes \mathbb{Q}.$$

Supplement (in case it is not clear from the context): Throughout the argument, we have identified $R \otimes \mathbb{Q}$ as a subspace of $\mathbb{K}$ via

$$(r, q) \in R \times \mathbb{Q} \to qr \in \mathbb{K}.$$

This map is clearly $\mathbb{Z}$-bi-linear and just induces a map ψ from $R \otimes \mathbb{Q}$ to $\mathbb{K}$. The map ψ is $\mathbb{Q}$-linear. We have to check that ψ is injective which is straightforward if we use the fact that $\mathbb{Q}$ is a flat $\mathbb{Z}$-module. We can also check this directly (in fact, following this step we can prove the flatness of $\mathbb{Q}$). For this, let $(r_i, q_i), i \in F$ where F is a collection of some finite indices. Suppose that as an element in $\mathbb{K}$,

$$(*) \qquad \sum_i q_i r_i = 0.$$

This is precisely saying that $\psi(\sum_i r_i \otimes q_i) = 0$. Now by multiplying a suitable non-zero integer M in (*), we can find integers z_i so that

$$\sum_i z_i r_i = 0.$$

Therefore we see that

$$M(\sum_i r_i \otimes q_i) = \sum_i r_i \otimes z_i = (\sum_i z_i r_i) \otimes 1 = 0.$$

This says that $\sum_i r_i \otimes q_i$ is a torsion element. In particular, this says that $R \otimes \mathbb{Q}$ contains non-trivial torsions if $\sum_i r_i \otimes q_i$ is not the zero elements. Among all the non-trivial torsion elements. Find one, say a, with the smallest possible expansion, i.e.

$$a = \sum_{i \in I} r_i \otimes q_i$$

where $\#I > 0$ is as small as possible. This is possible to be done because all elements in $R \otimes \mathbb{Q}$ can be written with a sum of finite terms of elements like $r \otimes q$. There is a non-zero integer N so that

$$0 = Na = N \sum_i r_i \otimes q_i.$$

(as we did above) We can find a non-zero integer N' so that $N'q_i$ are integers. Then $NN'a = 0$. Therefore

$$0 = NN'a = N(\sum_i r_i \otimes N'q_i) = N(\sum_i N'q_i r_i) \otimes 1.$$

This says that $\#I$ must be one and $a = r \otimes q$ for some $(r, q) \in R \times \mathbb{Q}$. Then $Na = 0$ means that $r \otimes Nq = 0$. Then by finding another integer if necessary, we can assume that $Nq \in \mathbb{Z} \setminus \{0\}$. Thus we see that $Nqr \otimes 1 = 0$ and this must happen if $Nqr = 0$. Since R is clearly torsion-free because the addition structure are those from $\mathbb{C}$ and $\mathbb{C}$ is obviously torsion-free. Thus we must have $r = 0$. Thus we must have $a = 0$. Thus $R \otimes \mathbb{Q}$ is torsion-free. Thus ψ must be injective.

□

Definition 3.33. *Let Λ, Λ' be two lattices:*

$$Hom(\Lambda_1, \Lambda_2) = \{\alpha \in \mathbb{C} : \alpha\Lambda_1 \subset \Lambda_2\}.$$

If there is an α such that $\alpha\Lambda_1 = \Lambda_2$ then we say that Λ_1, Λ_2 are equivalent.

$Sym(\Lambda)$ tells us how Λ is kept somehow unchanged under the multiplications of some complex numbers. Such symmetries should be seen from the elliptic curve w.r.t. Λ.

Definition 3.34. *Let $E = E(\mathbb{C})$ be an elliptic curve. An element ψ of $End(E)$ is a rational map*

$$\psi : E \to E$$

which is at the same time a group homomorphism. Such a map ψ is an isogeny of E. More generally, given two elliptic curves E_1, E_2. A rational map $\psi : E_1 \to E_2$ which is also a group homoporphism is an isogeny from E_1 to E_2.

Theorem 3.35. *Let E_1, E_2 be two regular elliptic curves over $\mathbb{C}$. Let ψ be a non-constant rational map from E_1 to E_2 mapping ∞ to ∞. Then ψ is a group homomorphism and thus it is an isogeny.*

Remark 3.36. *This result holds for elliptic curves over general fields. However, as one might have guessed, we need to replace all pole-zero counting arguments (e.g. Liouville's Theorems) with Riemann-Roch.*

Proof. Let Λ_1, Λ_2 be the lattices associated to E_1, E_2. Then let ϕ_1, ϕ_2 be the corresponding Weierstrass maps (from tori to elliptic curves). Next, consider the map

$$T = \phi_2^{-1} \circ \psi \circ \phi_1 : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2.$$

The fact that ψ is rational implies that T is meromorphic. We can extend T as $T : \mathbb{C} \to \mathbb{C}/\Lambda$ by periodicity. Then we can lift this map and obtain

$$T : \mathbb{C} \to \mathbb{C}, T(0) = 0.$$

This can be done via path extension and the simply connectedness of $\mathbb{C}$. T is now a meromorphic map but it may not be periodic w.r.t Λ_1. As it is lifted from a map to $\mathbb{C}/\Lambda_2$, we see that there are no poles of T in $\mathbb{C}$. Therefore T is analytic on $\mathbb{C}$. Let $\lambda \in \Lambda_1$. Consider the function

$$T_\lambda(z) = T(z + \lambda) - T(z).$$

This function is continuous. As T is lifted from a function between tori, the image set of T_λ is contained in Λ_2 which is discrete. Thus T_λ must be constant. Thus T'_λ is zero. Thus T' is elliptic w.r.t Λ_1. Since T' is also analytic, we see that T' must be constnat. This implies that T is linear. Since $T(0) = 0$ we see that $T(z) = \alpha z$ for some $\alpha \in \mathbb{C}^*$. This linear map must send Λ_1 to a subset of Λ_2. Therefore we see that

$$\alpha \in Hom(\Lambda_1, \Lambda_2).$$

Thus $T : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$ is a group homomorphism. Next, ϕ_1, ϕ_2 establish group isomorphisms between $\mathbb{C}/\Lambda_i$ and E_i for $i \in \{1, 2\}$. Therefore we see that

$$\psi = \phi_2 \circ T \circ \phi_1^{-1}$$

establishes a group homomorphism from E_1 to E_2. Since it is rational, we see that it is an isogeny. □

In particular, we see that for an elliptic curve E over $\mathbb{C}$,

$$End(E) \subset Sym(\Lambda).$$

The set $End(E)$ is naturally a ring whose multiplication is the composition of functions and whose addition is the pointwise group addition of functions. For example, if $\phi_1, \phi_2 \in End(E)$ we have

$$(\phi_1 + \phi_2)(P) = \phi_1(P) + \phi_2(P)$$

where the latter '+' is the group operation for E. W.r.t. this addition, we see that $End(E)$ is a subgroup of $Sym(\Lambda)$. Similarly, compositions of maps in $End(E)$ are multiplications in $Sym(\Lambda)$. Thus $End(E) \cong Sym(\Lambda)$ as rings.

Let Λ be a lattice. Observe that for each $\alpha \neq 0$, the ring $Sym(\alpha\Lambda)$ is the same as $Sym(\Lambda)$. Thus equivalent lattices give the same Sym ring and this gives a well-defined map

$$K : \text{Equivalent classes of lattices} \to Sym \text{ rings}.$$

In particular, we can rotate and scale Λ so that $1 \in \Lambda$. W.o.l.g. we have $\Lambda = \mathbb{Z}+\mathbb{Z}\tau$ for some $\tau \in \mathbb{H}$. Let $\alpha \in Sym(\Lambda) \subset \mathcal{O}_{\mathbb{K}}$. Then we have $\alpha \cdot 1 \in \Lambda$. Namely, for some integers a, b,

$$\alpha = a + b\tau.$$

Therefore $b\tau \in \mathcal{O}_{\mathbb{K}}$. As a $\mathbb{Z}$-module, $R = Sym(\Lambda)$ is a submodule of $\mathcal{O}_{\mathbb{K}}$ with finite index, say k. This is because $R \otimes \mathbb{Q} = \mathbb{K} = \mathcal{O}_{\mathbb{K}} \otimes \mathbb{Q}$. Then $kb\tau \in k\mathcal{O}_{\mathbb{K}}$ which is contained in $R = Sym(\Lambda)$. From here we see that the lattice

$$\Lambda' = kb\Lambda \subset Sym(\Lambda).$$

Clearly, Λ' is a subgroup. Next, for each $\alpha \in Sym(\Lambda)$, and $\lambda' \in \Lambda'$ we see that for $\lambda = (kb)^{-1}\lambda'$,

$$\alpha\lambda' = kb\alpha\lambda \in kb\alpha\Lambda \subset kb\Lambda = \Lambda'.$$

Thus Λ' is an ideal of $Sym(\Lambda)$. Thus if $K([\Lambda]) = R$ then $[\Lambda]$ can be represented by an ideal of R. Namely, for some non-zero γ, $\gamma\Lambda \subset R$ as an ideal. Of course, there are other choices of γ. Thus there are different ideals we can obtain. Those ideals are equivalent in the sense that they differ only by multiplying a non-zero

complex number. In terms of algebraic number theory, this is the equivalence class defining the class group.

Lemma 3.37. *Let Λ be a lattice. Let $R = Sym(\Lambda)$. Then Λ is equivalent to an ideal of R.*

Proof. We proved this already. □

In fact, in our situation, ideals of R are themselves lattices.

Lemma 3.38. *Let R be an order of a quadratic number field $\mathbb{K}$. Let Λ be an ideal of R. Then Λ is a lattice. Consider the elliptic curve E w.r.t. Λ. Let $R' = End(E) = Sym(\Lambda)$. Then $R \subset R'$ as a subring.*

Proof. We need to check that $R\Lambda \subset \Lambda$ which is trivial as Λ is an ideal of R. We also need to check that Λ is a lattice. First, as we have $R \subset \mathbb{C}$, Λ is clearly an additive group. As $R \subset \mathcal{O}_{\mathbb{K}}$ which is discrete, Λ must be discrete(*). Next, we need to check that the $\mathbb{Z}$-rank of Λ is two. We have for each $\lambda \in \Lambda$,

$$R\lambda \subset \Lambda \subset R.$$

As $R\lambda, R$ are $\mathbb{Z}$-modules of rank two ($R \otimes \mathbb{Q}$ has $\mathbb{Q}$-dimension two). We see that Λ must have rank two. This proves the result.

Supplement: For (*), a general fact is that $\mathcal{O}_{\mathbb{K}}$ for a general number field $\mathbb{K}$ is always discrete w.r.t. a certain topology. For each number field $\mathbb{K}$, there are $\deg \mathbb{K}$ many field embeddings $\sigma_i : \mathbb{K} \to \mathbb{C}$. We have the map

$$\Delta : \mathbb{K} \to \mathbb{C}^{\deg \mathbb{K}}$$

as

$$\Delta(x) = (\sigma_i(x)).$$

Some of the embeddings are real (with real images), and some of the embeddings are not real (in this case the conjugate of those embeddings are also embeddings). The set $\Delta(\mathcal{O}_{\mathbb{K}})$ is a discrete subset of $\mathbb{C}^{\deg \mathbb{K}}$. For more details, check the topic Geometry of Numbers. □

Let $\mathbb{K}$ be a quadratic number field and let R be an order. For each class of ideals of R (i.e. lattices Λ with $R\Lambda \subset \Lambda$), we find uniquely a class of lattices (ideals of R are themselves lattices). Such correspondence is a bijection. For

example, if $R = \mathcal{O}_{\mathbb{K}}$, then each ideal of R corresponds to a lattice Λ which corresponds to an elliptic curve E with $End(E) = R$. How many non-isogenious curves E have $End(E) = R$? Answer: the class number $h_{\mathbb{K}}$. In this way, the theory of isogeny classes of elliptic curves is related to the class number of quadratic number fields. Check the excellent note by S-W Zhang (https://web.math.princeton.edu/ shouwu/publications/elc.pdf).

Theorem 3.39. *Let $\mathbb{K}$ be a quadratic number field. Let $\tau \in \mathbb{H}$ and consider the lattice $\Lambda = \mathbb{Z} + \mathbb{Z}\tau$. Then $Sym(\Lambda) \subset \mathcal{O}_{\mathbb{K}}$ is an order if and only if $\tau \in \mathbb{K}$. In this case, $j(\tau)$ is algebraic.*

Proof. If $Sym(\Lambda)$ is an order, then it contains non-real elements. If $\alpha \in Sym(\Lambda) \subset \mathcal{O}_{\mathbb{K}}$ is not a real number, then $\alpha = a+b\tau$ for integers a, b where $b \neq 0$. This implies that $\tau \in \mathbb{K}$.

Conversely, if $\tau \in \mathbb{K}$ is not real, then then for an integer $k > 0$ we have $k\tau \in \mathcal{O}_{\mathbb{K}}$. W.o.l.g. we can assume that $\tau \in \mathcal{O}_{\mathbb{K}}$. Suppose that $\alpha\Lambda \subset \Lambda$, then there are integers a, b with

$$\alpha = a + b\tau.$$

This implies that $\alpha \in \mathcal{O}_{\mathbb{K}}$. We see that $Sym(\Lambda) \subset \mathcal{O}_{\mathbb{K}}$. To show that it is an order, we need to find a non-real α and integers a, b, c, d so that

$$\alpha = a + b\tau, \alpha\tau = c + d\tau.$$

We can set $a = 0, b = 1$ and $\alpha = \tau$. Then

$$\alpha\tau = \tau^2.$$

Since $\tau \in \mathcal{O}_{\mathbb{K}}$, we see that for some integers c, d,

$$\tau^2 = c + d\tau.$$

This concludes the proof.

Now, we show that $j(\tau)$ is algebraic.

Let σ be a field automorphism of the extension $\mathbb{C}/\mathbb{Q}$, i.e. σ is a field isomorphism which fixes $\mathbb{Q}$. For example, $z \to \overline{z}$. Another slightly more involved example is as follows: Let ξ be a transcendental number. Let ξ' be another transcendental number so that $\mathbb{Q}[\xi, \xi']$ is isomorphic to the polynomial ring $\mathbb{Q}[X, Y]$. We can set $P(\xi) \to P(\xi')$ for each $\mathbb{Q}$-polynomial P. We can extend this map to a field

mapping between algebraic closure of $\mathbb{Q}(\xi)$ and of $\mathbb{Q}(\xi')$. By using Zorn's lemma, we can extend this map further to $Aut(\mathbb{C})$.

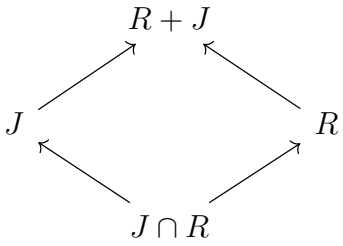
Returning to $j(\tau)$. Let $E : y^2 = 4x^3 - ax - b$ be the curve associated with the lattice $\mathbb{Z} + \mathbb{Z}\tau$. Then $\sigma(j(\tau))$ is the j-invariant of $E^\sigma : y^2 - 4x^3 - \sigma(a)x - \sigma(b)$. As groups, $E \cong E^\sigma$ and this implies that $End(E) \cong End(E^\sigma)$ as rings. We also know that $End(E), End(E^\sigma)$ are subsets of $\mathbb{C}$ which are orders of (possibly different) quadratic fields. Let $\psi : End(E) \cong End(E^\sigma)$ be the map establishing the ring isomorphism. Then $\psi(1) = 1$. Each $\alpha \in End(E)$ is an element of $\mathcal{O}_\mathbb{K}$. Thus α as well as $\psi(\alpha)$ satisfy the same quadratic equation. This implies that each $\alpha' \in End(E^\sigma)$ is also an element of $\mathcal{O}_\mathbb{K}$. As orders, $End(E)$ and $End(E^\sigma)$ are isomorphic. As subsets of $\mathbb{C}$, they may not be the same. Let w_1, w_2 $\mathbb{Z}$-generate $End(E)$. Then $End(E^\sigma)$ can be generated by one of the following sets

$$\{w_1, w_2\}, \{w_1, \overline{w_2}\}, \{\overline{w_1}, \overline{w_2}\}, \{\overline{w_1}, w_2\}.$$

Let Λ^σ correspond to E^σ. Then Λ^σ must be equivalent to an ideal I of an order containing $End(E^\sigma)$. Then $I \cap End(E^\sigma)$ is an ideal of $End(E^\sigma)$. In other words, $I \cap End(E^\sigma)$ is a sublattice of I. On the other hand, $I\mathcal{O}_\mathbb{K}$ is an ideal of $\mathcal{O}_\mathbb{K}$. We have the following inclusion relations of lattices

$$I \cap End(E^\sigma) \subset I \subset I\mathcal{O}_\mathbb{K}.$$

Let us write $R = End(E^\sigma)$ and $J = I\mathcal{O}_\mathbb{K}$. Then consider the following commutative diagram as $\mathbb{Z}$-modules



Then we count the index $[R + J : J \cap R]$ in two different ways and use the third isomorphism theorem to see that

$$[\mathcal{O}_\mathbb{K} : R] \geq [R + J : R] = [J : J \cap R].$$

In conclusion, Λ^σ is equivalent to a suplattice of an ideal $I \cap End(E^\sigma)$ and at the same time it is a sublattice of $I\mathcal{O}_\mathbb{K}$. For each fixed $I\mathcal{O}_\mathbb{K}$, there are only finitely

many sublattices with a bounded co-volume. Thus there are only finitely many such lattices Λ^σ modulo equivalence. Also, there are only finitely many equivalent classes of ideals $I\mathcal{O}_\mathbb{K}$. Thus there are at most finitely many possible Λ^σ modulo equivalence.

We proved that as σ ranges over $Aut(\mathbb{C})$, Λ^σ ranges over finitely many equivalence classes of lattices. Thus $\sigma(j(\tau))$ attains only finitely many values. This implies that $j(\tau)$ is algebraic. In fact, if $x = j(\tau)$ is transcendental, $\mathbb{Q}[x] \to \mathbb{Q}[x+1], x \to x+1$ is a field isomorphism. Such a field isomorphism extends to an element in $Aut(\mathbb{C})$ mapping x to $x+1$. Similarly, for each integer k, there is some map in $Aut(\mathbb{C})$ mapping x to $x+k$. Thus x has infinitely many images under maps in $Aut(\mathbb{C})$. Of course, there are so many other ways to make this argument work. □

For our curve $y^2 = 4x^3 - g_2(\tau)x^2 - g_3(\tau)$, it turns out that $1728j(\tau)$ is an algebraic integer rather than just an algebraic number. Proving such a result involves the theory of modular forms and we omit it. See Washington Section 10.3 for a direct proof which uses without mentioning the theory of modular forms. In fact, it is possible to explicitly compute the minimal polynomial of $j(\tau)$.

With $q = e^{2\pi i\tau}$, we have the following Fourier expansion

$$1728j(\tau) = \frac{1}{q} + 744 + 196884q + 21493760q^2 + \dots. \tag{*}$$

The point is that each coefficient of those q-powers is an integer. Let us see where does 1728 come from. First, we have (by writing cot as a fraction of sums of exponential functions and then using geometric sums)

$$\cot(\pi\tau) = i - 2i\sum_{j=0}^{\infty} q^j.$$

We also have (By observing that the LHS, RHS have exactly the same poles and the same residues and they are both 1-periodic. Then check that $LHS - RHS$ stays bounded as $\operatorname{Im}\tau \to \infty$ for $\operatorname{Re}\tau$ in a arbitrarily fixed bounded interval)

$$\cot(\pi\tau) = \frac{1}{\pi\tau} + \frac{1}{\pi}\sum_{j=1}^{\infty}\left(\frac{1}{\tau - n} + \frac{1}{\tau + n}\right).$$

We can differentiate ($2k$ times) the two expressions for $\cot(\pi\tau)$ and obtain for following equation

$$\sum_{j\geq 1}(2\pi i)^{2k}j^{2k-1}q^j = (-1)^{2k}(2k-1)!\sum_{n\in\mathbb{Z}}\frac{1}{(\tau+n)^{2k}}.$$

The LHS is a Fourier sum and the RHS is a partial fraction sum. Manipulations sum expressions allows us to obtain Fourier expansions of many other modular forms. In particular, we have for integers $k \geq 2$

$$G_{2k} - 2\zeta(2k) = 2\frac{(2\pi i)^{2k}}{(2k-1)!}\sum_{m\geq 1}\sum_{j\geq 1}j^{2k-1}q^{mj} = 2\frac{(2\pi i)^{2k}}{(2k-1)!}\sum_{n\geq 1}\sigma_{2k-1}(n)q^n,$$

where $\sigma_{2k-1}(n) = \sum_{d|n} d^{2k-1}$. From here, one can obtain

$$j(\tau) = (60G_4)^3/((60G_4)^3 - 27(140G_6)^2)$$

as a sum of powers of q. It is not hard to check directly that some coefficients multiplying 1728 are integers. For example, the coefficient of $1/q$ for $j(\tau)$ is precisely $1/1728$. This direct method can be used to show that all the other coefficients are integers. However, as one might have guessed, this approach is too complicated. For example, it is tedious to directly compute $(60G_4)^3 - 27(140G_6)^2$. A better way to show this is via the theory of modular forms. We sketch the ideas. First, consider the discriminant function ($g_2 = 60G_4, g_3 = 140G_6$)

$$\Delta(\tau) = g_2^3 - 27g_3^2.$$

This is a modular form of weight 12. Consider the following magical function with integer coefficients

$$\Delta_m(\tau) = (2\pi)^{12}q\prod_{k\geq 1}(1-q^k)^{24}.$$

It can be checked that Δ_m is a modular form with weight 12 as well. A difficult step is to check that under $\tau \to -1/\tau$, $\Delta_m(q)$ transforms correctly. To do this, one can use Jacobi Theta series. More details can be obtained in books/lectures on Modular Forms. After this, we see that Δ/Δ_m is an analytic modular form of weight 0. Then we can use the $k/12$-formula to conclude that Δ/Δ_m should be

constant. Then matching their q coefficient allows us to see that this constant is 1. Our task is almost over. Observe that

$$\frac{1}{1-q^k} = \sum_{j\geq 0} q^k.$$

Now it should be easier to find the Fourier coefficients of $1728j(\tau)$.

Having the expansion (*) at hand, there are funny applications. Let $d > 0$ be a square-free integer and consider the field $\mathbb{K} = \mathbb{Q}(\sqrt{-d})$. Suppose that the class number of this field is 1. Let $1, \tau$ $\mathbb{Z}$-generate $\mathcal{O}_{\mathbb{K}}$. Any automorphism ($Aut(\mathbb{C})$) σ sends $j(\tau)$ to $\sigma(j(\tau))$. This j-invariant comes from an elliptic curve with a lattice which is equivalent to an ideal of an order of $\mathbb{K}$ containing $\mathcal{O}_{\mathbb{K}}$. As the class number of $\mathbb{K}$ is 1, we see that there is only one possible equivalence class of Λ we can find. Thus $\sigma(j(\tau)) = j(\tau)$ for all automorphism. Thus $1728j(\tau)$ must be a rational algebraic integer. Thus it must be in $\mathbb{Z}$. For example,

$$1728j(\sqrt{-43}) = -884736000, 1728j(\sqrt{-67}) = -147197952000.$$

Using (*), we see that for $q = e^{-\pi\sqrt{43}}$

$$e^{\pi\sqrt{43}} = (integer) - 744 - 196884q - 21493760q^2 - \cdots = Integer + 0.999775...$$

Similarly,

$$e^{\pi\sqrt{67}} = Integer + 0.999817...$$

and

$$e^{\pi\sqrt{163}} = Integer + 0.99999999999250072597...$$

Since it is easy to perform numerical computation for $j(\tau)$, it is easy to check whether or not $1728j(\tau)$ is an integer for each given quadratic τ. Therefore it is simple to confirm that a quadratic field $\mathbb{Q}(\sqrt{-d})$ is NOT of class number one. The converse is not true. The $1728j$ function is surjective and it can attain any integer value. However, there are only finitely many square-free d with $\mathbb{Q}(\sqrt{-d})$ having class number one.

Concluding remark: The connections between elliptic curves, modular forms, and quadratic number fields go much beyond this point. A good starting point to look at those connections is Heegner's proof of Gauss' class number one conjecture.

3.13. **Elliptic curves and modular forms.** As we already saw, elliptic curves and modular forms are closely related. Many key algebraic signatures of elliptic curves, e.g. j-invariant, discriminant, are modular forms. The study of modular forms is unfortunately too large to be covered in this module.

Another level of this entanglement is the modularity of elliptic curves established by A. Wiles et.al. asserting that L-series w.r.t. elliptic curves are modular via the Hecke correspondence. This result was conjectured by Weil and Taniyama-Shimura. Beyond this conjecture (now theorem), we now have a more ambitious set of conjectures which is known as Langland's program. A part of this program (Langland's reciprocity conjecture) asserts that all 'naturally formed' L-series are modular.

3.14. **Appendix: Singular Curves.** In our consideration for elliptic curves over $\mathbb{C}$, we identified a group structure on $E(\mathbb{C})$ which is isomorphic to a natural additive group on some torus. All complex elliptic curves from tori are regular. In this short section, we explore the group structures for singular curves. Recall that our algebraic proof for the group law of elliptic curves deals with general degree three curves other than the regular ones!

Theorem 3.40. *Consider the 'ugly' singular curve $C : y^2 = x^3$ with the point at infinity $(0, 1, 0)$. Then the map*

$$(x, y) \in C \to x/y \in \mathbb{K}, \infty \to 0$$

establishes the isomorphism

$$E(\mathbb{K}) \setminus \{(0, 0)\} \cong (\mathbb{K}, +).$$

Proof. It is easy to check that the map is bijective. We now verify that it is a group isomorphism. Consider the relation $(x_1, y_1) + (x_2, y_2) + (x_3, y_3) = 0$ in $E(\mathbb{K})$. We want to show that

$$\frac{x_1}{y_1} + \frac{x_2}{y_2} + \frac{x_3}{y_3} = 0.$$

Assume that $x_1 \neq x_2$. The fact that $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ are colinear on C implies that

$$x_1 + x_2 + x_3 = \left(\frac{y_1 - y_2}{x_1 - x_2}\right)^2.$$

From $y^2 = x^3$ on C we see that

$$(y/x)^2 = x, (y/x)^3 = y.$$

Write $y_i/x_i = k_i$. Then we see that

$$k_1^2 + k_2^2 + k_3^2 = \left(\frac{k_1^3 - k_2^3}{k_1^2 - k_2^2}\right)^2.$$

This is

$$k_3^{-2} = (k_1^{-1} + k_2^{-1})^2.$$

Also for the y-coordinate, we have

$$y_3 = \left(\frac{y_1 - y_2}{x_1 - x_2}\right)(x_3 - x_2) + y_2$$

which is

$$k_3^3 = \left(\frac{k_1^3 - k_2^3}{k_1^2 - k_2^2}\right)(k_3^2 - k_2^2) + k_2^3$$

which is

$$k_3^{-3} = -(k_1^{-1} + k_2^{-1})^3.$$

Overall we see that

$$k_1^{-1} + k_2^{-1} + k_3^{-1} = 0.$$

This is what we wanted to prove. The case for $x_1 = x_2$ is similar but simpler. □

Theorem 3.41. *Consider the 'bad' singular curve $C : y^2 = x^2(x+a)$ with $a \neq 0$. Consider $E(\mathbb{K})$. Let $\alpha^2 = a$ and*

$$\psi_a : (x, y) \to \frac{y + \alpha x}{y - \alpha x}, \infty \to 1.$$

If $a \in \mathbb{K}^2$, then ψ_a defines an isomorphism between $E(\mathbb{K}) \setminus \{(0,0)\}$ and $(\mathbb{K}^*, *)$.

If $a \notin \mathbb{K}^2$ then ψ_a defines an isomorphism between $E(\mathbb{K}) \setminus \{(0,0)\}$ and the multiplicative group

$$\{u + \alpha v : u, v \in \mathbb{K}, u^2 - av^2 = 1\}.$$

Here $\mathbb{K}^2$ denotes the arithmetic square of $\mathbb{K}$, not the Cartesian product.

Remark 3.42. *Consider the quadratic extension $\mathbb{K}(\sqrt{a})$ for $a \notin \mathbb{K}$. Then $\sqrt{a} \to -\sqrt{a}$ establishes a Galois map. In this context, $(u, v) \to u^2 - av^2$ is the norm of $u + \alpha v \in \mathbb{K}(\sqrt{a})$.*

Proof. The proof can be done with a direct but lengthy computation. We omit the details. □

Now the motivation for the terminologies in the following definition should be clear.

Definition 3.43. *Consider any elliptic curve over $\mathbb{Q}$ (or $\mathbb{Z}$). Then it is possible to reduce the coefficients by $\mod p$ to obtain an elliptic curve over $\mathbb{F}_p$. The following three cases can occur:*

1. 'ugly': The reduced curve has a cusp point (as in $y^2 = x^3$). In this case, we call the reduction to be additive.

2. 'bad'(1): The reduced curve is singular but not 'ugly'. Moreover, the curve is equivalent to $y^2 = x^2(x + a)$ where a is not a square in $\mathbb{F}_p$. In this case, we call the reduction to be non-split multiplicative.

3. 'bad'(2): The reduced curve is singular but not 'ugly'. Moreover, the curve is equivalent to $y^2 = x^2(x + a)$ where a is a square in $\mathbb{F}_p$. In this case, we call the reduction to be split multiplicative.

4. 'Good': The reduced curve is regular. In this case, the reduction is good.

4. Torsion points

4.1. **The main result.** Most of our study of elliptic curves over $\mathbb{C}$ directly gives us the corresponding theory over $\mathbb{Q}$. This follows from the fact that $E(\mathbb{Q}) \subset E(\mathbb{C})$ is a subgroup which can be checked directly by looking at the explicit addition formula (as functions, they are rational over the ring $\mathbb{Z}$). Since we also proved the addition formula outside of the field $\mathbb{C}$, in this section, we do not have any presumed restrictions on the base field $\mathbb{K}$ although for some proofs we will only work with Weierstrass equations and in order for those proofs to work for general elliptic curves, we require that $char(\mathbb{K}) \notin \{2, 3\}$.

With more effort, it is possible to study the addition formula over $\mathbb{Z}$ or other rings rather than fields. The following result will be used.

Theorem 4.1. *Let $E(\mathbb{Z}_n)$ be an elliptic curve defined over the ring $\mathbb{Z}_n$. Suppose that $n = n_1 n_2$ with $\gcd(n_1, n_2) = 1$. Then*

$$E(\mathbb{Z}_n) \cong E(\mathbb{Z}_{n_1}) \oplus E(\mathbb{Z}_{n_2}).$$

Remark 4.2. *This can be viewed as the Chinese remainder theorem for elliptic curves. If n is a prime, then $E(\mathbb{Z}_n)$ is of course $E(\mathbb{F}_n)$.*

Proof. The proof is simple but computationally complicated. One has to define $E(\mathbb{Z}_n)$ in a proper way, figure out the addition formula and then check (via lengthy computations) that the formula is all polynomials over $\mathbb{Z}$ so that $\mod n$ can be well defined. After that, a simple computation can give us the result. We omit the proof. Read Washington Section 2.10. □

The main result in this section concerns the torsion point of $E(\overline{\mathbb{K}})$. Let $n > 0$ be an integer. The n-torsion subgroup $E(n)$ is the set of all pints P with $nP = \infty$. Note that $E(n)$ may not be contained in $E(\mathbb{K})$.

Theorem 4.3. *If $n > 0$ is an integer which is not a multiple of $char(\mathbb{K})$. Then*

$$E(n) \cong \mathbb{Z}_n \oplus \mathbb{Z}_n.$$

Otherwise, write $n = p^r n'$ with $\gcd(p, n') = 1$. Then

$$E(n) \cong \mathbb{Z}_{n'} \oplus \mathbb{Z}_{n'}$$

or

$$E(n) \cong \mathbb{Z}_{n'} \oplus \mathbb{Z}_n.$$

Remark 4.4. *If $\overline{\mathbb{K}} = \mathbb{C}$, then we know that $E(\mathbb{C}) \cong \mathbb{C}/\Lambda$ as a group. Thus the assertion of the n-torsion points is almost trivial. This works for fields such as $\mathbb{Q}, \overline{\mathbb{Q}}, \mathbb{R}$ or any number fields.*

If $char(\mathbb{K}) = 2$ or 3 we cannot work with Weierstrass equations. Nonetheless, it is fairly straightforward to prove this result directly for $E(2), E(3)$ in those cases. We will prove the theorem with the condition that $char(\mathbb{K}) \notin \{2, 3\}$ so that we have the Weierstrass equation at hand.

4.2. **Multiplication by integers: preparation in $\mathbb{C}$.** We want to know what is $n(x, y)$ for (x, y) on an elliptic curve. This can be done explicitly in $\mathbb{C}$ (or in general fields with Riemann-Roch). We fix a lattice Λ and consider the corresponding curve

$$E = E(\mathbb{C}) : y^2 = x^3 - g_2 x - g_3.$$

Given $(x, y) \in E$, to find $n(x, y) = (x_n, y_n)$, we need to find $x_n = \wp(nz), y_n = \wp'(nz)$ in terms of $x = \wp(z), y = \wp'(z)$. Let us first consider $\wp(nz)$. We now find its poles. This is not hard,

$$P_n = \text{Poles}(z \to \wp(nz)) = \{z \in \mathbb{C}/\Lambda : nz = 0\} = E(n).$$

The set P_n has n^2 many points and at each point in P_n, the pole has order two. Next, consider the polynomial

$$K_n(z) = \prod_{w \in E(n)\setminus\{0\}} (z - \wp(w)).$$

The function $k_n(z) = K_n(\wp(z))$ has zeros at $E(n) \setminus \{0\}$. The order of those zeros are 2. The poles of k_n are only at 0, which is of order $2(n^2 - 1)$. It not yet clear what are the coefficients of $K_n(z)$ (or $k_n(\wp(z))$ of powers of $\wp(z)$). It is simple to check (by using the Laurent expansion of $\wp(z)$) that those coefficients are in $\mathbb{Q}[G_4, G_6, G_8, \dots]$. Next, higher Eisenstein series are in $\mathbb{Q}[G_4, G_6]$. For example, for G_8 we check G_8/G_4^2 which is a modular form of weight 0 whose value at ∞ is $2\zeta(8)/(2\zeta(4))^2 \in \mathbb{Q}$. Thus the coefficients of $K_n(z)$ are actually in $\mathbb{Q}[G_4, G_6]$.

Now we consider $\phi_n(z) = \wp(nz)K_n(\wp(z))$. There are no poles in $\mathbb{C}/\Lambda$ except a pole of order of $2n^2$ at 0. Since ψ_n is clearly even and elliptic, we see that there is a polynomial Φ_n of degree n^2 so that $\Phi_n(\wp(z)) = \phi_n(z)$. The coefficients of Φ_n are again in $\mathbb{Q}[G_4, G_6]$.

We obtain that

$$\wp(nz) = K_n(\wp(z))/\Phi_n(\wp(z)).$$

It is not clear if K_n, Φ_n can have common roots. If so, then

$$\wp(nz)\Phi_n^{red}(\wp(z)) = K_n^{red}(\wp(z))$$

for a polynomial Φ_n^{red} of degree strictly smaller than n^2-1 and a polynomial K_n^{red} of degree strictly smaller than n^2. The RHS has no poles at non-lattice points. Thus the LHS also has no such poles. However, there are $2(n^2-1)$ many poles for $\wp(nz)$. A polynomial $\Phi_n^{red}(\wp(n))$ cannot produce enough zeros if $\deg \Phi_n^{red} < n^2-1$. Thus we see that K_n, Φ_n cannot have common roots.

Similarly, we can find $\wp'(nz)$ in terms of $\wp(z), \wp'(z)$. We have proved the following result.

Theorem 4.5. *There are polynomials $K_n[X], \Phi_n[X], L_n[X,Y], \Psi_n[X,Y]$ with coefficients in $\mathbb{Q}[G_4, G_6]$ such that*

$$n(x,y) = \left(\frac{\Phi_n(x)}{K_n(x)}, \frac{\Psi_n(x,y)}{L_n(x,y)}\right).$$

Moreover, K_n, Φ_n are coprime and $\deg K_n = n^2 - 1, \deg \Phi_n = n^2$.

Remark 4.6. *It is possible to compute explicitly those polynomials and gain some more information. For example, let the curve be $y^2 = x^3 + Ax + B$ so that A, B are in $\mathbb{Q}[G_4, G_6]$, then*

$$2(x,y) = \left(\frac{x^4 - 2Ax^2 - 8Bx + A^2}{4x^3 + 4Ax + 4B}, \textit{something not very important}\right).$$

The reason that we book-kept the coefficients for all polynomials above is to automatically obtain the same formulae for $E(\mathbb{Q})$. It is also possible to directly apply those polynomials in finite fields $E(\mathbb{F}_q)$. For finite fields, one technical point is to check that the leading coefficients of the polynomials we obtained are not zero so that the degrees of those polynomials do not decrease in finite fields. This is indeed the case as long as the characteristic is not dividing n (for $n(x,y)$). Another technical point is that our coefficients are in $\mathbb{Q}[G_4, G_6]$ and some rationals may not be defined in finite fields. For example, $2/3$ is not defined in $\mathbb{F}_3$. Luckily, those coefficients are indeed in $\mathbb{Z}[60G_4, 140G_6]$. This can be shown easily (but lengthily) by direct computations of those coefficients.

4.3. **Proof of Theorem 4.3.** Suppose that n is not a multiple of the field characteristic p. This is the only case we will consider and we omit the proof otherwise. Then consider the formula

$$n(x,y) = (x_n, y_n)$$

with

$$x_n = r(x) = \frac{\Phi_n(x)}{K_n(x)} = \frac{x^{n^2} + \dots}{n^2 x^{n^2-1} + \dots}.$$

For each fixed x_n there should be exactly n^2 many solutions for $r(x) = x_n$ unless there are multiple roots for $r(x) - x_n$. Our non-division condition for the field

characteristic forbids the existence of multiple roots. Thus there are exactly n^2 many points (x, y) with $n(x, y) = \infty$. We therefore know that

$$\#E(n) = n^2.$$

Now as a finite abelian group, $E(n)$ is a direct sum of cyclic groups with orders $n_1|n_2|n_3 \dots$

$$E(n) = Z_{n_1} \oplus Z_{n_2} \oplus Z_{n_3} \dots$$

For each divisor d of n, $E(d)$ is an abelian group of order d^2. Let $l|n_1$ be a prime number. Then $E(l)$ has order l^2. However, $l|n_1|n_2|n_3 \dots$ thus

$$E(l) = Z_{l^{k_1}} \oplus Z_{l^{k_2}} \oplus Z_{l^{k_3}} \oplus \dots$$

The only possibility is that there are only two terms in the direct sum. Thus we have

$$E(n) = Z_{n_1} \oplus Z_{n_2}.$$

Next, n is the order of each element in $E(n)$. Thus $n_1|n_2|n$. As $n_1 n_2 = n^2$, we can only have $n_1 = n_2 = n$. This proves the result for the case when p is not a divisor of n.

4.4. **Weil pairing/bilinear form.** Later, we want to count points of elliptic curves over finite fields. In order to do this, a standard strategy is to use some sort of Fourier analysis. We start by introducing a certain bilinear form on $E(n)$ that turns out to be useful for counting points in finite fields.

Theorem 4.7 (Weil Pairing)**.** *Let $E(\mathbb{K})$ be a regular elliptic curve. Let n be a positive integer. Assume that $\mathrm{char}(\mathbb{K})$ is not a divisor of n such that $E(n) \subset E(\overline{\mathbb{K}})$ is $\mathbb{Z}_m \oplus \mathbb{Z}_n$. There is a bilinear form $e_n : E(n) \times E(n) \to \mathbb{K}^*$ with the following properties (without explicit mentioning, X, Y, Z can be any elements in $E(n)$):*

1. *(Bilinear) $e_n(X+Y, Z) = e_n(X, Z)e_n(Y, Z), e_n(X, Y+Z) = e_n(X, Y)e_n(X, Z)$;*
2. *(non-degenerate) $e_n(X, Y) = 1$ for all $Y \in E(n)$ implies that $X = \infty$;*
3. *(quasi-symmetric) $e_n(X, Y) = (e_n(Y, X))^{-1}$;*
4. *(normalised) $e_n(X, X) = 1$;*
5. *(Galois) $e_n(\sigma X, \sigma Y) = \sigma(e_n(X, Y))$ for each Galois map of $\overline{\mathbb{K}}$ over $\mathbb{K}$.*
6. *(rational) $e_n(\alpha(X), \alpha(Y)) = e_n(X, Y)^{\deg \alpha}$ for each isogeny α (rational endomorphism).*

This bilinear form is called the Weil pairing.

Remark 4.8. *As $nX = nY = \infty$, we see that $e_n^n(X, Y) = e_n(nX, Y) = e_n(X, nY) = 1$. Thus in fact, the values of e_n are n-th roots of unity forming a multiplicative group. Properties 1-4 are standard. Properties 5-6 are special in our algebraic setting and make this bilinear form distinguished.*

We will only prove this result for $\mathbb{C}$. The results hold generally, in particular, for $\mathbb{F}_q$. In order to prove such a result, we need some Riemann-Roch type arguments. Before that, there are some consequences.

Corollary 4.9. *Let E be an elliptic curve over $\mathbb{Q}$, i.e. $60G_4, 140G_6$ are rational numbers. Then for each $n \geq 3$, $E(n) \not\subset E(\mathbb{Q})$.*

Proof. Assume the contrary. For each $\sigma \in Aut(\mathbb{C})$ and $(X, Y) \in E(n) \subset E(\mathbb{Q})$, we have $\sigma(e_n(X, Y)) = e_n(\sigma(X), \sigma(Y)) = e_n(X, Y)$. Clearly, e_n is surjective because for each basis X, Y of $E(n)$, $e_n(X, Y)$ must be a primitive root of unity of order n. Thus we see that all n-th roots of unity are rational. This is not the case for $n \geq 3$. □

Let X, Y $\mathbb{Z}$-generate $E(n)$. For each isogeny $\alpha \in End(E)$, $\alpha(E(n)) \subset E(n)$. Thus we have

$$(\alpha(X), \alpha(Y))^T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} (X, Y)^T$$

for $\alpha_n = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(\mathbb{Z})$.

Corollary 4.10. *With notations as above, if $char(\mathbb{K})$ is not dividing n, then $\det \alpha_n = \deg(\alpha) \mod n$.*

Proof. We have $e_n^{\deg \alpha}(X, Y) = e_n(\alpha(X), \alpha(Y)) = e_n(aX + bY, cX + dY) = e_n^{\det \alpha_n}(X, Y)$. Therefore we see that $\det \alpha_n = \deg(\alpha) \mod n$. □

4.5. **Riemann-Roch on $\mathbb{C}/\Lambda$: bookkeeping functions by poles and zeros.** We can identify points on elliptic curves with points on torus via the Weierstrass map. Thus, we can simply consider that elliptic curves are tori. However, there are some technical issues: We want to study Galois maps, and rational maps on elliptic curves. We can use the identification to obtain maps between tori. It is not clear what are those maps between tori. For Galois maps, we know that the corresponding maps on tori must be group isomorphism. However, as Galois

maps are not continuous in general, the tori isomorphism may not be continuous. For the rational map case, the situation is much better. We know that rational endomorphisms are linear maps between tori.

Consider the field of elliptic functions on $\mathbb{C}/\Lambda$. For each function f, we associate an object $div(f)$ (called a divisor) of f which is simply the formal summation

$$div(f) = \sum_{z \in zeros} [z] - \sum_{z \in poles} [z]$$

where $[z]'s$ indicates that they are not treated as complex numbers nor elements in $\mathbb{C}/\Lambda$, they are purely treated as symbols. Thus we have a formal map

$$f \to div(f) \in \bigoplus_{z \in \mathbb{C}/\Lambda} \mathbb{Z}[z] = A_{\mathbb{Z}}(\mathbb{C}/\Lambda) = D(\mathbb{C}/\Lambda).$$

Similar objects can be found in Algebraic Topology when people formally define complexes for the study of homologies. The notation $A_{\mathbb{Z}}$ indicates the terminology 'Adele' which we do not use here. The notation D indicates the name 'divisor' which we will use here. The $\bigoplus$ notation indicates that the elements we are considering are finite summations (rather than infinite sums). Thus $D(\mathbb{C}/\Lambda)$ is a partially ordered group with $d \geq 0$ if and only all coefficients of d are ≥ 0.

Next, for each divisor $d \in D = D(\mathbb{C}/\Lambda)$, we write

$$d = \sum_z o_z z$$

where o_z are integers. We define

$$\deg d = \sum_{z, o_z \neq 0} o_z,$$

$$T(d) = \sum_{z, o_z \neq 0} o_z z.$$

All of them are finite sums and they define $\mathbb{Z}$-homomorphisms from D to $\mathbb{Z}$ and $\mathbb{C}/\Lambda$. The notation $T()$ is chosen to indicate that T maps a divisor to a point in the torus $\mathbb{C}/\Lambda$. We translate some Liouville Theorems and Abel-Jacobi Theorem here.

Theorem 4.11 (Liouville, Abel-Jacobi)**.** *If f is an elliptic function then*

$$\deg div(f) = 0, T(d) = 0.$$

If f is elliptic and $div(f) = 0$ then f is constant.
Let d be a divisor, if $T(d) = 0$ and $\deg d = 0$, then there is an elliptic function f so that $div(f) = d$.

From f to $div(f)$ we lose some information like residues. However, this is not a significant loss.

Lemma 4.12. *Let f, g be non-zero elliptic functions with $div(f) = div(g)$. Then $f = cg$ for a complex number c.*

Proof. We see that f/g has no poles, nor zeros. In order words, $div(f/g) = 0$. Liouville's Theorem tells that f/g must be a non-zero constant. □

Divisors that come from functions are called principal divisors (like principal ideals in number fields). Thus the map T maps degree zero divisors onto $\mathbb{C}/\Lambda$ whose kernel is precisely the set of principal divisors.

Definition 4.13. *For each $d \in D$, let $L(d)$ be the set*

$$\{f : f = 0 \text{ or } div(f) + D \geq 0\}.$$

Let $l(d)$ be the dimension of $L(d)$ as a vector space over $\mathbb{C}$.

We have the following result.

Theorem 4.14 (Riemann-Roch for torus)**.** *On $\mathbb{C}/\Lambda$, there is a divisor c such that for each divisor d we have*

$$l(d) - \deg d = l(c - d).$$

Remark 4.15. *A general result holds for an arbitrary Riemann surface(algebraic curve over $\mathbb{C}$),*

$$l(d) - \deg d = l(c - d) - g + 1$$

where g is the genus of the surface. Our proof is more complicated than necessary. Read any book on algebraic curves for concise proofs.

Proof. Let c be the zero divisor. Then we want to show that for each divisor d, $l(d) - \deg d = l(-d)$. Assume that the non-zero coefficients of d are either 1 or -1 and they are all at different places. If $\deg d < 0$, then $l(d) = 0$ as our function f should have zeros to match the poles of d and by doing this f need to have the same amount of poles and those poles cannot be matched by zeros of d because d does not have enough zeros. Similarly, if $\deg d > 0$, then $l(-d) = 0$. If $\deg d = 0$, then the situation is symmetric under $d \to -d$ and we obtain that $l(d) = l(-d)$. Thus, we need to show that as long as $\deg d > 0$,

$$l(d) = \deg d.$$

In this case, our f need to have zeros at the poles of d. Then the poles of f need to be at the zeros of d. As there are more zeros of d than poles of d, we see that $l(d) \geq 1$. Suppose that d has Z zeros and P poles. Then we can choose any set of P zeros of d. In such a way, we fixed P many zeros and poles of f. In order to find such an f, we need the Abel-Jacobi condition which may not hold for these poles and zeros. However, if the A-J condition holds, then we can already find a desired function. Otherwise, as $Z > P$, we have at least one place to accommodate a pole of f. Then we have one additional quota for one zero of f. We locate the zero to be somewhere such that the Abel-Jacobi condition holds. Then we can find such an f. This implies that

$$l(d) \geq 1.$$

If $Z = P + 1$, then the function f either has poles at all zeros of d or else it has poles only at P out of Z zeros of d. The two cases cannot co-exist because if there is a function f_1 for the first case, f_2 for the second case, then f_1/f_2 would be an elliptic function with only one pole and one zero and this is not possible by Liouville's Theorem. If we use A-J's theorem, we see that this one pole and one zero must be at the same place which is again not possible. In all these cases, it is possible to see that $l(d) = 1 = \deg d$.

In general, suppose that some choices of P-zeros of d with the P-poles of d do NOT have the A-J condition. Then we can fix any such P-zeros of d, say $\{z_1, \ldots, z_P\}$. For each one of the rest of $Z - P$ zeros (z, say), we can find a function with a pole at z and other poles among $z_1, \ldots, z_P$. We can find functions $f_1, \ldots, f_{Z-P}$. Clearly those functions are $\mathbb{C}$-linearly independent as they have different poles.

Otherwise, all choices of P-zeros have the A-J condition. This is not possible if $Z > P$.

In any case, we see that $l(d) \geq \deg d$. Let f be any function with $div(f)+d \geq 0$. In particular, f has at least P many zeros and thus at least P many poles. If any of the poles are not in $z_1, \ldots, z_P$ we can subtract the corresponding $f_1, \ldots, f_{Z-P}$ to cancel those poles. As at least P zeros of f persist during this procedure, after all these steps the result is a function with at least P zeros and at most P poles. Thus if this function is not the zero function then it has exactly P poles precisely at $z_1, \ldots, z_P$. By A-J, we see that $z_1, \ldots, z_P$ together with the P poles of d satisfy the A-J condition which contradict our assumption. From here we see that $l(d) = \deg d$. □

4.6. **Construct the Weil pairing.** Let $n > 1$ be an integer. We consider $E(n) \subset \mathbb{C}/\Lambda$ as a subgroup. Let $x, y \in E(n)$. Suppose that $x \neq 0$. Choose a non-zero point x' in $E(n^2)$ with $nx' = x$. By Abel-Jacobi, we can find a function f such that

$$div(f) = n[x] - n[0].$$

We can also find a function g_x such that

$$div(g_x) = \sum_{r \in E(n)} [r + x'] - \sum_{r \in E(n)} [r].$$

The function $f \circ n : z \to f(nz)$ has divisor

$$div(f \circ n) = n \sum_{r \in E(n)} [r + x'] - n \sum_{r \in E(n)} [r].$$

In the above, the choice of x' does not affect g_x nor the expression of $div(f \circ n)$. We see that g^n and $f \circ n$ have the same divisor and by Liouville Theorem we conclude that $g^n = cf \circ n$ and we can assume $c = 1$.

We define $e_n(x, y) = g_x(z + y)/g_x(z)$ for any $z \in \mathbb{C}/\Lambda$ so that $g_x(z)$ is not zero nor ∞. We can do this as $ny = 0$ and

$$(g_x(z + y)/g_x(z))^n = f(nz + ny/f(nz)) = f(nz)/f(nz) = 1.$$

Thus $g_x(z + y)/g_x(z)$ has discrete values. It is also continuous. We see that it is constant. For ease of notation, we write this constant simply as

$$e_n(x, y) = g_x(y)/g_x(0).$$

If any of x, y is zero, we simply have $e_n(x, y) = 1$. Properties 1-4 are simple to show. For Property 5, we need to use the Weierstrass function to identify $\mathbb{C}/\Lambda$ and the elliptic curve. As g_x is a polynomial in $\wp, \wp'$ we see that $e_n(.,.)$ reacts to Galois map as stated. For Property 6, we know that any rational map α between elliptic curves is a linear map (also denoted as α) between tori which is a group homomorphism whose kernel is of size $\deg \alpha$. We can find a function f_α with

$$div(f_\alpha) = n[\alpha(x)] - n[0].$$

A technical point here is that $\alpha(x)$ should not be 0. We assume this. Then the function $f_\alpha \circ \alpha$ has divisor

$$div(f_\alpha \circ \alpha) = n \sum_{x' \in x + Ker(\alpha)} [x'] - n \sum_{x' \in Ker(\alpha)} [x'].$$

We also have (choose x', x'' with $nx'' = \alpha(x), \alpha(x') = x''$)

$$div(g_{\alpha(x)}) = \sum_{r \in E(n)} [r + x'] - \sum_{r \in E(n)} [r].$$

This shows that

$$div(g_{\alpha(x)} \circ \alpha) = \sum_{r \in E(n)} \sum_{t' : \alpha(t') = r + x''} [t'] - \sum_{r \in E(n)} \sum_{t' : \alpha(t') = r} [t'].$$

Consider the first double sum

$$\sum_{r \in E(n)} \sum_{t' : \alpha(t') = r + x''} [t']$$

The range of the sum is the set

$$\alpha^{-1}(E(n)) + x'.$$

The set $\alpha^{-1}(E(n))$ is a finite subgroup of $\mathbb{C}/\Lambda$. Clearly $E(n) \subset \alpha^{-1}(E(n))$. The quotient space $\alpha^{-1}(E(n))/E(n)$ contains $\deg \alpha$ many elements. Thus $\alpha^{-1}(E(n))$ is a disjoint union of $\deg \alpha$ many cosets of $E(n)$. We write this union as

$$\alpha^{-1}(E(n)) = \bigcup_{i=1}^{\deg \alpha} (s_i + E(n))$$

for some $s_1, \ldots, s_{\deg \alpha} \in \alpha^{-1}(E(n))$. Thus we have

$$div(g_{\alpha(x)} \circ \alpha) = \sum_{i=1}^{\deg \alpha} \sum_{r \in E(n)} [s_i + r + x'] - \sum_{i=1}^{\deg \alpha} \sum_{r \in E(n)} [r + s_i]$$
$$= \sum_{i=1}^{\deg \alpha} div(g_i)$$

where $g_i(z) = g_x(z - s_i)$. Thus we have for some non-zero number c

$$g_{\alpha(x)} \circ \alpha(z) = c \prod_{i=1}^{\deg \alpha} g_i(z).$$

This implies that for $y \in E(n)$

$$e_n(\alpha(x), \alpha(y)) = g_{\alpha(x)} \circ \alpha(z + y)/g_{\alpha(x)} \circ \alpha(z) = \frac{c}{c} \prod_{i=1}^{\deg \alpha} \frac{g_i(z + y)}{g_i(z)} = e_n^{\deg \alpha}(x, y).$$

For the last equality, observe that the value of $g_x(z + y)/g_x(z)$ does not depend on z. This is what we want to show.

4.7. **Conclusion.** Although almost all the arguments in this section have been based on $\mathbb{C}$, all of them hold for general algebraically closed fields. The proofs are almost identical once some results on transcendental degree one extensions over algebraically closed fields have been established. Read any book on algebraic curves/Riemann Surfaces/algebraic geometry on this topic.

5. Elliptic curves over $\mathbb{Q}$

After planting the seeds of elliptic curves over $\mathbb{C}$, we can now harvest our sweet fruits of elliptic curves over $\mathbb{Q}$. In fact results in this section hold for elliptic curves over general number fields with almost the same proof modulo some technical issues like the finiteness of class fields, the finite generatability of the ring of units, the finiteness of integers with a given norm, etc...all of which are trivial for $\mathbb{Q}$. Actually, we can even study elliptic curves over elliptic curves! In this case, we do not have any of the required finiteness so the study is much more complicated.

5.1. **p-adic height and Lutz-Nagell Theorem.** Let p be a prime number. For each rational number x, we can write it uniquely as $x = p^k y$ for rational $y = r/s, \gcd(r,s) = \gcd(r,p) = \gcd(s,p) = 1$. We define

$$\nu_p(x) = k, |x|_p = p^{-k}.$$

In such a way, we defined a norm on $\mathbb{Q}$. The completion of $\mathbb{Q}$ w.r.t. this norm is the field of p-adic numbers, $\mathbb{Q}_p$. The ring $\mathbb{Z}_p$ are the elements in $\mathbb{Q}$ with non-negative ν_p value. In other words, $\mathbb{Q}_p$ is a p-adic valuation field with the evaluation ring $\mathbb{Z}_p$. Notice that $\mathbb{Z}_p$ is a local ring whose unique maximal ideal is $(p)\mathbb{Z}_p$. This ideal is called the p-adic valuation ideal.

The point of introducing p-adic valuation is to book-keep p factors through algebraic manipulations. For example, let us consider the curve

$$E : y^2 = x^3 + Ax + B$$

for some integers A, B. Let p be a prime number not dividing A, B. Let $k = \nu_p(x)$. If $k < 0$, then the LHS is

$$\frac{(integer) + p^2(integer) + p^3(integer)}{p^3(integer)},$$

where numbers in $(integer)$ are not dividable by p. Therefore the numerator is not a multiple of p. Thus $\nu_p(y^2) = 3k$. This implies that $\nu_p(y) = 3k/2$ and k is even. Similarly, if $k > 0$, then $\nu_p(y) = 0$.

Definition 5.1. *Let E be an elliptic curve. We define $h_p(x,y)$ to be $-\nu_p(x)$ and $H_p(x,y)$ to be $p^{-\nu_p(x)}$.*

Thus a high point has high powers of p in the denominators of the coordinates. We consider points on E with certain heights.

Definition 5.2. *For each prime p and integer $r > 0$, we define*

$$E_{(p),r} = \{(x,y) \in E(\mathbb{Q}) : \nu_p(x) \leq -2r\}.$$

When p is clear from the context, we write $E_r = E_{(p),r}$. Thus points in E_r for large r are high. Thus we have the following relations

$$E_1 \supset E_2 \supset E_3 \supset \dots.$$

There exist points with height exactly $2r$ for each positive integer r so the above relations are all strict. We also put ∞ in all those sets.

We want to see that each E_r is a subgroup rather than just a subset.

Theorem 5.3 (Height cannot be reduced)**.** *For each $r > 0$, E_r is a subgroup of $E(\mathbb{Q})$. In other words, elliptic curve arithmetic cannot strictly decrease the height of points.*

Remark 5.4. *This result can be interpreted as summing points on elliptic curves gives us no simpler points.*

Proof. We first give an easy proof for the case when P_1, P_2 have different heights. Consider the sum $P_1(x_1, y_1)+P_2(x_2, y_2)$. First let $x_1 \neq x_2$. The line through P_1, P_2 is

$$y = k(x - x_1) + b, k = \frac{y_2 - y_1}{x_2 - x_1}, b = y_1.$$

Consider the equation

$$(kx - x_1k + b)^2 = x^3 + Ax + B.$$

The three solutions x_1, x_2, x_3 should have

$$-x_1x_2x_3 = B - (b - x_1k)^2.$$

In order to find the height of x_3 we need to find the height of $(b - x_1k)^2$. Here, adding integers will not change the height as long as the height is positive. So the integer B is not making any difference. Now consider

$$b - x_1k = y_1 - x_1\frac{y_2 - y_1}{x_2 - x_1} = \frac{y_1x_2 - x_1y_2}{x_2 - x_1}.$$

Suppose that x_1, x_2 have different heights and suppose that x_1 is higher. Then $\nu_p(x_2 - x_1) = \nu_p(x_1)$. Observe that $\nu_p(y_1x_2) < \nu_p(y_2x_1)$ and this implies that

$$\nu_p(y_1x_2 - x_1y_2) = \nu_p(y_1x_2) = \nu_p(y_1) + \nu_p(x_2).$$

Thus we see that

$$\nu_p((b - x_1k)^2) = 2(\nu_p(y_1) + \nu_p(x_2)) - 2\nu_p(x_1) = \nu_p(x_1) + 2\nu_p(x_2).$$

Then we see that

$$\nu_p(x_3) = \nu_p(x_1) + 2\nu_p(x_2) - \nu_p(x_1) - \nu_p(x_2) = \nu_p(x_2).$$

As x_1, x_2 have symmetric roles, we see that if $\nu_p(x_1) \neq \nu_p(x_2)$,

$$\nu_p(x_3) = \max\{\nu_p(x_1), \nu_p(x_2)\}.$$

□

Proof. We prove the general case. The difficulty appears when P_1, P_2 have equal height. In this case, not only do we need to keep the p-powers but also the other non-p factors. Observe that if $\nu_p(x) = -2r < 0$ then $\nu_p(y) = -3r$. Then $\nu_p(x/y) = r$ and $p^{-r}(x/y)$ has 0 height. This p-adic unit carries rich information.

We want to show that $\lambda_r : E_r/E_{5r} : (x, y) \to p^{-r}(x/y) \mod p^{4r}$ is an injective group homomorphism after we declare $\lambda_r(\infty) = 0$. To show this, we need to show that whenever P_1, P_2, P_3 are collinear and $P_1, P_2 \in E_r$, then $P_3 \in E_r$ and $\lambda_r(P_1) + \lambda_r(P_2) + \lambda_r(P_2)$ is 0 in $\mod p^{4r}$. For this task, it is more convenient to use a new variable x/y. For example, write $t = x/y$ and $s = 1/y$. Then our curve E is

$$s = t^3 + As^2t + Bs^3.$$

In fact, this curve is the same as our old curve but now we look at it from ∞. Given a line $y = kx + b$ we have this line in the new set of coordinates as

$$1 = kt + bs.$$

This is still a line. This is significant as our coordinate change is non-linear. We just calibrated our viewing point from $(0, 0, 1)$ to $\infty = (0, 1, 0)$ in the projective plane. Why do we want to do this? We defined the height of points so that high points have large powers of p-factors in the denominator. In terms of projective coordinates, $(x, y, 1)$ is (xz, yz, z). Thus a high point (x, y) corresponds to a point (xz, yz, z) with z having large power of p-factors and xz, yz not having large p-factors. Dividing yz, then we have $(x/y, 1, 1/y)$ so that $1/y$ has a large p-factor. The effect of looking at ∞ turns high points to points whose coordinates have large positive power p-factors. Those points are "close" to the "origin" (which is $\infty = (0, 1, 0)$). So a geometric intuition is that high points in terms of the original (x, y) coordinate is 'far away' from $(0, 0)$. 'Far away' points are 'close' to ∞. Then around ∞, those 'far away' points are near the new origin. The metric (distance) here is the p-adic norm.

Our points P_1, P_2, P_3 are in E_r. If $(x, y) \in E$, then we see that $\nu_p(t) \geq r, \nu_p(s) \geq 3r$. We can eliminate s and obtain an cubic equation for t,

$$0 = t^3 + G(A, B, k, b)t^2 + \dots$$

$$G(A,B,k,b)=\frac{\frac{A}{b^2}(-2k)+\frac{B}{b^3}(3k^2)}{1+\frac{A}{b^2}k^2+B\frac{-k^3}{b^3}}.$$

We see that

$$t_1+t_2+t_3=-G(A,B,k,b).$$

In order to proceed further, we need to understand p-factors of k,b. For example, we have

$$\frac{k}{b}=\frac{s_2-s_1}{t_2-t_1}.$$

Since t_1,t_2,s_1,s_2 have large p-heights, so that in terms of $|.|_p$, they are small. It is plausible that $(s_2-s_1)/(t_2-t_1)$ is close to the tangent line at $(s,t)=(0,0)$. This tangent line is special because $(0,0)$ is inflexion. Thus s/t should be $O(t^2)$. In terms of our p-adic analysis (close to 0 means divisible by high powers of p), s/t should be divisible by p^{2r} if t is divisible by p^r. This is actually intuitive in the ∞-adic norm (which is the norm in $\mathbb{Q}$ giving the standard topology on $\mathbb{R}$). Consider a smooth curve C passing through $(0,0)$. Suppose that $\{y=0\}$ is the tangent line at $(0,0)$. Then the slope k_x of the tangent line via $(x,y)\in C$ should be in general $O(x)$. If this tangent line is tangent to C of a higher order (inflexion), then this $O(x)$ can be upgraded to $O(x^2)$ or even $O(x^k), k\geq 3$. For the p-adic norm, the situation is similar. We now make this point clear. It is not surprising that our arguments will be completely similar to those you would normally have in your first-year Calculus course, where the analysis of the ∞-adic norm was studied. Observe that

$$s_2-s_1=(t_2^3-t_1^3)+A(s_2^2t_2-s_1^2t_1)+B(s_2^3-s_1^3).$$

We can write (Lebniz rule)

$$s_2^2t_2-s_1^2t_1=s_2^2(t_2-t_1)+t_1(s_2^2-s_1^2).$$

We see that

$$(s_2-s_1)(1-At_1(s_1+s_2)-B(s_2^2+s_1^2+s_1s_2))=(t_2^3-t_1^3)+As_2^2(t_2-t_1).$$

This implies that (Such argument can be seen as the L'Hospital's rule for implicit functions. If we are considering ∞-valuation (as we will do later), then this is

exactly the L'Hospital's rule in analysis.)

$$\frac{s_2 - s_1}{t_2 - t_1} = \frac{t_1^2 + t_2^2 + t_1 t_2 + A s_2^2}{1 - A t_1 (s_1 + s_2) - B(s_2^2 + s_1^2 + s_1 s_2)}.$$

The denominator is 1 plus something with a non-trivial p-factor. This means that the denominator has zero p-height. The numerator, on the other hand, has p-factor of power at least $2r$. This implies that $k/b = (s_2 - s_1)/(t_2 - t_1)$ has ν_p value at least $2r$.

For $1/b$, we can perform the same argument. Or, we can use the information about k/b. Observe that $1/b = kt/b + s$. Since kt/b has p-factor with power at least $3r$ and s has p-factor with power at least $3r$ as well, we see that $1/b$ has p-factor at least $3r$.

Back to $G(A, B, k, b)$, we see that the denominator is 1 plus some p multiples. The numerator has p-factor with power at least $\nu_p(k/b^2)$ which is at least $5r$. As t_1, t_2 have ν_p value at least r and $t_1 + t_2 + t_3 = -G(A, B, k, b)$, we see that $\nu_p(t_3) \geq r$. This implies that $P_3 \in E_r$. Next, we showed that if P_1, P_2, P_3 are collinear points in E_r then $\lambda_r(P_1) + \lambda_r(P_1) + \lambda_r(P_1)$ has p-factor with power at least $5r - r = 4r$. Thus $\lambda_r(P_1) + \lambda_r(P_1) + \lambda_r(P_1)$ is zero in $\mod p^{4r}$.

We proved that λ_r is a group homomorphism with kernal containing E_{5r}. On the other hand, if $p^{-r}(x/y)$ is divisible by p^{4r} then $t = x/y$ is divisible by p^{5r}. As $\nu_p(t) = \nu_p(x) - \nu_p(y) = -\nu_p(x)/2$ we see that $\nu_p(x) \leq -10r$ and this means that $(x, y) \in E_{5r}$. This finishes the proof. □

Let $P = (x, y) \in E_r$ but not E_{r+1}. Then $\lambda_r(P) = p^{-r}(x/y)$ has ν_p value equal to $-r + \nu_p(x) - \nu_p(y) = 0$. Therefore $\lambda_r(P) \neq 0$.

Let P be a torsion point in $E_r \setminus E_{r+1}$ for some $r > 0$. Then for some integer $n > 0$ we have

$$nP = \infty.$$

Then $0 = \lambda_r(\infty) = \lambda_r(nP) = n\lambda_r(P)$. This implies that

$$\lambda_r(P) = 0.$$

Thus $P \in E_{r+1}$. This contradiction shows that P cannot be in any E_r with $r > 0$. Thus P must be an integer point. We have proved the first part of the following result.

Theorem 5.5 (Lutz-Nagell)**.** *Let $E : y^2 = x^3 + Ax + B$ be an elliptic curve with $A, B \in \mathbb{Z}$. Then rational torsion points on E have integer coordinates. Moreover, if (x, y) is a rational torsion point, then $y^2|\Delta = 4A^3 + 27B^2$.*

Proof. We have obtained the fact that $(x, y) \in \mathbb{Z}^2$. Observe that $2(x, y)$ is torsion as well. Recall that we already have an explicit formula for $2(x, y)$,

$$2(x, y) = (x_2, y_2),$$
$$x_2 = \frac{x^4 - 2Ax^2 - 8Bx + A^2}{4y^2}.$$

Thus $y^2|x^4 - 2Ax^2 - 8Bx + A^2$. We obviously have $y^2|x^3 + Ax + B$. Therefore we have

$$y^2|Resultant(x^4 - 2Ax^2 - 8Bx + A^2, x^3 + Ax + B).$$

For computational polynomial arithmetic, resultant is more convenient to use than that GCD. The resultant between $x^4 - 2Ax^2 - 8Bx + A^2$ and $x^3 + Ax + B$ is $(4A^3+27B^2)^2$. This already tells us that $y|4A^3+27B^2$. This is already satisfactory as there are only finitely many integers y to be checked.The stronger statement helps us to greatly reduce the amount of y to be checked. The stronger result can be checked by observing

$$4A^3+27B^2 = (3x^2+A)(x^4-2Ax^2-8Bx+A^2)-(3x^3-5Ax-27B)(x^3+Ax+B).$$

We actually want to find the generators of the ideal $(x^4 - 2Ax^2 - 8Bx + A^2, x^3 + Ax + B)$ in $\mathbb{Z}[x, A, B]$. This can be done by finding Groebner basis. Nowadays, there are efficient algorithms for this task, e.g. the "GroebnerBasis" function in Mathematica. It is possible to find that $4A^3 + 27B^2$ is in one set of Groebner basis. This means that we can find polynomials f, g in $\mathbb{Z}[A, B, x]$ such that $f(x)(x^4 - 2Ax^2 - 8Bx + A^2) + g(x)(x^3 + Ax + B) = 4A^3 + 27B^2$. Of course, we have found one such expression explicitly. □

With Lutz-Nagell, it is possible to find all torsion points in $E(\mathbb{Q})$ for elliptic curves defined over $\mathbb{Z}$. For elliptic curves defined over $\mathbb{Q}$, it is possible to rationally transform it into one over $\mathbb{Z}$. Notice that rational points are still rational points under rational transformations. Also, as rational transformations are isogenies, we see that torsion points are still torsion points under rational transformations. Thus we have an effective method for obtaining rational torsion points of any elliptic curves over $\mathbb{Q}$.

Warning: Lutz-Nagell theorem gives a necessary condition for torsion points. This condition is not sufficient. There can be integer non-torsion points with $y^2|\Delta$.

5.2. **Another proof of Lutz-Nagell.** The proof of Theorem 5.3 shows much more than what is stated in the theorem. We have constructed an injective homomorphism $\lambda_r : E_r/E_{5r} \to \mathbb{Z}/p^{4r}\mathbb{Z}$. We now give a slightly more complicated but logically more straightforward proof.

Lemma 5.6. *Let $(x, y) \in E_r \setminus E_{r+1}$ for some $r \geq 1$. Then $n(x, y) \in E_r$ for all $n \geq 1$.*

Proof. Recall that $n(x, y) = (x_n, y_n)$ has

$$x_n = \frac{x^{n^2} + \dots}{n^2 x^{n^2-1} + \dots}.$$

We proved this for elliptic curves over $\mathbb{C}$. The rational map $x_n(x)$ has in fact $\mathbb{Z}$ coefficients. Thus it is well defined in $E(\mathbb{Q})$. If $p \not| n$, then we have

$$x_n = \frac{p^{-2rn^2}}{p^{-2r(n^2-1)}}(sth),$$

where sth is a p-adic unit. Thus $\nu_p(x_n) = p^{-2r}$. Thus $x_n \in E_r$. If $p|n$, then

$$x_n = p^{-2r}(sth)$$

where $\nu_p(sth) < 0$. Thus $\nu_p(x_n) < -2r$ and $x_n \in E_r$. □

Now we can proof Lutz-Nagell. Let $P = (x, y) \in E(\mathbb{Q})$ be a torsion point. Consider the cyclic group $C_P = \mathbb{Z}P$. This group has finite order as P is a torsion point.

Suppose that $nP = \infty$. Then for each prime $l|P$ in C_P we can find Q with $lQ = \infty$. Then $\mathbb{Z}Q$ has l points on $E(\mathbb{Q})$, among which $l - 1$ are finite. $Q = (x_Q, y_Q)$ has the property that $x_Q \in \mathbb{Z}/l^2$. This is because it is a rational number and a algebraic number with equation

$$l^2 x_Q^{l^2-1} + \dots = 0.$$

Since $\nu_l(x_Q)$ must be even, we see that $x_Q = q_1/l^2$ with $l \not| q_1$ or otherwise $x_Q \in \mathbb{Z}$. We assume the former case. The polynomial

$$P_l(x) = l^2 x^{l^2-1} + \dots$$

has factor $(x - q_1/l^2)$. This holds for all other finite points in C_Q. Thus we can find numbers $q_2, \dots, q_{l-1}$ not being multiples of l such that $P_l(x)$ is also a multiple of $(x - q_2/l^2), \dots, (x - q_{l-1}/l^2)$. Thus $(x - q_1/l^2) \dots (x - q_{l-1}/l^2)$ divides $P_l(x)$ in $\mathbb{Q}[x]$. Thus $(l^2x - q_1) \dots (l^2x - q_{l-1})$ divides $P_l(x)$ in $\mathbb{Z}[x]$. This is not possible as long as $l - 1 > 1$ by the consideration of the leading coefficients. Thus we have only the following possibilities:

1. n is even.
2. n is odd and for each prime $l|n$, finite points in C_P with order l are integer points. Thus all finite points in C_P are integer points.

We now consider those cases.

If n is even, then we can find $Q \in C_P$ with order two. However, $Q \in E(\mathbb{Q})$ and $2Q = \infty$. This shows that $Q = (x', y')$ must be in $(\mathbb{Z}, 0)$. By the above lemma, we see that P must not be in E_1. Since this holds for all p, we see that P must be an integer point.

At this stage, we already proved the result that all rational torsion points are integer points. Then we can argue similarly as in the first proof with $P, 2P$ being integer points and show that $y^2|4A^3 + 27B^2$. However, we record the following argument which follows a different approach and carries a bit more intuition.

If n is odd, then for each prime divisor l of y, we consider $E(\mathbb{F}_l)$. Technically, we have to choose $l \notin \{2, 3\}$. However, l cannot be two and the $l = 3$ case creates no issues as we are considering the Weierstrass equation directly. If mod l is a good reduction, then $E(\mathbb{F}_l)$ is regular and we have the following homomorphism:

$$\psi : C_P \to E(\mathbb{F}_l).$$

Notice that $\psi(P) \in (\mathbb{F}_l, 0)$. Thus we see that $\psi(P)$ has order two. (We can only say this if $E(\mathbb{F}_l)$ is regular. If it is singular, it can also happen that $\psi(P)$ is the singular point.) Thus $\psi(C_P)$ is an abelian group with a non-trivial element of order two. Thus $\#\psi(C_P)$ is even. However, $\#\psi(C_P)$ must be a divisor of $\#C_P$ we see that $\#C_P$ must be even as well and this contradicts the assumption. Thus mod l must not be a good reduction and this says that $l|4A^3 + 27B^2$. As this holds for each prime divisor of y we almost have $y|4A^3 + 27B^2$. (To actually show

that $y|4A^3+27B^2$, we need to show reduction mod l^k for $l^k|y$. For this we need the theory of elliptic curves over rings.) Next, for each prime number l so that $\nu_l(y)<0$, we can still mod l. The image $\psi(P)$ is ∞.

5.3. **$\mathfrak{p}$-heights and Lutz-Nagell in number fields.** What can be said about curves over number fields? The $\mathbb{Q}$ theory can be extended to number fields, say $\mathbb{K}$. For this, we need to have p-adic analysis on number fields. This is $\mathfrak{p}$-adic analysis for prime ideals in a number field. Polynomial arithmetic over $\mathbb{Z}$ and over $\mathcal{O}_{\mathbb{K}}$ are completely similar. One issue is that if $\deg_{\mathbb{Q}}\mathbb{K}$ is large, then GroebnerBasis-related algorithms are slow. Nonetheless, it is possible to say that for elliptic curves defined over $\mathcal{O}_{\mathbb{K}}$, torsion points on $E(\mathbb{K})$ have coordinates in $\mathcal{O}_{\mathbb{K}}$ and y^2 is a factor of the discriminant.

5.4. **∞-height and Mordell-Weil Theorem.** Let $x=r/s$ be a rational number in the lowest form. We define $H(x)=\max\{r,s\}$ (∞-height) and $h(x)=\log H(x)$ (logarithmic ∞-height).

Definition 5.7. *Let E be an elliptic curve. We define $h_\infty(x,y)$ to be $h(x)$ and $H_\infty(x,y)$ to be $H(x)$.*

Thus a high point is complicated in the sense that to describe this point we need many decimal digits ($\approx h_\infty(x,y)/\log 10$ many decimal digits). As p-adic heights, the ∞-height is also useful. Our first step is to study how heights are transformed via the arithmetic of elliptic curves. Our intuition is that, as in the p-adic case, arithmetic on an elliptic curve can only make the points more and more complicated. This is indeed the case if we consider nP for a given rational non-torsion point P. As n grows, nP becomes very complicated and soon it is too expensive to compute and store the coordinates of nP. For example, consider the point $(-4,6)$ on $y^2=x^3-25x$. Try to compute $n(-4,6)$ for $n=2,3,4,\ldots$.

Theorem 5.8. *For each rational point P, the limit $\hat{h}(P)=2^{-1}\lim_{n\to\infty}4^{-n}h(2^nP)$ exists. We call $\hat{h}(P)$ the canonical height of P.*

1. *$\hat{h}(P)\geq 0$ for each $P\in E(\mathbb{Q})$.*
2. *There is a constant $c>0$ so that $|0.5h(P)-\hat{h}(P)|<c$. This c depends on the curve and can be estimated explicitly.*
3. *The number of points bounded by a given height is finite.*

4 $\hat{h}(nP) = n^2\hat{h}(P)$.
5 $\hat{h}(P+Q) + \hat{h}(P-Q) = 2\hat{h}(P) + 2\hat{h}(Q)$.
6 $\hat{P} = 0$ *if and only if* P *is torsion.*
7 $(P, Q) = \hat{h}(P+Q) - \hat{h}(Q) - \hat{h}(P)$ *is a* $\mathbb{Z}$*-bilinear form which is nondegenerate. In particular, given points* $P_1, \ldots, P_n$*, the matrix* $(P_i, P_j)_{i,j}$ *has zero determinant if and only if* $P_1, \ldots, P_n$ *are* $\mathbb{Z}$*-dependent.*

Thus it is not quite true that $P + Q$ is at least as complicated as P or Q. However, we see that either $\hat{h}(P+Q)$ or $\hat{h}(P-Q)$ must at least $\hat{h}(P) + \hat{h}(Q)$.

The result follows by showing that for some constant $c > 0$,

$$|h(P+Q) + h(P-Q) - 2h(P) - 2h(Q)| < c.$$

As we already have a formula for $P+Q, P-Q$ in terms of P, Q, the above inequality follows after some direct computations. We omit the full details. The proof is not so interesting although the result is very important. It turns $E(\mathbb{Q})$ into an inner product space.

We now prove a famous result of Mordell-Weil.

Theorem 5.9. *Elliptic curves over* $\mathbb{Q}$ *have finite rank as* $\mathbb{Z}$*-modules.*

To prove this result, we first prove the following 2-descent theorem.

Theorem 5.10. *Let* $E(\mathbb{Q})$ *be an elliptic curve. Then the quotient group* $E(\mathbb{Q})/2E(\mathbb{Q})$ *is a finite group.*

Theorem 5.10+ Theorem 5.8 $\implies$ *Theorem 5.9.* Since $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite, we can find a finite set of representatives $a_1, \ldots, a_n$. For each point P on $E(\mathbb{Q})$, we can find a point R and a point a_i so that

$$P + 2R = a_i.$$

Thus we see that

$$2R = a_i - P.$$

Then we have

$$4\hat{h}(R) = \hat{h}(2R) = \hat{h}(a_i - P) \leq 2\hat{h}(a_i) + 2\hat{h}(P).$$

Suppose that $\hat{h}(P) > \hat{h}(a_i)$, then

$$\hat{h}(R) < \hat{h}(P).$$

Let $c = \max\{\hat{h}(a_i)\}_i$. Then the set of all points with height at most c is finite. Let G be the group generated by those points. Suppose that G is not the entire $E(\mathbb{Q})$, then we can find some point P not in G. Then we can find all points with height at most $\hat{h}(P)$. Then among those points which are not in G, we can find one with the smallest height. Let this point be P. Clearly, $\hat{P} > c$ for otherwise $P \in G$ by construction. Then we see from

$$P + 2R = a_i$$

that $\hat{h}(R) < \hat{h}(P)$. Thus $R \in G$ for otherwise, we would not have chosen the point P. However, if $R \in G$ then certainly $P \in G$. Thus the contradiction says that G must be the whole of $E(\mathbb{Q})$. Thus $E(\mathbb{Q})$ is finitely generated as a $\mathbb{Z}$-module and therefore has a finite rank. □

Definition 5.11. *Let $E(\mathbb{Q})$ be an elliptic curve. We write the curve as*

$$y^2 = x^3 + Ax + B = (x - e_1)(x - e_2)(x - e_3).$$

Consider the number field $\mathbb{K} = \mathbb{Q}(e_1, e_2, e_3)$. Then the Mordell-Weil map ψ is defined to be

$$\psi : (x, y) \in E(\mathbb{K}) \to (x - e_1, x - e_2, x - e_3) \in (\mathbb{K}^*/(\mathbb{K}^*)^2)^3$$

for $x \notin \{e_1, e_2, e_3\}$. We also define

$$\psi(\infty) = (1, 1, 1),$$

$$\psi(e_1, 0) = ((e_1 - e_2)(e_1 - e_3), e_1 - e_2, e_1 - e_3),$$

$$\psi(e_2, 0) = (e_2 - e_1, (e_2 - e_1)(e_2 - e_3), e_2 - e_3),$$

$$\psi(e_3, 0) = (e_3 - e_1, e_3 - e_2, (e_3 - e_1)(e_3 - e_2)).$$

Lemma 5.12. *ψ is a group homomorphism and* $\ker\psi = 2E(\mathbb{K})$

Proof. We start with the first part. Let P_1, P_2, P_3 be points on $E(\mathbb{K})$. We want to show that if $P_1 + P_2 + P_3 = 0$ then $\psi(P_1)\psi(P_2)\psi(P_3)$ is in $(\mathbb{K}^*)^2$. Suppose that P_1, P_2 does not have zero y-coordinate. Then the fact that P_1, P_2, P_3 are collinear allows us to find a line over $\mathbb{K}$

$$l_{a,b} : y - (ax + b)$$

so that $l_{a,b}(P_1) = l_{a,b}(P_2) = l_{a,b}(P_3) = 0$. We also assume that P_1, P_2 have different x-coordinates so that a is finite. (Otherwise, $P_3 = \infty$. We also have $\psi(P_1) = \psi(P_2)$. Thus $\psi(P_1)\psi(P_2)\psi(P_3)$ has square coordinates and this what we wanted to show.) Therefore the x-coordinates of P_1, P_2, P_3 satisfy

$$(ax + b)^2 = x^3 + Ax + B.$$

We can make the above equation with terms $x - e_1, x - e_2, x - e_3$. For example, we have

$$(a(x - e_1) + b + ae_1)^2 = (x - e_1 + e_1)^3 + A(x - e_1 + e_1) + B.$$

It is possible to see that

$$(x_1 - e_1)(x_2 - e_1)(x_3 - e_1) = -(B + Ae_1 + e_1^3 - (b + ae_1)^2).$$

Notice that $e_1^3 + Ae_1 + B = 0$. Thus we see that

$$(x_1 - e_1)(x_2 - e_1)(x_3 - e_1) = (b + ae_1)^2.$$

Similar results holds for e_2, e_3 (in the places of e_1) as well. Therefore we have

$$\psi(P_1)\psi(P_2)\psi(P_3) = ((b + ae_1)^2, (b + ae_2)^2, (b + ae_3)^2) = 0 \in (\mathbb{K}^*/(\mathbb{K}^*)^2)^3.$$

This finishes the proof for generic cases. The boundary cases are left as exercises.

Next, we examine the kernel of ψ. Consider a finite point $(x, y) \in \ker\psi$. We must have

$$x - e_1, x - e_2, x - e_3$$

are all non-zero squares in $\mathbb{K}$. From here, we want to find $(x', y') \in E(\mathbb{K})$ such that $2(x', y') = (x, y)$.

The fact that $x - e_1, x - e_2, x - e_3$ are squares allows us to find non-zero v_1, v_2, v_3 with

$$x - e_i = v_i^2.$$

This v_i^2 should be $(x' - e_i)^2$. Thus we want to find a square root r of $x - e_i$ in $\mathbb{K}$. Then $r + e_i$ should be our x'. As e_1, e_2, e_3 are roots of the polynomial $T^3 + AT + B$, we can consider performing arithmetic in $\mathbb{K}[T]/(T^3 + AT + B)$. This is a standard circle of ideas in algebraic number theory. For any algebraic number α, the field $\mathbb{Q}(\alpha)$ and $\mathbb{Q}[T]/(f(T))$ are actually isomorphic for the minimal polynomial of α. Here we need to consider $\mathbb{K}$ instead of $\mathbb{Q}$. The issue is that $T_3 + AT + B$ split in $\mathbb{K}$ so the quotient ring is not a field. However, this issue cannot stop us from

considering the arithmetic on the ring $R = \mathbb{K}[T]/(T^3 + AT + B)$. We can find a polynomial f over $\mathbb{K}$ with degree at most 3 so that

$$f(e_i) = v_i.$$

Then we see that the polynomial $g(T) = x - T - f^2(T)$ has roots e_1, e_2, e_3. Thus $T^3 + AT + B | g(T)$ in $\mathbb{K}[T]$. In order words, $x - T - f^2(T) = 0 \in \mathbb{K}[T]/(T^3 + AT + B) = R$. Thus we see that $x - T$ is a square in the ring R. We can then find $r \in R$ with $x - T = r^2$. The most general elements in R can be written as $u_0 + u_1T + u_2T^2$ with $u_0, u_1, u_2 \in \mathbb{K}$. From here see that $x - T$ is equal to

$$(u_0+u_1T+u_2T^2)^2 = (u_0^2-2Bu_1u_2)+(2u_0u_1-2Au_1u_2-Bu_2^2)T+(u_1^2+2u_0u_2-Au_2^2)T^2$$

in R. Observe that $1, T, T^2$ are $\mathbb{K}$-linearly independent in R. We see that

$$(*) \qquad u_0^2 - 2Bu_1u_2 = x, 2u_0u_1 - 2Au_1u_2 - Bu_2^2 = -1, u_1^2 + 2u_0u_2 - Au_2^2 = 0.$$

Thus we found a square root of $x - e_i$ which is $u_0 + u_1e_i + u_2e_i^2$. Our next task is to show that $u_0 + u_1e_i + u_2e_i^2 + e_i$ is the x-coordinate of some point in $E(\mathbb{K})$, namely, if we put $x' = u_0 + u_1e_i + u_2e_i^2 + e_i$, then $(x')^3 + Ax' + B$ is a square in $\mathbb{K}$. This can be proved via the last two identities in $(*)$. The first identity exactly tells us that $2(x', y') = (x, y)$ for a carefully chosen y' (there are only two possible choices). Details are left as an exercise.

□

Theorem 5.10. Let the curve be $y^2 = (x - e_1)(x - e_2)(x - e_3)$. Assume that $e_1, e_2, e_3 \in \mathbb{Q}$. This is not a very restrictive assumption. The proof works for e_1, e_2, e_3 in any number fields. However, we need some tools from algebraic number theory to complete the proof. All we need to show now is that the image of ψ is finite. We can assume that e_1, e_2, e_3 are integers by using some rational if necessary.

Now we have the Mordell-Weil map from $E(\mathbb{Q})$ to squarefree rationals triples. Let the integer triple (a, b, c) represents such an image. We are going to do p-adic analysis for (a, b, c). First, observe that abc is a square because are are considering the equation $y^2 = (x-e_1)(x-e_2)(x-e_3)$ with $e_1, e_2, e_3 \in \mathbb{Q}$. Next, if $\nu_p(abc) > 0$, then $\nu_p(abc)$ must be 2. This is because $\nu_p(a), \nu_p(b), \nu_p(c)$ can only be $0, 1$. This means that at least two of $x - e_1, x - e_2, x - e_3$ must be multiples of p. Thus at

least one of $e_1 - e_2, e_e - e_3, e_2 - e_3$ must be a multiple of p. Thus we must have

$$p|(e_1 - e_2)(e_1 - e_3)(e_2 - e_3).$$

Thus $\nu_p(abc) > 0$ only for p in the set of prime factors of $(e_1-e_2)(e_1-e_3)(e_2-e_3)$. There are only finitely many such prime factors. Thus there are only finitely many possible a, b, c. This proves the result. □

The proof of Theorem 5.10 can be used for finding $E(\mathbb{Q})$. We know that $E(\mathbb{Q}) \cong \oplus\mathbb{Z}^r$ for a torsion group T and some integer $r \geq 0$ which is the rank of $E(\mathbb{Q})$. On the one hand, by using Lutz-Nagell (or many other methods), it is possible to identify the torsion group T. Thus we see that

$$E(\mathbb{Q})/2E(\mathbb{Q}) \cong T/2T \oplus \mathbb{Z}_2^r.$$

Thus once we find $E(\mathbb{Q})/2E(\mathbb{Q})$, we can determine r and then we can determine $E(\mathbb{Q})$. For $E(\mathbb{Q})/2E(\mathbb{Q})$, we can try to find the image of Mordell-Weil map. The proof of Theorem 5.10 leaves us only finitely many candidates that can be in the image. Then task is to check that each of them is either contained or not contained in the image. However, this is surprisingly difficult.

5.5. **Rank of elliptic curves.** It is computationally difficult to determine the rank of elliptic curves. One way is to use the 2-descent method. However, this method is only valid if the Tate-Shafarevich group of E is trivial. Another way is to use a result of Silverman which tells us that $\hat{h}(P)$ can be estimated by $h(P)$ effectively. Then we have an upper bound of the ∞-heights of the generators of $E(\mathbb{Q})$. Then we have an upper bound of the rank. We need to check the linear dependence of all subsets of those generators. Computationally, it is possible to confirm linear independence. However, to confirm a linear dependence, we actually need to perform an extensive search of all possible linear combinations.

Yet another not quite correct way is to use the Birth-Swinnerton-Dyer conjecture. On the one hand, for a given curve, we can compute explicitly the number of points in finite fields. Then we can approximate the L-series of this curve as well as we wish. Then we can approximate $L(0), L'(0), L^{(2)}(0)$ and so on. We can computationally confirm a non-zero. However, we cannot computationally confirm a zero. So we can find the smallest number k so that $L^{(k)}(0)$ is a confirmed non-zero. Then $k-1$ is an upper bound of the rank. To confirm that $k-1$ is the

rank, we need to find a set of $k-1$ linearly independent generators. This can be done by using the height pairing.

By using a combination of the above methods (and other methods), it is often possible to determine the rank of a given curve. Of course, given infinite computational power, it is always possible to determine the rank of all curves. The problem is that we only have finite computational power.

It is believed that almost all elliptic curves have rank 0 or 1. In fact, half of the curves have a rank 0 and the other half have a rank 1 and the other zero proportion has higher ranks. This is a difficult conjecture of Goldfeld as well as Katz-Sarnak.

5.6. **Elliptic curves over number fields.** As Lutz-Nagell, Mordell-Weil also holds for elliptic curves over number fields. The generalisation is straightforward but requires results in algebraic number theory. Basically, all arithmetic we can do in $\mathbb{Q}, \mathbb{Z}$ can be done (after some twists) in number fields. In particular, we no longer require that all roots e_1, e_2, e_3 of the X part are in $\mathbb{Q}$.

6. Elliptic curves over finite fields

Most of this Chapter is non-examinable except the statement of Hasse's theorem plus a few special and simple corollaries.

Elliptic curves are defined in finite fields as well. As long as the field characteristic is not $2, 3$, we only need to consider Weierstrass equations. In all cases, the group law for the curves is defined via line intersections. Since there are only finitely many points on curves over finite fields, elliptic curves are finite groups. Thus for each $E(\mathbb{F}_q)$, we want to completely determine it as a group. This task can be done with a computer efficiently. We will be interested in the study of general properties for all (regular) elliptic curves over $\mathbb{F}_q$.

Nowadays, there are efficient algorithms for the arithmetic of $E(\mathbb{F}_q)$. Basically, we understand them just as we understand $\mathbb{Z}$. The study of the arithmetic of finite structures (link finite fields, elliptic curves, etc.) is often studied in computer science as well. One of the most profound applications is in cryptography. Most of the cryptography schemes based on arithmetic on integers can be translated to elliptic curves (of course, also to algebraic integers). The advantage of dealing with elliptic curves is that we often do not need to introduce super large primes (as in integer RSA). Instead, we only need to find a small set of relatively small integers

(often primes) and we achieve an equivalent cryptography goal. The price is that the encoding-decoding procedure is more complicated. Cryptography schemes based on elliptic curves are proved to be less secure (than the RSA) under quantum algorithms. In fact, all cryptography schemes based on index arithmetic (or discrete logarithm) are not secure under Shur's factor algorithm. With practical quantum computers on the horizon, finding a quantum-safe cryptography scheme is nowadays a very active topic.

Our study of elliptic curves should be considered a very special case of the study of general curves or varieties. In particular, for finite fields, the study of arithmetic, combinatorial, and geometric properties of points on varieties is an extremely fascinating field. Read books on arithmetic geometry for more on this topic.

6.1. **Counting points: Hasse's theorem.** For curves over finite fields, the first natural question is to determine their cardinalities. In some cases, we can have a closed-form formula for counting points on curves. However, you should try to convince yourself, that it is hopeless to have a closed-form formula or even an effective algorithm for counting points on general curves. Of course, given a prime number q and a curve C, we can always run through all pairs $\mathbb{F}_q \times \mathbb{F}_q$ to find all points on C. As q becomes large, this algorithm is not efficient. However, as we will see shortly, that it is quite easy to estimate the number of points on each curve over $\mathbb{F}_q$ with a surprisingly high level of accuracy.

We first provide a simple result for counting points.

Theorem 6.1. *Let k be an integer and consider $E : y^2 = x^3 - kx$. Let q be a prime number which is $-1 \mod 4$. Then $E(\mathbb{F}_q)$ contains exactly $q+1$ many points including ∞.*

Remark 6.2. *The case when -1 is not a square is consider to be the 'real' case. If -1 is a square, then we consider the case as being 'complex'.*

Proof. The question is to check for each $x \in \mathbb{F}_q$, whether or not $x(x^2-k)$ is a square in $\mathbb{F}_q$. Since -1 is not a square in $\mathbb{F}_q$, we see that for each $x \neq 0$ exactly one $x(x^2-k)$ or $-x(x^2-k)$ is a square. For each square $x(x^2-k)$, there are two values for y. Thus there are exactly q many finite points. Thus there are $q+1$ points. □

From the proof, we have a general strategy for counting points. For each element $a \in \mathbb{F}_q$, we write

$$\left(\frac{a}{\mathbb{F}_q}\right) = 1/-1/0$$

if a is a square/non-square/zero. Then we have

$$\#E(\mathbb{F}_q) = 1 + q + \sum_{a \in \mathbb{F}_q} \left(\frac{a^3 + Aa + B}{\mathbb{F}_q}\right).$$

6.2. **A simple proof of Hasse's theorem.** We record a simple proof of the following result.

Theorem 6.3 (Hasse)**.** *Let $y^2 = x^3 + Ax + B$ be an elliptic curve over $\mathbb{F}_{q=p^n}$. Then*

$$|q + 1 - \#E(\mathbb{F}_q)| \leq 2\sqrt{q}.$$

In $\overline{\mathbb{F}_q}$, the map (Frobenius map)

$$\phi_q : x \to x^q$$

is a field endormorphism. We have that $Fix(\phi_q) = F_q$. More generally, for each $n \geq 1$, $Fix(\phi_q^n) = F_{q^n}$.

We can define $\phi_q(x, y) = (x^q, y^q)$ on $E(\mathbb{F}_q)$. This map is an isogeny of degree q. However, $\ker\phi_q < q^2$ because the field characteristic p divides q.

Lemma 6.4. *For each integer $n > 1$, $Ker(\phi_q^n - 1) = E(\mathbb{F}_{q^n})$. The map ϕ_q^n is separable.*

Proof. $\phi_q^n(x, y) = (x, y)$ if and only if that $(x, y) \in E(\mathbb{F}_{q^n})$. We can explicitly write down the expression of $(x^{q^n}, y^{q^n}) - (x, y)$. Then the separability is easy to show. We omit the lengthy details. □

We now want to estimate $\deg(\phi_q - 1)$. This will help us as

$$\#E(\mathbb{F}_q) = \deg(\phi_q - 1).$$

For this we need to use Weil pairing: For integers r, s with $p \not| s$, for each $n > 1$ by considering Weil pairing for $E(n)$, we have

$$\deg(r\phi_q - s) \equiv \det(\alpha_n) \mod n.$$

The map α_n is the linear map $r\phi_q - s$ on $E(n)$. Let $\phi_q, 1$ denote the linear map on $E(n)$ as well. Then we have in $\mathbb{Z}/n\mathbb{Z}$,

$$\begin{aligned}
&\deg(r\phi_q - s) = \det(\alpha_n) = \det(r\phi_q - s)\\
&= r^2 \det \phi_q + s^2 \det(1) - rs(\det(\phi_q - 1) - \det(\phi_q) - \det(1))\\
&= r^2 \deg \phi_q + s^2 \deg(1) - rs(\deg(\phi_q - 1) - \deg(\phi_q) - \deg(1))\\
&= r^2 q + s^2 - rs(\deg(\phi_q - 1) - q - 1).
\end{aligned}$$

Thus we have $(a = q + 1 - \#E(\mathbb{F}_q))$,

$$\deg(r\phi_q - s) = r^2 q + s^2 - rsa.$$

We should have

$$r^2 q + s^2 - rsa \geq 0.$$

This implies that $qx^2 - ax + 1 \geq 0$ for all real numbers x. This implies that $a \leq 2\sqrt{q}$.

The big problem of this very elegant proof is that unless you are already an expert in arithmetic geometry, there is no easy way to explain how to come up with such a proof. The crucial ingredient is the Weil pairing. For this, we need Riemann-Roch. Even if we invest time and energy in those topics, the study of $\deg(r\phi_q - s)$ is still out of nowhere. Weil in 1948 proved a much deeper result that deals with all algebraic curves over finite fields. The proof shares a similar taste to the proof we presented here. Later, in the 1970s, Stepanov, Bombieri, and Schmidt found a different approach that gives the same result. That proof used ideas of what can be now referred to as the polynomial method (see Larry Guth's nice book). We shall see very soon that these counting results have profound links in trigonometric sums. Hopefully, we can say more about the intuition behind this proof in a different context.

6.3. **Multiplicative and additive characters.** We see that $\left(\frac{\cdot}{\mathbb{F}_q}\right)$ is a group character. Namely, it is a group homomorphism from the multiplicative group $\mathbb{F}_q^*$ to $\{-1, 1\} \subset S^1$. For this reason, we also call it a multiplicative character. In analytic number theory, you also see the name Dirichlet character. A warning is that the finite field $\mathbb{F}_{q^2}$ is not $\mathbb{Z}/q^2\mathbb{Z}$. Thus multiplicative characters on $\mathbb{F}_{q^2}^*$ are not Dirichlet characters on $(\mathbb{Z}/q^2\mathbb{Z})^*$.

Definition 6.5. *Let G be a group. A character ψ is a homomorphism from G to the multiplicative group $\mathbb{C}^*$. The set of group characters is denoted as $\hat{G}$. If G is an abelian group, then $\hat{G}$ is an abelian group. If G is finite, then $\psi(G) \subset S^1$.*

Remark 6.6. *If G is a locally compact abelian group, a classical result of Pontryagin says that $\hat{\hat{G}} \cong G$ in a natural way.*

Remark 6.7. *$\mathbb{F}_q$ is an additive group with additive characters. $\mathbb{F}_q^*$ is a multiplicative group with multiplicative characters. We have the intuition that addition and multiplication are quite different. Thus additive characters are strongly not multiplicative in some sense and vice versa.*

The notion of character is a bit strange. A more accurate name should be a one-dimensional linear representation. Given a field $\mathbb{K}$. A n-dimensional linear representation of G on $\mathbb{K}$ is a group homomorphism from G to $GL_n(\mathbb{K})$. Characters are traces of linear representations. For abelian groups, the only non-trivial representations have dimension one. Therefore the notion of character and linear representation coincide. In the analytic world, the study of representations of G is usually called Harmonic Analysis (or Fourier Analysis) on G. The theory of representations is extremely rich. For finite groups, Lie Groups (e.g. $SL_2(\mathbb{R})$), and possibly infinite Galois groups, the theory is well-developed with vast applications.

Here we have some basic properties of multiplicative characters of $\mathbb{F}_q^*$.

- There are $q-1$ many different multiplicative characters.

  This is because $\mathbb{F}_q^*$ is a cyclic group. For each finite cyclic group G, we have a primitive element g that generates G. For each integer k, we can map $\psi_k(g) = e^{2\pi ik/\#G}$ and extend this map as a group homomorphism. Clearly, $\psi_k, \psi_{k'}$ are equal if and only if $k \equiv k' \mod \#G$. Thus we found $\#G$ many characters. Let ψ be any character, then $\psi(g) \in S^1$ must be a $\#G$th root of unity because $g^{\#G} = 1_G$. Thus $\psi(g) = e^{2\pi ik/\#G}$ for some integer k.
- Let ψ be a multiplicative character. We have

$$\sum_{t \in \mathbb{F}_q^*} \psi(t) = q - 1$$

  if ψ is trivial and otherwise the sum is 0.

If ψ is trivial, then we just need to add $q-1$ many ones. If ψ is non-trivial, then we see that for each $a \in \mathbb{F}_q^*$,

$$\psi(a)\sum_{t\in\mathbb{F}_q^*}\psi(t)=\sum_{t\in\mathbb{F}_q^*}\psi(at)=\sum_{t\in\mathbb{F}_q^*}\psi(t)$$

We just choose a so that $\psi(a)\neq 1$. Then we see that $\sum_{t\in\mathbb{F}_q^*}\psi(t)=0$.

- For each $a\in\mathbb{F}_q^*$, the sum

$$\sum_{\psi}\psi(a)=q-1$$

if $a=1$ and otherwise it is zero.

Let ψ' be a non-trivial character so that $\psi'(a)\neq 1$. Then we can perform the same argument as in the previous result. The point is that there exist multiplicative characters with $\psi'(a)\neq 1$. This is true as we can directly check by examining the set of all multiplicative characters.

Since $\mathbb{F}_q$ is an additive group. We can also study the set of additive characters. We denote additive characters as ϕ. Then there are q many additive characters. We also have

$$\sum_{a}\phi(a)=q$$

if ϕ is trivial and otherwise the sum is zero. Likewise,

$$\sum_{\phi}\phi(a)=q$$

if $a=0$ and otherwise the sum is zero.

We now have the result recording the 'statistical' independence between additive and multiplicative characters.

Theorem 6.8 (square root cancellation of Gaussian sums)**.** *Let ψ,ϕ be non-trivial multiplicative and additive characters in a finite field $\mathbb{F}_q$. Then we have the following result for the Gaussian sum*

$$|G(\psi,\phi)|=|\sum_{x\in\mathbb{F}_q}\psi(x)\phi(x)|=\sqrt{q}.$$

Remark 6.9. *For the special case when q is a prime, it is possible to determine the sum (not only the norm) explicitly via elementary methods.*

Proof. Consider the sum

$$\begin{aligned}
&|\sum_{x\in\mathbb{F}_q}\psi(x)\phi(x)|^2\\
=&\sum_{x,y\in\mathbb{F}_q}\psi(x)\phi(x)\psi(y^{-1})\phi(-y)\\
=&\sum_x\sum_t\psi(x)\phi(x)\psi(t^{-1}x^{-1})\phi(-tx)\\
=&\sum_x\sum_t\psi(t^{-1})\phi((1-t)x)\\
=&\sum_t\sum_x\psi(t^{-1})\phi((1-t)x).
\end{aligned}$$

If $t=1$ then $\psi(t^{-1})\phi((1-t)x)=1$. If $t\neq 1$, then $\sum_x\phi((1-t)x)=0$. Thus we see that

$$|\sum_{x\in\mathbb{F}_q}\psi(x)\phi(x)|^2=q.$$

□

Why this square root cancellation is such a big deal? Intuitively speaking, a 'truly random' trigonometric sum should have square root cancellation. More precisely, let $w_i, i=1,\dots,q$ be a set of independently randomly chosen vectors on the unit circle. For a large range of probability distributions, $|\sum_i w_i| = O(\sqrt{q})$ with a high probability. With this idea in mind, we should consider a set of vectors w_i not being randomly chosen if their sum is either too large or too small. Two extreme non-random cases are that those vectors are collinear (sum to norm q) or balanced (sum to zero).

Next, fix a set of unit vectors $w_i, i=1,\dots,q$. Consider another set of randomly chosen unit vectors $v_i, i=1\dots,q$. Then again with high probability, $|\sum_i w_i v_i| = O(\sqrt{q})$. Thus if the sum is too large or too small, then we expect that w_i, v_i are related.

In many other places, this square root cancellation intuition appears. One such example is the distribution of ± 1 of the famous Möbius function μ, Liouville function λ, and the prime indicator function P. Those functions are defined via prime numbers which have a clear multiplicative structure. Thus, on the set of integers ordered in an additive way (the normal way, $0, 1, 2, \dots$), we expect that those functions behave like a certain sequence of random variables. In fact, the square root cancellation of the partial sums of μ is equivalent to the Riemann Hypothesis. So the Riemann Hypothesis can also be considered a strong assertion of the randomness of prime numbers.

6.4. **Trigonometric sums.** After Theorem 6.1, we provide a formula relating the number of points on certain elliptic curves on finite fields and a certain character sum. From the previous section, we can consider character sums as trigonometric sums (i.e. sums of form $\sum_k e^{2\pi i a_k}$). This counting method is so standard that it deserves a special name: Fourier analysis. In different places, we can also see names such as the circle method, cyclotomic analysis, linear representations, etc.

In our counting problem in finite fields, we would like to consider the following sum

$$\sum_{x \in \mathbb{F}_q} \psi(f(x))\phi(g(x)),$$

where f, g are polynomials over $\mathbb{F}_q$. The idea is that this sum should have square root cancellation (as Gaussian sums where f, g are linear) except for some obvious reasons. To illustrate those obvious reasons, let us consider only the sum

$$\sum_{x \in \mathbb{F}_q} \psi(f(x)).$$

If f has some obvious multiplicative structure, then under some multiplicative character ψ, then values $\psi(f(x))$ for different x may not appear to be 'random' enough. For example, $f(x) = h(x)^d$ for some $d > 1$ and ψ has order d. We have the following result proved by Weil using algebraic geometry. At that time, the framework of modern algebraic geometry was not available, Weil has to invent his own framework. This made Weil's original work somehow difficult to understand. Luckily, now we have many modern expositions of Weil's work. In fact, we have much more than that. Weil conjectured that his result could be extended to

deal with much more general situations (e.g. with varieties rather than just with curves). He stated his conjecture in four parts. Grothendieck proved 3/4 of this conjecture of Weil and Deligne proved the last 1/4. The proof of this conjecture uses the modern algebraic geometry framework which is more accessible to non-experts. Nowadays, there are many expositions of this result. See for example https://www.jmilne.org/math/CourseNotes/lec.html

Theorem 6.10. *Let q be a power of a prime number p. Let f, g be polynomials over $\mathbb{F}_q$. If either of the following holds:*

- *ψ is not trivial and $(d, \deg f) = 1$ where $d > 1$ is the order of ψ,*
- *ϕ is not trivial and $(\deg g, q) = 1$,*

then we have

$$|\sum_{x \in \mathbb{F}_q} \psi(f(x))\phi(g(x))| \leq (\deg f + \deg g - 1)q^{1/2}.$$

If ψ or ϕ is trivial, then we can ignore the degree of f or g in above.

Remark 6.11. *The conditions on f, g can be generalised a bit. For f we can require that $y^d = f(x)$ be irreducible over $\overline{\mathbb{F}_q}$. For g, we can require that $y^q - y - g(x)$ be irreducible over $\overline{\mathbb{F}_q}$.*

We can use this result for ϕ being trivial and ψ being the quadratic residue symbol $(./\mathbb{F}_q)$. Then we see that

$$\left|\sum_{x \in \mathbb{F}_q} \left(\frac{x^3 + Ax + B}{\mathbb{F}_q}\right)\right| \leq 2q^{1/2}.$$

From here we obtain Hasse's theorem as a special case of Theorem 6.10.

Weil's proof and the later Deligne's proof used an idea that is now called *etale cohomology*. In some way, it is a (very high level) Fourier analytic method. We want to discuss a different proof of Theorem 6.10 discovered by Stepanov, Schmidt, Bombieri, etc. (See the book of Schmidt: Equations over Finite Fields)

(1) First, we can easily establish a relation between points counting and trigonometric sums just as we did after Theorem 6.1.
(2) The usual strategy is to control those trigonometric sums. There are many brilliant ideas including Mordell's method, Weyl's method, Vinogradov's

method, Hardy-Littlewood method, and so on. However, Stepanov developed a polynomial method that can be used to provide a coarse estimate of $C(\mathbb{F}_q)$:

$$\#C(\mathbb{F}_q) = q + O(q^{1/2}).$$

The $O(.)$ term here is far from being optimal. However, we already have the $q^{1/2}$ error term. This is usually extremely hard to establish by only controlling the trigonometric sums.

(3) By using the relation between points counting and trigonometric sums, we obtain a non-optimal control of those trigonometric sums.

(4) We can use a relation established by Hasse and Davenport. This relation tells us once we know some trigonometric sum such as

$$g_1 = \sum_{x \in \mathbb{F}_q} \psi(f(x))\phi(g(x))$$

over $\mathbb{F}_q$, then we can extend the characters ψ, ϕ to $\mathbb{F}_{q^n}$ and compute

$$g_n = \sum_{x \in \mathbb{F}_{q^n}} \psi(f(x))\phi(g(x)).$$

(5) We can now apply the coarse counting result (2) and obtain estimates of $|g_n|$ for all n. Now there are two pieces of information for g_n. One from g_1 and Hasse-Davenport, the other one from using Stepanov's polynomial method. From here, we see that Stepanov's coarse bound for $|g_n|$ impacts what we know for $|g_1|$. The final result follows.

The key step in the above strategy is Hasse-Davenport's relation which allows us to consider curves in extensions of finite fields. A similar idea also appeared in our earlier simple proof of Hasse's theorem. We considered elliptic curves over extensions of finite fields and then used Weil pairing to obtain some relations that hold in such extensions. From those relations, we obtain our estimate for $E(\mathbb{F}_q)$. Although we do not show the discussed proof of Theorem 6.10. From here, we at least know some intuitions behind the simple proof of Hasse's theorem.

Stepanov's polynomial method basically uses the fact that high-degree polynomials are complicated enough to vanish on a chosen set that is not too complicated. For example, in the one-variable case, we can find polynomials of degree $n > 1$ that vanish on any chosen set of n points. We can perform such arguments

for polynomials with more than one variable and require the target polynomial to vanish at some points with prescribed multiplicity. Somehow, this is a Riemann-Roch type theory. This polynomial method has many other applications including proving Roth's theorem and Schmidt's subspace theorem, solving Erdős' distance problem (Guth and Katz), proving Baker's theorem on linear forms of logarithms, etc.

7. Zeta functions, RH and BSD

Consider a curve (e.g. an elliptic curve) C over $\mathbb{F}_q$ for some $q = p^k$, p prime. Then this curve is defined as $C(\mathbb{F}_{q^n})$ in any extension $\mathbb{F}_{q^n}$, $n \geq 1$. Denote $N_n = \#C(C(\mathbb{F}_{q^n}))$. We can consider the following formal power series

$$f_C(T) = \sum_{n=1}^{\infty} \frac{N_n}{n} T^n.$$

Then we can define the following zeta function for C over $\mathbb{F}_q$,

$$\zeta_{q,C}(s) = \exp(f_C(q^{-s})).$$

This is the local zeta function for C at q. Here we do not prove that $\zeta_{q,C}$ is actually a well-defined function. We borrow this result from Weil. The famous Weil's theorem says that $\zeta_{q,C}$ is a rational function in q^{-s} (i.e. $\exp(f_C(T))$ is a rational function in T). Moreover, $\zeta_{q,C}(s) = 0$ only if $\operatorname{Re} s = 1/2$. This is often stated as the Riemann Hypothesis for curves over finite fields.

7.1. RH for elliptic curves over finite fields.

Let E be a regular elliptic curve over $\mathbb{F}_q$. We know from Hasse's theorem that $a_q = \#E(\mathbb{F}_q) - (q+1)$ satisfies

$$|a_q| \leq 2\sqrt{q}.$$

Next, it is possible to show that there are numbers α, β such that for all $n \geq 1$,

$$a_{q^n} = \alpha^n + \beta^n.$$

Such a result can be obtained by using Hasse-Davenport relation. For the case of elliptic curves, we can also deduce this result directly from the consideration of $\ker(\phi_q^n - 1)$ which is the det of a certain matrix representing a $\mathbb{Z}$-linear map over $E(l) \subset E(\overline{\mathbb{F}_q})$ for integers l. From this method, we can moreover obtain that $|\alpha| = |\beta| = q^{1/2}$.

We can then insert the knowledge of a_{q^n} into the definition $\zeta_{q,E}$. Observe that for each number α, we have

$$\sum_{n=1}^{\infty} \frac{\alpha^n}{n} T^n = -\log(1-\alpha T).$$

Then we obtain that

$$\zeta_{q,C}(s) = \frac{(1-\alpha q^{-s})(1-\beta q^{-s})}{(1-q^{-s})(1-q^{1-s})}.$$

As we know that α, β has norm $q^{1/2}$. From this result, we see that if $\zeta_{q,C}(s) = 0$ then have $\operatorname{Re} s = 1/2$. This is the Riemann Hypothesis for elliptic curves over finite fields.

7.2. **Local-Global relations: L-series and the BSD conjecture.** We can interpret $\mathbb{Z}$ as $\mathbb{Z}_\infty$ (global) or as $\mathbb{Z}_p$ (local) for each prime p. There are different intuitions for the terminologies as "Global, Local". In our case, it is useful to think of the global object as a giant plank with many small nails. Those nails are local objects.

Let E be an elliptic curve over $\mathbb{Z}$. Then we can consider $E(\mathbb{Q}_p)$ for all primes p and for ∞. For some p, $E(\mathbb{Q}_p)$ is not regular, but this is not an issue. Notice that once we have a point on $E(\mathbb{Q})$, we can also consider this point as in $E(\mathbb{Q}_p)$ for all p. Thus we see that in one direction

$$\text{Global point} \to \text{Local points}.$$

The reverse direction is more difficult. Given a point on $E(\mathbb{Q}_p)$, we want to find a global point on a specific $E(\mathbb{Q})$ that gives us this point. This is only possible if we are sure that for all other local curves $E(\mathbb{Q}_p)$, we can identify the corresponding point. Even if this is the case, it is still possible that we can not find a global point. Basically, once we have a large plank, we automatically obtain all the nails on this plank. However, it can happen that we have all the nails but those nails are not on a plank. In a non-precise way, if the direction

$$\text{Local points} \to \text{Global point}$$

goes through, we say that Hasse's principle holds.

We can push this local-global consideration to zeta functions. We already studied the local ζ functions for elliptic curves. Those functions carry information about elliptic curves over certain finite fields.

Let E be an elliptic curve over $\mathbb{Z}$. Then for each prime p, we have $\zeta_{p,E}(s)$. For $E(\mathbb{Q})$, we simply define

$$L_E(s) = \text{sth} \times \prod_{p:good} \frac{1}{(1-\alpha_p p^{-s})(1-\beta_p p^{-s})}.$$

Here α_p, β_p are so that $\#E(\mathbb{F}_{p^n}) - (p^n+1) = \alpha_p^n + \beta_p^n$. The "sth" is a finite product regarding bad primes p for E. Such definition makes the following product of zeta functions

$$L_E(s) \prod_{p<\infty} \zeta_{p,E}(s)$$

extremely simple. (Compute it!)

The function L_E carries arithmetic information about $E(\mathbb{Q})$. We do not prove here that L_E is a nice meromorphic function. We borrow this result from Wiles who proved that if E is modular w.r.t. to a certain modular group and L_E is meromorphic (at least around 1).

Conjecture 7.1 (BSD). *$Ord_{s=1} L_E = rank(E(\mathbb{Q}))$.*

This conjecture bridges local and global information of E. It is extremely simple to determine α_p, β_p for E. We just need to compute several $\#E(\mathbb{F}_{p^n})$. Once we have ALL such information we can determine L_E. Or, once we have computed $\zeta_{p,E}$ for sufficiently many p, we can approximate L_E reasonably well. Then it is just a matter of time to determine (or find a good upper bound of) $Ord_{s=1} L_E$. This will release information of $rank(E(\mathbb{Q}))$ which is a global feature.

7.3. **the 2-descend method for elliptic curves.** The "residue" of L_E at $s = 1$ should also carry arithmetic information. (Actually, all analytic features of L_E should reflect arithmetic features of E.) In order to state this conjectural information, we first discuss a 2-descend method for elliptic curves.

Recall that the Mordell-Weil map ψ sends $E(\mathbb{Q})/2E(\mathbb{Q})$ to $(\mathbb{K}/\mathbb{K}^2)^3$. By examining the image of ψ, we can completely determine $E(\mathbb{Q})$.

How to examine the image of ψ if we do not know $E(\mathbb{Q})$? We only have finitely many possibilities for squarefree integral representations of triples $(a, b, c) \in$

$(\mathbb{K}/\mathbb{K}^2)^3$ where $\mathbb{K}$ is a number field. We need to examine each of those possibilities.

For example, consider $Y^2 = (X-e_1)(X-e_2)(X-e_3)$ with $e_1, e_2, e_3 \in \mathbb{Q}$. Then a hypothetical (x, y) gives squarefree integers a, b, c such that abc is a square and

$$x - e_1 = au^2, x - e_2 = bv^2, x - e_3 = cw^2$$

for rational numbers u, v, w. Then we should have

$$au^2 - bv^2 = e_2 - e_1, cw^2 - bv^2 = e_2 - e_3.(*)$$

Sometimes, it can be checked that there are no rational solutions to the above pair of equations. This then shows that the hypothetical (x, y) cannot be on $E(\mathbb{Q})$.

One way to check the non-existence of (u, v, w) is via local check. Namely, if we can find a prime p so that (*) does not have solutions in $\mathbb{Q}_p$, then we are sure that (*) cannot have solutions in $\mathbb{Q}$. However, it can happen that (*) has solutions in all $\mathbb{Q}_p$ but no solution in $\mathbb{Q}$. (Namely, Hasse's principle does not hold.) In this case, we encode this phenomenon in a group known as the Tate-Shafarevich group.

7.4. **Tate-Shafarevich group and the refined BSD conjecture.** Consider the curve $C_{a,b,c}$:

$$au^2 - bv^2 = e_2 - e_1, cw^2 - bv^2 = e_2 - e_3.$$

What we did in the 2-descend method is to check that $C(\mathbb{Q}_q)$ is empty for some p (possibly ∞) in order to eliminate the possibility that (a, b, c) is in the image of the Mordell-Weil map. Consider the following set

$$S_2 = \{(a, b, c) : C_{a,b,c}(\mathbb{Q}_p) \neq \emptyset \text{ for all } p \leq \infty\}.$$

We see that S_2 is a subgroup of

$$(\mathbb{K}^*/\mathbb{K}^{*2})^3.$$

It is the 2-Selmer group. It contains points that cannot be eliminated by the p-adic method. Consider the following quotient group

$$T_2 = S_2/\operatorname{Im}(\psi).$$

This is the 2-Tate-Shafarevitch group. If it is trivial, then we have a kind of Hasse's principle. Namely, if we want to check a point (a, b, c), we are sure that

$C_{a,b,c}(\mathbb{Q}_p) \neq \emptyset$ for all $p \leq \infty$ implies that (a, b, c) is in the image of the Mordell-Weil map. Thus T_2 is the obstruction for Hasse's principle for our 2-descend method.

It is possible to generalize the notion of 2-Selmer group, 2-Tate-Shafarevitch group to n-Selmer group, n-Tate-Shafarevitch group. They are related to the n-descend method which we do not discuss in this lecture. Consider the group T_∞ generated by all the n-Tate-Shafarevitch groups $T_2, T_3 \ldots$. For this group, we have the following conjecture.

Conjecture 7.2 (Tate-Shafarevitch)**.** *T_∞ is finite.*

This conjecture is open although we do know that it holds for certain special classes of elliptic curves.

Now we can state the refined version of the BSD conjecture.

Conjecture 7.3 (Strong BSD)**.** *Let E be an elliptic curve over $\mathbb{Q}$. Let L_E be its L-series. Let r be the rank of E. Let $P_1, \ldots, P_r$ generate the infinite part of E. Then we have*

$$L_E(s) = (s-1)^r \frac{A_E \# T_\infty \det(P_i, P_j)}{\#Tor(E)^2} + \textit{higher order terms},$$

where A_E is a number depending on E .

Here A_E is easy to compute. Note that it is not known whether or not $T_\infty < \infty$.

8. Appendix: A problem on 🍎, 🍌, 🍍

Consider the following equation,

$$\frac{🍎}{🍌 + 🍍} + \frac{🍌}{🍍 + 🍎} + \frac{🍍}{🍎 + 🍌} = 4.$$

Find 🍎, 🍌, 🍍 $\in \mathbb{Q}$.

This is a problem of solving a Diophantine equation. We can first transform the problem to obtain an easier-to-read equation,

$$\frac{a}{b+c} + \frac{b}{c+a} + \frac{c}{a+b} = 4.$$

Then we see that

$$a^3+b^3+c^3-3(a^2b+ab^2+a^2c+ac^2+b^2c+bc^2)-5abc=0.$$

This is a degree three homogeneous polynomial. We cannot say that it defines an elliptic curve yet unless we know that this polynomial is irreducible (over your favourite field). We also need to check that this curve is regular if we want to use the theory of elliptic curves over $\mathbb{Q}$.

However, from our knowledge of algebraic curves, we can try to transform any degree 3 curves to Weierstrass form. If we fail, it means that the curve is not irreducible. Luckily, in the fruit case, we can perform the following explicit transformation

$$a=\frac{56-x+y}{56-14x}, b=\frac{56-x-y}{56-14x}, c=\frac{-28-6x}{28-7x}$$

to obtain the elliptic curve

$$E: y^2=x^3+109x^2+224x.$$

To find such a transformation, one way is to find an inflection point and put this point at ∞. Then we should have a nice Weierstrass equation. We can obtain a short Weierstrass equation, but we need to work on $\mathbb{Q}(\sqrt{65})$.

We still have Lutz-Nagell and Mordell-Weil. Let us first find the rational torsion points. Even though our elliptic curve is not in the short Weierstrass form, Lutz-Nagell and Mordell-Weil are still true. The discriminant of the curve is $551183360 = 2^{10}5^17^213^3$. From here we can try the square divisors of the discriminant and locate all the rational torsion points. They are

$$Tor=\{(56,728),(4,52),(0,0),(4,-52),(56,-728),\infty\}.$$

This torsion group is a cyclic group of order 6.

Next, we can use Mordell-Weil to conclude that the elliptic curve $E(\mathbb{Q})$ has a finite rank. Finding the rank is not an easy task but luckily it is possible (by using a computer). We know that $rank(E(\mathbb{Q}))=1$. Observe that $(-100,260)$ is on the curve. That point is not a torsion point. So we see that

$$E(\mathbb{Q})=Tor+\mathbb{Z}(-100,260).$$

To find rational solutions of a,b,c we just need to transform (x,y) back to (a,b,c).

9. Appendix: A family of curves

Let p be a prime number and consider $E_p : y^2 = x^3 - p^2x$. The problem is to determine $E_p(\mathbb{Q})$. The difficult part is to determine the rank. For a specific p, we can use the usual Mordell-Weil map and try the l-adic method to find an upper bound of the rank. There are many other numeric methods. Once again, recall that determining the rank of any given elliptic curve could be very hard.

9.1. **The congruent number problem.** The congruent number problem for an integer n is to find a right triangle with rational sides whose area is n. Such a triangle is called a congruent number solution to n. It is a difficult problem to determine whether or not a given n has a congruent number solution. This problem is more than 1000 years old and it is still open. The best general result in this direction is proved by Tunnell.

Theorem 9.1. *If p is an odd prime and has a congruent number solution, then the integer solutions of*

$$2x^2 + y^2 + 8z^2 = p$$

can be divided equally according to the parity of z. Namely, the number of integers solutions with even z is equal to the number of solutions to odd z.

The converse is true if we accept the BSD conjecture.

Theorem 9.2. *Conditioned on the BSD conjecture, if p is an odd prime and the integer solutions of*

$$2x^2 + y^2 + 8z^2 = p$$

can be divided equally according to the parity of z, then n has congruent number solutions.

Remark 9.3. *Tunnell's result deals with all squarefree integers rather than just primes.*

The congruent number problem is linked to the study of elliptic curves. Let us observe that for p to be the area of a right triangle with rational sides, we need to find rational numbers a, b, c with

$$a^2 + b^2 = c^2, ab = 2p.$$

Then we see that

$$(a+b)^2/4 = (c/2)^2 + p, (a-b)^2/4 = (c/2)^2 - p.$$

Thus we can find a rational square x such that

$$x = (c/2)^2$$

and $x-p, x+p$ are rational squares. If we consider the curve

$$E_p : y^2 = x^3 - p^2 x$$

then we see that a congruent number solution to p corresponds to a rational point on E_p. The converse is not true in general, e.g. $\infty, (\pm p, 0), (0,0)$ are all rational points on E_p but they do not give congruent number solutions to p. Not all is lost, if we consider the Mordell-Weil map ψ_p for the curve $E_p(\mathbb{Q})$, we see that if there is a non-trivial point P on $2E(\mathbb{Q})$, then $\psi_p(P)$ should give a triple of rational squares and we can find a congruent number solution to p. Thus we see that

Theorem 9.4. *If p is an odd prime, then p has a congruent number solution if and only if*

$$rank(E_p(\mathbb{Q})) > 0.$$

From here, the connection of the congruent number problem with the BSD conjecture is clear.

9.2. **the 2-descend method for E_p.** Let us see how the 2descend method can be used to get some information on the congruent number problem. Consider the curve $E_p : y^2 = x^3 - p^2 x$. By Lutz-Nagell, we can see that the torsion part of E_p is precisely the 2-torsions $E(2)$. There are four of them. The non-torsion part has a finite rank by Mordell-Weil. We need to determine this rank.

Observe that the image of ψ_p can be represented by triples (a,b,c) with a,b,c being squarefree integers and abc is a square. Moreover, prime factors of abc must divide the discriminant of E_p. Thus we see that $a,b,c \in \{\pm 1, \pm 2, \pm p, \pm 2p\}$. We

construct the following table

$$\begin{bmatrix} 1 & YES & (1,1) & (-1,-p) & (p,2) & (-p,-2p) \\ 2 & TODO & (1,p) & (-1,-1) & (p,2p) & (-p,-2) \\ 3 & TODO & (2,p) & (-2,-1) & (2p,2p) & (-2p,-2) \\ 4 & TODO & (1,2p) & (-1,-2) & (p,p) & (-p,-1) \\ 5 & TODO & (2,2p) & (-2,-2) & (2p,p) & (-2p,-1) \\ 6 & TODO & (1,2) & (-1,-2p) & (p,1) & (-p,-p) \\ 7 & TODO & (2,1) & (-2,-p) & (2p,2) & (-2p,-2p) \\ 8 & TODO & (2,2) & (-2,-2p) & (2p,1) & (-2p,-p) \end{bmatrix}.$$

This table contains all possible pairs (a,b) from (a,b,c) that can be the image of ψ_p. The first line corresponds to the torsion points. Consider lines 3,5,7,8. They can be treated via 2-adic analysis. The result is that they cannot appear from the image of ψ_p. From here we see that

$$rank(E_p(\mathbb{Q})) \leq 3.$$

For lines $2, 4, 6$, we cannot eliminate them in general. However (exercise), for specific values for p, it is possible to eliminate some of them. In fact, if we want to keep line 2, then -1 must be square in $\mathbb{F}_p$. If we want to keep line 6, then 2 must be square in $\mathbb{F}_p$. If we want to keep line 4, then $-1, -2$ must be squares in $\mathbb{F}_p$. Therefore we have the following result.

Theorem 9.5. *If $p \equiv 3 \mod 8$, then $rank(E_p(\mathbb{Q})) = 0$. If $p \not\equiv 1 \mod 8$, then $rank(E_p(\mathbb{Q})) \leq 1$.*

On the other hand, for $p \equiv 5, 7 \mod 8$, it is possible that $rank(E_p(\mathbb{Q})) = 1$. Therefore, we cannot eliminate lines $2, 6$ in general. Although we cannot reproduce Tunnell's theorem with our 2-descend method, we can at least say that if $p \equiv 3 \mod 8$, then there is no congruent number solution to p.

References

Han Yu,
Email address: Han.Yu.2@warwick.ac.uk