Modularity and Fermat's Last Theorem

Akshay Manoj Sonali Sant

December, 2023

A thesis presented for the degree of
M.Sc. Mathematics in the program Mathematics International
under the supervision of Prof. Dr. Claus Fieker and Dr. Jeroen Hanselman

Fachbereich Mathematik
Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

Declaration of independent writing

I hereby affirm that I have independently written this thesis and have not used any sources and aids other than those indicated, and that I marked quotations accordingly.

To be extra cautious, I would like to add that I have closely followed the book [DS05], [Mas15], [DDT95]. for majority part of my thesis and the due credit is given at the start of every chapter.

I declare that I have also used and cited softwares or databases like Magma, Desmos, LMFDB.

Lastly I would like to declare that sometimes we will closely follow proofs from either of these sources for the sake of completeness. Due to the extreme complexity of the overall topic and to demonstrate all of this within 6 months, it was inevitable to closely follow some of the proofs but written in my own words. Having said that, I must add that in many proofs, remarks, propositions etc, I have expanded on my own, giving more details to maintain and add a touch of the originality.

The style that many of these sources follow which is expected from a mathematical text is that they start the proofs but leave out many details as exercises. This contributed to most parts of my thesis which led to filling in the gaps and details. Sometimes one might notice some similarities between some proofs in one of the resources and the proof presented here, especially with the flow of the proof. This is not a mere coincidence but rather a consequence of closely following some of the proofs and in such instances or in any case I do not want to claim the originality of the proofs except for the proofs, which I did on my own from scratch, that includes also filling in some gaps of the original proofs in the sources.

Introduction

The Modularity Theorem and Fermat's Last Theorem are two of the most celebrated and important results in the field of mathematics. The Modularity Theorem, also known as the Taniyama-Shimura-Weil Conjecture, establishes a deep connection between elliptic curves and modular forms. This theorem played a crucial role in Andrew Wiles' proof of Fermat's Last Theorem, which had remained an unsolved problem for over 350 years.
Fermat's Last Theorem states that there are no non-zero solutions to the equation $a^n + b^n = c^n$ for any values of a, b, c, and $n \in \mathbb{N}$ greater than 2. This theorem was first proposed by Pierre de Fermat in 1637 and became one of the most famous and challenging problems in the history of mathematics. It was not until 1994, when Andrew Wiles presented his proof of Fermat's Last Theorem that this longstanding problem was finally solved. The Modularity Theorem and Fermat's Last Theorem have profoundly impacted the field of mathematics, leading to new insights and discoveries in algebraic geometry, number theory, and representation theory. They have also inspired new research directions, such as studying modularity lifting theorems and the Langlands program. Let us start with brief history of Fermat's last theorem. We will follow some parts from [Rib79] and [DDT95].

Pierre de Fermat:

Pierre de Fermat, a 17th-century French jurist residing in Toulouse, was a polymath with interests spanning from poetry and Greek studies to law. However, he was predominantly passionate about mathematics, particularly the integer solutions of equations.
Fermat closely examined equations formulated as:

$$X^2 - dY^2 = \pm 1$$

Here, d stands for a positive integer devoid of square factors other than 1 (termed square-free). He revealed the presence of an infinite set of solutions to such equations.
In the margin of his copy of Bachet's edition of the complete works of Diophantus, Fermat wrote:

> *" It is impossible to separate a cube into two cubes, or a biquadrate into two biquadrates, or in general any power higher than the second into powers of like degree; I have discovered a truly remarkable proof which this margin is too small to contain."*

The 1670 edition of Fermat's works, curated by his son Samuel in Toulouse, includes a notable comment that originally dates back to around 1637, according to Dickson's volume II on the history of number theory. Tannery refers to an 1883 letter from Fermat to Mersenne, dated June 1638, seeking to solve specific mathematical problems involving cubes and biquadrates. This letter is also documented in a 1962 volume of Mersenne's correspondence and corroborated by Itard in 1948. Similar mathematical challenges were posed to Frénicle de Bessy in 1640 and to Wallis and Brouncker in 1657, although the acclaimed proof Fermat claimed to have found is not detailed.

In modern language, we state the theorem as follows:

Fermat's Last Theorem:

If n is an integer greater than 2, then there are no three non-zero integers a, b, c that satisfy the equation $a^n + b^n = c^n$.

No proof of this statement was ever found among Fermat's papers. He did, however, write a proof that the equations $X^4 - Y^4 = Z^2$ and $X^4 + Y^4 = Z^4$ have no solutions in integers all different from 0 . This is one of two proofs by Fermat in number theory that have been preserved. With very few exceptions, all Fermat's other assertions have now been confirmed. So, this problem is usually called Fermat's last theorem, although it was not proved until 1994.

One of Fermat's most notable miscalculations was his assumption that the numbers $F_n = 2^{2n} + 1$ are invariably prime. However, Euler demonstrated that this isn't the case, particularly showing that F_5 is not a prime number. In addition, mathematicians like Sierpiński and Schinzel have highlighted other inaccuracies in Fermat's claims.
The question of whether Fermat had a valid proof for his theorem is a subject of ongoing discussion among mathematicians. It's possible that he thought he had a proof at some point but was mistaken. Even though Fermat had a reputation for being open about his errors, it remains puzzling why other great mathematicians haven't been able to reconstruct a proof, assuming one ever existed.

To illustrate Fermat's candor, we quote a part from his letter dated October 18, 1640 to Frénicle de Bessy:

> **French:** *"Mais je vous advoue tout net (car par advance je vous advertis que comme je suis pas capable de m'attribuer plus que je ne sçay, je dis avec même franchise ce que je ne sçay pas) que je n'ay peu encore démonstrer l'exclusion de tous diviseurs en cette belle proposition que je vous avois envoyée, et que vous m'avez confirmée touchant les nombres* $3, 5, 17, 257, 65537$ *& c. Car bien que je réduise l'exclusion à la plupart des nombres, et que j'aye même des raisons probables pur le reste, je n'ay peu encore démonstrer nécessairement la vérité de cette proposition, de laquelle pourtant je ne doute non plus à cette heure que je faisois auparavant. Si vous en avez la preuve assurée, vous m'obligerez de me la communiquer: car après cela rien ne m'arrestera en ces matières"*
> **English Translation:** *But I must candidly admit to you (for I want to make it clear in advance that just as I don't claim to know more than I do, I also freely admit what I don't know) that I have not yet been able to prove conclusively the exclusion of all divisors in that beautiful proposition I had sent you, and which you confirmed to me concerning the numbers* $3, 5, 17, 257, 65537$ *and so on. Although I can eliminate most numbers as divisors and even have convincing reasons for the rest, I have not yet been able to necessarily prove the truth of this proposition, which I nonetheless don't doubt any more now than I did before. If you have a certain proof, you would do me a favor by sharing it with me; for after that, nothing will hold me back in these matters.*

Another instance of his honesty and his collaborative mind can be found in his correspondence with Pascal dated August 29th, 1654 where he proposed the same problem:

> **French:** "Au *reste, il n'est rien à l'avenir que je ne vous communique avec toute franchise. Songez cependant, si vous le trouvez à propos, à cette proposition: les puissances carrées de 2, augmentées*

> *de l'unité, sont toujours des nombres premiers:* $2^2+1=5, 2^{2^2}+1=17, 2^{2^3}+1=257, 2^{2^4}+1=65537$*, sont premiers, et ainsi à l'infini. C'est une proposition de la verité de laquelle je vous répond. La démonstration en est trés malaisée, et je vous avoue que je n'ai pu encore la trouver pleinement; je ne vous la proposerois pas pour la chercher si j'en étois venu à bout."*
> **English translation:** *Furthermore, rest assured that I will share everything with you openly in the future. However, please consider this proposition if you find it appropriate: The square powers of 2, when increased by one, are always prime numbers. For example,* $2^2+1=5$, $2^{2^2}+1=17$, $2^{2^3}+1=257$, $2^{2^4}+1=65537$*, and so on to infinity, are all prime. I can vouch for the truth of this proposition. However, proving it is very challenging, and I must admit that I have not yet fully discovered the proof. I wouldn't propose it for your consideration if I had already figured it out.*

Shifting our focus back to Fermat's last theorem, it is believed to be also highly improbable that Fermat would have claimed to have proved his last theorem just because he succeeded in proving it for a few small exponents.

In contrast, Gauss believed Fermat's assertions were mostly extrapolations from particular cases. To attest to this we look at one of the quotes from Gauss which states:

> *Higher arithmetic has this special feature that many of its most beautiful theorems may be easily discovered by induction, while any proof can be only obtained with the utmost difficulty. Thus, it was one of the great merits of Euler to have proved several of Fermat's theorems which he obtained, it appears, by induction.*

Even though he himself gave a proof for the case of cubes, Gauss did not hold the problem in such high esteem. On March 21, 1816, he wrote about the recent mathematical contest of the Paris Academy on Fermat's last theorem:

> *I am very much obliged for your news concerning the Paris prize. But I confess that Fermat's theorem as an isolated proposition has very little interest for me, because I could easily lay down a multitude of such propositions, which one could neither prove nor dispose of.*

Pythagoras and $n = 2$

Pythagoras: If we just look at the integer solutions of the equation we are considering,

$$x^n + y^n = z^n$$

but this time letting $n = 2$, it looks quite similar to the problem we are dealing with but has been figured out by Pythagoras many centuries ago.

Pythagoras was an ancient Greek mathematician and philosopher who lived around 570-495 BCE. Pythagoras is considered one of the founding figures of Western mathematics and philosophy, and his work laid the foundation for many subsequent developments in these fields. His life and work are surrounded by legend and myth, but he is most famous for the Pythagorean Theorem, which is named after him and is a famous result in geometry that states that in a right-angled triangle, the square of

the length of the hypotenuse (the side opposite the right angle) is equal to the sum of the squares of the lengths of the other two sides. Mathematically, it can be expressed as:

$$a^2 + b^2 = c^2$$

This theorem as mentioned above deals with integer solutions (e.g., the well-known 3-4-5 triangle) and has been known and proven for thousands of years.
Unlike Pythagoras' Theorem, Fermat's Last Theorem is a statement about the non-existence of certain integer solutions. In summary, while both Pythagoras' Theorem and Fermat's Last Theorem deal with integer solutions, they are very different in terms of their content, complexity, and the nature of the problems they address. Pythagoras' Theorem is a classical result in geometry, while Fermat's Last Theorem is a profound statement in number theory and algebraic geometry that remained unproven for centuries.

In attempting to validate Fermat's theorem for all positive integers $n \geq 3$, one straightforward observation emerges. If the theorem is valid for a certain integer m, and if $n = l \times m$, then the theorem is equally applicable to n. This is because, assuming x, y, z are nonzero integers satisfying $x^n + y^n = z^n$, we can deduce $(x^l)^m + (y^l)^m = (z^l)^m$, which conflicts with the original assumption. Given that any integer $n \geq 3$ is either a multiple of 4 or a non-2 prime p, it's sufficient to establish the validity of Fermat's theorem for $n = 4$ and for each prime $p \neq 2$. Additionally, I may intermittently discuss proofs for exponents in the form $2p$ or p^n, where p is an odd prime.
Fermat's Last Theorem can be dissected into two specific scenarios:
The first scenario is applicable when, for a given exponent p, no integers x, y, z can be found such that $p \nmid xyz$ and $x^p + y^p = z^p$. The second scenario is applicable when, for the same exponent p, there are no nonzero integers x, y, z with $p \mid xyz$, $\gcd(x, y, z) = 1$, and $x^p + y^p = z^p$.

Smaller cases

Euler and Gauss: Fermat's assertion for biquadrates (fourth powers) used the method of infinite descent. By assuming a solution exists, this method then derives successively smaller solutions, eventually leading to a contradiction and therefore invalidating the initial assumption.
Euler turned his attention to the case of cubes (third powers). He began by making a change of variables to represent x and y as $x = a - b$ and $y = a + b$, respectively. This led him to study numbers of the form $a^2 + 3b^2$. Euler's approach, although intricate, effectively demonstrated that no three integers could satisfy the equation $x^3 + y^3 = z^3$.
Gauss's proof for the cube case ventured into the realm of complex numbers, specifically those from the quadratic field $\mathbb{Q}(\sqrt{-3})$. He introduced the concept of integers in this field and explored their divisibility properties, laying the foundation for future investigations into number fields.

Sophie Germain:
Sophie Germain (1776-1831) stands as one of the luminaries in the realm of mathematics, having made profound contributions to number theory, elasticity, and acoustics. Among her correspondents was the renowned mathematician Gauss, who discovered her true identity only after extolling her genius to her

friend, Joseph Louis Lagrange.

Central to the history of Fermat's Last Theorem is Germain's pivotal contribution, commonly referred to as "Germain's Theorem" or the "Sophie Germain Identity". Germain's Theorem posits that if there exists an auxiliary prime p such that:

1. p doesn't divide the product xyz,
2. x, y, z aren't divisible by p,
3. $x^p + y^p + z^p$ is divisible by p,

, then x, y, z cannot be solutions to Fermat's equation for the exponent p.
Armed with this theorem, Germain demonstrated the absence of solutions for primes less than 100, under the conditions her theorem outlined.
In addition, the "Sophie Germain Identity" is elegantly captured as:

$$a^4 + 4b^4 = (a^2 + 2b^2 + 2ab)(a^2 + 2b^2 - 2ab)$$

While this identity isn't directly tied to her work on Fermat's Last Theorem, it underscores her invaluable contributions to algebra and number theory.
Sophie Germain's enduring legacy is twofold: her indelible mark on mathematics and the inspiration she provides as a testament to perseverance and passion in the face of societal constraints.

Kummer and regular primes:
Ernst Eduard Kummer's approach to Fermat's Last Theorem signaled a groundbreaking shift in understanding. He connected the theorem to intricate questions about class numbers of cyclotomic fields, paving the way for the introduction of advanced algebraic number theory concepts.
At the heart of Kummer's work was the factorization of Fermat's equation over the ring $\mathbb{Z}[\zeta_\ell]$, which encompasses the ℓth roots of unity:

$$(x + y)\,(x + \zeta_\ell y) \cdots \left(x + \zeta_\ell^{\ell-1} y\right) = z^\ell$$

This expression illuminated the nature of factorization within $\mathbb{Z}[\zeta_\ell]$. While it might not always exhibit the unique factorization property as integers do, Kummer demonstrated it always allowed unique factorization into prime ideals.
He introduced the notion of the ideal class group, a concept that measures the failure of the ring $\mathbb{Z}[\zeta_\ell]$ to satisfy unique factorization. If the class group's order, denoted h_ℓ, was not divisible by ℓ, then the ring fulfilled a property he termed UF_ℓ:

$$ab = z^\ell, \text{ and } \gcd(a, b) = 1 \Rightarrow a \text{ and } b \text{ are } \ell \text{ th powers up to units of } R$$

Using this, Kummer introduced the concept of "regular" primes: if ℓ was regular, Fermat's theorem held true for exponent ℓ.
Kummer further delved deep into the relationship between the class number h_ℓ and the Bernoulli numbers B_n, providing an explicit formula for the class number h_ℓ^-. If ℓ didn't divide any of the numerators

of B_{2i} for $1 \leq i \leq (\ell - 3)/2$, then ℓ was regular.

Historical remark concerning Kummer:
A well-known story concerning a wrong proof of Fermat's theorem submitted by Kummer, originates with Hensel. Specifically, in his address to commemorate the first centennial of Kummer's birth, Hensel (1910) stated:

> *"Although it is not well known, Kummer at one time believed he had found a complete proof of Fermat's theorem. (This is attested to by reliable witnesses including Mr. Gundelfinger who heard the story from the mathematician Grassmann.) Seeking the best critic for his proof, Kummer sent his manuscript to Dirichlet, author of the insuperably beautiful proof for the case $\lambda = 5$. After a few days, Dirichlet replied with the opinion that the proof was excellent and certainly correct, provided the numbers in α could not only be decomposed into indecomposable factors, as Kummer proved, but that this could be done in only one way. If however, the second hypothesis couldn't be satisfied, most of the theorems for the arithmetic of numbers in α would be unproven and the proof of Kummer's theorem would fall apart. Unfortunately, it appeared to him that the numbers in α didn't actually possess this property in general."*

This is confirmed in a letter, which is not dated (but likely from the summer of 1844), written by Eisenstein to Stern, a mathematician from Göttingen. In a recent paper, Edwards (1975) analyzes this information, in the light of a letter from Liouville to Dirichlet and expresses doubts about the existence of such a "false proof" by Kummer.
One can see from its brief history, how the so simply stated problem remained unsolved for so long.
The main idea that finally did the job is to construct a Galois representation associated with an elliptic curve, and then to show that this representation is modular, i.e., it can be realized as a representation of the group of automorphisms of a modular curve.
In this master's thesis, we will explore the Modularity Theorem and its role in the proof of Fermat's Last Theorem.
Furthermore, this thesis will delve into the technical details of the Modularity Theorem and discuss the key objects involved in the discussion of the modularity theorem.

We begin with elliptic curves in chapter 1.
The statement that Andrew wiles proved in [TW95] was

All Semistable elliptic curves over Q are modular.

Further building blocks include modular forms, Hecke operators, Eichler-Shimura relation and Galois representations.
Chapter 2 discusses modular forms and the discussion about modular forms continues in chapter 3 via Hecke operators. Think of this chapter more or less as linear algebra but at the level of modular forms.
Chapter 4 and 5 discusses different aspects of curves in general and further goes on to discuss Jacobians and Abelian varieties.
Chapter 6 discuss brief history of Galois representations and then introduces basic terminology required to discuss the modularity theorem.
Lastly, chapter 7 discusses brief account of proof of Fermat's last theorem.

We will also discuss the various approaches taken by mathematicians over the centuries to solve Fermat's Last Theorem, and how the Modularity Theorem provided the key breakthrough that ultimately led to its solution. Over all, this thesis aims to provide a comprehensive overview of the Modularity Theorem and Fermat's Last Theorem and to explore their historical, technical, and philosophical aspects. We aim to make this thesis as accessible as possible to graduate as well as undergraduate students.
There are many good books on this topic, some of them specialising for a particular topic, for an example [Sil13], for elliptic curves.
For a detailed account on the proof of Taniyama-Shimura conjecture, we refer the reader to [DDT95] or [CS99], that is of course, apart from the main papers [TW95],[Wil95a].
I am deeply indebeted to these papers, and I cannot express my gratitude enough.
The main reference for our thesis will be the Book by Fred Diamond and Jerry Shurman, [DS05]. We closely follow this book. Other main references for the thesis which we have followed are [DDT95], [DDT97], [Bos03], [Mas15], [HS00], [Sai14], [Lan95], [Rib79],[Vis18].

Acknowledgements:

First of all, I would like to thank my parents Manoj Sant and Sonali Sant for raising me the way they did. Next, I would like to immensely thank the Department of Mathematics, RPTU Kaiserslautern-Landau, for supporting me for most of Masters studies with a scholarship and/or a HiWi Job. This took off financial pressure greatly, and I could focus on my studies comfortably. Thanks to Deutschlandstipendium and DAAD Scholarship for the funding. Next, I would like to express my huge gratitude towards my Thesis advisors Dr. Jeroen Hanselman and Prof. Fieker. Without them, I wouldnt have learnt so much. Thanks to Dr. Hanselman for answering all my questions with grace and patience. Thanks to Prof. Fieker for supporting me throughout the thesis, when it came to Mathematics or even for recommending to PhD positions or Scholarships etc. I would like to thank my friends Soham and Cedric for so many helpful mathematical discussions. Special thanks to my parents and to Soham and Manuja for being great friends and always looking after me and proving the meaning of actually being there for someone. A special mention to Manuja for the great years spent during my Bachelors studies and for shaping me the way I am today being a true friend to me and lastly for motivating me to come to Germany for my masters. If it wasn't for her, I would not have considered coming to Germany in the first place. Lastly, I thank my inner self for being strong during hardships to finally see myself sail through.

Contents

1 Elliptic Curves

As discussed in the introduction, the aim of the thesis is to give a formal overview of the proof of one of the most celebrated problems(now a Theorem) in Mathematics, Fermat's last theorem, due to Prof. Dr. Andrew Wiles in 1995. Elliptic curves have been one of the central and key objects in the proof of Fermat's last theorem. Sir Andrew Wiles proved in his paper that all semistable elliptic curves over the set of rational numbers $\mathbb{Q}$ are modular. Fermat's Last theorem follows as a corollary by virtue of previous work by Frey, Serre and Ribet. Thus it is important that we begin by introducing the building blocks and the objects of the proof and study certain aspects revolving around them which are important to completely understand the proof.
We will closely follow the book [DS05] and Lecture notes from Marc Masdeu, [Mas15] , [Bos03], [Sil13]. We will sometimes closely follow proofs from either of these sources for the sake of completeness. At times the ideas of the proofs are not original, neither I want to claim so, although, at some places, I have expanded on my own giving more details.
We will start with defining double coset operators, a notion that lies as a key concept in the background of theory. At times, the ideas are inevitably not original but closely followed.

1.1 Weierstrass Equations

The Weierstrass equation is a fundamental equation in the theory of elliptic curves. It provides a standard form for elliptic curves over a field (often over the complex numbers, but also over other fields like the reals or rationals).

Definition 1.1.1
The general Weierstrass equation over a field k is given by:

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

where, a_1, a_2, a_3, a_4, and a_6 are coefficients lying in the given field k.

Remark 1.1.2
The equation defines a cubic curve in the projective plane.

Definition 1.1.3 (Elliptic Curves)
An elliptic curve (E, O) over a field k is a smooth, projective curve E of genus 1 along with a specified point O known as the base point, all defined over k.

Remark 1.1.4
O is usually understood as the base point, and thus we will denote an elliptic curve defined over a field k by E instead of (E, O). Furthermore, the non-singularity or smoothness just means that the curve has no cusps or self-intersections.

The following proposition gives the exact connection between Weierstrass equations and elliptic curves.

Proposition 1.1.5
Let E be an elliptic curve. Then there exist coordinate functions $x, y \in k(E)$ and constants $a_1, a_2, a_3, a_4, a_6 \in k$ such that the following condition holds:

If $f : E \to \mathbb{P}^2_k$ is the map given by $P \mapsto [x(P) : y(P) : 1]$, then f gives an isomorphism of elliptic curves(understanding where the base point is mapped to) between E and the curve $Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$. Conversely, every such equation gives an elliptic curve with base point $O = [0 : 1 : 0]$ whenever it is smooth.

Proof:
See, in section 4.2 in this thesis. □

First, let us restrict to the affine case and consider a Weierstrass equation,

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6. \tag{1}$$

1.2 Simplified Weierstrass equation, Examples

In this section, we'll focus on the affine case, using the non-homogeneous coordinates x and y. It's important to remember that in this setting, there's always a point at infinity, denoted as $O = [0 : 1 : 0]$. Let's take a closer look at a Weierstrass equation, which we'll call E. If the characteristic of the field k is not 2, we can simplify the equation by completing the square. This is done by introducing a new variable, η, defined as:

$$\eta = y + \frac{a_1x + a_3}{2}.$$

Using this transformation, E can be rewritten as:

$$\eta^2 = x^3 + \frac{b_2}{4}x^2 + \frac{b_4}{2}x + \frac{b_6}{4},$$

where the coefficients b_2, b_4, and b_6 are defined as $b_2 = a_1^2 + 4a_2$, $b_4 = a_1a_3 + 2a_4$, and $b_6 = a_3^2 + 4a_6$, respectively.
Similarly, if the characteristic of k is not 3, we can complete the cube by introducing another variable, ξ, defined as $\xi = x + \frac{b_2}{12}$. This transforms the equation into:

$$\eta^2 = \xi^3 - \frac{c_4}{48}\xi - \frac{c_6}{864},$$

where $c_4 = b_2^2 - 24b_4$ and $c_6 = -b_2^3 + 36b_2b_4 - 216b_6$.
Additionally, we define two more quantities:

$$b_8 = a_1^2a_6 - a_1a_3a_4 + 4a_2a_6 + a_2a_3^2 - a_4^2$$

and

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6.$$

Definition 1.2.1
The quantity Δ is called the discriminant of a Weierstrass equation.
The j-invariant of the elliptic curve E is defined as:

$$j = \frac{c_4^3}{\Delta}.$$

The discriminant Δ is a crucial parameter in understanding the nature of the elliptic curve E, and the

j-invariant is a key characteristic that helps classify the curve.
Furthermore, when the characteristic of k is neither 2 nor 3, using suitable transformations, any Weierstrass equation in the form of $\eta^2 = x^3 + \dots$ can be simplified to a more manageable form:

$$y^2 = x^3 - 27c_4x - 54c_6.$$

This leads us to what is known as the simplified Weierstrass form:

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

This simplified form is easier to work with and forms the basis for many further investigations in the study of elliptic curves.

Next, we give a proposition that characterises the discriminant of a Weierstrass equation and provides a lovely connection between j-invariants and elliptic curves.

Proposition 1.2.2
Let E be a curve given by a Weierstrass equation. Then,
E is non-singular if and only if $\Delta \neq 0$.

Proof:
Consider an elliptic curve E defined by the Weierstrass equation:

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6.$$

Our first goal is to demonstrate that the point at infinity is not a singular point on this curve. To do this, we extend the equation into the projective plane $\mathbb{P}^2$ by introducing a homogenized form of the equation:

$$F(X, Y, Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 = 0.$$

We focus on the point at infinity, $O = [0 : 1 : 0]$. Calculating the partial derivative of F with respect to Z at O gives us:

$$\frac{\partial F}{\partial Z}(O) = 1 \neq 0.$$

This result, according to the Projective Jacobi criterion, confirms that O is indeed a non-singular point on E.
Now, let's suppose that E has a singular point, denoted as $P_0 = (x_0, y_0)$. By shifting coordinates:

$$x = x' + x_0, \quad y = y' + y_0,$$

we note that after doing tedious and straight forward computations the discriminant Δ and c_4 remain unchanged. Therefore, we can assume, without loss of generality, that E is singular at the origin $(0, 0)$. This leads us to deduce that:

$$a_6 = f(0, 0) = 0, \quad a_4 = \frac{\partial f}{\partial x}(0, 0) = 0, \quad a_3 = \frac{\partial f}{\partial y}(0, 0) = 0.$$

Consequently, the equation for E simplifies to:

$$E : y^2 + a_1xy - a_2x^2 - x^3 = 0,$$

from which it follows that $\Delta = 0$.
Conversely, assuming that the characteristic of the field k is not 2, let's consider a Weierstrass equation in the form:

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_0.$$

It can be shown that the curve associated with this equation is singular if and only if it has a double point of the form $(x_0, 0)$. This condition is equivalent to the vanishing of the discriminant, which is 16Δ in this case.
A similar approach can be employed for charachteristic 2 and 3. For more details see, [Sil13] □

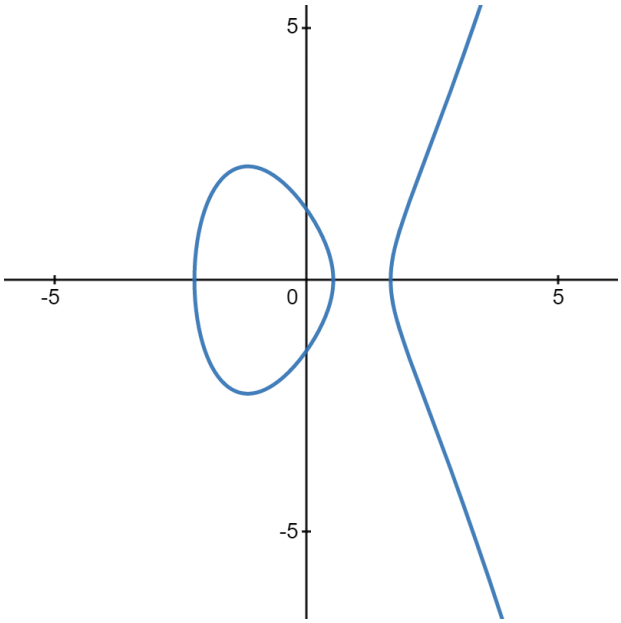
Proposition 1.2.3
Two elliptic curves E, E' are isomorphic (over $\bar{k}$) if and only if they have the same ***j*-invariant**.

Proof:
It is clear that if E and E' are isomorphic over $\bar{k}$, we get that their *j-invariants* are equal. This is due to the fact that, we can get weierstrass equations assosciated with them and establish a transformation, which maps one equation to the other. It can be noted that, since the quantities c_4, Δ remain invariant under such transformations and so do their j-invariants.

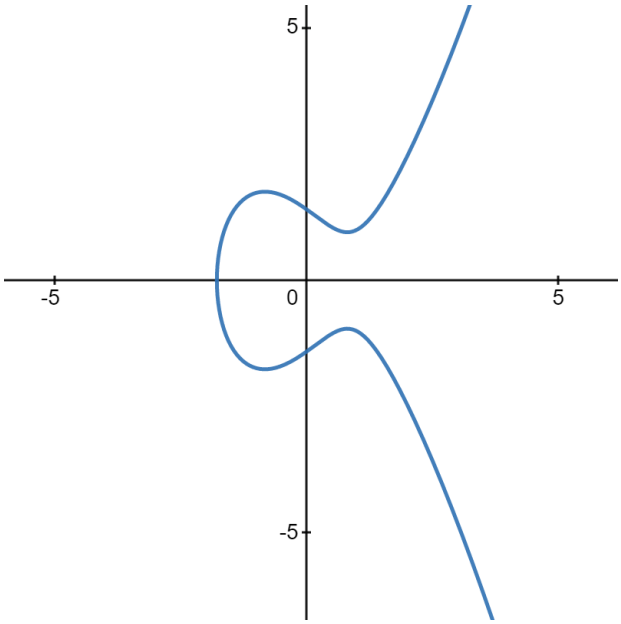
Conversely, again for simplicity, assume $char(k) \neq 2, 3$, and thus as per the discussion above, we get a simplified Weierstrass equation of the form $y^2 = x^3 + ax + b$.
In this case, we get that $\Delta = -16(4a^3 + 27b^2)$ and $j = -1728\frac{64a^3}{\Delta}$.
Let $E : y^2 = x^3 + ax + b$
$E' : y^2 = x^3 + Ax + B$. Assume that their j-invariants are equal, but then by the above definitions, we get that $a^3B^2 = A^3b^2$. Now, we make a case distinction:
case 1) If $a = 0$, it is clear that since an Elliptic curve is non-singular and by part (a) $\Delta \neq 0$ forces that $A = 0$ and thus we get that $(x, y) \mapsto (c^2x, c^3y)$ where $c = (\frac{b}{B})^{\frac{1}{6}}$ gives an isomorphism between E and E' (over $\bar{k}$). Note here we use that the isomorphism is over $\bar{k}$ to get such a c.
Case 2) By symmetry, the case $b = 0$ can be handled analogously.
Case 3) When $ab \neq 0$, then, it is clear that $AB \neq 0$, thus we have that $(\frac{a}{A})^{\frac{1}{4}} = (\frac{b}{B})^{\frac{1}{6}}$, thus chose $c = (\frac{a}{A})^{\frac{1}{4}}$ as in case 1), we get an isomorphism between E and E' (over $\bar{k}$)
For more details and for proof in $char(k) = 2$, see [Sil13] □

Now that we have seen important quantities associated with the Weierstrass equation and elliptic curves we give some examples.
Note that for plotting curves, I used DESMOS

Example 1.2.4
1) Consider the Weierstrass equation $y^2 = x^3 - 4x + 2$, with the real plot below,



Let us calculate some of the quantities mentioned above: $\Delta = -16(4a^3 + 27b^2)$ and thus, substituting values of a, b we get that,
$\Delta = -16(4(-4)^3 + 27(2)^2)$, i.e,
$\Delta = 2368$.
Similarly, we know that $j = -1728\frac{64a^3}{\Delta}$,
thus again substituting values of a, b, Δ we get that,
$j = -1728\frac{64(-4)^3}{2368}$.
2) Consider the Weierstrass equation $y^2 = x^3 - 2x + 2$, with the real plot below:



1.3 Complex Tori and Elliptic curves

This section defines basic terminologies like Lattices, Complex tori, etc, and states and proves some results related to them. Furthermore, the main aim of this section is to state the correspondence

between Complex Tori and Elliptic curves.

Definition 1.3.1 (Complex lattice)
A *complex lattice* Λ is a set of the form $\mathbb{Z}\Lambda_1 \oplus \mathbb{Z}\Lambda_2$ such that
I) $\Lambda_1, \Lambda_2 \in \mathbb{C}$,
II) The set $\{\Lambda_1, \Lambda_2\}$ forms a $\mathbb{R}$- basis of $\mathbb{C}$.

For example, the set of Gaussian integers $\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$ is a complex lattice.

Let $\mathcal{H} = \{z \in \mathbb{C} : Im(z) > 0\}$ denote the complex upper half plane. We begin with a small lemma.

Lemma 1.3.2
Consider two lattices $\Lambda = \Lambda_1\mathbb{Z} \oplus \Lambda_2\mathbb{Z}$ and $\Lambda' = \Lambda_1'\mathbb{Z} \oplus \Lambda_2'\mathbb{Z}$ with $\Lambda_1/\Lambda_2 \in \mathcal{H}$ and $\Lambda_1'/\Lambda_2' \in \mathcal{H}$. Then $\Lambda' = \Lambda$ if and only if

$$\begin{bmatrix} \Lambda_1' \\ \Lambda_2' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix} \quad \text{for some } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

Proof:
Suppose there exists a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $\mathrm{SL}_2(\mathbb{Z})$ such that,

$$\begin{bmatrix} \Lambda_1' \\ \Lambda_2' \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix}.$$

This just implies that each basis vector of Λ' can be expressed as an integer linear combination of the basis vectors of Λ. Since the matrix is in $\mathrm{SL}_2(\mathbb{Z})$, its determinant is 1, and it is invertible with its inverse also having integer entries. This means that there exists a matrix $A^{-1} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$ in $\mathrm{SL}_2(\mathbb{Z})$ such that,

$$\begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \begin{bmatrix} \Lambda_1' \\ \Lambda_2' \end{bmatrix}.$$

Similarly, as above this just means that each basis vector of Λ can be expressed as an integer linear combination of the basis vectors of Λ'.
Thus, we get both the containments, $\Lambda \subset \Lambda'$ and $\Lambda' \subset \Lambda$, proving that, $\Lambda' = \Lambda$.

Now, assume $\Lambda' = \Lambda$. This means every vector in Λ' can be expressed as a linear combination of the basis vectors of Λ, and vice versa. In terms of matrices, this means that there exist matrices A, B with entries in the ring of integers $\mathbb{Z}$, such that

$$\begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix} = A \begin{bmatrix} \Lambda_1' \\ \Lambda_2' \end{bmatrix},$$

$$\begin{bmatrix} \Lambda_1' \\ \Lambda_2' \end{bmatrix} = B \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix}.$$

From, this we get that,

$$\begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix} = AB \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix}.$$

We claim that $AB = I$, where I denote the $2X2$ identity matrix.
Since A, B are matrices with entries from $\mathbb{Z}$, it is clear that AB has integer entries. Let

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Then, from above, we have,
$(a-1)\Lambda_1 + b\Lambda_2 = 0$ and similarly, $c\Lambda_1 + (d-1)\Lambda_2 = 0$.
Recall that Λ_1, Λ_2 generate a lattice and thus they are linearly independent over the real numbers. This implies that, $a - 1 = b = c = d - 1 = 0$. Thus $AB = I$.
This gives that det(A)=det(B)=$\pm$.1

But, it can be shown that

$$\mathrm{Im}(\frac{\Lambda_1'}{\Lambda_2'}) = k.\det(A)\,\mathrm{Im}.(\frac{\Lambda_1}{\Lambda_2}),$$

for some positive constant k.
Due to our assumption that $\Lambda_1/\Lambda_2 \in \mathcal{H}$ and $\Lambda_1'/\Lambda_2' \in \mathcal{H}$, we have that $\det(A) = 1$, giving us the desired matrix in $\mathrm{SL}_2(\mathbb{Z})$.

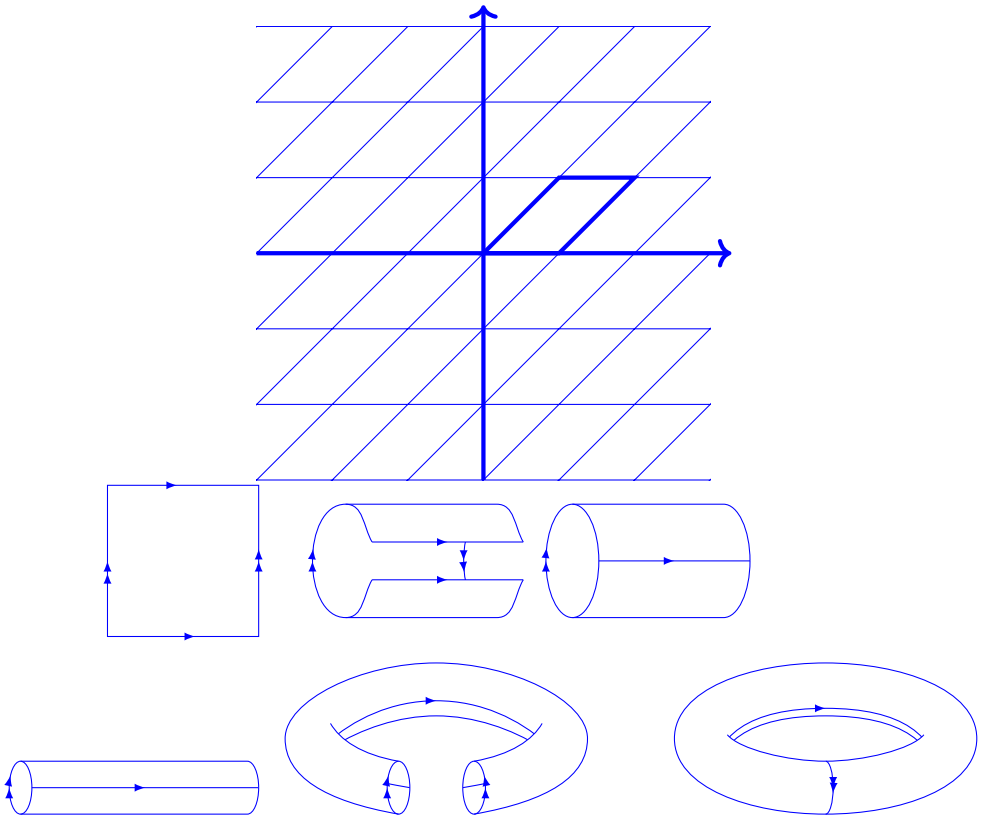
□

Remark 1.3.3
Note that, the assumption that $\Lambda_1/\Lambda_2 \in \mathcal{H}$ and $\Lambda_1'/\Lambda_2' \in \mathcal{H}$, used above in the proof is really necessary for the statement above. For example, consider, the lattices spanned by $1, i$ and $1, -i$. They are equal, but the only matrix with integer entries A satisfying the above condition is $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, which has determinant -1. Note this happened since, $-i$ does not lie in the upper half plane as that i.

Definition 1.3.4 (Complex Torus)
A complex torus is a quotient of the complex plane by a lattice,

$$\mathbb{C}/\Lambda = \{z + \Lambda : z \in \mathbb{C}\}.$$

Remark 1.3.5
A complex torus algebraically is an Abelian group under the addition it inherits from $\mathbb{C}$ as its quotient. Let us try to visualise a complex torus geometrically. Consider the lattice $\Lambda = \Lambda_1\mathbb{Z} \oplus \Lambda_2\mathbb{Z}$ and the assosciated torus $\mathbb{C}/\Lambda$. It is a parallelogram spanned by $\{\Lambda_1, \Lambda_2\}$ with its sides identified in opposing pairs. Identifying one pair of sides rolls the parallelogram into a tube, and then identifying the other pair bends the tube into a torus. The following diagram explains this visually.



I learned how to draw this on this stack exchange page.(Hyperlink included in digital version)
Topologically, A complex torus is a Riemann surface. We will discuss the topological aspects in greater depth in the next chapter.

Definition 1.3.6 (Isogeny)
A nonzero holomorphic homomorphism between *complex tori* is called an **isogeny**.

We now state a couple of results without proof. They are essential for further discussion of isogenies.

Proposition 1.3.7
Suppose $\varphi : \mathbb{C}/\Lambda \longrightarrow \mathbb{C}/\Lambda'$ is a holomorphic map between *complex tori*. Then there exist complex numbers m, b with $m\Lambda \subset \Lambda'$ such that $\varphi(z + \Lambda) = mz + b + \Lambda'$. The map is a bijection if and only if $m\Lambda = \Lambda'$.

Corollary 1.3.8
Suppose $\varphi : \mathbb{C}/\Lambda \longrightarrow \mathbb{C}/\Lambda'$ is a holomorphic map between complex tori, $\varphi(z + \Lambda) = mz + b + \Lambda'$ with $m\Lambda \subset \Lambda'$. Then the following are equivalent:
(1) φ is a group homomorphism,
(2) $b \in \Lambda'$, so $\varphi(z + \Lambda) = mz + \Lambda'$,
(3) $\varphi(0) = 0$.
In particular, there exists a nonzero holomorphic group homomorphism between the complex tori $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ if and only if there exists some nonzero $m \in \mathbb{C}$ such that $m\Lambda \subset \Lambda'$, and there exists a holomorphic group isomorphism between the complex tori $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ if and only if there exists some $m \in \mathbb{C}$ such that $m\Lambda = \Lambda'$.

See, [DS05], chapter 1 for more details.

Remark 1.3.9
Consider a generic lattice Λ, which is generated by two complex numbers Λ_1 and Λ_2, where Λ_1 and Λ_2 are such that their ratio Λ_1/Λ_2 is in the upper half-plane,$\mathcal{H}$. Define τ as this ratio, Λ_1/Λ_2, and construct another lattice, Λ_τ, using τ and 1 as its basis.
The key insight here is that scaling the original lattice Λ by $1/\Lambda_2$ transforms it into Λ_τ. This scaling leads us to an interesting map, φ_τ, which takes a point z in the complex plane modulo the original lattice Λ and maps it to the scaled point z/Λ_2 modulo the new lattice Λ_τ. By 1.3.8, φ_τ, is an isomorphism, meaning it preserves the structure between these two complex tori.
This result implies that every complex torus can be equivalently represented by a torus whose lattice is generated by a complex number τ in the upper half-plane $\mathcal{H}$ and the number 1. While this representation is not unique for each torus, any other representative τ' in $\mathcal{H}$ can be related to τ through via the action of an element in $\mathrm{SL}_2(\mathbb{Z})$. In essence, this means that each complex torus is uniquely associated with a point τ in the upper half-plane, modulo the action of $\mathrm{SL}_2(\mathbb{Z})$. We will see this again when we discuss fundamental domains.

Let us now come to the aim of this section. We will establish correspondence between elliptic curves over $\mathbb{Q}$ and complex tori and thus we can interchangeably use both the terms and use their properties to connect the Complex analytic and algebraic worlds.

We start with discussing Eisenstein Series, the Weierstrass function and the standard results around them. In the end, we establish the desired correspondence.

Definition 1.3.10 (Eisenstein Series)
Let $k \in \mathbb{N}$. The Eisenstein series of weight k with respect to a Complex lattice Λ is

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

Proposition 1.3.11
The Eisenstein series $G_k(\lambda)$ converges absolutely for all $k \geq 3$.

Proof:
For any natural number m, let's look at two specific finite sums. The first sum, S_m, is defined as:

$$S_m = \sum_{\substack{\lambda = m_1\Lambda_1 + m_2\Lambda_2 \\ -m \leq m_1, m_2 \leq m}} \frac{1}{|\lambda|^k},$$

and the second sum, T_{m+1}, is essentially the difference between two consecutive sums S_m:

$$T_{m+1} = S_{m+1} - S_m.$$

$$\sum_{\lambda \in \Lambda \setminus \{0\}} \frac{1}{|\lambda|^k} = \sum_{m=1}^{\infty} T_m.$$

To establish the convergence of this series, we focus on the sum on the right-hand side.
Take any element $\lambda = m_1\Lambda_1 + m_2\Lambda_2$ within the bounds $-m \leq m_1, m_2 \leq m$. We define two parameters, a and b, as the minimum and maximum, respectively, of the absolute values of Λ_1, Λ_2, $\Lambda_1 + \Lambda_2$, and $\Lambda_1 - \Lambda_2$:

$$a = \min\{|\Lambda_1|, |\Lambda_2|, |\Lambda_1 + \Lambda_2|, |\Lambda_1 - \Lambda_2|\}$$

and

$$b = \max\{|\Lambda_1|, |\Lambda_2|, |\Lambda_1 + \Lambda_2|, |\Lambda_1 - \Lambda_2|\}.$$

Using these, we can estimate the bounds for $|\lambda|$:

$$ma \leq |\lambda| \leq mb.$$

This leads to the inequalities:

$$\frac{1}{(ma)^k} \geq \frac{1}{|\lambda|^k} \geq \frac{1}{(mb)^k}.$$

Considering that each T_m comprises $8m$ terms, we derive:

$$\frac{8m}{(ma)^k} \geq T_m \geq \frac{8m}{(mb)^k}.$$

Thus, we arrive at the following bounds for our series:

$$\frac{8}{b^k}\sum_{m=1}^{\infty}\frac{1}{m^{k-1}} \leq \sum_{m=1}^{\infty} T_m \leq \frac{8}{a^k}\sum_{m=1}^{\infty}\frac{1}{m^{k-1}}.$$

Consequently, we conclude that the Eisenstein series converges for $k \geq 3$. □

Given a lattice Λ, the meromorphic functions $f : \mathbb{C}/\lambda \longrightarrow \widehat{\mathbb{C}}$ on the torus are naturally identified with the Λ-periodic meromorphic functions $f : \mathbb{C} \longrightarrow \widehat{\mathbb{C}}$ on the plane. We now see an important example of such a function.

Definition 1.3.12 (Weierstrass $\wp$-function)
Let Λ be a complex lattice. The function given by

$$\wp_\lambda(z) = \frac{1}{z^2} + \sum_{\lambda\in\Lambda}{}' \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right), \quad z \in \mathbb{C}, z \notin \Lambda.$$

is called the Complex Weierstrass $\wp$-function associated to the lattice Λ.
Usually, we will for simplicity skip the notation $\wp_\Lambda$ if the lattice we are working with is clear and rather use $\wp$ and call this as the Weierstrass function associated with the lattice Λ.

Proposition 1.3.13
The Weierstrass $\wp$-function associated with a lattice Λ is meromorphic.

Proof:

Let λ be a lattice and let $R > 0$ be arbitrary. Let is define two functions, $g(z)$ and $f(z)$, using λ defined as follows:

$$g(z) = \frac{1}{z^2} + \sum_{\substack{\lambda \in \Lambda \setminus \{0\} \\ |\lambda| \leq 2R}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right),$$

and $f(z)$ is given by:

$$f(z) = \sum_{\substack{\lambda \in \Lambda \setminus \{0\} \\ |\lambda| > 2R}} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right).$$

Let us quickly note that with these newly defined functions, the Weierstrass $\wp$-function can be expressed as $\wp(z) = g(z) + f(z)$.
The function g is essentially a finite sum because it is summed over all the lattice points inside a bounded disc, so it is meromorphic on the open disc centered at zero with radius R, D_R. It remains to show that f is holomorphic on D_R to imply that $\wp$ is meromorphic on D_R. Furthermore, once achieving this it can be seen that since this is valid for any $R > 0$, $\wp$ is meromorphic on the entire complex plane $\mathbb{C}$.
For absolute and uniform convergence of f within D_R, consider a point z in D_R and an element λ from λ where $|\lambda| > 2R$.

We then have the following inequilities by simple application of triangle inequility for the usual absolute value on the complex plane:

$$|2\lambda - z| \leq 3|\lambda|$$

and

$$|z - \lambda| \geq \frac{|\lambda|}{2}.$$

A simple corollary of these inequilities is that $f(z)$ can be simplified as:

$$\left| \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right| = \frac{|z||2\lambda - z|}{|z-\lambda|^2|\lambda|^2} \leq \frac{3R|\lambda|}{\frac{|\lambda|^2}{4}|\lambda|^2} = 12R\frac{1}{|\lambda|^3}.$$

This gives,

$$\sum_{\substack{\lambda \in \Lambda \setminus \{0\} \\ |\lambda| > 2R}} \left| \frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right| \leq 12R \cdot G_3(\lambda).$$

This with the application of 1.3.11 show the absolute convergence of f. The uniformity of the convergence is assured since the upper bound involved does not depend on z.
Therefore, we have shown that $\wp$ is a meromorphic function on the complex plane. □

Proposition 1.3.14
$\wp$ is an even function and that $\wp'$ is Λ-periodic.

Proof:
Indeed, for all $z \in \mathbb{C}$ and all $\lambda \in \Lambda$ we have

$$\frac{1}{(z-\lambda)^2} = \frac{1}{(-z+\lambda)^2}.$$

Since $\wp$ is a meromorphic function, we may reorder the series by interchanging Λ with $-\Lambda$, and thus obtain $\wp(-z) = \wp(z)$ for all $z \in \mathbb{C}$.
Next, let us compute the first derivative $\wp'$ of $\wp$ by differentiating the summands of the series:

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

Since $\wp$ is meromorphic, $\wp'$ is also meromorphic, and hence the series on the right is absolutely convergent. Thus, we reorder the summands, to obtain

$$\sum_{\lambda \in \Lambda} \frac{-2}{(z-\Lambda)^3} = \sum_{\lambda \in \Lambda} \frac{-2}{(z-\Lambda+\Lambda')^3}$$

for any element $\lambda' \in \Lambda$. This implies

$$\wp'(z) = \wp'\left(z+\Lambda'\right)$$

Thus, $\wp'$ is Λ-periodic

□

Corollary 1.3.15
$\wp$ is Λ-periodic.
Proof:
Consider the function $(\wp(z) - \wp\left(z+\Lambda\right))$ for $\lambda \in \Lambda$. We have that,

$$\left(\wp(z) - \wp\left(z+\Lambda_0\right)\right)' = 0 \text{ for all } z \in \mathbb{C}\setminus\Lambda.$$

This is in particular true for the two generators Λ_1, Λ_2 of Λ.
Thus there exist constants $c_1, c_2 \in \mathbb{C}$ such that

$$\wp(z) = \wp\left(z+\Lambda_i\right) + c_i \text{ for all } z \in \mathbb{C}\setminus\Lambda, i = 1,2$$

The function $\wp$ is periodic if and only if $c_1 = c_2 = 0$.
To compute c_1, it suffices to consider one point of $\mathbb{C}\setminus\Lambda$.
We choose $z := -\frac{\Lambda_1}{2}$.
Since Λ_1 is a generator of Λ, we have $z \in \mathbb{C}\setminus\Lambda$. We compute

$$\wp\left(-\frac{\Lambda_1}{2}\right) = \wp\left(\frac{\Lambda_1}{2}\right) + c_1 = \wp\left(-\frac{\Lambda_1}{2}\right) + c_1$$

For the last equality, we use that $\wp$ is an even function from 3. From this immediately follows $c_1 = 0$, and analogously for c_2.□ □

Proposition 1.3.16 (The differential equation)

Let Λ be a lattice and let $G_k(\Lambda), \wp(z)$ denote the Eisenstein series and Weierstrass function assosciated to the lattice Λ defined as above.
Then we have that,

$$\big(\wp'(z)\big)^2 = 4(\wp(z))^3 - g_2(\Lambda)\wp(z) - g_3(\Lambda)$$

where $g_2(\Lambda) = 60G_4(\Lambda)$ and $g_3(\Lambda) = 140G_6(\Lambda)$.

Proof:
Let's consider a meromorphic function defined by:

$$f(z) := \wp'(z)^2 - 4\wp(z)^3 + g_2\wp(z) + g_3.$$

This function, f, belongs to the field of meromorphic functions on the lattice Λ. It's interesting to note that the only potential poles of f are the points of Λ.
To delve deeper, we look at the power series expansion of $\wp(z)$ around 0:

$$\wp(z) = \frac{1}{z^2} + 3G_4(\Lambda)z^2 + 5G_6(\Lambda)z^4 + \dots$$

which gives,

$$4\wp(z)^3 = \frac{4}{z^6} + \frac{36}{z^2}G_4(\Lambda) + 60G_6(\Lambda) + 36G_4(\Lambda)^2z^2 + \dots.$$

Similarly, differentiating $\wp(z)$ gives us:

$$\wp'(z) = -\frac{2}{z^3} + 6G_4(\Lambda)z + 20G_6(\Lambda)z^3 + \dots$$

and consequently:

$$\wp'(z)^2 = \frac{4}{z^6} - \frac{24}{z^2}G_4(\Lambda) - 80G_6(\Lambda) + \dots$$

Combining these power series and arranging them by powers of z, we obtain:

$$f(z) = \begin{cases} (4-4)\frac{1}{z^6} \\ (-24G_4(\Lambda) - 36G_4(\Lambda) + g_2)\,\frac{1}{z^2} \\ (-80G_6(\Lambda) - 60G_6(\Lambda) + g_3) + \dots \end{cases}$$

From this, it becomes clear that we can extend f at the point $z_0 = 0$ by simply setting $f(0) := 0$.
Due to the periodic nature of f, this extension is valid for all points in Λ. This implies that f is a constant function. Since we know that $f(0) = 0$, it follows that f must be identically zero everywhere. Therefore, this concludes that the differential equation, as represented by $f(z)$, is indeed valid. □

We will now make use of this differential equation to achieve the correspondence between Complex Tori and Elliptic curves. But first we need this lemma.

Lemma 1.3.17
The first derivative $\wp'$ of the Weierstrass $\wp$-function has exactly three zeroes each of order one on the semi-open parallelogram of periods P_{Λ_1,Λ_2}, which are

$$\rho_1 = \frac{\Lambda_1}{2}, \quad \rho_2 = \frac{\Lambda_2}{2}, \quad \text{and} \quad \rho_3 = \frac{\Lambda_1 + \Lambda_2}{2},$$

and the values $e_i := \wp(\rho_i), i = 1, 2, 3$ are pairwise different.

Proposition 1.3.18
Given an elliptic curve

$$y^2 = 4x^3 - a_2x - a_3 \quad a_2, a_3 \in \mathbb{C},$$

there exists a lattice Λ such that $a_2 = g_2(\Lambda)$ and $a_3 = g_3(\Lambda)$.

Proof:
The proof of this proposition is partially left as an exercise in [DS05] as it assumes a_2, a_3 being non-zero. Let us try to fill in the details for these cases.
Let $\Lambda = \mathbb{Z}[i]$. Then, for the lattice Λ, it holds that $G_6(\Lambda) = 0$.

Let us consider the function $G_6(\Lambda)$ defined as

$$G_6(\Lambda) = \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^6}.$$

We apply the automorphism given by the multiplication by the imaginary unit i, That is , $\varphi(\omega) = i\omega$ to each term in $G_6(\Lambda)$:

$$\varphi(G_6(\Lambda)) = \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{(i\omega)^6} = \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{i^6\omega^6} = -G_6(\Lambda).$$

However, the sum $G_6(\Lambda)$ is defined over the lattice Λ and φ is an automorphism, thus it remains invariant under lattice automorphism.

Therefore, $\varphi(G_6(\Lambda)) = G_6(\Lambda)$.
Equating the two expressions for $G_6(\Lambda)$, we have:

$$G_6(\Lambda) = -G_6(\Lambda) \implies G_6(\Lambda) = 0.$$

Next, let $\zeta = e^{(\frac{2\pi i}{3})}$. For the lattice Λ generated by 1 , ζ, we have $G_4(\Lambda) = 0$. Let us onsider the function $G_4(\Lambda)$ defined by:

$$G_4(\Lambda) = \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^4}.$$

Recall that ζ satisfies $\zeta^2 + \zeta + 1 = 0$ and $\zeta^3 = 1$. Thus we get an automorphism $\varphi : \Lambda \to \Lambda$ defined by $\varphi(\omega) = -\zeta\omega$. Note that, φ being a homomorphism and injective is really trivial. The only challenging part is the surjectivity. But given any $a + b\zeta$, $a, b \in \mathbb{Z}$, consider $(b - a) + a\zeta \in \Lambda$. We have that,

$$\varphi((-b + a) + a\zeta) = -\zeta\left[(-b + a) + a\zeta\right].$$

Expanding this expression:

$$\begin{aligned} -\zeta\left[(-b + a) + a\zeta\right] &= -\zeta(-b + a) - \zeta^2 a \\ &= \zeta(b - a) - \zeta^2 a. \end{aligned}$$

Here, ζ is a primitive 3rd root of unity, satisfying $\zeta^2 + \zeta + 1 = 0$. Thus, $\zeta^2 = -\zeta - 1$. Substituting this into our expression:

$$\zeta(b-a) - \zeta^2 a = \zeta(b-a) - (-\zeta - 1)a = \zeta b - \zeta a + \zeta a + a = a + \zeta b.$$

This gives subjectivity of φ and that it is an automorphism of Λ.

Applying φ to $G_4(\Lambda)$, we get:

$$\varphi(G_4) = \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{(-\zeta\omega)^4} = \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\zeta^4 \omega^4}.$$

Simplifying further, since $\zeta^4 = \zeta$, we obtain:

$$\varphi(G_4) = \zeta^2 \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^4} = \zeta^2 G_4.$$

As before,since G_4 is defined over the lattice Λ, it is invariant under automorphisms of Λ. Thus, $\varphi(G_4) = G_4$. Comparing the two results for G_4, we have:

$$G_4 = \zeta^2 G_4.$$

Since $\zeta^2 \neq 1$ as it's a primitive 3rd root of unity, the only solution to this equation is $G_4 = 0$. To finish the proof for the cases $a_2 = 0$ or $a_3 = 0$, we claim the following:

An elliptic curve defined over the complex numbers with $a_2 = 0$ or $a_3 = 0$ can be represented as a quotient of the complex plane by a scaled lattice, specifically $c\mathbb{Z}[\zeta]$ (for $a_2 = 0$) or $c\mathbb{Z}[i]$ (for $a_3 = 0$).

Let's consider a non-zero complex number c and a lattice Λ in $\mathbb{C}$. We define the scaled lattice $c\Lambda$ as $\{c\omega \mid \omega \in \Lambda\}$. It can be easily seen that

$$G_4(c\Lambda) = c^{-4} G_4(\Lambda),$$
$$G_6(c\Lambda) = c^{-6} G_6(\Lambda).$$

It is also not difficult to see that, $G_4(\Lambda)$ (respectively $G_6(\Lambda)$) are non-zero real numbers when $\Lambda = \mathbb{Z}[\zeta]$ (for $a_2 = 0$) or respectively $\Lambda = \mathbb{Z}[i]$ (for $a_3 = 0$).

For an elliptic curve with $a_2 = 0$, we consider a curve represented by the equation $y^2 = 4x^3 - ax$ with $a \neq 0$. Assume that for the lattice $\mathbb{Z}[i]$, $G_4(\mathbb{Z}[i]) = t \neq 0$. From the scaling property of G_4, we have $G_4(c\mathbb{Z}[i]) = c^{-4}t$. To align this with our curve, we seek a c such that $c^{-4}t = a$. This choice of c will ensure that the scaled lattice $c\mathbb{Z}[i]$ corresponds to the given elliptic curve. It is also noteworthy that $g_3(c\mathbb{Z}[i]) = 0$, as required.
Therefore, the claim follows that every elliptic curve with either $a_2 = 0$ or $a_3 = 0$ can indeed be expressed in the form of $\mathbb{C}/\Lambda$ with Λ being a suitably scaled lattice, either $c\mathbb{Z}[\zeta]$ or $c\mathbb{Z}[i]$, respectively. For, other cases, See, [DS05] Proposition 1.4.3. □

1.4 Group structure

It can be seen that being a complex tori, it carries a group structure since it can be seen as a quotient of $\mathbb{C}^g$ and a lattice Λ. We shall see in the next section, that Complex Abelian varieties of dimension 1 are precisely the Complex Elliptic curves. Topologically, one-dimensional complex torus is homeomorphic to a torus.

The group Law

As mentioned above, the set of points on an elliptic curve has a natural abelian group structure on it. The main behind the geometrical idea is Bezout's theorem for projective curves over an algebraically closed field. For the simplicity of the discussion, we restrict to the elliptic curves given by Weierstrass equations. In fact, using Riemann-Roch theorem, it can be proved that every elliptic curve is given by a Weierstrass equation.

We will discuss the group law both algebraically and geometrically.
For convenience, we restrict to the case when $char(k) \neq 2, 3$, i.e $y^2 = x^3 + ax + b$.

Proposition 1.4.1 (The group law Geometrically)
Let $P, Q \in E, L$ the line connecting P and Q (tangent line to E if $P = Q$), and R the third point of intersection of L with E (by Bezout's Theorem). Let L' be the line connecting R and O. Then $P \oplus Q$ is the unique point in $L' \cap E \neq R, O$. The following diagrams illustrates this rule. We claim that this defines the group law that is under the operation $\oplus$ we have that E is an abelian group. This can be broken down into following pieces.
(a) If a line L intersects E at the (not necessarily distinct) points P, Q, R, then

$$(P \oplus Q) \oplus R = O.$$

(b) $P \oplus O = P$ for all $P \in E$. (c) $P \oplus Q = Q \oplus P$ for all $P, Q \in E$. (d) Let $P \in E$. There is a point of E, denoted $\ominus P$, so that

$$P \oplus (\ominus P) = O.$$

(e) Let $P, Q, R \in E$. Then

$$(P \oplus Q) \oplus R = P \oplus (Q \oplus R)$$

Proof:
See, [Sil13]. □

Proposition 1.4.2 (The group law Algorithm)
Let P and Q be two points on our elliptic curve $E : y^2 = x^3 + Ax + B$. We want to compute the point $R = P + Q$ by expressing the coordinates of R as rational functions of the coordinates of P and Q. If either P or Q is the point O at infinity, then R is simply the other point, so we assume that P and Q are affine points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$. There are two cases.

Case 1. $x_1 \neq x_2$. The line $\overline{PQ}$ has slope $m = (y_2 - y_1) / (x_2 - x_1)$, which yields the linear equation

$y - y_1 = m(x - x_1)$ for $\overline{PQ}$. This line is not vertical, so it intersects the curve E in a third affine point $-R = (x_3, -y_3)$. Plugging the equation for the line $\overline{PQ}$ into the equation for the curve E yields the co-ordinates of R as,

$$m = (y_2 - y_1) / (x_2 - x_1)$$
$$x_3 = m^2 - x_1 - x_2$$
$$y_3 = m(x_1 - x_3) - y_1$$

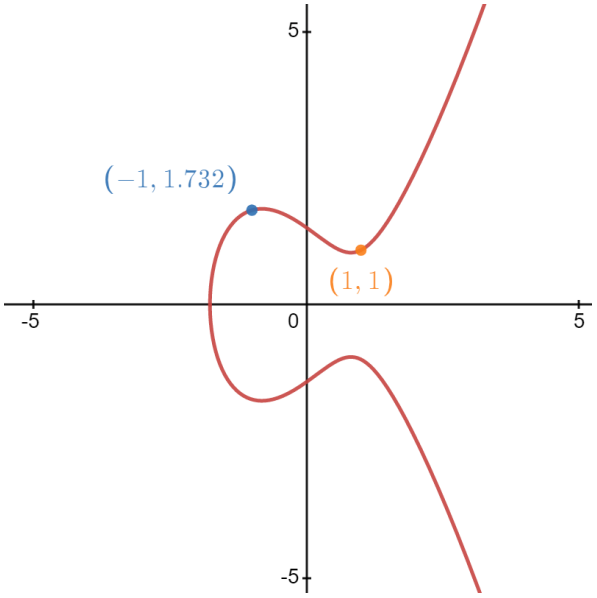
Case 2. $x_1 = x_2$. We must have,the slope of the tangent line

$$m = \frac{3x_1^2 + A}{2y_1},$$

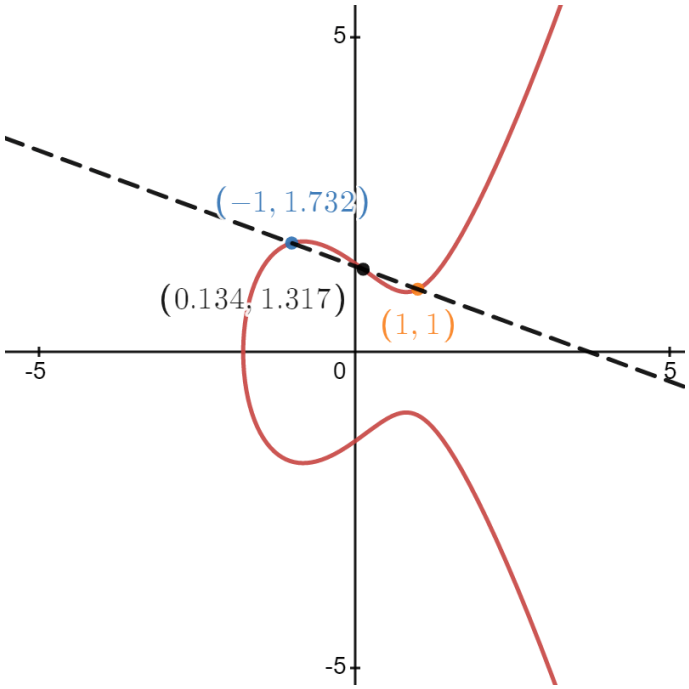
and once we know m we havem, x_3 and y_3 as in case 1) .

Proof:
See [Sil13]. □

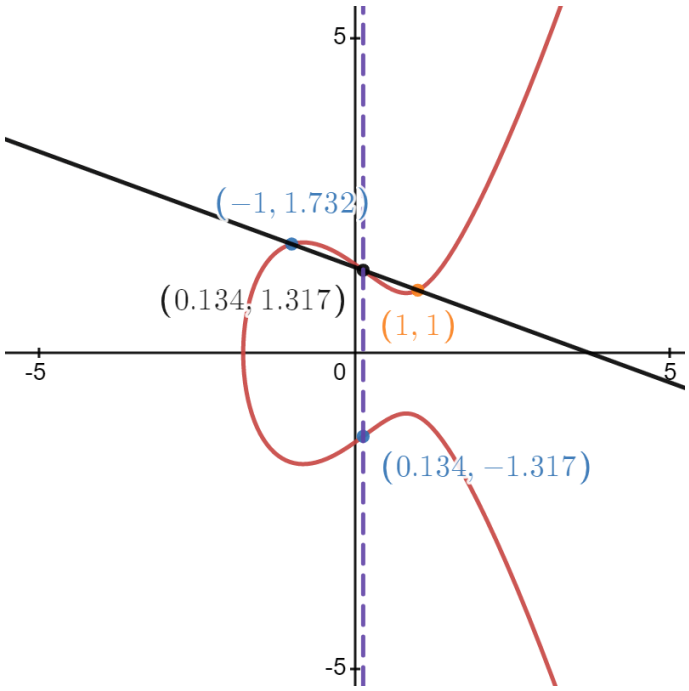
Example 1.4.3
Consider the Weierstrass equation $y^2 = x^3 - 2x + 2$, as in example above. Let us select two points randomly on the elliptic curve plotted below.



Let A denote the red point and B denote the blue point.
Let us follow the geometric steps. By Bezout's theorem, if we join the line joining A, B we must get a third point. Let us join these two points and get a point on E.



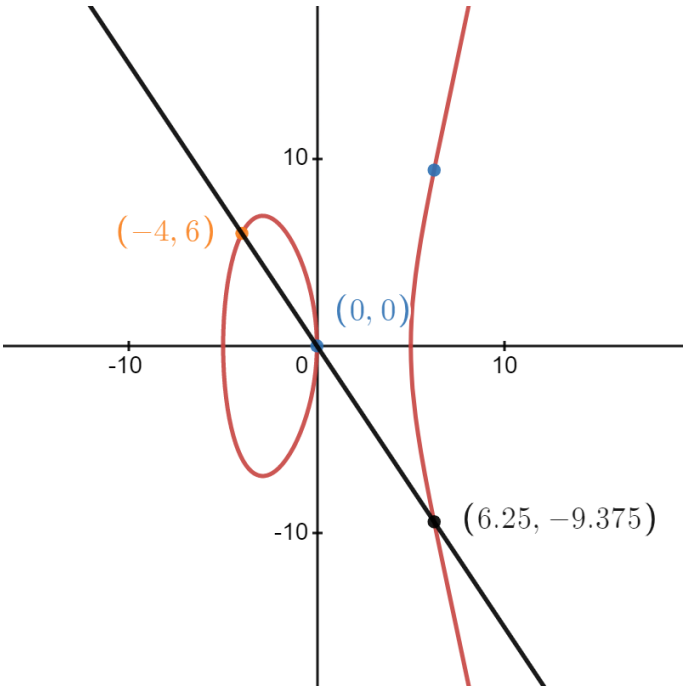
Now, lastly as per the fixed notations and considering the point at infinity in the y- direction, we draw a line parallel to y-axis through the point C of intersection of the line joining A, B and E. This can be seen easily in the following plot.



Another interesting example would be to consider $y^2 = x^3 - 25x$. Look at the real plot below. Since $x^3 - 25x$ has distinct roots, E is nonsingular, so E really is an elliptic curve. The line L through $P := (-4, 6)$ and $Q := (0, 0)$ has the equation $y = (-3/2)x$. We compute $L \cap E$ by substitution:

$$\begin{aligned} ((-3/2)x)^2 &= x^3 - 25x \\ 0 &= (x+4)x(x-25/4) \end{aligned}$$

and find $L \cap E = \{P, Q, R\}$ where $R := (25/4, -75/8)$. Thus $P + Q + R = O$ in the group law, and $P + Q = -R = (25/4, 75/8)$.



1.5 Elliptic curves and their reductions

In this section, we aim to discuss elliptic curves over $\mathbb{Q}$ and their reductions modulo primes. The key point is to introduce some terminology which then will be used to discuss the proof of Fermat's last theorem.

For each prime p, let $\nu_p(E)$ denote the smallest power of p appearing in the discriminant of any integral Weierstrass equation equivalent to E. This is essentially the minimum of a set of nonnegative integers, given by:

$$\nu_p(E) = \min \left\{ \nu_p(\Delta(E')) : E' \text{ integral, equivalent to } E \right\}.$$

Definition 1.5.1

The global minimal discriminant of E is defined as:

$$\Delta_{\min}(E) = \prod_p p^{\nu_p(E)}.$$

Note that this is a finite product since $\nu_p(E) = 0$ for all p such that $p \nmid \Delta(E)$. It can be shown that the p-adic valuation of the discriminant can be simultaneously minimized to $\nu_p(E)$ for all primes p under admissible changes of variables. In other words, E is isomorphic over $\mathbb{Q}$ to an integral model E' with discriminant $\Delta(E') = \Delta_{\min}(E)$. This integral model E' is referred to as the global minimal Weierstrass equation of E, and it serves as the model for reducing E modulo primes.

From now on, it is assumed without loss of generality that elliptic curves over $\mathbb{Q}$ are given in this global minimal form.

Let us consider the usual projection map from the ring of integers onto $\mathbb{Z}/p\mathbb{Z}$. With this map, we reduce a global minimal Weierstrass equation E to a Weierstrass equation $\widetilde{E}$ over $\mathbf{F}_p = \mathbb{Z}/p\mathbb{Z}$. It is worth noting that thus $\widetilde{E}$ defines an elliptic curve over $\mathbf{F}_p$ if and only if $p \nmid \Delta_{\min}(E)$.
We now define several types of reduction of an elliptic curve modulo a prime p.

Definition 1.5.2 (Good reduction)
We say that an elliptic curve E has good reduction at a prime p, if $\widetilde{E}$ is also an elliptic curve. This can be further classified into 2 types:
Ordinary: If $\widetilde{E}[p] \cong \mathbb{Z}/p\mathbb{Z}$.
Supersingular: If $\widetilde{E}[p] = \{0\}$.

Definition 1.5.3 (Bad reduction)
We say that an elliptic curve E has bad reduction at a prime p, if $\widetilde{E}$ is not an elliptic curve. and it can be further classified into:
Multiplicative or Semistable: If $\widetilde{E}$ has a node.
Additive or stable: If $\widetilde{E}$ has a cusp.

Remark 1.5.4
An elliptic curve with good reduction at 2 cannot admit a minimal Weierstrass model of the form say $E' : y^2 = x^3 + ax + b$. This is because E has good reduction at 2 so, E' is an elliptic curve modulo 2. But then, 2 doesn't divide the discriminant of E' that equals $-16(4a^3 + 27b^2)$, a contradiction.

Definition 1.5.5
The algebraic conductor of an elliptic curve is given by $N_E = \prod_p p^{f_p}$ where

$$f_p = \begin{cases} 0 & \text{if } E \text{ has good reduction at } p \\ 1 & \text{if } E \text{ has multiplicative reduction at } p \\ 2 & \text{if } E \text{ has additive reduction at } p \text{ and } p \notin \{2,3\} \\ 2+\delta_p & \text{if } E \text{ has additive reduction at } p \text{ and } p \in \{2,3\} \end{cases}$$

Furthermore, note that, $\delta_2 \leq 6$ and $\delta_3 \leq 3$.

Definition 1.5.6
Let E be an elliptic curve over $\mathbb{Q}$. Assume E is in reduced form. Let p be a prime and let $\widetilde{E}$ be the reduction of E modulo p. Then

$$a_p(E) = p + 1 - \left|\widetilde{E}\left(\mathbf{F}_p\right)\right|.$$

Note that If E is an elliptic curve over $\mathbb{Q}$ and p is prime then the previous definition extends to

$$a_{p^e}(E) = p^e + 1 - \left|\widetilde{E}\left(\mathbf{F}_{p^e}\right)\right|, \quad e \geq 1,$$

An interesting property that we have related to the solution counts modulo prime is that we have a recurrence relation on prime powers which will be again discussed when we discuss forward and reverse maps on the Picard group.

Proposition 1.5.7

$$a_{p^e}(E) = a_p(E)a_{p^{e-1}}(E) - \mathbf{1}_E(p)pa_{p^{e-2}}(E) \quad \text{for all } e \geq 2.$$

Here $\mathbf{1}_E$ is the trivial character modulo the algebraic conductor N_E of E, so $\mathbf{1}_E(p)$ is 1 for primes of good reduction and 0 for primes of bad reduction.

This has a very nice corollary which relates to Elliptic curves with good reduction at some prime p.

Corollary 1.5.8
Let E be an elliptic curve over $\mathbb{Q}$, and let p be a prime such that E has good reduction at p. Then the reduction is

$$\begin{cases} \text{ordinary} & \text{if } a_p(E) \not\equiv 0 (\text{mod} p), \\ \text{supersingular} & \text{if } a_p(E) \equiv 0 (\text{mod} p). \end{cases}$$

Proof:
See, [DS05], chapter 8. □

2 Modular forms

Modular forms are special complex-valued functions that have several important properties. They are defined on the upper half-plane and they have a transformation behavior under the modular group (a group of linear fractional transformations that preserve the upper half-plane). This transformation behavior allows modular forms to encode information about the geometry and topology of modular curves, which are essential objects in number theory. Modular forms are used in many areas of mathematics, including number theory, algebraic geometry, and representation theory. They have applications in physics, including string theory and conformal field theory.
In this chapter, we aim to introduce and discuss some results about two central objects, namely **Modular forms** and **Modular curve**. As discussed in the introduction, the statement that ultimately led to proving Fermat's last theorem is

All semistable elliptic curves are modular.

We will try to make this precise and understand what it actually means. We will closely follow the book [DS05], lecture notes from Marc Masdeu [Mas15] and [Vis18]. We will sometimes closely follow proofs from either of these sources for the sake of completeness. At times the ideas of the proofs are not original but inspired and even closely followed at times. Although, in some places, I have expanded on my own giving more details. Besides these references, there are many good references on modular forms for example, [Lan95],[Miy06] .

2.1 Basic defintions

Definition 2.1.1 (Modular group)
The modular group is the group of 2-by-2 matrices with integer entries and determinant 1. The Modular group in set form is

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{Z}, ad - bc = 1 \right\}.$$

Proposition 2.1.1
The modular group is generated by the two matrices

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Let $\mathcal{H}$ denote the Poincare upper half plane which as a set is given by

$$\mathcal{H} = \{\tau \in \mathbb{C} : \mathrm{Im}(\tau) > 0\}$$

Each element of the modular group is also viewed as an automorphism of the Riemann sphere $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, the fractional linear transformation

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} (\tau) = \frac{a\tau + b}{c\tau + d}, \quad \tau \in \widehat{\mathbb{C}}.$$

For $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, we denote by $\gamma(\tau)$ a fractional transformation on $\widehat{\mathbb{C}}$ given by, $\gamma(\tau) = \frac{a\tau+b}{c\tau+d}, \quad \tau \in \widehat{\mathbb{C}}$.

Remark 2.1.2

• Here we understand that if $c \neq 0$ then $-d/c$ maps to ∞ and ∞ maps to a/c, and if $c = 0$ then ∞ maps to ∞.

• The identity matrix I and its negative $-I$ both give the identity transformation as for $\tau \in \widehat{\mathbb{C}}$ we have ,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}(\tau) = \frac{1\tau+0}{0\tau+1} = \tau$$

and

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}(\tau) = \frac{-1\tau+0}{0\tau+1} = \tau.$$

• In general each pair $\pm\gamma$ of matrices in $\mathrm{SL}_2(\mathbb{Z})$ gives a single transformation, since for $\tau \in \widehat{\mathbb{C}}$, we have that,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}(\tau) = \frac{a\tau+b}{c\tau+d}$$

,

$$\begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}(\tau) = \frac{-a\tau-b}{-c\tau-d} = \frac{a\tau+b}{c\tau+d}$$

• The group of transformations defined by the modular group is generated by the maps described by the two matrix generators, namely for transformations given by,

$$\tau \mapsto \tau + 1 \quad \text{and} \quad \tau \mapsto -1/\tau.$$

It can be

Proposition 2.1.3

If $\gamma \in \mathrm{SL}_2(\mathbf{Z})$ and $\tau \in \mathcal{H}$ then also $\gamma(\tau) \in \mathcal{H}$. In other words, the modular group maps the upper half plane to itself.

Proof:

Let $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbf{Z})$ be arbitrary. From above we have, $\gamma(\tau) = \frac{a\tau+b}{c\tau+d}$. Now, multiplying and dividing by complex conjugate of $c\tau + d$ we get that,

$$\gamma(\tau) = \frac{(a\tau+b)(c\overline{\tau}+d)}{\mid c\tau+d \mid^2}.$$

Thus we have,

$$\gamma(\tau) = \frac{ad\tau+bc\overline{\tau}}{\mid c\tau+d \mid^2} + \frac{ac\tau\overline{\tau}+bd}{\mid c\tau+d \mid^2}.$$

From this expression, it is clear that $\mathrm{Im}(\gamma(\tau)) = (ad-bc)\mathrm{Im}(\tau)$. The claim follows since, $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. □

2.2 Weakly Modular Functions and Modular Forms

Let us now move ahead toward introducing one of the key objects to understand the proof of Fermat's last theorem, namely Modular forms. We will first introduce weakly modular functions and then special kinds of weakly modular functions, namely Modular forms, which will be the central object of study from this point onward.

Definition 2.2.1 (Weakly modular functions)
Let k be an integer. A meromorphic function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is weakly modular of weight k if

$$f(\gamma(\tau)) = (c\tau + d)^k f(\tau) \quad \text{for } \gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \text{ and } \tau \in \mathcal{H}. \tag{2}$$

Remark 2.2.2
1. We will soon establish a key lemma which states that if a function f satisfies the weak modularity condition 2 stated in 2.2.1 when we substitute γ as $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the function f satisfies the weak modularity condition 2 for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.
2. This reduces our work immensely to check the weak modularity condition for all the matrices in the modular group to check it only on two matrices.

Observe that for $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, we have $T(\tau) = \tau + 1$ and for $S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ we have $S(\tau) = \frac{-1}{\tau}$

In other words, we get an alternative definition for a weakly modular function: A meromorphic function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is weakly modular of weight k if

$$f(\tau + 1) = f(\tau) \quad \text{and} \quad f(-1/\tau) = \tau^k f(\tau).$$

Definition 2.2.3 (**Modular forms**)
Let k be an integer. A function $f : \mathcal{H} \longrightarrow \mathbf{C}$ is a modular form of weight k if
(1) f is holomorphic on $\mathcal{H}$,
(2) f is weakly modular of weight k,
(3) f is holomorphic at ∞.

Remark 2.2.4 (q-expansions and holomorphicity at ∞)
Let us make the notion (3) in the definition 2.2.3 precise. We begin by noting that modular form for $\mathrm{SL}_2(\mathbb{Z})$ satisfies $f(\tau + 1) = f(\tau)$. Thus f is always periodic with period 1. We can therefore describe f in terms of a new variable $q := e^{2\pi i\tau}$. This is how we get a so-called q-expansion.
We define $\tilde{f} : D' \longrightarrow \mathbb{C}$ $\tilde{f}(q) = f(\log(q)/(2\pi i))$.
Well-definedness is clear due to $\mathbb{Z}$-periodicity of f and the fact that log is well-defined up to $2\pi i\mathbb{Z}$. Moreover, we have that $\tilde{f}(q) = f(\tau)$. This is also well-defined. Let $q = e^{2\pi i\tau}$ and $q' = e^{2\pi i\tau'}$. Then,

$$\begin{aligned} q = q' &\implies e^{2\pi i\tau} = e^{2\pi i\tau'} \implies \tau = \tau' + n \quad \text{for some } n \in \mathbb{Z} \\ &\implies f(\tau) = f(\tau' + n) \implies f(\tau) = f(\tau') \text{ (by periodicity)} \\ &\implies \tilde{f}(q) = \tilde{f}(q') \end{aligned}$$

Furthermore we note that, if $\tau = x + iy \in \mathcal{H}$, then $|q| = |e^{2\pi i\tau}| = |e^{2\pi i\tau x}||e^{-2\pi y}| = |e^{-2\pi y}| \leq 1$. The

last inequality follows since $y > 0$. Thus, the change of variables sends the upper half plane $\mathcal{H}$, to the punctured open unit disc $D' = \{q \in \mathbb{C} \mid 0 < |q| < 1\}$.
If f is holomorphic on the upper half plane then the composition $\tilde{f}$ is holomorphic on the punctured disk. Since the logarithm can be defined holomorphically about each point, $\tilde{f}$ has a Laurent expansion $\tilde{f}(q) = \sum_{n \in \mathbf{Z}} a_n q^n$ for $q \in D'$.
We also note that as $\mathrm{Im}(\tau) \to \infty$, we have $q \to 0$. Hence the condition that $f(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$ implies that $\tilde{f}(q)$ is bounded as $q \to 0$. Therefore, by the Riemann Removable Singularity Theorem, we have that $\lim_{q \to 0} \tilde{f}(q)$ exists and furthermore, that the extension of $\tilde{f}$ to the point $q = 0$ defined by $\tilde{f}(0) := \lim_{q \to 0} \tilde{f}(q)$ is analytic on the unit disc $D = \{q \in \mathbb{C} \mid 0 \leq |q| < 1\}$. In other words, this means that we can think of ∞ as lying far in the imaginary direction, and one can define f to be holomorphic at ∞ if $\tilde{f}$ extends holomorphically to the puncture point $q = 0$, *i. e.* , the Laurent series sums over $n \in \mathbb{N}$.
Therefore, f has a Fourier expansion, given as

$$f(\tau) = \tilde{f}(q) = \sum_{n=0}^{\infty} a_n(f) q^n, \qquad \text{where } q = e^{2\pi i \tau}.$$

Also, to show that a weakly modular holomorphic function $f : \mathcal{H} \longrightarrow \mathbb{C}$ is holomorphic at ∞ one doesn't need to compute its Fourier expansion and one can simply show either $\lim_{\mathrm{Im}(\tau) \to \infty} f(\tau)$ exists or even $f(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$ since $q \to 0$ if and only if $\mathrm{Im}(\tau) \to \infty$.

Definition 2.2.5
The set of modular forms of weight k is denoted $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$.

Proposition 2.2.6
$\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ forms a vector space over $\mathbb{C}$.

Proof:
Let $f, g \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$, then by linearity of the weak modularity condition, we have $\alpha f + \beta g$ is weakly modular of weight k. The holomorphy condition on $\mathcal{H}$ and at ∞ is clear. Thus we have $\alpha f + \beta g \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ for any $\alpha, \beta \in \mathbb{C}$. Thus, $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ forms a vector space over $\mathbb{C}$. □

Lemma 2.2.7
Let f_1 be a modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$ and let f_2 be a modular form of weight k_2 for $\mathrm{SL}_2(\mathbb{Z})$. Then $f_1 f_2$ is a modular form of weight $k_1 + k_2$ for $\mathrm{SL}_2(\mathbb{Z})$.

Proof:
We note that the product $f_1 f_2$ is clearly holomorphic on $\mathcal{H}$ and that $f_1 f_2(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$. Thus, all that remains is to prove the weak modularity condition (2).
Let $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$.
We end the proof by finally noting that

$$f_1 f_2(\tau) = f_1(\tau) f_2(\tau) = (c\tau + d)^{k_1} f_1(\tau)(c\tau + d)^{k_2} f_2(\tau) = (c\tau + d)^{k_1 + k_2} f_1 f_2(\tau)$$

□

In particular

$$\mathcal{M}\left(\mathrm{SL}_2(\mathbf{Z})\right) = \bigoplus_{k\in\mathbf{Z}} \mathcal{M}_k\left(\mathrm{SL}_2(\mathbf{Z})\right)$$

forms a graded ring.

Proposition 2.2.8

Let f be a complex-valued function that is holomorphic on $\mathcal{H}$ and is bounded as $\mathrm{Im}(\tau) \to \infty$. Then f is a modular form of weight k for $\mathrm{SL}_2(\mathbb{Z})$ if and only if it satisfies the following equations

$$f(\tau+1) = f(\tau) \quad \text{and} \quad f\left(\frac{-1}{\tau}\right) = \tau^k f(\tau) \tag{3}$$

Proof:

We first note that substituting T and S into the modularity condition yields the two equations respectively given in 3. Thus, if f is a modular form, then f satisfies 3.

Conversely, assume f satisfies both equations given in 3.

It remains to show that if the weak modularity condition 2 is satisfied for some two matrices $\gamma_1, \gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$, then it is satisfied for $\gamma_1\gamma_2$ and for γ_1^{-1}.

Let $\gamma_1 = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ and $\gamma_2 = \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$.

Note that $\gamma_1^{-1} = \begin{bmatrix} d_1 & -b_1 \\ -c_1 & a_1 \end{bmatrix}$, as $\det(\gamma_1) = 1$. We thus have,

$$\begin{aligned} f(\gamma_1\gamma_2\tau) &= (c_1(\gamma_2\tau) + d_1)^k f(\gamma_2\tau) \\ &= \left(c_1\left(\frac{a_2\tau + b_2}{c_2\tau + d_2}\right) + d_1\right)^k (c_2\tau + d_2)^k f(\tau) \\ &= (c_1(a_2\tau + b_2) + d_1(c_2\tau + d_2))^k f(\tau) \\ &= ((c_1a_2 + d_1c_2)\tau + c_1b_2 + d_1d_2)^k f(\tau) \end{aligned}$$

It is clear that since $\gamma_1\gamma_2 = \begin{bmatrix} a_1a_2 + b_1c_2 & a_1b_2 + b_1d_2 \\ c_1a_2 + d_1c_2 & c_1b_2 + d_1d_2 \end{bmatrix}$, the weak modularity condition (2) is satisfied for $\gamma_1\gamma_2$.

Similarly checking γ_1^{-1}, we note

$$\begin{aligned} & f\left(\gamma_1\left(\gamma_1^{-1}\tau\right)\right) = \left(c_1\left(\gamma_1^{-1}\tau\right) + d_1\right)^k f\left(\gamma_1^{-1}\tau\right) \\ \Rightarrow \quad & f(\tau) = \left(c_1\left(\frac{d_1\tau - b_1}{-c_1\tau + a_1}\right) + d_1\right)^k f\left(\gamma_1^{-1}\tau\right) \\ \Longrightarrow \quad & (-c_1\tau + a_1)^k f(\tau) = (c_1(d_1\tau - b_1) + d_1(-c_1\tau + a_1))^k f\left(\gamma_1^{-1}\tau\right) \\ \Longrightarrow \quad & (-c_1\tau + a_1)^k f(\tau) = (a_1d_1 - b_1c_1)^k f\left(\gamma_1^{-1}\tau\right) \end{aligned}$$

As $a_1d_1 - b_1c_1 = \det(\gamma_1) = 1$ is one, we obtain: $f\left(\gamma_1^{-1}\tau\right) = (-c_1\tau + a_1)^k f(\tau)$ which proves that the weak modularity condition is satisfied for γ_1^{-1}.

□

Next, the modular forms with constant term 0 in their Fourier expansions are of great importance. This

leads to studying cusp forms.

Definition 2.2.9
A cusp form of weight k is a modular form of weight k whose fourier expansion has leading coefficient $a_0 = 0$, i.e.,

$$f(\tau) = \sum_{n=1}^{\infty} a_n q^n, \quad q = e^{2\pi i \tau}$$

2.3 Examples

Let us now look at particular cases of weak modular functions of weight k.

- **Weight 0**: For $k = 0$, weak modularity condition reduces to $f(\gamma(\tau)) = (c\tau + d)^0 f(\tau)$ for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathcal{H}$. That is for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathcal{H}$, $f(\gamma(\tau)) = f(\tau)$. This means that weight 0 weakly modular functions are $\mathrm{SL}_2(\mathbb{Z})$-invariant.
- **Weight k when k is odd**: Let f be a weakly modular function. We observe that $-I = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$. Now, substituting $\gamma = -I$ in weak modularity condition (2.2.1) we get that, $f(\frac{-\tau}{-1}) = (-1)^k f(\tau)$. Thus we have that $f(\tau) = (-1)^k f(\tau)$. Since k is odd, f is an identically zero function. In other words, the zero function is the only weakly modular function of weight k when k is odd.
- The zero function is a modular form of every weight,
- Constant functions are modular forms of weight 0.

Note that, a good reason to consider working with weakly modular functions is that even though we don't necessarily have weakly modular functions fully $\mathrm{SL}_2(\mathbb{Z})$-invariant, we have that $f(\tau)$ and $f(\gamma(\tau))$ always have the same zeros and poles since the factor $c\tau + d$ doesn't have any on $\mathcal{H}$.
For a more non-trivial example, we consider the Eisenstein series with respect to a lattice Λ and as we have seen, a lattice Λ has a basis generated by $\tau, 1$ for some $\tau \in \mathcal{H}$, we can consider the same Eisenstein series and instead consider it as a function on $\mathcal{H}$, namely consider, for an even integer $k > 3$

$$G_k(\tau) = \sum_{\substack{(m,n)\in\mathbb{Z}^2 \\ (m,n)\neq(0,0)}} \frac{1}{(m\tau + n)^k} \tag{4}$$

Note that it essentially is the Eisenstein series $G_k(\Lambda)$ for a lattice Λ generated by $1, \tau$ denoted differently. By our discussion in Chapter 1, we get that $G_k(\tau)$ converges absolutely on $\mathcal{H}$ for $k > 2$. Hence the order of the summation does not matter, and thus the definition given in the equation 4 is well-defined. We also note that the sum converges uniformly on compact subsets of $\mathcal{H}$, and thus G_k is holomorphic.
It remains to see that $Gk_(\tau)$ is indeed a modular form. We will make use of 2.2.8.
It is thus sufficient to check weak modularity condition 2 for generators of the full modular group S, T. Also, for simplicity let us denote by $Z' = \mathbb{Z}^2 \backslash \{(0,0)\}$ as before. Now, we check weak modularity condition for S and T. The absolute convergence of Eisenstein series come in handy.

$$G_k(\tau+1) = \sum_{(m,n)\in Z'} \frac{1}{(m\tau+m+n)^k} = \sum_{(m,n)\in Z'} \frac{1}{(m'\tau+n')^k} = G_k(\tau)$$

The second equality follows from absolute convergence, thus we may rearrange terms, noting that $(m,n) \mapsto (m, m+n)$ is a bijection from $\mathbb{Z}^2\backslash\{(0,0)\}$ to $\mathbb{Z}^2\backslash\{(0,0)\}$. Now,

$$\begin{aligned} G_k\left(\frac{-1}{\tau}\right) &= \sum_{(m,n)\in Z'} \frac{1}{\left(m\left(\frac{-1}{\tau}\right)+n\right)^k} = \sum_{(m,n)\in Z'} \frac{\tau^k}{(-m+n\tau)^k} \\ &= \tau^k \sum_{(m,n)\in Z'} \frac{1}{(n\tau-m)^k} = \tau^k \sum_{(m',n')\in Z'} \frac{1}{(m'\tau+n')^k} \\ &= \tau^k G_k(\tau) \end{aligned}$$

where, as before, the second last equality follows from rearranging terms.
Thus G_k satisfies the weak modularity condition for both S and T, and thus satisfies the condition for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ by 2.2.8. Finally, it can be proved that

$$\lim_{\mathrm{Im}(\tau)\to\infty} G_k(\tau) = 2\sum_{n=1}^{\infty} \frac{1}{n^k} = 2\zeta(k),$$

where ζ denotes the Riemann-zeta function. More details can be found in [DS05].
In particular, we have that $G_k(\tau)$ is bounded as $\mathrm{Im}(\tau) \to \infty$, and thus we get that the Eisenstein series is a modular form.
Moreover, One can prove that the q-expansion of $G_k(\tau)$ can be given as [DS05][Page 5]

$$\tilde{G}_k(q) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty} \sigma_{k-1}(n)q^n$$

where $\sigma_{k-1}(n)$ denotes the sum of the $(k-1)$-th powers of the positive divisors of n

$$\sigma_{k-1}(n) = \sum_{\substack{d|n \\ d>0}} d^{k-1}$$

We also note the normalised Eisenstein series, defined as $E_k(\tau) = G_k(\tau)/(2\zeta(k))$, where the constant coefficient in the q-expansion normalises to 1 . Noting that ζ evaluated at the even integers gives [DS05][Page 10]

$$\zeta(k) = (-1)^{k/2+1}\frac{B_k(2\pi)^k}{2(k!)} \quad \text{for } k \text{ even}, k \geq 2$$

we thus obtain

$$\begin{aligned} \tilde{E}_k(q) &= 1 + \frac{(2\pi i)^k}{\zeta(k)(k-1)!}\sum_{n=1}^{\infty} \sigma_{k-1}(n)q^n \\ &= 1 - \frac{2k}{B_k}\sum_{n=1}^{\infty} \sigma_{k-1}(n)q^n \end{aligned}$$

Therefore, the coefficients in the q-expansion of the normalised Eisenstein series E_k are all rational with a common denominator and moreover the first few coefficients for E_4 and E_6 are given as

$$\tilde{E}_4(q) = 1 + 240q + 2160q^2 + 6720q^3 + \mathcal{O}\left(q^4\right)$$
$$\tilde{E}_6(q) = 1 - 504q - 16632q^2 - 122976q^3 + \mathcal{O}\left(q^4\right)$$

2.4 Congruence Subgroups

Replacing the modular group $\mathrm{SL}_2(\mathbb{Z})$ in the weak modularity condition by a subgroup Γ generalizes the notion of weak modularity, allowing more examples of weakly modular functions. In this more general sense, we will especially see examples of modular forms of odd weight. We mainly want to focus on the subgroups of $\mathrm{SL}_2(\mathbb{Z})$, which are of finite index. We now see an important class of such subgroups: principal congruence subgroups of level N.

Definition 2.4.1
Let N be a positive integer. The principal congruence subgroup of level N is

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

Note that the matrix congruence is defined entry-wise. That is, $a, c \equiv 1(\mathrm{mod} N)$ and $b, c \equiv 0 \pmod{N}$Furthermore, as mentioned earlier, we want to focus on finite index subgroups of the modular group. Thus it is essential to check that the principal subgroups of level N are of finite index in the modular group.

Proposition 2.4.2
Let N be a positive integer. Then $\Gamma(N)$ has finite index in $\mathrm{SL}_2(\mathbb{Z})$.

Proof:
We define a natural homomorphism $\varphi : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix}$$

where $\bar{x}$ denotes the congruence class of x modulo N. It can easily be seen that φ is a homomorphism. Now let us calculate the kernel,

$$\varphi(\gamma) = I \iff \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix} = I$$

Now, by defintion of $\Gamma(N)$ we get that

$$\gamma \in ker(\varphi) \iff \gamma \in \Gamma(N)$$

Thus $ker(\varphi) = \Gamma(N)$, and since kernel of group homomorphisms are normal in the domain we get that $\Gamma(N)$ is a normal subgroup in $\mathrm{SL}_2(\mathbb{Z})$.
Next, we show that φ is surjective.
Firstly, if $N = 1$, then $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ consists of a single element whereby φ is a constant map and thus

trivially surjective. We therefore now assume $N > 1$.

Indeed, let $\gamma' = \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix}$ be an arbitrary element in $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Thus, we have

$$ad - bc \equiv 1 \pmod{N} \implies ad - bc + qN = 1$$

for some $q \in \mathbb{Z}$. Therefore $\gcd(a, b, N) = 1$ and in particular, by Bezout's theorem, we have $\gcd(a, b)$ is coprime to N.

Now, we may assume without loss of generality that a, b, c, d are all positive integers, as the class of x equals class of $x + N$ modulo N for any integer x. We define $g = \gcd(a, b)$. Let t be an integer satisfying the following congruence conditions:

• for all primes p dividing g

$$t \equiv 0 \,(\mathrm{mod} p)$$

• for all all primes q dividing a and not dividing g(note that either or both sets may be empty).

$$t \equiv 1 \,(\mathrm{mod} q)$$

The Chinese Remainder theorem gives the existence of such a $t \in \mathbb{Z}$ since the primes concerned for the congruences are coprime by construction. Now, define $b' = b + tN$. We now claim that $\gcd(a, b') = 1$. Indeed, assume for contradiction that there exists some prime q such that $q \mid a$ and $q \mid b'$. We consider two cases:

Case 1: q divides g. Thus, $q \mid \gcd(a, b)$ which implies $q \mid b$, hence $q \mid b' - b = tN$. However, by construction of t we have $t \equiv 1 (\mathrm{mod} q)$. Thus $N \equiv tN \equiv 0 (\mathrm{mod} q)$ and hence q divides N. Therefore $q \mid \gcd(a, b, N)$ which contradicts $\gcd(a, b, N) = 1$.

Case 2: q does not divide g. As $q \mid a$, we have by construction of t that $q \mid t$. Thus $q \mid b' - tN = b$. Therefore, as $q \mid a$ and $q \mid b$, this implies $q \mid \gcd(a, b)$, contradicting our assumption that $q \nmid g$.

Therefore, we conclude that $\gcd(a, b') = 1$.

Thus by Bézout's identity, there exist $x, y \in \mathbb{Z}$ such that $ax - b'y = 1$. Now, define $c \in \mathbb{Z}$ and $d \in \mathbb{Z}$ by:

$$\begin{aligned} c' &= c + y\left(1 - \left(ad - b'c\right)\right) \\ d' &= d + x\left(1 - \left(ad - b'c\right)\right) \end{aligned}$$

and define $\gamma = \begin{bmatrix} a & b' \\ c' & d' \end{bmatrix}$. We prove that $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Indeed

$$\begin{aligned} \det(\gamma) &= ad' - b'c' \\ &= a\left(d + c\left(1 - \left(ad - b'c\right)\right)\right) - b'\left(c + y\left(1 - \left(ad - b'c\right)\right)\right) \\ &= ad - b'c + \left(ax - b'y\right)\left(1 - \left(ad - b'c\right)\right) \\ &= ad - b'c + \left(1 - \left(ad - b'c\right)\right) = 1 \end{aligned}$$

and thus $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Furthermore, we note

$$\begin{aligned} c' = c + y\left(1 - \left(ad - b'c\right)\right) &\equiv c + yqN \equiv c \pmod{N} \\ d' = d + x\left(1 - \left(ad - b'c\right)\right) &\equiv d + xqN \equiv d \pmod{N}. \end{aligned}$$

Thus we finally obtain

$$\varphi(\gamma) = \varphi \begin{bmatrix} a & b' \\ c' & d' \end{bmatrix} = \varphi \begin{bmatrix} a & b + tN \\ c' & d' \end{bmatrix} = \begin{bmatrix} \bar{a} & \bar{b + tN} \\ \bar{c'} & \bar{d'} \end{bmatrix} = \gamma'$$

Thus, this proves φ is surjective.
We, therefore, obtain by the first isomorphism theorem that φ induces the isomorphism:

$$\mathrm{SL}_2(\mathbb{Z})/\Gamma(N) \cong \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$$

□

Definition 2.4.3
A subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup if $\Gamma(N) \subset \Gamma$ for some $N \in \mathbb{N}$, in which case Γ is a congruence subgroup of level N.

Remark 2.4.4
Besides the principal congruence subgroups, the most important congruence subgroups are

- $$\Gamma_0(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} (\mathrm{mod} N) \right\}$$

  (where "*" means unspecified or entries can be arbitrary and the only important entry is that we want c to be divisible by N)

- $$\Gamma_1(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbf{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} (\mathrm{mod} N) \right\}$$

We now collect some properties of these congruence subgroups in a single proposition.

Proposition 2.4.5

1. $$\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset \mathrm{SL}_2(\mathbf{Z})$$

2. The map

   $$\Gamma_1(N) \longrightarrow \mathbf{Z}/N\mathbf{Z}, \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto b (\mathrm{mod} N)$$

   is a surjection with kernel $\Gamma(N)$.

3. The map

   $$\Gamma_0(N) \longrightarrow (\mathbf{Z}/N\mathbb{Z})^*, \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto d (\mathrm{mod} N)$$

   is a surjection with kernel $\Gamma_1(N)$ so that $\Gamma_1(N) \triangleleft \Gamma_0(N)$

Proof:

1. This is clear just by definition of the congruence subgroups and principal congruence subgroups.

2. The surjectivity of the map is clear. This is because for any class $\bar{x} \in \mathbb{Z}/N\mathbb{Z}$, $\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \mapsto \bar{x}$. Besides, by definition of congruence subgroup $\Gamma(N)$, we have that it is precisely the inverse image of the class of 0 in $\mathbb{Z}/N\mathbb{Z}$. This completes 2). Analogoulsly, 3) can be done. □

Remark 2.4.6

First, just by definitions of congruence subgroup and principal congruence subgroups, we get that $\Gamma(N) \triangleleft \Gamma_1(N)$.

Apply first isomorphism theorem to part 2 and 3 in Proposition 2.4.5, we get that

$$\Gamma_1(N)/\Gamma(N) \xrightarrow{\sim} \mathbb{Z}/N\mathbb{Z}, \quad [\Gamma_1(N) : \Gamma(N)] = N$$

$$\Gamma_0(N)/\Gamma_1(N) \xrightarrow{\sim} (\mathbb{Z}/N\mathbb{Z})^*, \quad [\Gamma_0(N) : \Gamma_1(N)] = \varphi(N),$$

Where φ is the Euler totient function.

Furthermore, by tower of law of group indices together with 2.4.2 shows,

$$[\mathrm{SL}_2(\mathbf{Z}) : \Gamma_0(N)] = N \prod_{p|N} (1 + 1/p),$$

the product taken over all primes dividing N.

We are almost there to define modular forms in a more general sense, that is with respect to congruence subgroups of the modular group, but before we need to define two important notions.

Definition 2.4.7 (Factor of Automorphy)

For a matrix $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ dthe factor of automorphy $j(\gamma, \tau) \in \mathbb{C}$ for $\tau \in \mathcal{H}$ is given by

$$j(\gamma, \tau) = c\tau + d$$

Definition 2.4.8 (Weight k-Operator)

Let $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and k be any integer. The weight- k operator $[\gamma]_k$ on functions $f : \mathcal{H} \longrightarrow \mathbf{C}$ is given by

$$(f[\gamma]_k)(\tau) = j(\gamma, \tau)^{-k} f(\gamma(\tau)), \quad \tau \in \mathcal{H}$$

We now note important properties of these newly defined notions.

Proposition 2.4.9

For all $\gamma, \gamma' \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathcal{H}$

(a) $j(\gamma\gamma', \tau) = j(\gamma, \gamma'(\tau))\, j(\gamma', \tau)$,

(b) $(\gamma\gamma')(\tau) = \gamma(\gamma'(\tau))$

(c) $[\gamma\gamma']_k = [\gamma]_k [\gamma']_k$ (this is an equality of operators),

(d) $\mathrm{Im}(\gamma(\tau)) = \frac{\mathrm{Im}(\tau)}{|j(\gamma,\tau)|^2}$, (e) $\frac{d\gamma(\tau)}{d\tau} = \frac{1}{j(\gamma,\tau)^2}$.

Proof:

Let's start with the basic action of any matrix γ in $\mathrm{SL}_2(\mathbb{Z})$ on a point in the complex upper half-plane $\mathcal{H}$. The action can be represented as a transformation of the complex point τ, which results in a new point $\gamma(\tau)$, accompanied by a scalar multiplier $j(\gamma, \tau)$.
This transformation can be written in vector-matrix form as:

$$\gamma \begin{bmatrix} \tau \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma(\tau) \\ 1 \end{bmatrix} j(\gamma, \tau).$$

Next, consider the composite action of two matrices γ and γ'. Applying the transformation rule iteratively, we have:

$$\gamma\gamma' \begin{bmatrix} \tau \\ 1 \end{bmatrix} = \gamma \left(\begin{bmatrix} \gamma'(\tau) \\ 1 \end{bmatrix} j(\gamma', \tau) \right) = \begin{bmatrix} \gamma(\gamma'(\tau)) \\ 1 \end{bmatrix} j(\gamma, \gamma'(\tau)) j(\gamma', \tau).$$

Since the above equality holds for the respective vector-matrix multiplication, it implies that the scalar multipliers must also be equal, i.e., $j(\gamma\gamma', \tau) = j(\gamma, \gamma'(\tau)) j(\gamma', \tau)$. This establishes the relationship between the actions of γ, γ', and their composite on τ.
For a function $f : \mathcal{H} \to \mathbb{C}$, we can express the action of γ and γ' as fractional linear transformations:

$$\begin{aligned} \left(f \left[\gamma\gamma'\right]_k\right)(\tau) &= j(\gamma\gamma', \tau)^{-k} f\left((\gamma\gamma')(\tau)\right), \\ \left(\left(f[\gamma]_k\right)\left[\gamma'\right]_k\right)(\tau) &= j(\gamma', \tau)^{-k} \left(f[\gamma]_k\right)(\gamma'(\tau)) \\ &= j(\gamma', \tau)^{-k} j(\gamma, \gamma'(\tau))^{-k} f(\gamma(\gamma'(\tau))). \end{aligned}$$

The part c) follows from equility of both the right hand sides which in turn follows from part a) and b).
Let us now derive the properties (d) and (e). Let us begin by noting that,

$$\mathrm{Im}(\gamma(\tau)) = \mathrm{Im}\left(\frac{a\tau + b}{c\tau + d}\right) = \mathrm{Im}\left(\frac{(a\tau + b)(\bar{c}\bar{\tau} + \bar{d})}{|c\tau + d|^2}\right).$$

Expanding and simplifying this expression, we get:

$$\mathrm{Im}(\gamma(\tau)) = \frac{(ad - bc)\,\mathrm{Im}(\tau)}{|c\tau + d|^2} = \frac{\mathrm{Im}(\tau)}{|c\tau + d|^2},$$

where we used $ad - bc = 1$ as $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Noting that $j(\gamma, \tau) = c\tau + d$, we arrive at the desired result:

$$\mathrm{Im}(\gamma(\tau)) = \frac{\mathrm{Im}(\tau)}{|j(\gamma, \tau)|^2}.$$

To derive $\frac{d\gamma(\tau)}{d\tau} = \frac{1}{j(\gamma,\tau)^2}$, we take the derivative of $\gamma(\tau)$ with respect to τ. The derivative of the fractional linear transformation is:

$$\frac{d\gamma(\tau)}{d\tau} = \frac{d}{d\tau}\left(\frac{a\tau + b}{c\tau + d}\right) = \frac{a(c\tau + d) - c(a\tau + b)}{(c\tau + d)^2} = \frac{ad - bc}{(c\tau + d)^2}.$$

Again we use that, $ad - bc = 1$ for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, this simplifies to:

$$\frac{d\gamma(\tau)}{d\tau} = \frac{1}{(c\tau + d)^2} = \frac{1}{j(\gamma, \tau)^2}.$$

□

Equipped with all the notions required, let's define the term 'modular form' in the context of a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$. Consider an integer k and a congruence subgroup Γ of $\mathrm{SL}_2(\mathbf{Z})$. A complex function f defined on the upper half-plane $\mathcal{H}$ is termed a modular form of weight k for Γ if it satisfies two key criteria: it is weakly modular of weight k with respect to Γ, and it adheres to a specific holomorphy condition. Let us make this precise.
Each congruence subgroup Γ of $\mathrm{SL}_2(\mathbf{Z})$ contains a translation matrix of the form

$$\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix},$$

where h is a positive integer. This is by definition of a congruence subgroupΓ containing $\Gamma(N)$ for some N, though h might be a proper divisor of N.
Consequently, any function $f : \mathcal{H} \to \mathbf{C}$ which is weakly modular with respect to Γ must exhibit a periodicity with respect to $h\mathbf{Z}$, leading to an associated function $g : D' \to \mathbf{C}$, where D' is the punctured disk, and $f(\tau) = g(e^{2\pi i\tau/h})$.
If f is also holomorphic on $\mathcal{H}$, then g extends holomorphically on the punctured disk and consequently has a Laurent series expansion.

Definition 2.4.10
The function f is said to be holomorphic at infinity if g can be holomorphically extended to $q = 0$, giving a Fourier expansion for f:

$$f(\tau) = \sum_{n=0}^{\infty} a_n e^{2\pi i n\tau/h}.$$

Definition 2.4.11 (Modular form of weight k with respect to a congruence subgroup Γ)
Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$ and let k be an integer. A function $f : \mathcal{H} \longrightarrow \mathrm{C}$ is a modular form of weight k with respect to Γ if
(1) f is holomorphic,
(2) f is weight- k invariant under Γ,
(3) $f[\alpha]_k$ is holomorphic at ∞ for all $\alpha \in \mathrm{SL}_2(\mathbf{Z})$.

Definition 2.4.12 (Cusp form of weight k with respect to a congruence subgroup Γ)
If f is a modular form of weight k with respect to a congruence subgroup Γ, and satisfies $a_0 = 0$ in the Fourier expansion of $f[\alpha]_k$ for all $\alpha \in \mathrm{SL}_2(\mathbf{Z})$, then f is called a cusp form of weight k with respect to Γ.

Remark 2.4.13
1. As mentioned before, we would like to have that the space of Modular forms is finite dimensional. It is thus natural to expect that they are holomorphic not only on $\mathcal{H}$ but also at certain limit points. In the case of a congruence subgroup Γ, this amounts to adjoining the rational numbers $\mathbf{Q}$ and the point at infinity to $\mathcal{H}$, and recognizing equivalence under Γ-action. A Γ-equivalence class of points in $\mathbf{Q} \cup \{\infty\}$ is termed a cusp of Γ. The number of such cusps is finite, given the finiteness of the index of Γ in $\mathrm{SL}_2(\mathbf{Z})$. We will see this later.

2. A modular form with respect to Γ must be holomorphic at all these cusps. For any rational number

s, represented as $s = \alpha(\infty)$ for some $\alpha \in \mathrm{SL}_2(\mathbf{Z})$, holomorphy at s translates to holomorphy at infinity under the action of the $[\alpha]_k$ operator.

3. Another interesting feature of modular forms with respect to this generality is the existence of non-zero modular forms of odd weight is when $-I$ is an element of Γ, contrary to the situation within the full modular group, $\mathrm{SL}_2(\mathbf{Z})$.

2.5 Fundamental Domains

From this point onwards, it is clear that we want to work with a modular form f for Γ. We note that the function value at some point $z \in \mathbb{C}$ determines the function value at the set of all points $\Gamma z = \{\gamma z \mid \gamma \in \Gamma\}$ by the weak modularity condition. Thus it is natural to be interested in and work with a minimal set of points on the complex plane that fully determines f on the upper half plane $\mathcal{H}$. This leads us to the notion of a fundamental domain.
Since we are considering the action of congruence subgroups on the upper half plane, to define a minimal set of points we naturally want that every orbit induced under the group action of Γ on $\mathcal{H}$ has at least one representative element in such a set $\mathcal{F}$. Furthermore, it is also a natural expectation that no two elements in the interior of our minimal set $\mathcal{F}$ should belong to the same orbit (i.e. the only possible exceptions are the boundaries).
Summarising, we define...

Definition 2.5.1 (**Fundamental Domains**)
Let f be a modular form of weight k for some finite index subgroup Γ. We define a **fundamental domain** of f as a subset $\mathcal{F} \subseteq \mathbb{C}$ satisfying the following two conditions:

1. $\Gamma\mathcal{F} = \{\gamma z \mid \gamma \in \Gamma \text{ and } z \in \mathcal{F}\} = \mathcal{H}$
2. $\mathrm{int}(\mathcal{F}) \cap \mathrm{int}(\gamma\mathcal{F}) = \emptyset \quad$ for all $\gamma \in \Gamma - \{I, -I\}$

Proposition 2.5.2
The subset $\mathcal{F} = \left\{\tau \in \mathcal{H} : \ |\mathrm{Re}(\tau)| \leq \frac{1}{2} \text{ and } |\tau| \geq 1\right\}$ of $\mathcal{H}$ is a (connected) fundamental domain for $\mathrm{SL}_2(\mathbb{Z})$. Let $\zeta_3 = e^{\frac{2\pi i}{3}}$ be the primitive third root of unity.
Moreover the stabilizer H_τ of a point $\tau \in \mathcal{F}$ in $\mathrm{SL}_2(\mathbb{Z})$ is

$$H_\tau = \begin{cases} C_6 = \langle ST \rangle = \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \right\rangle & \tau = \zeta_3, \\ C_6' = \langle TS \rangle = \left\langle \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle & \tau = \zeta_3 + 1, \\ C_4 = \langle S \rangle = \left\langle \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle & \tau = i, \\ C_2 = \langle -I \rangle = \left\langle \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\rangle & \text{else.} \end{cases}$$

Proof:
Let z be an element of the upper half-plane $\mathcal{H}$. As previously established, when γ belongs to the special linear group $\mathrm{SL}_2(\mathbb{Z})$, the following relationship holds:

$$\mathrm{Im}(\gamma z) = \frac{\mathrm{Im}(z)}{|cz+d|^2}, \quad \text{where} \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

There exist a finite number of pairs $(c,d) \in \mathbb{Z}^2$ for which $|cz+d| < 1$. Notably, it is possible to select a matrix $\gamma \in \langle S, T\rangle \subseteq \mathrm{SL}_2(\mathbb{Z})$ such that

$$\mathrm{Im}(\gamma z) \geq \mathrm{Im}\left(\gamma' z\right), \quad \text{for all} \quad \gamma' \in \langle S, T\rangle \subseteq \mathrm{SL}_2(\mathbb{Z})$$

We will now multiply γ by an appropriate power of T. This operation does not alter the imaginary part, allowing us to assume that $|\,\mathrm{Re}(\gamma z)| \leq \frac{1}{2}$.
Our next objective is to demonstrate that $|\gamma z| \geq 1$:

$$\mathrm{Im}(\gamma z) \geq \mathrm{Im}(S\gamma z) = \mathrm{Im}\left(-\frac{1}{\gamma z}\right) = \frac{\mathrm{Im}(\gamma z)}{|\gamma z|^2}$$

This implies $|\gamma z| \geq 1$, and consequently, γz belongs to the fundamental domain $\mathcal{F}$.
Suppose $z' = \gamma z$, and both z and z' are elements of $\mathcal{F}$. Without loss of generality, let us assume that $\mathrm{Im}(\gamma z) \geq \mathrm{Im}(z)$. Let us write $z = x + iy$. Then, $\mathrm{Im}(\gamma z) \geq \mathrm{Im}(z)$ is equivalent to

$$|cz+d|^2 = |cx+d|^2 + |cy|^2 \leq 1. \quad .$$

Given that $y > 1/2$, we deduce that $|c| \leq 1$. The case $c = 0$ implies $|d| \leq 1$. Furthermore, since $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$, this means $\gamma = \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$, representing a translation matrix. Therefore, $z' = z \pm 1$.
Let's assume $c = 1$ (the case $c = -1$ is analogous). In this scenario, the condition $|z+d|^2 \leq 1$ is only satisfied when $|z| = 1$ (for $d = 0$), when $z = \zeta_3$ (for $d = 1$), or when $z = \zeta_3 + 1$ (for $d = -1$).
For the analysis of stabilizers of points $z \in \mathcal{F}$, we can use the earlier calculations. If $\gamma z = z$, it necessarily follows that $c = \pm 1$, and by negating γ, we can assume $c = 1$. The quadratic equation resulting from $\gamma z = z$ leads to $|a+d| < 2$, which further implies $|a+d| \leq 1$. Concurrently, the condition that $z \in \mathcal{F}$ enforces $|a-d| \leq 1$. Together, these two inequalities yield $|a| \leq 1$.
We finish by analyzing the stabilizers in $\mathrm{SL}_2(\mathbb{Z})$
Case 1. Translation Matrices:
($\gamma = \pm I$): Here, γ is either $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ or $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. These matrices leave all points in $\mathcal{H}$ unchanged, hence fixing all points in $\mathcal{F}$.
Case 2. Horizontal Shifts:
Consider, ($\gamma = \pm \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $\pm \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$). Note that these matrices correspond to shifting points by ± 1 along the real axis. Since no point in $\mathcal{F}$ remains within $\mathcal{F}$ after such a shift, these matrices do not fix any points.
3. Inversion and Rotation:
($\gamma = \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$): This case includes matrices that essentially rotate and invert points in the complex

plane. The only fixed point under this transformation is i, as it is the unique point of intersection of the unit circle and the imaginary axis within $\mathcal{F}$.

4. The transformations, $\gamma = \pm\begin{pmatrix}-1 & -1\\ 1 & 0\end{pmatrix}, \pm\begin{pmatrix}0 & -1\\ 1 & 1\end{pmatrix}, \pm\begin{pmatrix}1 & -1\\ 1 & 0\end{pmatrix}, \pm\begin{pmatrix}0 & -1\\ 1 & -1\end{pmatrix}$ lead to the fixed points $\zeta_3 = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$ and $\zeta_3 + 1$. The former corresponds to one of the vertices of the fundamental domain, while the latter represents a point adjacent to ζ_3. The stabilizer subgroup calculations show that these specific transformations result in either ζ_3 or $\zeta_3 + 1$ being fixed.

□

Corollary 2.5.3

The group $\mathrm{SL}_2(\mathbb{Z})$ is generated by the matrices T and S.

Proof:

Let z_0 be a point in the interior of $\mathcal{F}$. Given $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, by proof of proposition 2.5.2, there exists a matrix, $\delta \in \langle T, S\rangle$ such that $\delta\gamma^{-1}z_0 \in \mathcal{F}$. Since z_0 is an interior point of $\mathcal{F}$, we have $\delta\gamma^{-1}z_0 = z_0$ and hence $\delta\gamma^{-1} = \pm I$. If $\delta\gamma^{-1} = I$, we are done. If $\delta\gamma^{-1} = -I$, we premultiply by $S^2 = -I$ to end the proof.

□

It is clear that there are infinitely many choices for a fundamental domain of the modular group $\mathrm{SL}_{2(}\mathbb{Z})$, but like many standard texts on Modular forms, we will mainly work with

$$\mathcal{F} = \left\{\tau \in \mathcal{H} : \ |\operatorname{Re}(\tau)| \leq \frac{1}{2} \text{ and } |\tau| \geq 1\right\}.$$

2.6 Moduli Spaces

Consider the set S of isomorphism classes of elliptic curves. Every complex elliptic curve is isomorphic to $\mathbb{C}/\Lambda$ for some lattice Λ, and in fact it is isomorphic to $\mathbb{C}/\Lambda_\tau$ for some $\tau \in \mathbb{H}$. Moreover,

$$\mathbb{C}/\Lambda_\tau \cong \mathbb{C}/\Lambda_{\tau'} \iff \mathrm{SL}_2(\mathbb{Z})\tau = \mathrm{SL}_2(\mathbb{Z})\tau'$$

Therefore there is a natural bijection

$$S \longleftrightarrow \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}, \quad [\mathbb{C}/\Lambda_\tau] \mapsto \mathrm{SL}_2(\mathbb{Z})\tau$$

The quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ is thus called the moduli space for isomorphism classes of elliptic curves.
Let now $f \in M_k(\mathrm{SL}_2(\mathbb{Z}))$ be a modular form of weight k. We define the following function F on the set of complex tori:

$$F(\mathbb{C}/\Lambda_\tau) = f(\tau)$$

This is well defined, because if $\Lambda_\tau = \Lambda_{\tau'}$ then $\tau = \tau' + b$ for some $b \in \mathbb{Z}$, and $f(\tau + b) = f(\tau)$. Moreover, suppose that $m\Lambda_\tau = \Lambda_{\tau'}$. Then

$$\tau = \begin{pmatrix} a & b\\ c & d\end{pmatrix}\tau', \quad m = c\tau' + d.$$

Then we may compute:

$$F\left(\mathbb{C}/m\Lambda_\tau\right) = F\left(\mathbb{C}/\Lambda_{\tau'}\right) = f\left(\tau'\right) = \left(c\tau'+d\right)^{-k} f(\tau) = F\left(\mathbb{C}/\Lambda_\tau\right) m^{-k}$$

From this we deduce that

$$F(\mathbb{C}/m\Lambda) = m^{-k} F(\mathbb{C}/\Lambda)$$

We could thus define modular forms as functions on complex tori satisfying the above relations. This prototype can be pushed to work for other congruence subgroups, although isomorphism classes of elliptic curves will have to be replaced by objects carrying more data.

Definition 2.6.1
An enhanced elliptic curve for $\Gamma_0(N)$ is a pair (E, C), where E is an elliptic curve and C is a cyclic subgroup of order N in $E[N]$. Two enhanced elliptic curves (E, C) and (E', C') are equivalent if there exists an isomorphism $\varphi : E \xrightarrow{\sim} E'$ such that $\varphi(C) = C'$.
We write $S_0(N)$ for the set of equivalence classes of enhanced elliptic curves.

Proposition 2.6.2 1. Each class in $S_0(N)$ has a representative of the form $\left(\mathbb{C}/\Lambda_\tau, \left\langle \frac{1}{N} + \Lambda_\tau \right\rangle\right)$, for some $\tau \in \mathbb{H}$.

2. Two pairs $\left(\mathbb{C}/\Lambda_\tau, \left\langle \frac{1}{N} + \Lambda_\tau \right\rangle\right)$ and $\left(\mathbb{C}/\Lambda_{\tau'}, \left\langle \frac{1}{N} + \Lambda_{\tau'} \right\rangle\right)$ are equivalent if and only if $\Gamma_0(N)\tau = \Gamma_0(N)\tau'$. Therefore the map $\tau \mapsto \left(\mathbb{C}/\Lambda_\tau, \left\langle \frac{1}{N} + \Lambda_\tau \right\rangle\right)$ induces a bijection of $Y_0(N) = \Gamma_0(N)\backslash\mathbb{H} \cong S_0(N)$

Proof:
Consider an enhanced elliptic curve $(\mathbb{C}/\Lambda, C)$. We have already seen that there is an isomorphism $\varphi : \mathbb{C}/\Lambda \cong \mathbb{C}/\Lambda_{\tau'}$ for some $\tau' \in \mathbb{H}$. Since C is cyclic of order N, the same is true for $\varphi(C)$. Therefore $(\mathbb{C}/\Lambda, C)$ is equivalent to $\left(\mathbb{C}/\Lambda_{\tau'}, \left\langle \frac{c\tau'+d}{N} + \Lambda_{\tau'} \right\rangle\right)$ for some integers c and d coprime to each other and to N. Since reduction modulo N gives a surjection $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N/ZZ)$, one can find a matrix

$$\gamma = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

such that $c' \equiv c (\mathrm{mod}\, N)$ and $d' \equiv d (\mathrm{mod}\, N)$. Set now $\tau = \gamma\tau'$ and $m = c'\tau' + d'$, so $m\Lambda_\tau = \Lambda_{\tau'}$ and, as we wanted to show,

$$m\left(\frac{1}{N} + \Lambda_\tau\right) = \frac{c'\tau'+d'}{N} + \Lambda_{\tau'} = \frac{c\tau'+d}{N} + \Lambda_{\tau'}$$

As for the second part, for an isomorphism between $\mathbb{C}/\Lambda_\tau$ and $\mathbb{C}/\Lambda_{\tau'}$ to exist there needs to exist $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$\left(c\tau'+d\right)\Lambda_\tau = \Lambda_{\tau'}$$

Moreover, for the corresponding isomorphism to respect the cyclic subgroups one needs to have

$$\left\langle \left(c\tau'+d\right)\left(\frac{1}{N} + \Lambda_\tau\right) \right\rangle = \left\langle \frac{1}{N} + \Lambda_{\tau'} \right\rangle$$

That is, γ satisfies

$$\left\langle \frac{c\tau'+d}{N} + \Lambda_{\tau'} \right\rangle = \left\langle \frac{1}{N} + \Lambda_{\tau'} \right\rangle$$

which is equivalent to $N \mid c$ (and then d is necessarily coprime to N). This last condition is precisely saying that γ must belong to $\Gamma_0(N)$

□

Definition 2.6.3
An enhanced elliptic curve for $\Gamma_1(N)$ is a pair (E, P), where E is an elliptic curve and P is a point of exact order N in $E[N]$. Two enhanced elliptic curves (E, P) and (E', P') are equivalent if there exists an isomorphism $\varphi : E \xrightarrow{\sim} E'$ such that $\varphi(P) = P'$.
We write $S_1(N)$ for the set of equivalence classes of enhanced elliptic curves for $\Gamma_1(N)$.

Proposition 2.6.4 1. Each class in $S_1(N)$ has a representative of the form $\left(\mathbb{C}/\Lambda_\tau, \frac{1}{N} + \Lambda_\tau\right)$, for some $\tau \in \mathbb{H}$.

2. Two pairs $\left(\mathbb{C}/\Lambda_\tau, \frac{1}{N} + \Lambda_\tau\right)$ and $\left(\mathbb{C}/\Lambda_{\tau'}, \frac{1}{N} + \Lambda_{\tau'}\right)$ are equivalent if and only if $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$. Therefore the map $\tau \mapsto \left(\mathbb{C}/\Lambda_\tau, \frac{1}{N} + \Lambda_\tau\right)$ induces a bijection of $Y_1(N) = \Gamma_1(N)\backslash\mathbb{H} \cong S_1(N)$.

Proof:
Let (E, Q) be any point in $S_1(N)$. Since E is isomorphic to $\mathbb{C}/\Lambda_{\tau'}$ for some $\tau' \in \mathbb{H}$, we may take $E = \mathbb{C}/\Lambda_{\tau'}$, and hence $Q = (c\tau' + d)/N + \Lambda_{\tau'}$ for some $c, d \in \mathbb{Z}$. The fact that the order of Q is exactly N means that $\gcd(c, d, N) = 1$, and therefore there exists $a, b, k \in \mathbb{Z}$ such that

$$ad - bc - kN = 1$$

Note that this means that the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has determinant $1 (\bmod N)$. Using that $\mathrm{SL}_2(\mathbb{Z})$ surjects into $\mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$ and the fact that c and d only matter modulo N, we find a matrix $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ with lower low (c, d). Let $\tau = \gamma\tau'$, and let $m = c\tau' + d$. Then we obtain $m\tau = a\tau' + b$, which implies that $m\Lambda_\tau = \Lambda_{\tau'}$. Moreover,

$$m\left(1/N + \Lambda_\tau\right) = \frac{c\tau'+d}{N} + \Lambda_{\tau'} = Q$$

Therefore the class $[E, Q]$ is the same as $[\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau]$.
Finally, given two points $\tau, \tau' \in \mathbb{H}$ such that $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$, we may write $\tau = \gamma\tau'$ for some $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1(N)$. Letting $m = c\tau' + d$, then:

$$m\Lambda_\tau = \Lambda_{\tau'}, \quad m\left(1/N + \Lambda_\tau\right) = \frac{c\tau'+d}{N} + \Lambda_{\tau'}$$

Since $(c, d) \equiv (0, 1) (\bmod N)$, the last term is just $1/N + \Lambda_{\tau'}$, as we wanted to show. □

2.7 Cusps and Elliptic Points

Definition 2.7.1 (**Modular Curve**)
For any congruence subgroup Γ in $\mathrm{SL}_2(\mathbb{Z})$, we define the modular curve $Y(\Gamma)$ as the set of orbits induced

by the action of Γ on $\mathcal{H}$.

$$Y(\Gamma) = \Gamma\backslash\mathcal{H} = \{\Gamma\tau \mid \tau \in \mathcal{H}\}$$

One can prove that $Y(\Gamma)$ forms a Riemann surface which is Haussdorf.(find a nice reference of state the proof). However, $Y(\Gamma)$ on its own is not compact. To compactify this surface, we need to extend the action of Γ on $\mathcal{H}$ to include the projective line over the the rationals $\mathbb{Q}$.

Definition 2.7.2
The **projective line** over $\mathbb{Q}$ is the set

$$\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$$

Remark 2.7.3
We know that the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$ gives fractional linear transformations on $\mathcal{H}$. We can extend this action to $\mathbb{P}^1(\mathbb{Q})$ as well, where we define

$$\gamma\left(\frac{p}{q}\right) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{p}{q} := \frac{ap+bq}{cp+dq} \quad \text{for all } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$$

and we define

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \infty = \frac{a}{c} \quad \text{and} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{-d}{c} = \infty$$

We therefore have that $\mathrm{SL}_2(\mathbb{Z})$ also acts on $\mathbb{P}^1(\mathbb{Q})$. Indeed, for an arbitrary subgroup Γ, we let Γ act on $\mathcal{H}\cup\mathbb{P}^1(\mathbb{Q}) = \mathcal{H}\cup\mathbb{Q}\cup\{\infty\}$. Taking the extended quotient induced by this action give us the compactified Riemann surface $X(\Gamma)$, See ([DS05],p.58),defined as

$$X(\Gamma) = \Gamma\backslash\mathcal{H} \cup \mathbb{P}^1(\mathbb{Q}) = Y(\Gamma) \cup \Gamma\backslash\mathbb{P}^1(\mathbb{Q}) = Y(\Gamma) \cup \Gamma\backslash\mathbb{Q} \cup \{\infty\}\big)$$

We, therefore, note that $X(\Gamma)$ consists of adjoining the points $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ to the modular curve $Y(\Gamma)$, thus compactifying the Riemann surface. We define these points $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ as the cusps of $X(\Gamma)$, which we shall also denote as Cusps(Γ).
The cusps are hence simply the set of orbits in $\mathbb{P}^1(\mathbb{Q})$ induced by the action of Γ. We first determine the cusps for $\mathrm{SL}_2(\mathbb{Z})$.

Proposition 2.7.4
$X(\mathrm{SL}_2(\mathbb{Z}))$ contains a single cusp.

Proof:
We simply prove that $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$. Indeed, let $c \in \mathbb{Q}$ be given, and denote $c = \frac{p}{q}$ where p and q are coprime. Thus, by Bézout's identity, there exists $r, s \in \mathbb{Z}$ such that $pr - qs = 1$. Now, define $\gamma := \begin{bmatrix} p & s \\ q & r \end{bmatrix}$. By definition of r and s we have that $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Note

$$\gamma(\infty) = \begin{bmatrix} p & s \\ q & r \end{bmatrix} \infty = \frac{p}{q} = c$$

Thus, c is in the orbit of ∞, for all $c \in \mathbb{Q}$. Hence, the action of $\mathrm{SL}_2(\mathbb{Z})$ induces a single orbit on $\mathbb{P}^1(\mathbb{Q})$,

and thus $\mathrm{SL}_2(\mathbb{Z})$ acts transitively.

□

We can furthermore prove that for any finite-index subgroup $\Gamma, X(\Gamma)$ consists of finitely many cusps.

Proposition 2.7.5
Let Γ be a finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Then $X(\Gamma)$ contains finitely many cusps.

Proof:
We first prove that the stabiliser of $\mathrm{SL}_2(\mathbb{Z})$ with respect to ∞ is

$$\mathrm{SL}_2(\mathbb{Z})_\infty = \left\{ \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbb{Z} \right\}$$

Indeed, let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})_\infty$. Then

$$\gamma\infty = \infty \implies \begin{pmatrix} a & b \\ c & d \end{pmatrix} \infty = \infty \implies \frac{a}{c} = \infty \implies c = 0$$

Thus, as $\det(\gamma) = 1$, we have $ad = 1$ and hence $a = d = 1$ or $a = d = -1$. We therefore have $\gamma = \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$. Conversely, it can easily be seen that any such γ of that form stabilizes ∞, thus proving the claim.
Thus, by the Orbit-Stabilizer theorem, we obtain a bijection between $\mathrm{SL}_2(\mathbb{Z})\backslash\mathrm{SL}_2(\mathbb{Z})_\infty$ and $\mathbb{P}^1(\mathbb{Q})$ given by the map $\gamma\mathrm{SL}_2(\mathbb{Z})_\infty \mapsto \gamma(\infty)$.
We now consider the function $\varphi : \Gamma\backslash\mathrm{SL}_2(\mathbb{Z}) \to \Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ defined by

$$\Gamma\gamma \mapsto \Gamma\gamma(\infty)$$

Note that φ is well-defined by associativity of the group action induced by $\mathrm{SL}_2(\mathbb{Z})$. Furthermore, we have that φ is surjective, as $\mathrm{SL}_2(\mathbb{Z})$ acts transitively on $\mathbb{P}^1(\mathbb{Q})$. Thus the order of $\Gamma\backslash\mathbb{P}^1(\mathbb{Q})$ is no greater than the order of $\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$, and as Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$ this implies $X(\Gamma)$ contains finitely many cusps.

□

Example 2.7.6
[Vis18]
Consider $\Gamma = \Gamma_1(4)$. Notably, the matrix $-I$ is not an element of Γ. Using the index formula, of Γ in $\mathrm{SL}_2(\mathbb{Z})$ we have that the index is 12.
Coset Representatives in $\mathrm{SL}_2(\mathbb{Z})$:
A set R of coset representatives for Γ in $\mathrm{SL}_2(\mathbb{Z})$ is given by:

$$R = \{I, -I, S, -S, ST, -ST, ST^{-1}, -ST^{-1}, ST^2, -ST^2, ST^2S, -ST^2S\}.$$

Coset Representatives in $\mathrm{PSL}_2(\mathbb{Z})$:
Similarly, a set Q of coset representatives for Γ in $\mathrm{PSL}_2(\mathbb{Z})$ includes:

$$Q = \{\pm I, \pm S, \pm ST, \pm ST^{-1}, \pm ST^2, \pm ST^2S\}.$$

Finding All Cusps:
To identify all the cusps, we calculate $\gamma(\infty)$ for each γ in R. We observe:
- $I(\infty) = \infty$, - For γ in $\{\pm S, \pm ST, \pm ST^{-1}, \pm ST^2\}$, $\gamma(\infty) = 0$, - For γ in $\{\pm ST^2S\}$, $\gamma(\infty) = -\frac{1}{2}$.
Since $T \in \Gamma$, we have $\frac{1}{2} = T(-\frac{1}{2}) = TST^2S(\infty)$. Therefore, $\frac{1}{2}$ and $-\frac{1}{2}$ correspond to the same cusp under the action of Γ. Also, ∞ and $\frac{1}{2}$, as well as 0 and $\frac{1}{2}$, are not in the same orbit.
Visualizing the Cosets and Cusps:
Visual observation reveals that the six coset translates of I correspond to the six cosets of $\Gamma_1(4)$ in $\mathrm{PSL}_2(\mathbb{Z})$. The three cusps at ∞, 0, and $-\frac{1}{2}$ correspond to the sets of cosets $\{I\}$, $\{ST^{-1}, S, ST, ST^2\}$, and $\{ST^2S\}$, respectively.
From this analysis, it becomes evident that the modular curve $X(\Gamma_1(4))$ consists of exactly three distinct cusps:

$$\mathrm{Cusps}(\Gamma_1(4)) = \{\Gamma(\infty), \Gamma(0), \Gamma(\frac{1}{2})\}.$$

Next, we come to the notion of **Elliptic Points** we remarked earlier that matrices I and $-I$ act trivially at every point $\tau \in \mathcal{H}$. That is $I\tau = (-I)\tau = \tau$ for all $\tau \in \mathcal{H}$. We are furthermore interested in calculating the points $\tau \in \mathcal{H}$ for which there exist matrices $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ other than I and $-I$ such that $\gamma(\tau) = \tau$.

Definition 2.7.7
Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We define the stabiliser subgroup of Γ with respect to τ (also denoted the isotropy subgroup of τ) as

$$\Gamma_\tau = \{\gamma \in \Gamma : \gamma(\tau) = \tau\}$$

Remark 2.7.8 1. Γ_τ is a subgroup of $\mathrm{SL}_2(\mathbb{Z})_\tau$.

2. For any $\delta \in \mathrm{SL}_2(\mathbb{Z})$, we have $\mathrm{SL}_2(\mathbb{Z})_{\delta\tau} = \delta\mathrm{SL}_2(\mathbb{Z})_\tau\delta^{-1}$, and thus $|\mathrm{SL}_2(\mathbb{Z})_{\delta\tau}| = |\mathrm{SL}_2(\mathbb{Z})_\tau|$.

Let us now calculate for a point τ the stabilizer subgroup, $\mathrm{SL}_2(\mathbb{Z})_\tau$ and summarise the result in the following proposition.

Proposition 2.7.9

$$|\mathrm{SL}_2(\mathbb{Z})_\tau| = \begin{cases} 4 & \text{if } \gamma(\tau) = i \text{ for some } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\ 6 & \text{if } \gamma(\tau) = \omega \text{ for some } \gamma \in \mathrm{SL}_2(\mathbb{Z}) \\ 2 & \text{otherwise} \end{cases}$$

Proof:
We can simplify our work by focusing on points within the fundamental domain $\mathcal{F}$. Let's consider a point τ in $\mathcal{F}$. It's evident that the set $\{I, -I\}$ is a subset of $\mathrm{SL}_2(\mathbb{Z})_\tau$. Suppose there's another element γ_0 in Γ that belongs to $\mathrm{SL}_2(\mathbb{Z})_\tau$ but is not in $\{I, -I\}$. This leads to the following equations:

$$\gamma_0(\tau) = \tau \quad \Longrightarrow \quad \frac{a\tau + b}{c\tau + d} = \tau \quad \Longrightarrow \quad c\tau^2 + (d - a)\tau - b = 0.$$

Assuming $c = 0$ leads to a contradiction. Either $a = d$ or $\tau = \frac{b}{d-a}$, but since $\tau \in \mathcal{H}$, it is not a rational number, $a = d$, and thus $b = 0$. Considering the determinant condition, $ad = 1$, we find that $a = d = 1$ or $a = d = -1$, contradicting $\gamma_0 \notin \{I, -I\}$. So, $c \neq 0$.
This leads us to:

$$\tau = \frac{a - d \pm \sqrt{(a-d)^2 + 4bc}}{2c},$$

and, taking the determinant into account:

$$\tau = \frac{a - d \pm \sqrt{(a+d)^2 - 4}}{2c}.$$

Given $\tau \in \mathcal{H}$, we deduce:

$$(a+d)^2 - 4 < 0 \quad \Longrightarrow \quad a + d \in \{-1, 0, +1\}.$$

We now consider the different cases for $a + d$:
- Case 1 ($a + d = 0$): We get $\tau = \frac{a}{c} \pm \frac{i}{c}$. The condition $\mathrm{Im}(\tau) \geq \frac{\sqrt{3}}{2}$ implies $|c| = 1$, leading to $\tau = i$. - Case 2 ($a + d = \pm 1$): We find $\tau = \frac{2a \pm 1}{2c} \pm \frac{\sqrt{3}}{2c} i$. The same condition on the imaginary part of τ leads us to $\tau = \pm\frac{1}{2} + \frac{\sqrt{3}}{2} i$.
For the stabilizer subgroup calculations, let's take $\gamma \in \mathrm{SL}_2(\mathbb{Z})_i$ and not in $\{I, -I\}$. Through similar reasoning, we find that γ must be in $\{S, -S\}$. Hence,

$$\mathrm{SL}_2(\mathbb{Z})_i = \{I, -I, S, -S\}.$$

Similarly, for $\mathrm{SL}_2(\mathbb{Z})_\omega$, considering $\gamma \notin \{I, -I\}$, we explore the cases for $a + d$ and deduce:

$$\gamma \in \left\{ \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \right\}.$$

Therefore,

$$\mathrm{SL}_2(\mathbb{Z})_\omega = \{I, -I\} \cup \left\{ \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \right\}.$$

□

Definition 2.7.10
Let Γ be a subgroup of $\mathrm{SL}_2(\mathbb{Z})$. A point $\tau \in \mathcal{H}$ is an elliptic point if the stabiliser subgroup Γ_τ is non-trivial. That is, there exists $\gamma \in \Gamma_\tau$ such that $\gamma \notin \{I, -I\}$.

Proposition 2.7.11
Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$. For each elliptic point τ of Γ the isotropy subgroup Γ_τ is finite cyclic. Moreover, each point $\tau \in \mathcal{H}$ has an associated positive integer,

$$h_\tau = |\{\pm I\}\Gamma_\tau / \{\pm I\}| = \begin{cases} |\Gamma_\tau| / 2 & \text{if } -I \in \Gamma_\tau, \\ |\Gamma_\tau| & \text{if } -I \notin \Gamma_\tau. \end{cases}$$

Proof:
See, [DS05], chapter 2. □

Remark 2.7.12 1. h_τ is called the period of τ, with $h_\tau > 1$ only for the elliptic points.

2. If $\tau \in \mathcal{H}$ and $\gamma \in \mathrm{SL}_2(\mathbf{Z})$ then the period of $\gamma(\tau)$ under $\gamma\Gamma\gamma^{-1}$ is the same as the period of τ under Γ. In particular, h_τ depends only on $\Gamma\tau$, making the period well defined on $Y(\Gamma)$, and if Γ is normal in $\mathrm{SL}_2(\mathbf{Z})$ then all points of $Y(\Gamma)$ over a point of $Y(\mathrm{SL}_2(\mathbf{Z}))$ have the same period.

3. The space $Y(\Gamma)$ depends on Γ as a group of transformations acting on $\mathcal{H}$, and $-I$ acts trivially, so defining the period as we did rather than simply taking $h_\tau = |\Gamma_\tau|$ is natural. The definition correctly counts the τ-fixing transformations.

2.8 The Genus: An application

In this short section, we will discuss a short application of some of the theory introduced and some we will introduce along the way.

In particular, we choose a particular curve $X_0(38)$ curve$X_0(38)$ using the material we have covered up so far.

As with every geometric object, our first attempt will be about understanding the genus of the Modular curve $X_0(38)$. The usual method in Algebraic geometry is to use Riemann-Hurwitz formula.
Let us first setup some background with notations.

Let's consider a congruence subgroup Γ of $\mathrm{SL}_2(\mathbf{Z})$. In chapter 4, we will see that we get a compact Riemann surface, $X(\Gamma) = \Gamma\backslash\mathcal{H}^*$. Topologically, it can be visualized as a sphere with a certain number of handles, say g, where g represents the genus of $X(\Gamma)$.

Consider a map $f : X \to Y$ between two compact Riemann surfaces, where f is holomorphic and not constant. Such a map f is surjective. This map f is characterized by a well-defined degree d, a positive integer, satisfying $|f^{-1}(y)| = d$ for almost all points y in Y. For any point x in X, let's denote e_x, a positive integer, as the degree of ramification of f at x. Intuitively, this just means the multiplicity with which f takes 0 to 0 in local co-ordinates.

It can be observed that the f is unramified at all but finitely many points in X. Let us thus define the exceptional set $\mathcal{E}$ in X as those points where f ramifies. Then, by removing these exceptional points from X and their images from Y, we get two Riemann surfaces, X' and Y', respectively. Choosing any point y in Y', we can find neighborhoods around each pre-image of y in X' where f acts as a local bijection. The integer-valued function $y \mapsto |f^{-1}(y)|$ on Y' turns out to be constant due to the connectedness of Y'. Let d be the value of this constant, leading to the previously mentioned sum equation for all y in Y', and by continuity, this extends to all of Y. In summary, we have that, there exists a positive integer d such that:

$$\sum_{x \in f^{-1}(y)} e_x = d \quad \text{for all} \quad y \in Y$$

Revisiting the map $f : X \to Y$ of degree d between two compact Riemann surfaces, we denote the genera of X and Y by g_X and g_Y, respectively. To find the genus of $X(\Gamma)$, the Riemann-Hurwitz formula comes into play:

$$2g_X - 2 = d(2g_Y - 2) + \sum_{x \in X} (e_x - 1)$$

As stated earlier, our first goal is to compute the genus of $X(\Gamma)$. Let us consider Γ_1 as Γ and Γ_2 as $\mathrm{SL}_2(\mathbf{Z})$. Define the points y_2 as the image of i under $\mathrm{SL}_2(\mathbf{Z})$, y_3 as the image of μ_3 (representing an elliptic point of period 3), and y_∞ as the image of infinity, being the cusp of $X(1) = \mathrm{SL}_2(\mathbf{Z})\backslash\mathcal{H}^*$.

Let's denote by ε_2 and ε_3 the number of elliptic points in $X(\Gamma)$ that map to y_2 and y_3 respectively under a projection f, and ε_∞ the count of cusps in $X(\Gamma)$. Furthermore for $h = 2, 3$ we have,

$$\sum_{x \in f^{-1}(y_h)} (e_x - 1) = (h-1)(|f^{-1}(y_h)| - \varepsilon_h) = \frac{h-1}{h}(d - \varepsilon_h).$$

Similarly, we find that:

$$\sum_{x \in f^{-1}(y_\infty)} (e_x - 1) = d - \varepsilon_\infty.$$

Since $X(1)$ is of genus 0, the Riemann-Hurwitz formula yields the following theorem:

Theorem 2.8.1
Given Γ as a congruence subgroup of $\mathrm{SL}_2(\mathbf{Z})$ and $f : X(\Gamma) \to X(1)$ as the natural projection with degree d, and denoting ε_2, ε_3, and ε_∞ as the respective number of elliptic points of periods 2 and 3, and cusps in $X(\Gamma)$, the genus of $X(\Gamma)$ is determined by:

$$g = 1 + \frac{d}{12} - \frac{\varepsilon_2}{4} - \frac{\varepsilon_3}{3} - \frac{\varepsilon_\infty}{2}.$$

Proof:
Let g denote the genus of the curve $X(\Gamma)$.
Let us first note a couple of things.

1. The genus of the curve $X(1)$ is 0.
2. f is unramified at all but finitely many points and thus, $e_x - 1 = 0$ for all but exceptional points of $X(\Gamma)$.

Consequently, the Riemann-Hurwitz formula reduces to,

$$2g_X - 2 = -2d + \sum_{x \in \mathcal{E}} (e_x - 1)$$

As the execeptional points are split up in 3 categories, namely pre-images of elliptic points and cusps for $X(1)$, we can rewrite Riemann-Hurwitz formula as,

$$2g_X - 2 = -2d + \sum_{x \in f^{-1}(y_2)} (e_x - 1) + \sum_{x \in f^{-1}(y_3)} (e_x - 1) + \sum_{x \in f^{-1}(y_\infty)} (e_x - 1).$$

Using the expressions stated above, we have that,

$$2g_X - 2 = -2d + \frac{2-1}{2}(d - \varepsilon_2) + \frac{3-1}{3}(d - \varepsilon_3) + d - \varepsilon_\infty.$$

Simplifying we have,

$$2g_X - 2 = -2d + \frac{1}{2}(d - \varepsilon_2) + \frac{2}{3}(d - \varepsilon_3) + d - \varepsilon_\infty.$$

The claim follows by further simplifying and rearranging the terms.

□

Thus, what comes out as a conclusion, is that for us to find the genus of the Modular curve $X = X_0(38)$, we must find:

1. d: The degree of the projection map from the Modular curve X onto $X(1)$.
2: ε_2: The number of elliptic points of period 2.
3. ε_3: The number of elliptic points of period 3.
4. ε_∞: The number of cusps of X.

Theorem 2.8.2

The period 2 elliptic points of $\Gamma_0(N)$ are in bijective correspondence with the ideals J of $\mathbf{Z}[i]$ such that $\mathbf{Z}[i]/J \cong \mathbf{Z}/N\mathbf{Z}$. The period 3 elliptic points of $\Gamma_0(N)$ are in bijective correspondence with the ideals J of $\mathbf{Z}[\mu_6]$ (where $\mu_6 = e^{2\pi i/6}$) such that $\mathbf{Z}[\mu_6]/J \cong \mathbf{Z}/N\mathbf{Z}$.

Consequently, counting the ideals we have the number of elliptic points for $\Gamma_0(N)$ given by

$$\varepsilon_2(\Gamma_0(N)) = \begin{cases} \prod_{p|N} \left(1 + \left(\frac{-1}{p}\right)\right) & \text{if } 4 \nmid N \\ 0 & \text{if } 4 \mid N \end{cases}$$

where $(-1/p)$ is ± 1 if $p \equiv \pm 1 (\mathrm{mod} 4)$ and is 0 if $p = 2$, and

$$\varepsilon_3(\Gamma_0(N)) = \begin{cases} \prod_{p|N} \left(1 + \left(\frac{-3}{p}\right)\right) & \text{if } 9 \nmid N, \\ 0 & \text{if } 9 \mid N, \end{cases}$$

where $(-3/p)$ is ± 1 if , $p \equiv \pm 1 (\mathrm{mod} 3)$ and is 0 if $p = 3$.

Proof:

See, [DS05] Proposition 3.7.1. □

Corollary 2.8.3

For the modular curve $X = X_0(38)$, we have $\varepsilon_2 = \varepsilon_3 = 0$.

Proof:

From the formula, we have that

$$\varepsilon_2\left(\Gamma_0(38)\right) = \begin{cases} \prod_{p|38}\left(1+\left(\frac{-1}{p}\right)\right) & \text{if } 4 \nmid N \\ 0 & \text{if } 4 \mid N \end{cases}$$

Note that, $19 \equiv -1(\text{mod}4)$ and thus $1+(\frac{-1}{19}) = 0$. Since $19 \mid 38$, we have that $\varepsilon_2 = 0$.

Similarly, $2 \equiv -1(\text{mod}3)$ and $2 \mid 38$. Thus, the claim follows. □

Next, we compute, the degree of the projection map from X to $X(1)$.

Since $-I \in \text{SL}_2(\mathbf{Z})$ and $-I \in \Gamma(N)$ only for $N = 2$.Since $[\text{SL}_2(\mathbf{Z}) : \Gamma(N)] = N^3 \prod_{p|N}\left(1-1/p^2\right)$, the projection of modular curves $X(N) \longrightarrow X(1)$ has degree

$$d_N = [\text{SL}_2(\mathbf{Z}) : \{\pm I\}\Gamma(N)] = \begin{cases} (1/2)N^3 \prod_{p|N}\left(1-1/p^2\right) & \text{if } N > 2 \\ 6 & \text{if } N = 2 \end{cases}$$

In particular, for discussing the subgroup $\Gamma_0(N)$, let us consider without loss of generality that $N > 2$.This is because for $N = 2$, the group $\Gamma_0(2)$ is essentially the same as $\Gamma_1(2)$.
The negative identity matrix, denoted as $-I$ is not contained in $\Gamma_0(N)$, although it is included in $\text{SL}_2(\mathbf{Z})$. Given the index relation $[\Gamma_0(N) : \Gamma_1(N)] = \varphi(N)$, where φ represents Euler's totient function, it follows that the projection from the modular curve $X_0(N)$ to $X(1)$ has a specific degree. Specifically, for N exceeding 2, this degree is given by:

$$d(\Gamma_0(N)) = \frac{2d_N}{N\varphi(N)} \quad \text{for} \quad N > 2$$

Thus, for $N = 38$, degree $d = 60$, after substituting $N = 38$ in above formulas.

All that remains is to compute the number of cusps.
In general, we have that the number of cusps of $\Gamma_0(N)$ is given by,

$$\varepsilon_\infty\left(\Gamma_0(N)\right) = \sum_{d|N} \varphi(\gcd(d, N/d)).$$

Thus, for $N = 38$, degree $\varepsilon_\infty = 4$, after substituting $N = 38$ in above formula.

Thus, putting everything together, we have that the genus of the Modular curve $X_0(38)$ is given by

$$g = 1 + \frac{60}{12} - \frac{0}{4} - \frac{0}{3} - \frac{4}{2}.$$

Thus, $g = 1 + 5 - 2$, That is $g = 4$.

3 Hecke Operators

Hecke operators are fundamental mathematical tools that arise in the field of number theory, specifically in the study of modular forms and their associated L-functions. These operators were introduced by the German mathematician Erich Hecke in the early 20th century, and they have since played a central role in numerous areas of mathematics, including algebraic number theory, arithmetic geometry, and representation theory. Essentially, they provide a means to transform one modular form into another while preserving important properties. In this chapter, we will explore the theory of Hecke operators in depth, starting with their definition and basic properties. We will investigate their action on modular forms, establish the key algebraic properties of Hecke operators, and delve into their connection with Hecke eigenforms. We will closely follow the book [DS05] and Lecture notes from Marc Masdeu, [Mas15]. We will sometimes closely follow proofs from either of these sources for the sake of completeness. At times the ideas of the proofs are not original, neither I want to claim so, although, at some places, I have expanded on my own giving more details.
We will start with defining double coset operators, a notion that lies as a key concept in the background of theory.

Definition 3.0.1
Let Γ_1 and Γ_2 be two congruence subgroups, and let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$.
The **double coset** $\Gamma_1\alpha\Gamma_2$ is the set

$$\Gamma_1\alpha\Gamma_2 = \{\gamma_1\alpha\gamma_2 \mid \gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2\}$$

Proposition 3.0.2
If Γ is a congruence subgroup and $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$, Then:

1. $\alpha^{-1}\Gamma\alpha \cap \mathrm{SL}_2(\mathbb{Z})$ is also congruence subgroup
2. Any two congruence subgroups Γ_1, Γ_2 are commensurable. That is,

$$[\Gamma_1 : \Gamma_1 \cap \Gamma_2] < \infty \quad \text{and} \quad [\Gamma_2 : \Gamma_1 \cap \Gamma_2] < \infty$$

Proof:
By taking the least common multiple of the denominators of rational entries of α and α^{-1} we get a positive integer N_1 such that $N_1\alpha \in M_2(\mathbb{Z})$ and $N_1\alpha^{-1} \in M_2(\mathbb{Z})$. Furthermore, since Γ is a congruence subgroup, there is a positive integer N_2 such that $\Gamma(N_2) \subseteq \Gamma$. Let $N = \mathrm{LCM}(N_1, N_2)$. We know, from the fact that for positive integers N, M we have that $\Gamma(\mathrm{LCM}(N, M)) \subseteq \Gamma(N) \cap \Gamma(M)$. Thus, for given N we have that, $\Gamma(N) \subseteq \Gamma(N_1) \cap \Gamma(N_2) \subseteq \Gamma(N_2) \subseteq \Gamma$ and $N\alpha, N\alpha^{-1} \in M_2(\mathbb{Z})$.
Set $M = N^3$. Then we claim that $\alpha\Gamma(M)\alpha^{-1} \subseteq \Gamma(N)$, which implies that $\Gamma(M) \subseteq \alpha^{-1}\Gamma\alpha$. This is because any element of the set $\alpha\Gamma(M)\alpha^{-1}$ is of the form $\alpha\gamma\alpha^{-1}$ where $\gamma \equiv I \pmod{M}$. This means that we can write γ as $\gamma = I + N^3\gamma'$ for $\gamma' \in M_2(\mathbb{Z})$. Thus we have that, $\alpha\gamma\alpha^{-1} = I + N^3\alpha\gamma'\alpha^{-1}$. Using the fact that $N\alpha \in M_2(\mathbb{Z})$ and $N\alpha^{-1} \in M_2(\mathbb{Z})$, we get that $\alpha\gamma\alpha^{-1} = I + N\gamma''$, for some $\gamma'' \in M_2(\mathbb{Z})$ which gives us the claim. Since $\Gamma(M)$ is also contained in $\mathrm{SL}_2(\mathbb{Z})$, we have that $\Gamma(M) \subseteq \alpha\Gamma\alpha^{-1} \cap \mathrm{SL}_2(\mathbb{Z})$ which completes the proof of the first statement.
For the second assertion, let us begin by noting that there exists N, M with $\Gamma(N) \subseteq \Gamma_1$ and $\Gamma(M) \subseteq \Gamma_2$,

respectively. Then, we know that $\Gamma(\mathrm{LCM}(N,M)) \subseteq \Gamma(N) \cap \Gamma(M) \subseteq \Gamma_1 \cap \Gamma_2$. Thus, there is some $L = \mathrm{LCM}(N,M)$ such that $\Gamma(L) \subseteq \Gamma_1 \cap \Gamma_2$. Therefore the indices to compute are bounded above by $[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(L)]$, which is finite. □

Proposition 3.0.3
Let Γ_1 and Γ_2 be two congruence subgroups, and let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$. Let Γ_3 be the congruence subgroup defined as: $\Gamma_3 = \left[\alpha^{-1}\Gamma_1\alpha\right] \cap \Gamma_2$.
Then, the map

$$\Gamma_2 \to \Gamma_1 \backslash \left[\Gamma_1 \alpha \Gamma_2\right], \quad \gamma_2 \mapsto \Gamma_1 \alpha \gamma_2$$

induces a bijection between $\Gamma_3\backslash\Gamma_2$ and $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$.

Proof:
Consider the map

$$\Gamma_2 \to \Gamma_1 \backslash \left[\Gamma_1 \alpha \Gamma_2\right], \quad \gamma_2 \mapsto \Gamma_1 \alpha \gamma_2.$$

By definition of the map, it is surjective.
Moreover, two elements γ_2 and γ_2' get mapped to the same orbit if and only if:

$$\Gamma_1\alpha\gamma_2 = \Gamma_1\alpha\gamma_2' \iff \gamma_2'\gamma_2^{-1} \in \alpha^{-1}\Gamma_1\alpha$$

and the latter happens if and only if γ_2 and γ_2' are in the same coset for $\left[\alpha^{-1}\Gamma_1\alpha\right] \cap \Gamma_2 = \Gamma_3$.

□

Corollary 3.0.4
Let $\Gamma_2 = \cup \Gamma_3\gamma_j$ be a coset decomposition of $\Gamma_3\backslash\Gamma_2$. Then

$$\Gamma_1\alpha\Gamma_2 = \cup \Gamma_1\alpha\gamma_j$$

It is an orbit decomposition. In particular, the number of orbits of $\Gamma_1\alpha\Gamma_2$ under the action of Γ_1 is finite.

Definition 3.0.5
Let $f \in M_k[\Gamma_1]$ be a modular form of weight k for a congruence subgroup Γ_1. Let $\Gamma_1\alpha\Gamma_2$ be a double coset, where Γ_2 is a congruence subgroup and $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$. $\beta \in \mathrm{GL}_2^+(\mathbb{Q})$ and $k \in \mathbb{Z}$, the weight- $k\beta$ operator on functions $f : \mathcal{H} \longrightarrow \mathbb{C}$ is given by

$$[f[\beta]_k](\tau) = (\det\beta)^{k-1} j(\beta,\tau)^{-k} f(\beta(\tau)), \quad \tau \in \mathcal{H}.$$

The action of the double coset on f is defined as:

$$f\left[\Gamma_1\alpha\Gamma_2\right]_k = \sum_j f\left[\beta_j\right]_k$$

where $\Gamma_1\alpha\Gamma_2 = \cup\Gamma_1\beta_j$ is any orbit decomposition.

Proposition 3.0.6 1. The action is well-defined.

2. The double coset operator $[\Gamma_1\alpha\Gamma_2]_k : \mathcal{M}_k[\Gamma_1] \longrightarrow \mathcal{M}_k[\Gamma_2]$ takes modular forms with respect to Γ_1 to modular forms with respect to Γ_2. That is for each $f \in \mathcal{M}_k[\Gamma_1]$, the transformed $f[\Gamma_1\alpha\Gamma_2]_k$ is Γ_{2-} invariant and is holomorphic at the cusps.

3. The double coset operator $[\Gamma_1\alpha\Gamma_2]_k : \mathcal{S}_k[\Gamma_1] \longrightarrow \mathcal{S}_k[\Gamma_2]$ takes cusp forms to cusp forms, That is for each $f \in \mathcal{S}_k[\Gamma_1]$, the transformed $f[\Gamma_1\alpha\Gamma_2]_k$ vanishes at the cusps.

Example 3.0.7 1. $\Gamma_1 \supset \Gamma_2$. Taking $\alpha = I$ makes the double coset operator be $f[\Gamma_1\alpha\Gamma_2]_k = f$, the natural inclusion of the subspace $\mathcal{M}_k[\Gamma_1]$ in $\mathcal{M}_k[\Gamma_2]$, an injection.

2. As a more interesting example, given $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$ consider the conjugate $\Gamma' = \alpha^{-1}\Gamma\alpha$. Then $\Gamma\alpha\Gamma' = \Gamma\alpha$ is an orbit decomposition. This implies that acting by α induces a map

$$M_k(\Gamma) \to M_k\left[\alpha^{-1}\Gamma\alpha\right]$$

Since the inverse of this map is given by the action of α^{-1}, we conclude that $M_k(\Gamma)$ and $M_k\left[\alpha^{-1}\Gamma\alpha\right]$ are naturally isomorphic.

3. $\Gamma_1 \subset \Gamma_2$. Taking $\alpha = I$ and letting $\{\gamma_{2,j}\}$ be a set of coset representatives for $\Gamma_1\backslash\Gamma_2$ makes the double coset operator be $f[\Gamma_1\alpha\Gamma_2]_k = \sum_j f[\gamma_{2,j}]_k$, the natural trace map that projects $\mathcal{M}_k[\Gamma_1]$ onto its subspace $\mathcal{M}_k[\Gamma_2]$ by symmetrizing over the quotient, a surjection.

In fact, any double coset operator is a composition of these. Given Γ_1, Γ_2, and α, set $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2$ as usual and set $\Gamma_3' = \alpha\Gamma_3\alpha^{-1} = \Gamma_1 \cap \alpha\Gamma_2\alpha^{-1}$. Then $\Gamma_1 \supset \Gamma_3'$ and $\alpha^{-1}\Gamma_3'\alpha = \Gamma_3$ and $\Gamma_3 \subset \Gamma_2$, giving the three cases. The corresponding composition of double coset operators is

$$f \mapsto f \mapsto f[\alpha]_k \mapsto \sum_j f[\alpha\gamma_{2,j}]_k$$

3.1 The $\langle d\rangle$ and T_p operators

Let N be a positive integer. We start by recalling two congruence subgroups of the form

$$\Gamma_0(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} * & * \\ 0 & * \end{bmatrix} (\mathrm{mod} N) \right\}$$

and

$$\Gamma_1(N) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & * \\ 0 & 1 \end{bmatrix} (\mathrm{mod} N) \right\}$$

We have the containment $\Gamma_1(N) \subset \Gamma_0(N)$, and thus it is clear and expected that the smaller group has more modular forms associated to it, *i. e.* $\mathcal{M}_k[\Gamma_0(N)] \subseteq \mathcal{M}_k[\Gamma_1(N)]$.
We define two essential operators on the bigger space in this case, that is, on $\mathcal{M}_k[\Gamma_1(N)]$:
The $\langle d\rangle$ and T_p operators.

3.1.1 The diamond $\langle d \rangle$ operators

To define the diamond Hecke operators, take any $\alpha \in \Gamma_0(N)$, set $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$, and consider the weight- k double coset operator $[\Gamma_1 \alpha \Gamma_2]_k$. Since $\Gamma_1(N) \triangleleft \Gamma_0(N)$ this operator is case (2) from the example list above, taking each function $f \in \mathcal{M}_k[\Gamma_1(N)]$ to

$$f[\Gamma_1(N)\alpha\Gamma_1(N)]_k = f[\alpha]_k, \quad \alpha \in \Gamma_0(N)$$

again in $\mathcal{M}_k[\Gamma_1(N)]$. Consequently the group $\Gamma_0(N)$ acts on $\mathcal{M}_k[\Gamma_1(N)]$, and since its subgroup $\Gamma_1(N)$ acts trivially, we can see that the action is really the action of the quotient $\Gamma_0(N)/\Gamma_1(N) \cong (\mathbb{Z}/NZ)^\times$. With that motivation, we are ready to define the diamond operators.

Definition 3.1.1

Let $d \in \mathbb{Z}$ be an integer coprime to N. Let $\alpha = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z})$.

The **diamond operator** $\langle d \rangle$ is the operator on $M_k[\Gamma_1(N)]$. defined as the action of α determined by $d (\mathrm{mod} N)$ and denoted by

$$\langle d \rangle : \mathcal{M}_k[\Gamma_1(N)] \longrightarrow \mathcal{M}_k[\Gamma_1(N)]$$

given by

$$\langle d \rangle f = f[\alpha]_k \quad \text{for any } \alpha = \begin{bmatrix} a & b \\ c & \delta \end{bmatrix} \in \Gamma_0(N) \text{ with } \delta \equiv d (\mathrm{mod} N).$$

As discussed above since the action is really of the quotient, this makes the diamond operators well-defined and dependent only on the class of d modulo N.

Furthermore, it can be checked The operator $\langle d \rangle$ is a linear invertible map; thus, it makes sense to look at its eigenspaces and in turn, give a nice decomposition of $M_k[\Gamma_1(N)]$ into $\mathbb{C}-$Vector spaces. But first, we need some preliminary definitions.

Definition 3.1.2

A Dirichlet character modulo N is a group homomorphism

$$\chi : (\mathbb{Z}/N\mathbb{Z})^\times \to \mathbb{C}^\times.$$

Remark 3.1.3

It can be extended to a map $\chi : \mathbb{Z} \to \mathbb{C}$ as follows:

$$\chi(d) = \begin{cases} \chi(d \bmod N) & (d, N) = 1 \\ 0 & (d, N) \neq 1 \end{cases}$$

Note that since χ is a group homomorphism, and the identical zero function is totally multiplicative, the resulting function is totally multiplicative, *i. e.* it satisfies

$$\chi[d_1 d_2] = \chi[d_1]\chi[d_2] \quad \forall d_1, d_2 \in \mathbb{Z}.$$

Definition 3.1.4

The space of modular forms with character χ is

$$M_k\left[\Gamma_0(N),\chi\right]=\left\{f\in M_k\left[\Gamma_1(N)\right]:\quad f[\begin{bmatrix}a&b\\c&d\end{bmatrix}]_k=\chi(d)f,\begin{bmatrix}a&b\\c&d\end{bmatrix}\in\Gamma_0(N)\right\}.$$

Note that $M_k\left[\Gamma_0(N),\chi\right]$ can also be defined as

$$M_k\left[\Gamma_0(N),\chi\right]=\left\{f\in M_k\left[\Gamma_1(N)\right]\mid\langle d\rangle f=\chi(d)f,\quad d\in(\mathbb{Z}/N\mathbb{Z})^\times\right\}$$

Finally, we get the decomposition and state it in the following proposition.

Proposition 3.1.5

$$M_k\left[\Gamma_1(N)\right]=\bigoplus_{\chi \bmod N}M_k\left[\Gamma_0(N),\chi\right],$$

where the sum runs over the $\varphi(N)=\#(\mathbb{Z}/N\mathbb{Z})^\times$Dirichlet characters modulo N.

Proof:
When we select a basis for $M_k\left[\Gamma_1(N)\right]$, a representation ρ emerges as follows:

$$\rho:(\mathbb{Z}/N\mathbb{Z})^\times\to\mathrm{GL}_n(\mathbb{C}),\quad \rho(d)=\langle d\rangle,$$

Here, n is the dimension of $M_k\left[\Gamma_1(N)\right]$. Given that $(\mathbb{Z}/N\mathbb{Z})^\times$ is abelian, the representation ρ breaks down into a sum of irreducible representations, all of which must be one-dimensional. Consequently, we can choose a basis for $M_k\left[\Gamma_1(N)\right]$ such that:

$$\rho(d)=\mathrm{diag}\left[\chi_1(d),\ldots,\chi_n(d)\right]$$

This implies that $\langle d\rangle$ acts as the ith component as $\chi_i(d)$. To construct $M_k\left[\Gamma_0(N),\chi\right]$, we only need to gather the repeating χ values. □

3.1.2 The T_p operators

Definition 3.1.6
The second type of Hecke operator is also a weight- k double coset operator $\left[\Gamma_1\alpha\Gamma_2\right]_k$ where again $\Gamma_1=\Gamma_2=\Gamma_1(N)$, but now

$$\alpha=\begin{bmatrix}1&0\\0&p\end{bmatrix},\quad p\text{ prime.}$$

This operator is denoted T_p. Thus

$$T_p:\mathcal{M}_k\left[\Gamma_1(N)\right]\longrightarrow\mathcal{M}_k\left[\Gamma_1(N)\right],\quad p\text{ prime}$$

is given by

$$T_pf=f\left[\Gamma_1(N)\begin{bmatrix}1&0\\0&p\end{bmatrix}\Gamma_1(N)\right]_k$$

Remark 3.1.7

In order to describe the action of T_p more precisely, we need to understand the double coset $\Gamma_1(N)\begin{bmatrix}1&0\\0&p\end{bmatrix}\Gamma_1(N)$.

Note first that if $\gamma \in \Gamma_1(N)\begin{bmatrix}1&0\\0&p\end{bmatrix}\Gamma_1(N)$ then:

1. $\det\gamma = p$, and
2. $\gamma \equiv \begin{bmatrix}1&*\\0&p\end{bmatrix} (\mathrm{mod} N)$.

In fact, the converse is also true:

Proposition 3.1.8
We have that

$$\Gamma_1(N)\begin{bmatrix}1&0\\0&p\end{bmatrix}\Gamma_1(N) = \left\{\gamma \in M_2(\mathbb{Z}) \mid \det\gamma = p, \gamma \equiv \begin{bmatrix}1&0\\0&p\end{bmatrix} \quad (\mathrm{mod} N)\right\}$$

Proposition 3.1.9
Let $N \in \mathbb{Z}^+$, let $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$, and let $\alpha = \begin{bmatrix}1&0\\0&p\end{bmatrix}$ where p is prime. The operator $T_p = [\Gamma_1\alpha\Gamma_2]_k$ on $\mathcal{M}_k[\Gamma_1(N)]$ is given by

$$T_p f = \begin{cases} \sum_{j=0}^{p-1} f\left[\begin{bmatrix}1&j\\0&p\end{bmatrix}\right]_k & \text{if } p \mid N, \\ \sum_{j=0}^{p-1} f\left[\begin{bmatrix}1&j\\0&p\end{bmatrix}\right]_k + f\left[\begin{bmatrix}m&n\\N&p\end{bmatrix}\begin{bmatrix}p&0\\0&1\end{bmatrix}\right]_k & \text{if } p \nmid N, \text{ where } mp - nN = 1. \end{cases}$$

Proof:
Let's explore how the double coset operator functions by breaking down the coset decomposition of $\Gamma_3\backslash\Gamma_1(N)$. Here, Γ_3 can be understood through the formula:

$$\Gamma_3 = \begin{bmatrix}1&0\\0&p\end{bmatrix}^{-1}\Gamma_1(N)\begin{bmatrix}1&0\\0&p\end{bmatrix} \cap \Gamma_1(N).$$

We then introduce the set $\Gamma^0(p)$, which consists of matrices that are lower triangular when considered modulo p. This leads us to the realization that:

$$\Gamma_3 = \Gamma_1(N) \cap \Gamma^0(p).$$

Let us consider a series of matrices γ_j of the form $\begin{bmatrix}1&j\\0&1\end{bmatrix}$, where j ranges from 0 to $p-1$. These matrices are distinct when taken modulo $\Gamma_1(N) \cap \Gamma^0(p)$. For any matrix in $\Gamma_1(N)$ with the structure $\begin{bmatrix}a&b\\c&d\end{bmatrix}$, we can observe that:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & -j \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & -aj+b \\ c & -cj+d \end{bmatrix}.$$

When p does not divide a, the right side of this equation can be turned into an element of $\Gamma^0(p)$ for some value of j. If p is a factor of N, then it won't divide a, due to the determinant condition of the matrix. Adding in the matrix $\gamma_\infty = \begin{bmatrix} mp & n \\ N & 1 \end{bmatrix}$ from $\Gamma_1(N)$, we get:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \gamma_\infty^{-1} = \begin{bmatrix} * & -na+bmp \\ 0 & * \end{bmatrix}.$$

Given that p divides $-na+bmp$, the set $\{\gamma_j\} \cup \{\gamma_\infty\}$ forms a complete set of representatives. To obtain the representatives for the double coset, we simply multiply each γ_j by the fixed element $\alpha = \begin{bmatrix} 1 & 0 \\ 0 & p \end{bmatrix}$. This finishes the claim.

□

Let us now state the effect of the second kind of Hecke operator on Fourier coefficients of Modular forms.

Proposition 3.1.10
Let $f \in \mathcal{M}_k[\Gamma_1(N)]$ having a Fourier expansion

$$f(\tau) = \sum_{n=0}^{\infty} a_n(f) q^n, \quad q = e^{2\pi i \tau}$$

since since f has period 1 because $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \in \Gamma_1(N)$,
Then:
(a) Let $\mathbf{1}_N : (\mathbb{Z}/N\mathbb{Z})^* \longrightarrow \mathbb{C}^*$ be the trivial character modulo N. Then $T_p f$ has Fourier expansion

$$\begin{aligned}[T_p f](\tau) &= \sum_{n=0}^{\infty} a_{np}(f) q^n + \mathbf{1}_N(p) p^{k-1} \sum_{n=0}^{\infty} a_n(\langle p \rangle f) q^{np} \\ &= \sum_{n=0}^{\infty} \left[a_{np}(f) + \mathbf{1}_N(p) p^{k-1} a_{n/p}(\langle p \rangle f) \right] q^n\end{aligned}$$

That is,

$$a_n[T_p f] = a_{np}(f) + \mathbf{1}_N(p) p^{k-1} a_{n/p}(\langle p \rangle f) \quad \text{for } f \in \mathcal{M}_k[\Gamma_1(N)]$$

(Here $a_{n/p} = 0$ when $n/p \notin \mathbb{N}$, $\mathbf{1}_N(p) = 1$ when $p \nmid N$ and $\mathbf{1}_N(p) = 0$ when $p \mid N$.]
(b) Let $\chi : (\mathbb{Z}/N\mathbb{Z})^* \longrightarrow \mathbb{C}^*$ be a character. If $f \in \mathcal{M}_k(N, \chi)$ then also $T_p f \in \mathcal{M}_k(N, \chi)$, and now its Fourier expansion is

$$[T_p f](\tau) = \sum_{n=0}^{\infty} a_{np}(f) q^n + \chi(p) p^{k-1} \sum_{n=0}^{\infty} a_n(f) q^{np}$$
$$= \sum_{n=0}^{\infty} \left[a_{np}(f) + \chi(p) p^{k-1} a_{n/p}(f) \right] q^n$$

That is,

$$a_n [T_p f] = a_{np}(f) + \chi(p) p^{k-1} a_{n/p}(f) \quad \text{for } f \in \mathcal{M}_k(N, \chi)$$

Proof:
See, [DS05], proposition 5.2.2. □

Proposition 3.1.11
Let d and e be elements of $(\mathbb{Z}/N\mathbb{Z})^*$, and let p and q be prime. Then
(a) $\langle d \rangle T_p = T_p \langle d \rangle$
(b) $\langle d \rangle \langle e \rangle = \langle e \rangle \langle d \rangle = \langle de \rangle$
(c) $T_p T_q = T_q T_p$
Proof:
To begin, let's establish equations (b) and (c) under the assumption of equation (a). Observe that equation (a) implies that T_p keeps the spaces $M_k(\Gamma_0(N), \chi)$ invariant. Therefore, it suffices to prove equations (2) and (3) for modular forms $f \in M_k(\Gamma_0(N), \chi)$. This essentially confirms equation (2). For equation (3), we turn to the q-expansions of f, denoted as $f = \sum a_n q^n$.
The coefficients of $T_p f$ are then represented as:

$$a_n(T_p f) = a_{pn}(f) + \chi(p) p^{k-1} a_{n/p}(f)$$

Further, the coefficients of $T_p T_q f$ can be expressed as:

$$a_n(T_p T_q f) = a_{pn}(T_q f) + \chi(p) p^{k-1} a_{n/p}(T_q f)$$
$$= a_{pqn}(f) + \chi(q) q^{k-1} a_{pn/q}(f) + \chi(p) p^{k-1} (a_{nq/p}(f) + \chi(q) q^{k-1} a_{n/(pq)}(f))$$

This equation is symmetric in p and q.
To establish (a) we consider that Γ is a normal subgroup of $\Gamma_0(N)$. Then, it follows that

$$\Gamma \gamma \Gamma = \Gamma \gamma$$

Consequently, $\langle d \rangle f = f|_k \gamma$. Our aim is to demonstrate $\langle d \rangle^{-1} T_p \langle d \rangle = T_p$. Let's represent the double coset corresponding to T_p in terms of its orbit decomposition as $\Gamma \alpha \Gamma = \bigcup_j \Gamma \beta_j$. The goal then becomes to validate

$$\Gamma \alpha \Gamma = \bigcup_j \Gamma(\gamma \beta_j \gamma^{-1})$$

It can be observed that

$$\begin{aligned}\bigcup_j \Gamma(\gamma\beta_j\gamma^{-1}) &= \gamma(\bigcup_j \Gamma\beta_j)\gamma^{-1} \\ &= \gamma(\Gamma\alpha\Gamma)\gamma^{-1} \\ &= \Gamma(\gamma\alpha\gamma^{-1})\Gamma\end{aligned}$$

Upon examining this, it is clear that

$$\Gamma\alpha\Gamma = \Gamma(\gamma\alpha\gamma^{-1})\Gamma$$

This finishes the proof. □

3.2 The Petersson inner product and Adjoint operators

3.2.1 The Petersson inner product

Cusp forms play a significant role in the study of modular forms, and thus, studying the space of cusp forms $\mathcal{S}_k(\Gamma_1(N))$ is a natural discipline of work in Mathematics. One way to do so is to define an inner product on this space and make this space into an inner product space. We will define this inner product as an integral over the extended space of the fundamental domain which can be obtained by adding ∞ to $\mathcal{F}$ described in previous sections. We will also state some important results, namely that this inner product defined as an integral does not converge over the bigger space namely the space of Modular forms of weight k, $\mathcal{M}_k(\Gamma_1(N))$ and consequently showing that inner product structure is restricted to the cusp forms. As we just want to give a formal introduction to *The Petersson inner product*, we will mainly state the key results without proof.

Let $V \subseteq \mathbb{C}$. A 2-form on V is an expression of the form $\omega = f(z,\bar{z})dzd\bar{z}$. Note that

$$dzd\bar{z} = (dx + idy)\,(dx - idy) = -2idx\,dy.$$

Consider in particular the 2-form $\frac{dzd\bar{z}}{\mathrm{Im}(z)^2}$ and consider for $\alpha \in \mathrm{GL}_2^+(\mathbb{R})$, the change $z \mapsto \alpha z$. Then:

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

and also

$$d(\alpha z) = \frac{\det\alpha}{(cz+d)^2}dz, \quad \overline{d(\alpha z)} = \frac{\det\alpha}{(cz+d)^2}d\bar{z}$$

This gives that:

$$d(\alpha z)d(\overline{\alpha z}) = \frac{(\det\alpha)^2}{|cz+d|^4}dzd\bar{z}$$

Also, using the above dis, we can rewrite this as,

$$d(\alpha z)d(\overline{\alpha z}) = \frac{(\det\alpha)^2}{|cz+d|^4}dzd\bar{z} = \frac{dzd\bar{z}}{(\mathrm{Im}(\alpha z))^2}.$$

Thus, the 2-form $\frac{dzd\bar{z}}{\mathrm{Im}(z)^2}$ is invariant under the change $z \mapsto \alpha z$ or in another words invariant under the automorphism group $\mathrm{GL}_2^+(\mathbb{R})$. Also, observe that

$$\frac{-1}{2i}\frac{dzd\bar{z}}{\mathrm{Im}(z)^2} = \frac{dxdy}{y^2}.$$

Thus, we can of course, restrict to $V = \mathcal{H} \subset \mathbb{C}$, which gives us following definition.

Definition 3.2.1
The *hyperbolic measure* on the upper half plane is defined by

$$d\mu(\tau) = \frac{dxdy}{y^2}, \quad \tau = x + iy \in \mathcal{H}.$$

The hyperbolic measure defined above is invariant under the automorphism group $\mathrm{GL}_2{}^+(\mathbb{R})$. Thus in particular, the hyperbolic measure is $\mathrm{SL}_2(\mathbb{Z})$-invariant. Also, recall from Chapter 2 that a fundamental domain of $\mathcal{H}^*$ under the action of $\mathrm{SL}_2(\mathbb{Z})$ is

$$\mathcal{D}^* = \{\tau \in \mathcal{H} : \mathrm{Re}(\tau) \leq 1/2, |\tau| \geq 1\} \cup \{\infty\}$$

That is, every point $\tau' \in \mathcal{H}$ transforms under $\mathrm{SL}_2(\mathbb{Z})$ into the connected set $\mathcal{D}$, and barring certain cases on the boundary of $\mathcal{D}$ the transformation is unique; and every point $s \in \mathbb{Q} \cup \{\infty\}$ transforms under $\mathrm{SL}_2(\mathbb{Z})$ to ∞. Thus it suffices to integrate over the extended fundamental domain $\mathcal{D}$.

Now,Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup and let $\{\alpha_j\} \subset \mathrm{SL}_2(\mathbb{Z})$ represent the coset space $\{\pm I\}\Gamma\backslash\mathrm{SL}_2(\mathbb{Z})$, meaning that the union

$$\mathrm{SL}_2(\mathbb{Z}) = \bigcup_j \{\pm I\}\Gamma\alpha_j$$

is disjoint. If the function φ is Γ-invariant then the sum $\sum_j \int_{\mathcal{D}^*} \varphi\left[\alpha_j(\tau)\right] d\mu(\tau)$ is independent of the choice of coset representatives α_j. This is because $d\mu$ is $\mathrm{SL}_2(\mathbb{Z})$ invariant the sum is $\int_{\bigcup \alpha_j[\mathcal{D}^*]} \varphi(\tau) d\mu(\tau)$ and $\bigcup \alpha_j\left[\mathcal{D}^*\right]$ represents the modular curve $X(\Gamma)$ up to some boundary identification defined as at the end of section 2.3 of [DS05]. This quantity is naturally denoted $\int_{X(\Gamma)}$. Thus we have made the definition

$$\int_{X(\Gamma)} \varphi(\tau) d\mu(\tau) = \int_{\bigcup \alpha_j[\mathcal{D}^*]} \varphi(\tau) d\mu(\tau) = \sum_j \int_{\mathcal{D}^*} \varphi\left[\alpha_j(\tau)\right] d\mu(\tau).$$

In particular, setting $\varphi = 1$, we get the volume of $X(\Gamma)$ is

$$V_\Gamma = \int_{X(\Gamma)} d\mu(\tau)$$

Note that, the volume and index of a congruence subgroup have the following relation:

$$V_\Gamma = \left[\mathrm{SL}_2(\mathbb{Z}) : \{\pm I\}\Gamma\right] V_{\mathrm{SL}_2(\mathbb{Z})}.$$

Now, let us define: **Petersson inner product**.

Definition 3.2.2
Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup. The Petersson inner product,

$$\langle , \rangle_\Gamma : \mathcal{S}_k(\Gamma) \times \mathcal{S}_k(\Gamma) \longrightarrow \mathbb{C}$$

is given by

$$\langle f, g \rangle_\Gamma = \frac{1}{V_\Gamma} \int_{X(\Gamma)} f(\tau)\overline{g(\tau)}(\mathrm{Im}(\tau))^k d\mu(\tau)$$

Proposition 3.2.3
The Petersson inner product is well-defined and convergent.

Proof:
See, ([DS05]Page,182,183) □

Proposition 3.2.4
The Petersson inner product is a positive-definite hermitian product on the $\mathbb{C}$-vector space $S_k(\Gamma)$. That is:

1. $\langle a_1 f_1 + a_2 f_2, g \rangle_\Gamma = a_1 \langle f_1, g \rangle_\Gamma + a_2 \langle f_2, g \rangle_\Gamma$.
2. $\langle g, f \rangle_\Gamma = \overline{\langle f, g \rangle_\Gamma}$
3. $\langle f, f \rangle \geq 0$, with equality if and only if $f = 0$.

Proof:
See, [Mas15], proposition 4.4.5. □

3.2.2 Adjoint operators

The aim of this subsection is to calculate the adjoint operators of the hecke operators. We start by recalling the definition of an adjoint operator.

Definition 3.2.5
If $\langle \cdot, \cdot \rangle$ is an hermitian product on a $\mathbb{C}$-vector space V and $T : V \to V$ is a linear operator, the adjoint of T is defined as the operator T^* which satisfies:

$$\langle Tf, g \rangle = \langle f, T^* g \rangle$$

If $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ is a congruence subgroup and $\mathrm{SL}_2(\mathbb{Z}) = \bigcup_j \{\pm I\}\Gamma\alpha_j$ and $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$ then the map $\mathcal{H} \longrightarrow \mathcal{H}$ given by $\tau \mapsto \alpha(\tau)$ induces a bijection $\alpha^{-1}\Gamma\alpha\backslash\mathcal{H}^* \longrightarrow X(\Gamma)$. Thus the union $\bigcup_j \alpha^{-1}\alpha_j [\mathcal{D}^*]$ represents the quotient space $\alpha^{-1}\Gamma\alpha\backslash\mathcal{H}^*$ up to some boundary identification. For for continuous, bounded, $\alpha^{-1}\Gamma\alpha$-invariant functions $\varphi : \mathcal{H} \longrightarrow \mathbb{C}$ define

$$\int_{\alpha^{-1}\Gamma\alpha\backslash\mathcal{H}^*} \varphi(\tau) d\mu(\tau) = \sum_j \int_{\mathcal{D}^*} \varphi\left[\alpha^{-1}\alpha_j(\tau)\right] d\mu(\tau).$$

To proceed, we will need the following technical result.

Lemma 3.2.6
Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, and let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$

(a) If $\varphi : \mathcal{H} \longrightarrow \mathbb{C}$ is continuous, bounded, and Γ-invariant, then

$$\int_{\alpha^{-1}\Gamma\alpha\backslash\mathcal{H}^*} \varphi(\alpha(\tau))d\mu(\tau) = \int_{X(\Gamma)} \varphi(\tau)d\mu(\tau).$$

(b) If $\alpha^{-1}\Gamma\alpha \subset \mathrm{SL}_2(\mathbb{Z})$ then $V_{\alpha^{-1}\Gamma\alpha} = V_\Gamma$ and $\left[\mathrm{SL}_2(\mathbb{Z}) : \alpha^{-1}\Gamma\alpha\right] = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma]$.
(c) There exist $\beta_1, \ldots, \beta_n \in \mathrm{GL}_2^+(\mathbb{Q})$, where $n = \left[\Gamma : \alpha^{-1}\Gamma\alpha \cap \Gamma\right] = \left[\Gamma : \alpha\Gamma\alpha^{-1} \cap \Gamma\right]$, such that

$$\Gamma\alpha\Gamma = \bigcup \Gamma\beta_j = \bigcup \beta_j\Gamma$$

with both unions disjoint.

Proof:
[DS05], proposition 5.5.1. □

Next proposition acts as a tool to compute adjoints.

Proposition 3.2.7
Let $\Gamma \subset \mathrm{SL}_2(\mathbb{Z})$ be a congruence subgroup, and let $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$. Set $\alpha' = \det(\alpha)\alpha^{-1}$. Then
(a) If $\alpha^{-1}\Gamma\alpha \subset \mathrm{SL}_2(\mathbb{Z})$ then for all $f \in \mathcal{S}_k(\Gamma)$ and $g \in \mathcal{S}_k\left[\alpha^{-1}\Gamma\alpha\right]$,

$$\langle f[\alpha]_k, g\rangle_{\alpha^{-1}\Gamma\alpha} = \left\langle f, g\left[\alpha'\right]_k\right\rangle_\Gamma.$$

(b) For all $f, g \in \mathcal{S}_k(\Gamma)$,

$$\langle f[\Gamma\alpha\Gamma]_k, g\rangle = \left\langle f, g\left[\Gamma\alpha'\Gamma\right]_k\right\rangle$$

In particular, if $\alpha^{-1}\Gamma\alpha = \Gamma$ then $[\alpha]_k^* = [\alpha']_k$, and in any case $[\Gamma\alpha\Gamma]_k^* = [\Gamma\alpha'\Gamma]_k$.

Proof:
[DS05], proposition 5.5.2. □

Before, we state an important result and finally bear the fruit of our efforts, we recall normal operators.

Definition 3.2.8
A linear operator T is normal if it commutes with its adjoint:

$$TT^* = T^*T$$

Theorem 3.2.9
In the inner product space $\mathcal{S}_k\left[\Gamma_1(N)\right]$, the Hecke operators $\langle p\rangle$ and T_p for $p \nmid N$ have adjoints

$$\langle p\rangle^* = \langle p\rangle^{-1} \quad \text{and} \quad T_p^* = \langle p\rangle^{-1}T_p$$

Thus the Hecke operators $\langle n\rangle$ and T_n for n relatively prime to N are normal.

Proof:
See, [Mas15], Theorem 4.4.8 and corollary 4.4.9. □

To end the discussion of Petersson inner product and Adjoints of the operators, we state an important result that follows directly from the Spectral theorem and Theorem 3.2.9.
Recall that the spectral theorem states that if T is a normal operator on a finite dimensional $\mathbb{C}$-vector

space, Then T has an orthogonal basis of eigenvectors.
Applying this theorem multiple times we deduce that if a $\mathbb{C}$-vector space has a family of normal, pairwise commuting operators then it has a basis of simultaneous eigenvectors. Particularizing to our situation, we get the following result.

Theorem 3.2.10
The space $\mathcal{S}_k[\Gamma_1(N)]$ has an orthogonal basis of simultaneous eigenforms for the Hecke operators $\{\langle n\rangle, T_n : (n,N)=1\}$.

3.3 Eigenforms, newforms, oldforms, Atkin-Lehner Theory

So far, the theory that has been developed so far was mainly about one level N. It is sometimes important as we will see in the proof of Fermat's last theorem to results move between levels, that is by taking forms from lower levels $M \mid N$ up to level N, mostly with $M = Np^{-1}$ where p is some prime factor of N.
The following proposition gives one obvious way to move between levels.

Proposition 3.3.1
$M \mid N$ then $\mathcal{S}_k[\Gamma_1(M)] \subset \mathcal{S}_k[\Gamma_1(N)]$.

Proof:
This simply follows from $\Gamma_1(N) \subseteq \Gamma_1(M)$ and that cusps of $\Gamma_1(N)$ are contained in $\Gamma_1(M)$. □

Another way to embed $\mathcal{S}_k[\Gamma_1(M)]$ into $\mathcal{S}_k[\Gamma_1(N)]$ is by composing with the multiply-by-d map where d is any factor of N/M. For any such d, let

$$\alpha_d = \begin{bmatrix} d & 0 \\ 0 & 1 \end{bmatrix}$$

so that $[f[\alpha_d]_k](\tau) = d^{k-1} f(d\tau)$ for $f : \mathcal{H} \longrightarrow$ C. By Exercise 1.2.11, in [DS05] the injective linear map $[\alpha_d]_k$ takes $\mathcal{S}_k[\Gamma_1(M)]$ to $\mathcal{S}_k[\Gamma_1(N)]$, lifting the level from M to N. The weight k- operator is defined up to a scalar multiple(composition with the multiply-by-d map).
To normalize the scalar to 1 , we define ι_d

$$\iota_d = d^{1-k}[\alpha_d]_k : \mathcal{S}_k[\Gamma_1(M)] \longrightarrow \mathcal{S}_k[\Gamma_1(N)], \quad [\iota_d f](\tau) = f(d\tau),$$

acting on Fourier expansions as

$$\iota_d : \sum_{n=1}^{\infty} a_n q^n \mapsto \sum_{n=1}^{\infty} a_n q^{dn}, \quad \text{where } q = e^{2\pi i\tau}$$

This shows that if $f \in \mathcal{S}_k[\Gamma_1(N)]$ takes the form $f = \sum_{p|N} \iota_p f_p$ with each $f_p \in \mathcal{S}_k[\Gamma_1(N/p)]$. Furthermore, if the Fourier expansion of f is $f(\tau) = \sum a_n(f) q^n$, then $a_n(f) = 0$ for all n such that $(n,N) = 1$. The main lemma in the theory of newforms is that the converse holds as well. This is due to the celebrated theorem of Atkin and Lehner.

Theorem 3.3.2 (Atkin and Lehner)
If $f \in \mathcal{S}_k[\Gamma_1(N)]$ has a Fourier expansion $f(\tau) = \sum a_n(f) q^n$ with $a_n(f) = 0$ whenever $(n,N) = 1$, then f takes the form $f = \sum_{p|N} \iota_p f_p$ with each $f_p \in \mathcal{S}_k[\Gamma_1(N/p)]$.

Proof:
[DS05], section 5.7 and see the paper by Atkin and Lehner [**?**]. □

Summarising the observations it is natural to distinguish part of $\mathcal{S}_k\left[\Gamma_1(N)\right]$ coming from lower levels. This gives us the following definition.

Definition 3.3.3
For each divisor d of N, let i_d be the map

$$i_d : \left[\mathcal{S}_k\left[\Gamma_1\left[Nd^{-1}\right]\right]\right]^2 \longrightarrow \mathcal{S}_k\left[\Gamma_1(N)\right]$$

given by

$$(f, g) \mapsto f + g\left[\alpha_d\right]_k$$

The subspace of oldforms at level N is

$$\mathcal{S}_k\left[\Gamma_1(N)\right]^{\text{old}} = \sum_{\substack{p|N \\ \text{prime}}} i_p\left[\left[\mathcal{S}_k\left[\Gamma_1\left[Np^{-1}\right]\right]\right]^2\right]$$

and the subspace of newforms at level N is the orthogonal complement with respect to the Petersson inner product,

$$\mathcal{S}_k\left[\Gamma_1(N)\right]^{\text{new}} = \left[\mathcal{S}_k\left[\Gamma_1(N)\right]^{\text{old}}\right]^{\perp}.$$

Next, we state an important result.

Theorem 3.3.4
The subspaces $\mathcal{S}_k\left[\Gamma_1(N)\right]^{\text{old}}$ and $\mathcal{S}_k\left[\Gamma_1(N)\right]^{\text{new}}$ are stable under the Hecke operators T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}^+$.

Proof:
See, [DS05], Proposition 5.6.2 or [Mas15], 4.5.2. □

Let $M \mid N$ and let $d \mid (N/M), d > 1$. Thus $\Gamma_1(M) \supset \Gamma_1(N)$.

Definition 3.3.5
A nonzero modular form $f \in \mathcal{M}_k\left[\Gamma_1(N)\right]$ that is an eigenform for the Hecke operators T_n and $\langle n\rangle$ for all $n \in \mathbf{Z}^+$is a Hecke eigenform or simply an eigenform. The eigenform $f(\tau) = \sum_{n=0}^{\infty} a_n(f)q^n$ is normalized when $a_1(f) = 1$.

Note: A newform is essentially a normalized eigenform in $\mathcal{S}_k\left[\Gamma_1(N)\right]^{\text{new}}$
To provide an alternate proof for the properties of eigenforms in modular forms, using a structured and professional yet simpler mathematical language, we can proceed as follows:

Theorem 3.3.6
For an eigenform f in $M_k(\Gamma_1(N))$, with $T_n f = \lambda_n f$ for all n, the coefficients of the q-expansion of f at the cusp ∞ are the eigenvalues of the Hecke operators on f.

Proof:
Given f as an eigenform, we have for each n:

$$a_n(f) = a_1(T_n f) = \lambda_n a_1(f).$$

If $a_1(f) = 0$, then $a_n(f) = 0$ for all n, leading to $f = 0$. Thus, a non-constant, non-zero eigenform must satisfy $a_1(f) \neq 0$. Normalizing f results in $a_1(f) = 1$, and consequently, $a_n(f) = \lambda_n$. Thus, the eigenvalues λ_n of the Hecke operators on f are precisely the coefficients $a_n(f)$ of the q-expansion. □

Proposition 3.3.7
A modular form f in $M_k(\Gamma_0(N), \chi)$ with q-expansion $\sum_{n=0}^{\infty} a_n(f)q^n$ is a normalized eigenform if and only if it satisfies the following conditions:
1. $a_1(f) = 1$,
2. For all m, n with $(m, n) = 1$, $a_{mn}(f) = a_m(f)a_n(f)$,
3. For all primes p and $r \geq 2$, $a_{p^r}(f) = a_p(f)a_{p^{r-1}}(f) - p^{k-1}\chi(p)a_{p^{r-2}}(f)$.

Proof:
The forward implication follows from 3.3.6 and the definition of a normalized eigenform. For the reverse implication, assume f satisfies conditions 1, 2, and 3. We need to show f is an eigenform.
For any prime p and any $m \geq 1$, if $p \nmid m$, from the formula for T_m on q-expansions, $a_m(T_p f) = a_{pm}(f)$, which by condition 2 is $a_p(f)a_m(f)$. If $p \mid m$, write $m = p^r m'$ with $p \nmid m'$. Then $a_m(T_p f) = a_{p^{r+1}m'}(f) + \chi(p)p^{k-1}a_{p^{r-1}m'}(f)$. Using conditions 2 and 3, this can be expressed as $a_p(f)a_m(f)$. □

The space $S_k(\mathrm{SL}_2(\mathbb{Z}))$ has a basis comprised entirely of eigenforms for these Hecke operators. After normalising we have, $T_n f = a_n(f)f$ for all n. What is particularly noteworthy in this context is the concept of 'multiplicity one', which refers to the fact that each distinct system of eigenvalues $\{a_n(f)\}_{n\geq 1}$ uniquely corresponds to a specific eigenform f. This property essentially implies that $S_k(\mathrm{SL}_2(\mathbb{Z}))$ can be decomposed into a direct sum of one-dimensional eigenspaces, where each eigenspace corresponds to a unique eigenform.

For instance, both $\Delta(z)$ and $\Delta(2z)$ are cusp forms in $S_{12}(\Gamma_1(2))$, where

$$\Delta = \sum_{n\geq 1} \tau(n)q^n$$

Observe that, $T_p\Delta = \tau(p)\Delta$ for all p and,

$$T_p(\Delta(2z)) = \tau(p)\Delta(2z), \quad p \neq 2$$

Therefore $\Delta(z)$ and $\Delta(2z)$ have, when considered in $S_{12}(\Gamma_1(2))$, the same "system of eigenvalues" $\{\tau(n)\}_{(n,2)=1}$.
Therefore $S_{12}(\Gamma_1(2))$ does not satisfy multiplicity one. The following theorem is known as strong multiplicity one.

Theorem 3.3.8 (Strong multiplicity one)
Consider the space $S_k[\Gamma_1(N)]^{\text{new}}$ for $N \geq 1$.

1. The space $S_k[\Gamma_1(N)]^{\text{new}}$ has a basis of newforms.
2. If $f \in S_k[\Gamma_1(N)]^{\text{new}}$ is an eigenvector for $\{T_q\}_{q\nmid N}$ then f is a scalar multiple of a newform (hence

an eigenvector for all the Hecke operators.

3. If $f \in S_k\left[\Gamma_1(N)\right]^{\text{new}}$ and $g \in S_k\left[\Gamma_1(M)\right]^{\text{new}}$ are both newforms satisfying $a_q(f) = a_q(g)$ for all but finitely many primes q, then $N = M$ and $f = g$.

Proof:
See, [DS05] or [**?**]. □

A consequence of strong multiplicity one is the following result.

Theorem 3.3.9
The set

$$\mathcal{B}_k(N) = \{f(n\tau) : f \text{ is a newform of level } M \text{ and } nM \mid N\}$$

is a basis of $\mathcal{S}_k\left(\Gamma_1(N)\right)$.

Proof:
See, [DS05], 5.3.8. □

Remark 3.3.10
1. If f is a newform, then there is a Dirichlet character χ such that $f \in S_k\left(\Gamma_0(N), \chi\right)$. 2. If $\{\lambda_n\}_{(n,N)=1}$ is a system of eigenvalues for the T_n such that $(n, N) = 1$, then there exists unique newform $f \in S_k\left(\Gamma_1(M)\right)^{\text{new}}$ for some $M \mid N$, such that $T_n f = \lambda_n f$ for all n satisfying $(n, N) = 1$.

Finally, we see that the new subspaces give a complete description of $S_k\left[\Gamma_1(N)\right]$ and $S_k\left[\Gamma_0(N)\right]$.

Theorem 3.3.11
There are direct sum decompositions

$$S_k\left[\Gamma_1(N)\right] = \bigoplus_{M|N} \bigoplus_{dM|N} \alpha_d\left[S_k\left[\Gamma_1(M)\right]^{\text{new}}\right]$$

and

$$S_k\left[\Gamma_0(N)\right] = \bigoplus_{M|N} \bigoplus_{dM|N} \alpha_k\left[S_k\left[\Gamma_0(M)\right]^{\text{new}}\right]$$

Proof:
Decompose, the space of cusp forms into simultaenous eigenspaces. Focusing on each component, and each newform in that component, we get that by 3.3.10, 2nd part, that f comes from a unique newform of some level dividing N. This finishes the claim. □

4 Theory of curves and Modular curves

A critical perspective when looking at Modular curves is to see that it is a Riemann surface. We will briefly introduce what they are, and then we will further continue our journey by delving briefly into the fundamentals of algebraic curves, Riemann-Roch theory, and divisors, laying the groundwork for comprehending the concept of the Jacobian. Subsequently, having laid these foundations, we will introduce the Jacobian variety in the next chapter as a crucial geometric object associated with a given curve. We will closely follow the book [DS05] and Forster's book on Riemann surfaces [For81]. We will sometimes closely follow proofs from either of these sources for the sake of completeness. Due to the extreme complexity of the overall topic and to demonstrate all of this within 6 months, it was inevitable to closely follow some of the proofs but written in my own words. At times, the ideas are inevitably not original but closely followed. Having said that, I must add that in many places, I have expanded on my own, giving more details.

4.1 Riemann surfaces

Definition 4.1.1 (Manifold)
An n-dimensional real **Manifold** is a Hausdorff topological space such that every point $a \in X$ has an open neighborhood which is homeomorphic to an open subset of $\mathbb{R}^n$.

Definition 4.1.2
Let X be a two-dimensional manifold. A complex chart on X is a homeomorphism $\varphi : U \to V$ of an open subset $U \subset X$ onto an open subset $V \subset \mathbb{C}$.

Definition 4.1.3
We say two complex charts $\varphi_i : U_i \to V_i, i = 1, 2$ are said to be **holomorphically compatible** if the map

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1 (U_1 \cap U_2) \to \varphi_2 (U_1 \cap U_2)$$

is biholomorphic

Definition 4.1.4
A **complex atlas** on X is a collection $\mathfrak{I} = \{\varphi_i : U_i \to V_i, i \in I\}$ of charts which are holomorphically compatible and which cover X, i.e., $\bigcup_{i \in I} U_i = X$.

Definition 4.1.5
We say,two complex atlases $\mathfrak{J}$ and $\mathfrak{I}$ on X are called analytically equivalent if every chart of $\mathfrak{I}$ is holomorphically compatible with every chart of $\mathfrak{J}$.

Definition 4.1.6
By a complex structure on a two-dimensional manifold X we mean an equivalence class of analytically equivalent atlases on X.

Thus a complex structure on X can be given by the choice of a complex atlas.

Definition 4.1.7
A **Riemann surface** is a pair (X, Σ), where X is a connected two-dimensional manifold and Σ is a complex structure on X.

Example 4.1.8

1. The complex Plane $\mathbb{C}$. Its complex structure is defined by the atlas whose only chart is the identity $i : \operatorname{map} \mathbb{C} \to \mathbb{C}$.

2. The Riemann sphere $\mathbb{P}^1$. Let $\mathbb{P}^1 := \mathbb{C} \cup \{\infty\}$, where ∞. Introduce the following topology on $\mathbb{P}^1$. The open sets are the usual open sets $U \subset \mathbb{C}$ together with sets of the form $V \cup \{\infty\}$, where $V \subset \mathbb{C}$ is the complement of a compact set $K \subset \mathbb{C}$. With this topology $\mathbb{P}^1$ is a compact Hausdorff topological space, homeomorphic to the 2-sphere S^2. Set

$$U_1 := \mathbb{P}^1 \mid \{\infty\} = \mathbb{C}$$
$$U_2 := \mathbb{P}^1 \mid \{0\} = \mathbb{C}^* \cup \{\infty\}.$$

   Define maps $\varphi_i : U_i \to \mathbb{C}, i = 1, 2$, as follows. φ_1 is the identity map and

$$\varphi_2(z) := \begin{cases} 1/z & \text{for } z \in \mathbb{C}^* \\ 0 & \text{for } z = \infty \end{cases}$$

   These mappings are homeomorphisms, implying that $\mathbb{P}^1$ is a two-dimensional manifold. As both U_1 and U_2 are connected and share a non-empty intersection, $\mathbb{P}^1$ is connected as well. To define the complex structure on $\mathbb{P}^1$, we utilize the atlas comprising the charts $\varphi_i : U_i \to \mathbb{C}$, where $i = 1, 2$. The crucial task is to demonstrate that these two charts are holomorphically compatible. But $\varphi_1(U_1 \cap U_2) = \varphi_2(U_1 \cap U_2) = \mathbb{C}^*$, and ,

$$\varphi_2 \circ \varphi_1^{-1} : \mathbb{C}^* \to \mathbb{C}^*, \quad z \mapsto 1/z,$$

   is biholomorphic, thus we are done.

3. Another important example of a Riemann surface is that of a torus. Let Λ be a lattice and let $T = \mathbb{C}/\Lambda$ be the associated complex torus. Let $\pi : \mathbb{C} \to \mathbb{C}/\Lambda$ be the canonical projection. Introduce the following topology (the quotient topology) on $\mathbb{C}/\Lambda$. A subset $U \subset \mathbb{C}/\Lambda$ is open precisely if $\pi^{-1}(U) \subset \mathbb{C}$ is open. With this topology $\mathbb{C}/\Lambda$ is a Hausdorff topological space and the quotient map $\pi : \mathbb{C} \to \mathbb{C}/\Lambda$ is continuous. Since $\mathbb{C}$ is connected, $\mathbb{C}/\Lambda$ is also connected. As well $\mathbb{C}/\Lambda$ is compact, since it is covered by the image under π of the compact parallelogram.

$$P := \{\lambda\omega_1 + \mu\omega_2 : \lambda, \mu \in [0, 1]\}.$$

   We can easily see that the map π is open, i.e., the image of every open set $V \subset \mathbb{C}$ is open. We can see this as follows:

   We will show that show that $V' := \pi^{-1}(\pi(V))$ is open. But

$$V' = \bigcup_{\omega \in \Lambda} (\omega + V)$$

Since every set $\omega + V$ is open, so is V'. The complex structure on $\mathbb{C}/\Lambda$ is defined in the following way. Let $V \subset \mathbb{C}$ be an open set such that no two points in V are equivalent under Λ. Then $U := \pi(V)$ is open and $\pi \mid V \to U$ is a homeomorphism. Its inverse $\varphi : U \to V$ is a complex chart on $\mathbb{C}/\Lambda$.Let $\mathfrak{U}$ be the set of all charts obtained in this fashion. We have to show that any two charts $\varphi_i : U_i \to V_i, i = 1, 2$, belonging to $\mathfrak{U}$ are holomorphically compatible. Consider the map

$$\psi := \varphi_2 \varphi_1^{-1} : \varphi_1 (U_1 \cap U_2) \to \varphi_2 (U_1 \cap U_2) .$$

For every $z \in \varphi_1 (U_1 \cap U_2)$ one has $\pi(\psi(z)) = \varphi_1^{-1}(z) = \pi(z)$ and thus $\psi(z) - z \in \Gamma$. Since Γ is discrete and ψ is continuous, this implies that $\psi(z) - z$ is constant on every connected component of $\varphi_1 (U_1 \cap U_2)$. Thus ψ is holomorphic. Similarly ψ^{-1} is also holomorphic.

Lastly, let $\mathbb{C}/\Lambda$ have the complex structure defined by the complex atlas $\mathfrak{U}$. Thus we get that a torus is a Riemann surface.

Modular curve as a Riemann surface

Let us quickly recall that for any congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ the corresponding modular curve has been defined as the quotient space $\Gamma\backslash\mathcal{H}$, the set of orbits.

$$Y(\Gamma) = \{\Gamma\tau : \tau \in \mathcal{H}\}.$$

Our goal for further discussion would be to discuss that $Y(\Gamma)$ can be made into a Riemann surface that can be compactified. The resulting compact Riemann surface is denoted $X(\Gamma)$.
Let us define a topology on the Modular curve $Y(\Gamma)$.
The upper half-plane $\mathcal{H}$ inherits the Euclidean topology as a subspace of $\mathbb{R}^2$.
The natural surjection

$$\pi : \mathcal{H} \longrightarrow Y(\Gamma), \quad \pi(\tau) = \Gamma\tau$$

gives $Y(\Gamma)$ the quotient topology, meaning a subset of $Y(\Gamma)$ is open if its inverse image under π in $\mathcal{H}$ is open. Note that this just means tat π an open mapping, and the following equivalence holds.

Lemma 4.1.9

$$\pi (U_1) \cap \pi (U_2) = \emptyset \text{ in } Y(\Gamma) \quad \Longleftrightarrow \quad \Gamma (U_1) \cap U_2 = \emptyset \text{ in } \mathcal{H}$$

Proof:
The map $\pi : \mathcal{H} \longrightarrow Y(\Gamma)$ sends each point τ in $\mathcal{H}$ to its orbit $\Gamma\tau$ in $Y(\Gamma)$. This map is surjective and respects the group action, i.e., points in the same orbit under Γ are mapped to the same point in $Y(\Gamma)$. If $\pi(U_1) \cap \pi(U_2) = \emptyset$ in $Y(\Gamma)$, we have that $\{\Gamma\tau : \tau \in U_1\} \cap \{\Gamma\tau : \tau \in U_2\} = \emptyset$. Now assume for a contradiction that element $x \in \Gamma (U_1) \cap U_2$. Then, there exists $u \in U_1$ such that $\Gamma.u = x$. But then, group action just permutes the orbit, giving us that $\Gamma u = \Gamma x$. This gives us the contradiction to the fact that, $\{\Gamma\tau : \tau \in U_1\} \cap \{\Gamma\tau : \tau \in U_2\} = \emptyset$.
Conversely, we have $\Gamma(U_1) \cap U_2 = \emptyset$ in $\mathcal{H}$. Suppose we have some x in the intersection of $\pi(U_1)$ and $\pi(U_2)$, this means there exist $u \in U_1$ and $u' \in U_2$ such that $\pi(u) = \Gamma u$ and $\pi(u') = \Gamma u'$ are both equal to x in $Y(\Gamma)$. In other words, u and u' belong to the same orbit under Γ, which means there exists an

element $\gamma \in \Gamma$ such that $\gamma u = u'$. However this contradicts the assumption that $\Gamma(U_1) \cap U_2 = \emptyset$ in $\mathcal{H}$.

□

Since $\mathcal{H}$ is connected and π is continuous, the quotient $Y(\Gamma)$ is also connected.
We now show that the Modular curve $Y(\Gamma)$ is Hausdorff. But before that, we state an important lemma, which would be key to proving this. It will be furthermore useful when defining coordinate charts on Modular curve.$Y(\Gamma)$.

Lemma 4.1.10
The action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathcal{H}$ is properly discontinuous, i.e. for $\tau_1, \tau_2 \in \mathcal{H}$ given, there exist neighborhoods U_1 of τ_1 and U_2 of τ_2 in $\mathcal{H}$ with the property that

$$\text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}), \text{ if } \gamma(U_1) \cap U_2 \neq \emptyset \text{ then } \gamma(\tau_1) = \tau_2.$$

Note that τ_1 and τ_2 in the lemma can be equal.

Theorem 4.1.11
For any congruence subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$, the modular curve $Y(\Gamma)$ is Hausdorff.

Proof:
Let $\pi(\tau_1)$ and $\pi(\tau_2)$ be distinct points in $Y(\Gamma)$. Take neighborhoods U_1 of τ_1 and U_2 of τ_2 as in Lemma 4.1.10. Since $\gamma(\tau_1) \neq \tau_2$ for all $\gamma \in \Gamma$, the proposition says that $\Gamma(U_1) \cap U_2 = \emptyset$ in $\mathcal{H}$, and so equivalence given by 4.1.9 shows that $\pi(U_1)$ and $\pi(U_2)$ are disjoint supersets of $\pi(\tau_1)$ and $\pi(\tau_2)$ in $Y(\Gamma)$. They are neighborhoods since π is an open mapping.

□

Remark 4.1.12
Additionally, Since Euclidean space has a countable basis, so does the Modular curve being a quotient of euclidean space, which agrees with definition of a Riemann surface.

All remains is to put local coordinates on the modular curve $Y(\Gamma)$. This simply means that finding for each point $\pi(\tau) \in Y(\Gamma)$ a neighborhood $\widetilde{U}$ and a homeomorphism $\varphi : \widetilde{U} \longrightarrow V \subset \mathrm{C}$ such that the transition maps between the local coordinate systems are holomorphic.
Note: At a point $\pi(\tau)$ where $\tau \in \mathcal{H}$ is fixed only by the identity transformation in Γ, i.e., only by the matrices $\Gamma \cap \{\pm I\}$, this is simple: a small enough neighborhood U of τ in $\mathcal{H}$ is homeomorphic under π to its image $\pi(U)$ in $Y(\Gamma)$, as Lemma 4.1.10 guarantees such a neighborhood with no Γ-equivalent points. So a local inverse $\varphi : \pi(U) \longrightarrow U$ could serve as the local coordinate map. The problem occurs mainly at elliptic points, i.e at points where their isotropy subgroup is non-trivial. [DS05] uses proposition 2.7.11 to remedy this and to put local coordinates on $Y(\gamma)$. We encourage the readers to see [DS05] (Ch 2, Sections 2.1, 2.2) for more details about the complex structure on the modular curve $Y(\gamma)$.

Let Γ be a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. To compactify the modular curve $Y(\Gamma) = \Gamma\backslash\mathcal{H}$, define $\mathcal{H}^* = \mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$ and take the extended quotient

$$X(\Gamma) = \Gamma\backslash\mathcal{H}^* = Y(\Gamma) \cup \Gamma\backslash(\mathbb{Q} \cup \{\infty\}).$$

The points Γs in $\Gamma\backslash(\mathbb{Q} \cup \{\infty\})$ are also called the cusps of $X(\Gamma)$. For the congruence subgroups $\Gamma_0(N), \Gamma_1(N)$, and $\Gamma(N)$ we write $X_0(N), X_1(N)$, and $X(N)$.
Recall from 2.7.4, 2.7.5 that the modular curve $X(1) = \mathrm{SL}_2(\mathbb{Z})\backslash\mathcal{H}^*$ has one cusp. For any congruence

subgroup Γ of $\mathrm{SL}_2(\mathbb{Z})$ the modular curve $X(\Gamma)$ has finitely many cusps.
Topology: The topology on $\mathcal{H}^*$ consisting of its intersections with open complex disks (including disks $\{z : |z| > r\} \cup \{\infty\}$) contains too many points of $\mathbb{Q} \cup \{\infty\}$ in each neighborhood to make the quotient $X(\Gamma)$ Hausdorff. Instead, to put an appropriate topology on $X(\Gamma)$ start by defining for any $M > 0$ a neighborhood

$$\mathcal{N}_M = \{\tau \in \mathcal{H} : \mathrm{Im}(\tau) > M\}.$$

Adjoin to the usual open sets in $\mathcal{H}$ more sets in $\mathcal{H}^*$ to serve as a base of neighborhoods of the cusps, the sets

$$\alpha\left(\mathcal{N}_M \cup \{\infty\}\right) : M > 0, \alpha \in \mathrm{SL}_2(\mathbb{Z}),$$

and take the resulting topology on $\mathcal{H}^*$. Since fractional linear transformations are conformal and take circles to circles, if $\alpha(\infty) \in \mathbb{Q}$ then $\alpha\left(\mathcal{N}_M \cup \{\infty\}\right)$ is a disk tangent to the real axis.
Under this topology each $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ is a homeomorphism of $\mathcal{H}^*$. Finally, give $X(\Gamma)$ the quotient topology and extend natural projection to $\pi : \mathcal{H}^* \longrightarrow X(\Gamma)$.

Proposition 4.1.13
The modular curve $X(\Gamma)$ is Hausdorff, connected, and compact.

Proof:
See,[DS05], Proposition 2.4.2. □

Theorem 4.1.14 (**Modularity theorem**)
Let E be a complex elliptic curve with $j(E) \in \mathbb{Q}$. Then for some positive integer N there exists a surjective holomorphic function of compact Riemann surfaces from the modular curve $X_0(N)$ to the elliptic curve E,

$$X_0(N) \longrightarrow E$$

4.2 Divisors, Differentials, Riemann-roch theorem

Definition 4.2.1
Let C be a non-singular algebraic curve over a field k. We define the divisor group $\mathrm{Div}(C)$ for a curve C as the set of formal sums:

$$\mathrm{Div}(C) = \left\{ \sum_{P \in C} n_P(P) : n_P \in \mathbb{Z}, \text{ with almost all but finitely many } n_P = 0 \right\}.$$

Consider the localisation $\overline{\mathbb{Q}}[C]_P$ of $\overline{\mathbb{Q}}[C]$ at a point P. This is a local ring at P, and it has a unique maximal ideal M_P. Any generator of M_P is called a uniformizer at P. Let t be a uniformizer.
For any element F of this local ring we have that F takes the form $F = t^e u$ where $e \in \mathbb{N}$ and $u \in \overline{\mathbb{Q}}[C]_P^*$. This representation of F is unique, for if $F = t^e u = t^{e'} u'$ with $e \geq e'$ then $t^{e-e'} u = u'$, showing that $e = e'$ and $u = u'$. This defines a function on the co-ordinate ring of C known as valuation at P.

Definition 4.2.2

The valuation at P on the coordinate ring is the function

$$\nu_P : \overline{\mathbb{Q}}[C] \longrightarrow \mathbb{N} \cup \{+\infty\}, \quad \nu_P(f) = \begin{cases} +\infty & \text{if } f = 0 \\ e & \text{if } f = t^e u \end{cases}$$

This extends to the function field,

$$\nu_P : \overline{\mathbb{Q}}(C) \longrightarrow \mathbb{Z} \cup \{+\infty\}, \quad \nu_P(F) = \nu_P(f) - \nu_P(g), F = f/g$$

where it is well defined.

Definition 4.2.3
For a divisor $\mathcal{D} = \sum_{P \in C} n_P(P)$, we define the degree of a divisor $\deg(\mathcal{D}) = \sum_{P \in C} n_P$.

Remark 4.2.4
1. The notation (P) is used to distinguish the divisorial representation of a point P on the curve. For example, this distinction is particularly pertinent in the context of elliptic curves, where we distinguish between the divisor $\sum n_P(P)$ in $\mathrm{Div}(E)$ and the elliptic curve point $\sum [n_P]P$ in $\mathcal{E}$.
2. The coefficients n_P in a divisor can be negative, leading to a negative representation for $n_P < 0$ as $[n_P]\, P = -\,[-n_P]\, P$.
3. It is clear that since $n_P = 0$ for all but finitely many points P on a curve C, the degree of a divisor is well-defined.
4. To explain the intuition in a crude sense, divisors on a curve C encapsulate the distribution of zeros and poles of meromorphic functions defined on C and that divisors capture zeros and poles of a function on the curve C. Conventionally, the integers n_P denote the multiplicities of the zeroes and poles and that positive n_P denotes that P is a zero of multiplicity n_P and similarly, negative n_P denote P is a pole of multiplicity $-n_P$.
5. Geometrically, divisors provide a way to visualize the behavior of functions defined on C. Imagine walking along the curve: when you encounter a point in the divisor, you're either stepping on a zero or jumping over a pole, and the multiplicity tells you how significant that step or jump is.

For a general curve C that is a curve which is not necessarily an elliptic curve we also have following notions:

Definition 4.2.5
The subgroup of divisors of degree zero is defined as:

$$\mathrm{Div}^0(C) = \left\{ \sum n_P(P) \in \mathrm{Div}(C) : \sum n_P = 0 \right\}.$$

Definition 4.2.6
The divisor associated with a nonzero function F in the function field of C is given by:

$$\mathrm{div}(F) = \sum \nu_P(F)(P), \quad F \in \overline{\mathbb{Q}}(C)^*.$$

Remark 4.2.7
1. The divisors of this form are termed principal divisors, collectively denoted as $\mathrm{Div}^\ell(C)$.
A key property of principal divisors is that they are of degree zero.

This reminds us of Riemann surface theory which asserts the balance of zeroes and poles of F:

$$\operatorname{div}(F) = \sum_{P \in F^{-1}(0)} e_P(F)(P) - \sum_{P \in F^{-1}(\infty)} e_P(F)(P),$$

leading to the conclusion that $\deg(\operatorname{div}(F)) = 0$. The homomorphism property of the map div from the function field $\overline{\mathbb{1}}(C)^*$ to $\operatorname{Div}^0(C)$ is established in [DS05], Proposition 7.2.4.
Consequently, $\operatorname{Div}^\ell(C)$ forms a subgroup of $\operatorname{Div}^0(C)$.

Definition 4.2.8
The Picard group of C, specifically the degree zero Picard group $\operatorname{Pic}^0(C)$, is then defined as the quotient of these two groups:

$$\operatorname{Pic}^0(C) = \operatorname{Div}^0(C)/\operatorname{Div}^\ell(C).$$

Definition 4.2.9
If $h : C \longrightarrow C'$ is a nonconstant morphism. Then its induced forward and reverse maps of Picard groups and are given by,

$$h_* : \operatorname{Pic}^0(C) \longrightarrow \operatorname{Pic}^0\left(C'\right) \text{ and } h^* : \operatorname{Pic}^0\left(C'\right) \longrightarrow \operatorname{Pic}^0(C),$$

given by

$$h_*\left(\left[\sum_P n_P(P)\right]\right) = \left[\sum_P n_P(h(P))\right]$$

and

$$h^*\left(\left[\sum_Q n_Q(Q)\right]\right) = \left[\sum_Q n_Q \sum_{P \in h^{-1}(Q)} e_P(h)(P)\right]$$

Here the square brackets denote equivalence class modulo $\operatorname{Div}^\ell(C)$ or $\operatorname{Div}^\ell\left(C'\right)$ as appropriate.

Definition 4.2.10
A holomorphic (or meromorphic) differential form on an open set U of $\mathbb{C}$ is an expression of the form $f(z)dz$ with f holomorphic (or meromorphic).

Definition 4.2.11
Let $f : U \longrightarrow \mathbb{C}$ be a holomorphic map. Then $df := \frac{df}{dz}dz$ is called the associated differential form to f.

Definition 4.2.12
Let X be a compact Riemann Surface and $(U_i, z_i)_{i \in I}$ a complex structure. A holomorphic differential form on X is given by a family $(\alpha_i)_{i \in I}$ of differential forms $\alpha_i = f_i\left(z_i\right) dz_i$ on $z_i\left(U_i\right), \forall i \in I$ such that if $\omega_{ij} := z_i \circ z_j^{-1} : z_j\left(U_i \cap U_j\right) \to z_i\left(U_i \cap U_j\right)$ denote the holomorphic maps given by the definition of a complex structure. Then $\omega_{ij}^*\left(\alpha_i\right) = \alpha_j$ (or equivalently, $f_j\left(z_j\right) dz_j = f_i\left(\omega_{ij}\left(z_j\right)\right) \omega_{ij}'\left(z_j\right) dz_j$).

Note, that the theory of divisors also applies for the case of Riemann surfaces. See [DS05], Section 3.2, 3.3. There is also an analogous definition for 1-differential forms over a Variety X in [HS00], section A.1.4. This also helps us discussing differentials on curves. The discussion over curves is then analogous to what we do over Riemann surfaces.
The construction is analogous and similar properties are also true with some minor modifications here

and there.
We can also define a divisor of a differential form. Let ω be a differential form on a riemann surface X. Let $P \in X$ and (U, z) be a local chart at P. Then $\omega := f(z)dz$ on U. Note that $v_{z(P)}(f \circ z^{-1})$ is independent of local chart.

To see this, consider two local charts, (U_1, z_1) and (U_2, z_2), at a point P. Within these charts, let the differential forms be represented as $\omega_1 = f_1(z_1)dz_1$ on U_1 and $\omega_2 = f_2(z_2)dz_2$ on U_2.
We define a transition map between these charts by:

$$w_{1,2} : z_2(U_1 \cap U_2) \to z_1(U_1 \cap U_2),$$

where $w_{1,2} = z_1 \circ z_2^{-1}$. This map is biholomorphic, which implies that its derivative $w'_{1,2}$ does not vanish anywhere within its domain.
With this setup, we can relate ω_2 to ω_1 through the formula:

$$f_2(z_2)dz_2 = f_1(w_{1,2}(z_2))w'_{1,2}(z_2)dz_1.$$

This relationship suggests that the valuation of f_2 at $z_2(P)$, denoted $v_{z_2(P)}(f_2)$, is equal to the valuation of f_1 at $z_1(P)$, denoted $v_{z_1(P)}(f_1)$. This equality is fundamental as it allows for the definition of the valuation of a differential form ω at the point P, denoted as $v_P(\omega)$, to be independent of the local chart used.
Therefore, we define $v_P(\omega) = v_P(f)$, where f is the function representing ω in any local chart around P.
The divisor of the differential form ω is then defined as the formal sum of its valuations at each point on the Riemann surface, expressed as:

$$\operatorname{div}(\omega) := \sum_{P \in X} v_P(\omega) \cdot [P].$$

One nice property is the equivalence class of $\operatorname{div}(\omega)$ is independent of choice of differential ω. This gives us definition of a canonical divisor.

Definition 4.2.13
We write $K = \operatorname{div}(\omega)$, and we say the K is a canonical divisor.

Let X be a compact Riemann surface, $D \in \operatorname{Div}(X)$.
We define, $\mathscr{L}(D) = \{f \in \mathcal{C}(X) \mid \operatorname{div}(f) + D \geq 0\} \cup \{0\}$.
We now define a couple of results without proof. For more details see [DS05].

Theorem 4.2.14
The dimension of $\mathscr{L}(D)$ is finite for all $D \in \operatorname{Div}(X)$.
If $g \in \mathcal{M}(X)$ and $D' = D + \operatorname{div}(g)$. Then the map

$$\mathscr{L}(D) \longrightarrow \mathscr{L}\left(D'\right); f \mapsto fg^{-1}$$

is an isomorphism between $\mathscr{L}(D)$ and $\mathscr{L}\left(D'\right)$. Therefore, $\dim \mathscr{L}(D) = \dim \mathscr{L}\left(D'\right)$.

Theorem 4.2.15 (The Riemann-Roch theorem)
There exists $g = g_X \in \mathbb{N}$ such that for any $D \in \operatorname{Div}(X)$,

$$\ell(D) - \ell(K - D) = \deg D + 1 - g$$

where $\ell(D) = \dim \mathscr{L}(D), K$ is the canonical divisor, i.e, $K = \mathrm{div}(\omega)$ where ω is a differential form on X.

Remark 4.2.16
1. $\mathscr{L}(0) \cong \mathbb{C} \cong \{$ constant functions $f : X \longrightarrow \mathbb{C}\}$ as any holomorphic map on a compact Riemann surface is constant. Therefore, $\ell(0) = 1$.
2. $\deg K = 2g - 2$; This follows by applying Riemann-Roch theorem on the canonical divisor K.
3.$\ell(K) = g$. This means, the space of holomorphic differential forms on X is of dimension g. This follows by applying Riemann-Roch on the 0 divisor. Another interpretation of this result is that

$$\mathscr{L}(K) \cong \{ \text{ holomorphic differential forms on } X\}$$

Another application is that we get a Weierstrass equation for an elliptic curve. Recall 1.1.5. We finally give its proof sketch now.

Proof Sketch of Proposition 1.1.5. Consider k to be an algebraically closed field, and let n be a positive integer. According to the Riemann-Roch Theorem, the Riemann-roch space $\mathscr{L}(n(O))$ is of dimension n over k. Constant functions contribute to a one-dimensional subspace. Therefore, it is possible to find elements $u, v \in k(E)$ such that $\{1, u\}$ is a basis for $\mathscr{L}(2(O))$ and similary $\{1, u, v\}$ for $\mathscr{L}(3(O))$. It follows that u has a pole of order precisely two at O, and v has a pole of order exactly three at O. This infers that the collection $\{1, u, v, u^2, uv, v^2, u^3\}$ belongs to $\mathscr{L}(6(O))$, which defines a six-dimensional vector space, thereby indicating that these elements are linearly dependent. As such, there must be scalars $a_1, \ldots, a_7 \in k$, not all zero, for which $a_1 + a_2u + a_3v + a_4u^2 + a_5uv + a_6v^2 + a_7u^3 = 0$. Given that each term has a unique pole order at O apart from the last two, it must be that $a_6a_7 \neq 0$. By selecting appropriate substitutions for a_i, and due to the basis conditions on u and v, O is mapped to $[0 : 1 : 0]$ in the projective space.
Let this curve be denoted as C, and we have a rational map $f : E \to C$ which, due to the smooth nature of E, is a morphism. The function $u : E \to \mathbb{P}^1_k$ is of degree two as u has a pole of order two at O and no other, thus $[k(E) : k(u)] = 2$. In parallel, $[k(E) : k(v)] = 3$, hence $[k(E) : k(u, v)] = 1$. Therefore, f is a degree one morphism. If C is smooth, then f is an isomorphism being a degree one morphism between smooth curves. It remains to ascertain the behavior when C is singular. Employing Weierstrass equations, it can be demonstrated that under such circumstances, there is a rational map $g : C \to \mathbb{P}^1_k$ of degree one, and the composition $g \circ f$ yields a degree one morphism from E to $\mathbb{P}^1_k$. However, as E has genus one and $\mathbb{P}^1_k$ has genus zero, this results in a contradiction, hence completing the proof of the initial assertion.
For any smooth Weierstrass equation and the associated ecurve E, with O defined as $[0 : 1 : 0]$, we observe that (E, O) is an elliptic curve assuming E is of genus one. Indeed, the differential $\Lambda = du/(2v + a_1u + a_3)$ has no zeroes or poles on E, implying $\mathrm{div}\Lambda = 0$. The claim follows from Remark 4.3.16, Part 2. □
This proof also applies when k is not algebraically closed, suggesting that for any elliptic curve over any field k, a corresponding Weierstrass equation with coefficients in k can be formulated.
As a short application we will compute an equation of the modular curve $X_0(38)$. In chapter 2 we showed that the genus of this curve is 4.

Let $\omega_1, \ldots, \omega_4$ be a basis for $\Omega^1_{\text{hol}}(X_0(38))$. As the canonical divisor is very ample (and $X_0(38)$ can be shown to be non-hyperelliptic), the induced map

$$\begin{aligned} \varphi : C &\to \mathbb{P}^{g-1} \\ P &\mapsto [\omega_1(P) : \cdots : \omega_g(P)] . \end{aligned}$$

will embed C into P^3.

Note that their exists an isomorphism between the space of weight 2 cusp forms $\mathcal{S}_2(\Gamma_0(38))$ and the space differential 1-forms $\Omega^1_{\text{hol}}(X_0(38))$.
In Chapter 5 we will compute a basis f_1, f_2, f_3 and f_4 of weight 2 cusp forms of level 38. As any non-hyperelliptic genus 4 curve can be written down as a complete intersection of a cubic and a quadric (see, e.g. Example 5.5.2 from [Har77]), finding these kinds of relations between the f_i will give us an equation of the curve. (This method was cleverly used by Galbraith in [Gal].)
As per our computations, we have $f_1(\tau), f_2(\tau)$ newforms of level 38 and $g_1(\tau), g_1(2\tau)$, g_1 being newform of level 19.
We can just plug in the following commands in Magma:

```
M := ModularSymbols(38);
M_cusp := CuspidalSubspace(M);
M_dec := NewformDecomposition(M_cusp);
Relations(CuspidalSubspace(ModularForms(Gamma0(38))), 3, 20);
Relations(CuspidalSubspace(ModularForms(Gamma0(38))), 2, 20);
```

To get the relations of degree 2 and degree 3:
Degree 3: $a^2*c - a*b^2 - a*b*d - a*d^2 - b^2*c - b^2*d - b*c*d - b*d^2 - c^3 - 2*c^2*d - 2*c*d^2 - d^3$,

a$^2*d + a*d^2 - b^3 + 3*b^2*c + 2*b^2*d - 3*b*c^2 - 4*b*c*d - 2*b*d^2 + c^3 + 2*c^2*d + 2*c*d^2 + d^3$,

a*b*c - b$^3 - b^2*d - b*c^2 - b*c*d - b*d^2, a*c^2 - b^2*c - b*c*d - c^3 - c^2*d - c*d^2, a*c*d - b^2*d - b*d^2 - c^2*d - c*d^2 - d^3$

Degree 2: $a*c - b^2 - b*d - c^2 - c*d - d^2$
A quick check in Magma shows that the curve given by:

$$\begin{aligned} x^2w + xw^2 - y^3 + 3y^2z + 2y^2w - 3yz^2 - 4yzw - 2yw^2 + z^3 + 2z^2w + 2zw^2 + w^3, \\ xz - y^2 - yw - z^2 - zw - w^2 \end{aligned}$$

defines a curve of genus 4 which has bad reduction at the primes 19 and 2.

4.3 Algebraic Curves in arbitrary characteristic

Here in this section, the goal is to briefly define the setup and basic notions required to further discuss Algebraic curves in arbitrary characteristics. From this point onward, till the end of this chapter, we will very closely follow some sections of [DS05], chapter 8 to introduce the terminology and state some important results about curves and their reductions.
Algebraic curves are characterized using polynomials $\varphi_1, \ldots, \varphi_m$ to form an ideal I in the polynomial ring over the algebraically closed field $\overline{k}$. To be precise, Given polynomials $\varphi_1, \ldots, \varphi_m \in k\left[x_1, \ldots, x_n\right]$ such that the ideal

$$I = \langle \varphi_1, \ldots, \varphi_m \rangle \subset \overline{k}\left[x_1, \ldots, x_n\right]$$

is prime, let

$$C = \left\{ P \in \overline{k}^n : \varphi(P) = 0 \text{ for all } \varphi \in I \right\}.$$

The function field of this curve denoted as $\overline{k}(C)$, is derived by taking the quotient field of the coordinate ring $\overline{k}[C] = \overline{k}\left[x_1, \ldots, x_n\right]/I$.
The relationship between curves and their function fields given by Curves-Fields Correspondence, remains consistent as in the case over $\mathbb{C}$. See [DS05], chapter 7 for more details.
The map

$$C \mapsto k(C)$$

induces a bijection from the set of isomorphism classes over k of nonsingular projective algebraic curves over k to the set of conjugacy classes over k of function fields over k. And for any two nonsingular projective algebraic curves C and C' over k, the map

$$\left(h : C \longrightarrow C'\right) \mapsto \left(h^* : k\left(C'\right) \longrightarrow k(C)\right)$$

is a bijection from the set of surjective morphisms over k from C to C' to the set of k-injections of $k\left(C'\right)$ in $k(C)$. See [DS05],Chapter 7 for more details.

Let us, for now, shift our attention towards algebraic curves in characteristic p, where p is a prime number. Here, $\mathbb{F}_p$ denotes the field consisting of p elements and its algebraic closure is denoted $\overline{\mathbb{F}}_p$. Recall that, for every power q of p, there exists a unique field $\mathbb{F}_q$ contained in $\overline{\mathbb{F}}_p$.

Definition 4.3.1 (Frobenius map)
The Frobenius map on $\overline{\mathbb{F}}_p$ is

$$\sigma_p : \overline{\mathbb{F}}_p \longrightarrow \overline{\mathbb{F}}_p, \quad x \mapsto x^p$$

Remark 4.3.2
We recall some of the properties from basic algebra.

1. The inverse of the Frobenius map is an automorphism of $\overline{\mathbb{F}}_p$, but it isn't a polynomial function.
2. The fixed points of this inverse are the elements of $\mathbb{F}_p$, which are roots in $\overline{\mathbb{F}}_p$ of the polynomial

$x^p = x$.

3. In general, the fixed points of σ_p^e, where σ_p^e denotes the e-fold composition with itself, are given by $\mathbb{F}_q$, where $q = p^e$.

4. The group of automorphisms of $\overline{\mathbb{F}}_p$ is not cyclic.

Definition 4.3.3
The Frobenius map on $\overline{\mathbb{F}}_p^n$ is

$$\sigma_p : \overline{\mathbb{F}}_p^n \longrightarrow \overline{\mathbb{F}}_p^n, \quad (x_1, \ldots, x_n) \mapsto (x_1^p, \ldots, x_n^p)\,.$$

Remark 4.3.4
Due to the property $(x+y)^p = x^p + y^p$ in characteristic p, the Frobenius map is a field automorphism. Surjectivity of the Frobenius map is evident as we are working over a fixed algebraic closure of $\mathbb{F}_p$ and thus given any n-tuple in $\overline{\mathbb{F}}_p^n$, we can always find its inverse image under σ_p. The injectivity is also clear due to the property 1 mentioned at the beginning of this remark. In conclusion, This is a bijection, and its fixed points are $\mathbb{F}_p^n$.
Moreover, it induces a well defined bijection at the level of projective spaces,

$$\sigma_p : \mathbb{P}^n\left(\overline{\mathbb{F}}_p\right) \longrightarrow \mathbb{P}^n\left(\overline{\mathbb{F}}_p\right), \quad [x_0, x_1, \ldots, x_n] \mapsto [x_0^p, x_1^p, \ldots, x_n^p]$$

with fixed points $\mathbb{P}^n(\mathbb{F}_p)$.
Furthermore, consider a homogeneous polynomial $\varphi(x)$ defined as:

$$\varphi(x) = \sum_e a_e x^e$$

where x^e is a compact notation for $x_0^{e_0} \cdots x_n^{e_n}$. This polynomial belongs to the space $\overline{\mathbb{F}}_p[x_0, x_1, \ldots, x_n]$. If we apply the Frobenius map to the coefficients of this polynomial, we get another polynomial $\varphi^{\sigma_p}(x)$, written as:

$$\varphi^{\sigma_p}(x) = \sum_e a_e^{\sigma_p} x^e$$

An important relationship (See [DS05] Exercise 8.2.2)(Note that we are working over a field of characteristic p) between these two polynomials is:

$$\varphi^{\sigma_p}\left(x^{\sigma_p}\right) = \varphi(x)^{\sigma_p}$$

Now, let's imagine a projective curve C, defined over $\overline{\mathbb{F}}_p$ by a set of polynomials φ_i. The curve C^{σ_p} corresponds to the set of polynomials $\varphi_i^{\sigma_p}$. The Frobenius map, σ_p, establishes a transformation from curve C to C^{σ_p}. Specifically, for any point P in the projective space $\mathbb{P}^n\left(\overline{\mathbb{F}}_p\right)$ where all polynomials $\varphi_i(P)$ evaluate to zero, the transformed polynomials $\varphi_i^{\sigma_p}$ will also evaluate to zero at the point P^{σ_p}.
Additionally, when curve C is defined over $\mathbb{F}_p$, then C^{σ_p} coincides with C. This means that the Frobenius map σ_p defines a self-morphism on curve C.
This leads to the following defintion.

Definition 4.3.5

Let C be a projective curve over $\overline{\mathbb{F}}_p$. The Frobenius map on C is

$$\sigma_p : C \longrightarrow C^{\sigma_p}, \quad [x_0, x_1, \ldots, x_n] \mapsto [x_0^p, x_1^p, \ldots, x_n^p]$$

Example 4.3.6

- The Frobenius map, applied to the projective line $\mathbb{P}^1\left(\overline{\mathbb{F}}_p\right)$, sends t to t^p in its affine part. This gives rise to a function field extension, K/k, defined by:

  $$K = \mathbb{F}_p(t), \quad k = \mathbb{F}_p(s), \quad s = t^p$$

  So, $K = k(t)$. The polynomial $x^p - s$ is the minimal polynomial for t over k. Even though the Frobenius map is bijective, seemingly of degree of 1, the actual extension degree is p. This discrepancy occurs because the polynomial $(x-t)^p$ over K has the p^{th} root, t, repeated p times.

- In a similar manner, for an elliptic curve over $\mathbb{F}_p$, the Frobenius map sends (u, v) to (u^p, v^p) in the affine part. This leads to another function field extension, k/k, given by:

  $$K = \mathbb{F}_p(u)[v]/\langle E(u,v)\rangle, \quad k = \mathbb{F}_p(s)[t]/\langle E(s,t)\rangle, \quad s = u^p, t = v^p$$

  So, we have $K = k(u, v)$. The polynomial $x^p - s$ is the minimal polynomial of u in $k[x]$, which factors as $(x-u)^p$ over $k(u)$. Additionally, the minimal polynomial of v divides $E(u, y)$, which is quadratic in y. This results in:

  $$[k(u) : k] = p \quad \text{and} \quad [k(u, v) : k(u)] \in \{1, 2\}$$

  A similar approach yields:

  $$[k(v) : k] = p \quad \text{and} \quad [k(u, v) : k(v)] \in \{1, 3\}$$

  Consequently, $k = k(u) = k(v)$ and the extension degree $[k : k] = p$. Even in this case, the function field extension degree remains p, and the extension is determined by a p^{th} root repeating p times as the root of its minimal polynomial.

Let us now end this very brief section by introducing inducing forward and reverse maps of the Frobenius map and stating some of the properties without going into too many details but instead stating them for the sake of completeness and ease of discussing further things.

Given a projective curve C over $\mathbb{F}_p$, the forward induced map of σ_p on C transforms divisors as:

$$\sigma_{p,*} : (P) \mapsto (\sigma_p(P))$$

Given the bijective nature of σ_p and its ramification at all points with a degree p, its reverse induced map acts as:

$$\sigma_p^* : (P) \mapsto p\left(\sigma_p^{-1}(P)\right)$$

The p in the front comes due to the ramification. This is clear since the Frobenius map maps x to x^p.

If there's a map h from C to C' over $\mathbb{F}_p$, the Frobenius map commutes with h. Specifically, for the Frobenius map on C as $\sigma_{p,C}$ and on C' as $\sigma_{p,C'}$:

$$h \circ \sigma_{p,C} = \sigma_{p,C'} \circ h.$$

It's deduced that the forward induced map of the Frobenius map also commutes with the forward induced map of h:

$$h_* \circ (\sigma_{p,C})_* = \left(\sigma_{p,C'}\right)_* \circ h_*.$$

Since the Frobenius map commutes with h, its inverse does as well:

$$h \circ \sigma_{p,C}^{-1} = \sigma_{p,C'}^{-1} \circ h$$

Now, for any point P in C, we can compute:

$$\left(h_* \circ \sigma_{p,C}^*\right)(P) = \left(\sigma_{p,C'}^* \circ h_*\right)(P)$$

Consequently, the reverse induced map of the Frobenius map commutes with the forward induced map of h:

$$h_* \circ \sigma_{p,C}^* = \sigma_{p,C'}^* \circ h_*$$

To summarise, both the forward and reverse induced maps of the Frobenius map have commutative properties with the map h between projective curves over $\mathbb{F}_p$.

4.4 The reduction of algebraic curves

In this brief section, we aim to introduce essential notions like reduction of algebraic curves for example, at primes, etc. This will, in turn, facilitate the discussion for the reduction of Modular curves.
Let us begin with recalling the localization of the integers, $\mathbb{Z}$, at a prime p. This localized ring is denoted as:

$$\mathbb{Z}_{(p)} = \left\{ \frac{x}{y} : x, y \in \mathbb{Z}, y \notin p\mathbb{Z} \right\}$$

This is a subset of the rational numbers, $\mathbb{Q}$, which contains $\mathbb{Z}$ and is a local ring with the maximal ideal, $p\mathbb{Z}_{(p)}$. Furthermore, recall that There's a natural isomorphism between the quotient rings of $\mathbb{Z}$ and $\mathbb{Z}_{(p)}$ modulo p:

$$\mathbb{Z}/p\mathbb{Z} \simeq \mathbb{Z}_{(p)}/p\mathbb{Z}_{(p)}, \quad a + p\mathbb{Z} \mapsto a + p\mathbb{Z}_{(p)}$$

This leads to a well-defined reduction map from the localized integers to the field with p elements:

$$\sim : \mathbb{Z}_{(p)} \to \mathbb{F}_p, \quad \alpha \mapsto \tilde{\alpha} = \alpha + p\mathbb{Z}_{(p)}$$

Here, the map $\sim$ is surjective and has $p\mathbb{Z}_{(p)}$ as its kernel.

This structure becomes particularly useful when attempting to reduce an algebraic curve from the field of rationals $\mathbb{Q}$ to a curve over the finite field $\mathbb{F}_p$. The key tool in this reduction process is the localization $\mathbb{Z}_{(p)}$, which allows us to concentrate on the characteristics related to the prime p.

Let us begin with defining a few notions.

Definition 4.4.1
Let C be a nonsingular affine algebraic curve over $\mathbb{Q}$, defined by polynomials $\varphi_1, \ldots, \varphi_m \in \mathbb{Z}_{(p)}[x_1, \ldots, x_n]$. Then C has good reduction modulo p (or at p) if
(1) The ideal $I = \langle \varphi_1, \ldots, \varphi_m \rangle$ of $\mathbb{Z}_{(p)}[x_1, \ldots, x_n]$ is prime.
(2) The reduced polynomials $\tilde{\varphi}_1, \ldots, \tilde{\varphi}_m \in \mathbb{F}_p[x_1, \ldots, x_n]$ define a nonsingular affine algebraic curve $\widetilde{C}$ over $\mathbb{F}_p$.

Let us now quickly move on from affine curves to projective curves. Let k be any field. We will use the notation $x = (x_0, \ldots, x_n)$ and $x_{(i)} = (x_0, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n)$ for $i = 0, \ldots, n$, and $\varphi_{(i)} = \varphi\left(x_{(i)}\right)$ for $\varphi \in k[x]$.

Definition 4.4.2
For any homogeneous ideal $I = \langle \{\varphi\} \rangle \subset k[x]$
its ith dehomogenization for $i = 0, \ldots, n$ is

$$I_{(i)} = \left\langle \left\{ \varphi_{(i)} \right\} \right\rangle \subset k\left[x_{(i)}\right]$$

and its ith rehomogenization is

$$I_{(i),\,\mathrm{hom}} = \left\langle \left\{ \varphi_{(i),\,\mathrm{hom}} \right\} \right\rangle \subset k[x],$$

where $\varphi_{(i),\mathrm{hom}}$ means to multiply each term of $\varphi_{(i)}$ by the smallest power of x_i needed to make all the terms have the same total degree. Thus $\varphi = x_i^e \varphi_{(i),\mathrm{hom}}$ for some $e \geq 0$.

Lemma 4.4.3
Given a field k and I the homogenized version of a prime ideal $I_{(0)}$ in $k\left[x_{(0)}\right]$ associated with an affine algebraic curve,I itself is prime. Furthermore, for every i in the range 0 to n, if $I_{(i)}$ is not equal to $k\left[x_{(i)}\right]$, then $I_{(i)}$ is a prime ideal and the homogenization of $I_{(i)}$ is I and $I_{(i)}$ describes an affine algebraic curve.

Definition 4.4.4
Let C_{hom} be a nonsingular projective curve over $\mathbb{Q}$ represented by the homogenized ideal $I \subset \mathbb{Z}_{(p)}[x]$. This curve originates from a prime ideal $I_{(0)} \subset \mathbb{Z}_{(p)}\left[x_{(0)}\right]$ as described above. We say that C_{hom} has good reduction at p if, for each i ranging from 1 to n, the affine curve C_i defined by $I_{(i)}$ either has good reduction at p or $I_{(i)}$ reduces to all of $\mathbb{F}_p\left[x_{(i)}\right]$, implying that C_i has an empty reduction at p. The curve $\widetilde{C}_{\mathrm{hom}}$, defined by the homogenization $\left(\widetilde{I_{(0)}}\right)_{\mathrm{hom}} \subset \mathbb{F}_p[x]$, is then defined as the reduction of curve C at p.

Note:
The reduced curve $\widetilde{C}_{\mathrm{hom}}$ can be defined by any nonempty reduction $\widetilde{C}_i$ of an affine piece of C_{hom} , but this is not immediately obvious. See [DS05] Chapter 8, section 5 for more details. Furthermore, it is worth noting that Good reduction at p on one affine piece of a projective curve does not guarantee good

reduction of the curve.

Lastly, we quote an important theorem without proof before moving on to the discussion of morphisms.

Theorem 4.4.5
Let C be a projective algebraic curve over $\mathbb{Q}$ having a good reduction at a prime p. Then, the map from the curve to its reduced curve $f : C \longrightarrow \widetilde{C}$ is surjective.

Proof:
[DS05], Section 8.4,8.5.

□

It is also natural to think about the reduction of morphisms at primes.
Consider a morphism h between nonsingular projective algebraic curves C and C' over $\mathbb{Q}$ with good reduction at p:

$$h : C \longrightarrow C'$$

To deduce a morphism $\widetilde{h}$ for the reduced curves, let us first represent h as $[h_0, \ldots, h_r]$. The ideal $I_{(0)}$ in $\mathbb{Z}_{(p)}\left[x_{(0)}\right]$ defines C upon homogenization. It's assumed that each h_i belongs to the subring R of the coordinate ring $\mathbb{Q}[C_0]$.
The p-adic valuation of each h_i is:

$$\nu_p(h_i) = \max\{e : h_i \in p^e R\}$$

At least one $\nu_p(h_i)$ is finite. Consequently, the valuation of h is:

$$\nu_p(h) = \min\{\nu_p(h_i) : i = 0, \ldots, r\}$$

Rewriting h gives:

$$h = \left[h'_0, \ldots, h'_r\right]$$

where each h'_i belongs to R, and at least one entry in $R - pR$. This non-zero reduction forms the basis for $\tilde{h}$, a rational map from $\widetilde{C}_0$ to $\widetilde{C}'$.
For each point in $\widetilde{C}_0$, described as $\widetilde{P}$ with corresponding $P \in C_0$, $\tilde{h}(\widetilde{P})$ is non-zero for most points. If for instance $\tilde{h}'_i(\widetilde{P})$ isn't zero, it's shown that $\tilde{h}(\widetilde{P})$ lies in $\widetilde{C'_i}$. For any element of the defining ideal $\widetilde{I'_{(i)}}$, $\tilde{g}$ mirrors g in $I'_{(i)}$. Consequently, for $\widetilde{P}, i$, and $\tilde{g}$, $\tilde{g}(\tilde{h}(\widetilde{P}))$ is the reduction of $g(h(P))$, which is zero because $h(P)$ is in C'_i.
This argument proves that $\tilde{h}(\widetilde{P}) = \widetilde{h(P)}$ for all P in C that reduce outside a finite subset of $\widetilde{C}$. The statement remains consistent regardless of the chosen affine piece of C or the h_i representatives, even though the finite set might change.
In conclusion, $\tilde{h}$ is valid as a rational map from $\widetilde{C}$ to $\widetilde{C'}$. As usual, by theory of algebraic curves, this rational map extends to the final reduced morphism:

$$\tilde{h} : \widetilde{C} \longrightarrow \widetilde{C'}$$

Note:

For an algebraic curve C over $\mathbb{Q}$, its geometric genus is 0 if:
1. It's isomorphic over $\overline{\mathbb{Q}}$ to the projective line. 2. Its function field $\overline{\mathbb{Q}}(C)$ is isomorphic in $\overline{\mathbb{Q}}$ to a field $\overline{\mathbb{Q}}(t)$.
Curves with genus 0 are anomalies. However, if the target curve C' possesses a positive genus, both the morphism h and its reduction $\tilde{h}$ operate compatibly on its points and share the same degree.
The following theorem, further explains this:

Theorem 4.4.6
If C and C' are nonsingular projective algebraic curves over $\mathbb{Q}$ that have good reduction at p, and C' has a positive genus, then for any morphism h from C to C', there exists a commutative diagram with $deg(h) = deg(\widetilde{h})$.

$$\begin{array}{ccc} C & \xrightarrow{h} & C \\ \downarrow & & \downarrow \\ \widetilde{C} & \xrightarrow{\widetilde{h}} & \widetilde{C} \end{array}$$

Proof:
[DS05], Section 8.4,8.5.

□

It is also natura

Note: Since h is surjective by theorem 4.6.6, we get that $\widetilde{h}$ being the unique morphism that makes the diagram commute. For example,the map $h : \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ where $h[x, y] = [px, y]$ surjects over $\mathbb{Q}$ but it reduces at p to the zero map.

Corollary 4.4.7
Let's consider a curve C' with a positive genus. Then, the following is true:
(a) If a morphism $h : C \longrightarrow C'$ is surjective, then the reduced morphism $\tilde{h} : \widetilde{C} \longrightarrow \widetilde{C'}$ is also surjective.
(b) Given another morphism $k : C' \longrightarrow C''$ where C'' also has a positive genus, the composition of their reductions holds.

$$\widetilde{k \circ h} = \widetilde{k} \circ \widetilde{h}$$

(c) If the morphism h is an isomorphism, the reduced morphism $\tilde{h}$ retains this property and is also an isomorphism.

Proof:
[DS05], Section 8.4,8.5.

□

Next, we look at reduction and Picard Groups of Nonsingular Projective Curves.

Theorem 4.4.8
Given a nonsingular projective algebraic curve C over $\mathbb{Q}$ with good reduction at p, the reduction induces a mapping on degree-0 divisors:

$$\mathrm{Div}^0(C) \longrightarrow \mathrm{Div}^0(\widetilde{C}), \quad \sum n_P(P) \mapsto \sum n_P(\widetilde{P})$$

This mapping sends principal divisors to principal divisors, which further results in a surjective transformation of the Picard groups:

$$\mathrm{Pic}^0(C) \longrightarrow \mathrm{Pic}^0(\widetilde{C}), \quad \left[\sum n_P(P)\right] \mapsto \left[\sum n_P(\widetilde{P})\right]$$

Now, consider another curve C' — also a nonsingular projective algebraic curve over $\mathbb{Q}$ with good reduction at p — but with a positive genus. Given a morphism h from C to C' over $\mathbb{Q}$, there are induced forward maps:
$h_* : \mathrm{Pic}^0(C) \longrightarrow \mathrm{Pic}^0(C')$ and $\tilde{h}_* : \mathrm{Pic}^0(\widetilde{C}) \longrightarrow \mathrm{Pic}^0(C')$
These are induced by h and $\tilde{h}$ respectively. Moreover, the following diagram commutes:

$$\begin{array}{ccc} \mathrm{Pic}^0(C) & \xrightarrow{h_*} & \mathrm{Pic}^0(C') \\ \downarrow & & \downarrow \\ \mathrm{Pic}^0(\widetilde{C}) & \xrightarrow{\tilde{h}_*} & \mathrm{Pic}^0(\widetilde{C'}) \end{array}$$

Proof:
[DS05], Section 8.4,8.5.

□

4.5 Modular curves in characteristic p

Suppose N is a positive integer and p is a prime number that doesn't divide N. We will study the reduction of modular curves $X_1(N)$ and $X_0(N)$ at p.
First, we'll tackle the reduction of the moduli space at p. Let's use $\mathfrak{p}$ to represent a maximal ideal of $\overline{\mathbb{Q}}$ that sits above p. We know that if an elliptic curve E over $\overline{\mathbb{Q}}$ has good reduction at $\mathfrak{p}$, its j-invariant $j(E)$, belongs to $\overline{\mathbb{Z}}_{(\mathfrak{p})}$. This invariant will reduce at $\mathfrak{p}$ to $\widetilde{j(E)}$ in $\overline{\mathbb{F}}_p$. To make thing simpler, it is a convention to assume that j doesn't take the valeus 0 or 1728.
To express this, we'll adjust our notation with a prime symbol. So, the appropriate restriction of the moduli space $\mathrm{S}_1(N)$ over $\mathbb{Q}$ becomes:

$$\mathrm{S}_1(N)'_{\mathrm{gd}} = \left\{[E, Q] \in \mathrm{S}_1(N) : E \text{ has good reduction at } \mathfrak{p}, \widetilde{j(E)} \notin \{0, 1728\}\right\}.$$

Considering the characteristic p, the moduli space over $\overline{\mathbb{F}}_p$ is given by $\widetilde{\mathrm{S}}_1(N)$. This space comprises equivalence classes $[E, Q]$, where E is an elliptic curve over $\overline{\mathbb{F}}_p$ and Q is a point of order N on E. We'll also again make sure j doesn't take the values 0 or 1728:

$$\widetilde{\mathrm{S}}_1(N)' = \left\{[E, Q] \in \widetilde{\mathrm{S}}_1(N) : j(E) \notin \{0, 1728\}\right\}.$$

The reduction map is then given by:

$$\mathrm{S}_1(N)'_{\mathrm{gd}} \longrightarrow \widetilde{\mathrm{S}}_1(N)', \quad [E_j, Q] \mapsto \left[\widetilde{E}_j, \widetilde{Q}\right]$$

An interesting observation here is that any Weierstrass equation over $\overline{\mathbb{F}}_p$ can be lifted to $\overline{\mathbb{Z}}$. Moreover, the discriminant of this lift is another lift of the discriminant. This means that if the discriminant is non-zero in one space, it remains so in the other, ensuring elliptic curves in one space correspond to elliptic curves in the other. Another proposition ([DS05], Proposition 8.4.4(a)) confirms that each point of order N also has a corresponding lift. This ensures that the reduction map we derived is surjective.
When discussing the reduction of $X_1(N)$ at a prime p, we consider the universal elliptic curve $\widetilde{E}_j$ defined

over $\mathbb{F}_p(j)$. Due to our convention that j doesn't take the values 0 or 1728, we get a nice equation for the universal elliptic curve. This curve is given by:

$$\widetilde{E}_j : y^2 + xy = x^3 - \left(\frac{36}{j - 1728}\right) x - \left(\frac{1}{j - 1728}\right).$$

The discriminant for this curve is $j^2/(j-1728)^3$ and its invariant is j. When we speak of the points (x, y) on the curve $\widetilde{E}_j$, we are referring to those that lie within the projective plane $\mathbb{P}^2\left(\overline{\mathbb{F}_p(j)}\right)$.
Now, imagine a point Q on $\widetilde{E}_j$. If this point has an order N, it means that N times the point Q gives the zero point on the curve (notated as $[N]Q = 0_{\widetilde{E}}$). However, multiplying the point by any number less than N (between 0 and N) does not give the zero point.
The x-coordinate of the point Q, represented as $x(Q)$, has a minimal polynomial in the field $\mathbb{F}_p(j)[x]$ denoted by $\varphi_{1,N}$. Using this polynomial, we can define a new field:

$$\mathbb{K}_1(N) = \mathbb{F}_p(j)[x]/\left\langle \varphi_{1,N}(x) \right\rangle.$$

It's an interesting fact, which we accept without delving into the proof, that the intersection of the field $\mathbb{K}_1(N)$ and the algebraic closure of $\mathbb{F}_p$ is just $\mathbb{F}_p$. This implies that our new field $\mathbb{K}_1(N)$ can be thought of as a function field over $\mathbb{F}_p$.
The polynomial $\varphi_{1,N}$ also helps define a planar curve, which we'll call $\widetilde{X}_1(N)^{\text{planar}}$. This curve might not be smooth everywhere (it could be singular). The points on this curve can be visualized as pairs (j, x) within the plane $\overline{\mathbb{F}}_p^2$.
Lastly, there exists a birational equivalence over $\mathbb{F}_p$ connecting this planar curve to any non-singular projective algebraic curve with function field is isomorphic to $\mathbb{K}_1(N)$. Recall, the moduli space interpretation discussed in chapter 2. As discussed in the setup, we are essentially working over

Theorem 4.5.1 (Igusa's theorem)
For a positive integer N and a prime p (where $p \nmid N$), the modular curve $X_1(N)$ has good reduction at p. The function fields $\mathbb{F}_p\left(\widetilde{X}_1(N)\right)$ and $\mathbb{K}_1(N)$ are isomorphic. Additionally, we have the following commutative diagram:

$$\begin{array}{ccc} \mathrm{S}_1(N)'_{\mathrm{gd}} & \xrightarrow{\psi_1} & X_1(N) \\ \downarrow & & \downarrow \\ \widetilde{\mathrm{S}}_1(N)' & \xrightarrow{\widetilde{\psi_1}} & \widetilde{X_1}(N) \end{array}$$

Theorem 4.5.2
See, [DS05], section 8.6.

The diagram essentially dictates the compatibility between the reduction of modular curves and the reduction of moduli spaces.

The main takeaway is the surjective nature of the transformations, ensuring that all necessary points in the original space have corresponding points in the reduced space.
In the context of the Modularity Theorem, while we've primarily discussed $X_1(N)$, a similar line of reasoning is applicable for $X_0(N)$.

4.6 L-functions and Eicheler-Shimura relations

We briefly discuss the connection of modular forms and elliptic curves with L-functions.

Definition 4.6.1
Let Γ be a congruence subgroup. Let $f \in M_k(\Gamma_1(N))$ be a modular form, given by a q-expansion $f = \sum_{n=0}^{\infty} a_n q^n$.The L-function of f is the function of $s \in \mathbb{C}$ given formally as

$$L(f,s) = \sum_{n=1}^{\infty} a_n n^{-s}$$

Proposition 4.6.2
If $f \in S_k(\Gamma_1(N))$ is a cusp form then $L(f,s)$ converges absolutely for all s such that $\Re(s) > k/2+1$. If $f \in M_k(\Gamma_1(N))$ is not a cusp form then $L(f,s)$ converges absolutely for all s with $\Re(s) > k$.

Proof:
See, [Mas15], proposition 6.1.1. □

The following theorem gives a nice criterion for an eigenform to be a newform in terms of euler product expansions.

Theorem 4.6.3
Let $f \in M_k(\Gamma_0(N), \chi)$ be a modular form with q-expansion $f = \sum_{n>0} a_n q^n$. Then f is a normalized eigenform if and only if $L(f,s)$ has an Euler product expansion

$$L(f,s) = \prod_{p \text{ prime}} \left(1 - a_p p^{-s} + \chi(p) p^{k-1-2s}\right)^{-1}$$

Proof:
See, [Mas15], 6.1.2. □

Now, consider an elliptic curve E over $\mathbb{Q}$. Let N_E be the conductor of E defined in chapter 1.

Definition 4.6.4
We define the L-function attached to ane elliptic curve E via the following Euler product:

$$L(E,s) = \prod_{p | N_E} \left(1 - a_p(E) p^{-s}\right)^{-1} \prod_{p \nmid N_E} \left(1 - a_p(E) p^{-s} + p^{1-2s}\right)^{-1}, \quad \Re(s) > 3/2$$

where $a_p(E) = 1 + p - \#E(\mathbb{F}_p)$.

The Eichler-Shimura relations states that elliptic curves arise from modular forms.

Theorem 4.6.5
Let $f \in S_2(\Gamma_0(N))$ be a normalized eigenform whose Fourier coefficients $a_n(f)$ are all integers. Then there exists an elliptic curve E_f defined over $\mathbb{Q}$ such that $L(E_f, s) = L(f,s)$.

Idea behind constructing such an elliptic curve:
Consider the differential form $\omega_f = 2\pi i f(z) dz$. Define $\mathbb{H}^*$ as the union of the upper half-plane $\mathbb{H}$ and the projective line $\mathbb{P}^1(\mathbb{Q})$. For a point τ in $\mathbb{H}^*$, we associate a complex number $\varphi(\tau)$ defined by the integral:

$$\varphi(\tau) = \int_{\infty}^{\tau} \omega_f \in \mathbb{C}.$$

Relation with the Modular Group $\Gamma_0(N)$:
Given $\gamma \in \Gamma_0(N)$, consider the quantity β_γ defined as:

$$\beta_\gamma = \varphi(\gamma\tau) - \varphi(\tau) = \int_\tau^{\gamma\tau} \omega_f.$$

We can show that β_γ does not depend on the choice of τ through the following calculation:

$$\begin{aligned}\int_\tau^{\gamma\tau} \omega_f &= \int_\tau^{\infty} \omega_f + \int_\infty^{\gamma\infty} \omega_f + \int_{\gamma\infty}^{\gamma\tau} \omega_f \\ &= \int_\tau^{\infty} \omega_f + \int_\infty^{\gamma\infty} \omega_f + \int_\infty^{\tau} \omega_f \\ &= \int_\infty^{\gamma\infty} \omega_f.\end{aligned}$$

Define Λ_f as the set of complex numbers:

$$\Lambda_f = \left\{ \beta_\gamma = \int_\infty^{\gamma\infty} \omega_f \mid \gamma \in \Gamma_0(N) \right\} \subset \mathbb{C}.$$

With Λ_f, we have a well-defined mapping:

$$\Gamma_0(N)\backslash\mathbb{H}^* \to \mathbb{C}/\Lambda_f.$$

It can be demonstrated that Λ_f forms a lattice. We define E_f as the elliptic curve corresponding to the complex torus $\mathbb{C}/\Lambda_f$.
While constructing E_f is straightforward, proving that it is defined over $\mathbb{Q}$ and that its L-function $L(E_f, s)$ equals $L(f, s)$ requires more advanced techniques and deeper understanding of the relationship between elliptic curves and modular forms. This is beyond the scope of this thesis.
Section 8.7 [DS05], also describes Eichler-Shimura relation at the level of Picard groups. We encourage the readers to go through the section 8.7 for a more algebraic and geometric discussion of Eichler-Shimura relations.

5 Jacobian of Curves and Abelian Varieties

The study of algebraic curves and their properties has been a central theme in algebraic geometry for centuries. The concept of the Jacobian was first introduced in the mid-nineteenth century by mathematicians like Carl Gustav Jacobi and Bernhard Riemann as they sought to unify the theory of elliptic functions. We will briefly discuss Abelian Varieties over $\mathbb{C}$ and then move on to the discussion of Jacobians, describe what they are and in the end state another version of the Modularity theorem in terms of Jacobians.
We will closely follow the book [DS05] and [HS00]. We will sometimes closely follow proofs from either of these sources for the sake of completeness. Due to the extreme complexity of the overall topic and to demonstrate all of this within 6 months, it was inevitable to closely follow some of the proofs but written in my own words. Having said that, I must add that in many places, I have expanded on my own, giving more details.

5.1 Abelian Varieties over $\mathbb{C}$

Definition 5.1.1 (Abelian Variety)
An abelian variety is an algebraic group that is also a projective variety.

Definition 5.1.2
Consider a complex vector space $V = \mathbb{C}^g$, where $g \geq 1$, and let Λ be a lattice of full rank. This means that $\Lambda \cong \mathbb{Z}^{2g}$ as abstract group, and that $\mathbb{R} \otimes_{\mathbb{Z}} \Lambda = V$. The quotient

$$X = V/\Lambda$$

is a complex manifold of complex dimension g. Any such X is called a complex torus.

Remark 5.1.3
1. With the help of Lie theory, it can be shown that the set of complex points of an abelian variety forms a complex torus. In other words, the set of complex points of an abelian variety is isomorphic to $\mathbb{C}^g/\Lambda$ for some lattice Λ as complex analytic varieties.
2. Let $(\lambda_1, \ldots, \lambda_{2g})$ be a $\mathbb{Z}$-basis for Λ. Then, the set

$$\left\{ \sum_{i=1}^{2g} t_i \lambda_i \mid 0 \leq t_i \leq 1 \forall i \right\}$$

is a fundamental domain for the action of Λ; therefore, X is compact.

Definition 5.1.4
A Hermitian form is a map

$$H : \mathbb{C}^g \times \mathbb{C}^g \to \mathbb{C}$$

that is linear with respect to the first set of variables and satisfies

$$H(z, w) = \overline{H(w, z)}.$$

Definition 5.1.5

A Riemann form with respect to a lattice Λ is a Hermitian form on $\mathbb{C}^g \times \mathbb{C}^g$ whose imaginary part takes integer values when restricted to $\Lambda \times \Lambda$. A Riemann form is called nondegenerate if it is positive definite.

Theorem 5.1.6
Let Λ be a lattice in $\mathbb{C}^g$. The complex torus $\mathbb{C}^g/\Lambda$ is an abelian variety if and only if there exists a positive definite Hermitian form on $\mathbb{C}^g \times \mathbb{C}^g$ whose imaginary part takes integer values when restricted to $\Lambda \times \Lambda$.

Proof:
See, [HS00](Section A.5). □

Corollary 5.1.7
Every torus of dimension one is an abelian variety.

Proof:
Indeed, a lattice $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ be a lattice in $\mathbb{C}$. We define a function on $\mathbb{C} \times \mathbb{C}$ given by,

$$H(z, w) = \frac{z\bar{w}}{\text{Im}\,(\omega_1\bar{\omega}_2)}$$

We prove that H is indeed a Hermitian form.
Taking the conjugate of $H(w, z)$, we have:

$$\overline{H(w, z)} = \frac{\overline{w\bar{z}}}{\text{Im}(\omega_1\bar{\omega}_2)} = \frac{\bar{w}z}{\overline{\text{Im}(\omega_1\bar{\omega}_2)}}$$

Since $\overline{w\bar{z}} = \bar{w}z$ and $\text{Im}(\omega_1\bar{\omega}_2)$ is a real number, its conjugate is itself.
Thus,

$$\overline{H(w, z)} = \frac{\bar{w}z}{\text{Im}(\omega_1\bar{\omega}_2)} = H(z, w)$$

This shows that H is indeed Hermitian.
A Hermitian form H is positive definite if $H(z, z) > 0$ for all $z \neq 0$. Let's check this for our H:
For any non-zero z,

$$H(z, z) = \frac{z\bar{z}}{\text{Im}(\omega_1\bar{\omega}_2)} = \frac{|z|^2}{\text{Im}(\omega_1\bar{\omega}_2)}$$

Since $|z|^2$ is always positive for $z \neq 0$, the sign of $H(z, z)$ depends on the sign of $\text{Im}(\omega_1\bar{\omega}_2)$. As discussed in the first chapter, generators ω_1 and ω_2 of a lattice Λ, are typically chosen such that $\text{Im}(\omega_1\bar{\omega}_2) = \text{Im}(\frac{\omega_1}{\omega_2}) > 0$. Thus, H is positive definite, since $\text{Im}(\omega_1\bar{\omega}_2)$ is positive.

It remains to show that H is indeed a Riemann form.
Any two points z, ω in the lattice z can be expressed as:

$$z = m_1\omega_1 + m_2\omega_2$$

$$\omega = n_1\omega_1 + n_2\omega_2$$

where m_1, m_2, n_1, n_2 are integers.

Consider,

$$H(z,\omega) = \frac{(m_1\omega_1 + m_2\omega_2)(n_1\bar{\omega}_1 + n_2\bar{\omega}_2)}{\mathrm{Im}(\omega_1\bar{\omega}_2)}$$

Expanding this, we have:

$$H(z,\omega) = \frac{m_1n_1\omega_1\bar{\omega}_1 + m_1n_2\omega_1\bar{\omega}_2 + m_2n_1\omega_2\bar{\omega}_1 + m_2n_2\omega_2\bar{\omega}_2}{\mathrm{Im}(\omega_1\bar{\omega}_2)}$$

The imaginary part of $H(z,\omega)$ is given by:

$$\mathrm{Im}(H(z,\omega)) = \mathrm{Im}\left(\frac{m_1n_1\omega_1\bar{\omega}_1 + m_1n_2\omega_1\bar{\omega}_2 + m_2n_1\omega_2\bar{\omega}_1 + m_2n_2\omega_2\bar{\omega}_2}{\mathrm{Im}(\omega_1\bar{\omega}_2)}\right)$$

Note that as $\omega_1\bar{\omega}_1$ and $\omega_2\bar{\omega}_2$ are real numbers, they do not contribute to the imaginary part. Thus,

$$\mathrm{Im}(H(z,\omega)) = \mathrm{Im}\left(\frac{m_1n_2\omega_1\bar{\omega}_2 + m_2n_1\omega_2\bar{\omega}_1}{\mathrm{Im}(\omega_1\bar{\omega}_2)}\right)$$

The key point here again as before, $\omega_1\bar{\omega}_2$ and $\omega_2\bar{\omega}_1$ are conjugates of each other, so their imaginary parts are negatives of each other. Therefore, the imaginary part of $\omega_1\bar{\omega}_2$ is $-\mathrm{Im}(\omega_2\bar{\omega}_1)$.

$$\mathrm{Im}(H(z,\omega)) = \mathrm{Im}\left(\frac{m_1n_2\omega_1\bar{\omega}_2 - m_2n_1\omega_2\bar{\omega}_1}{\mathrm{Im}(\omega_1\bar{\omega}_2)}\right)$$

Thus, we have,

$$\mathrm{Im}(z,\omega) = m_1n_2 - n_1m_2.$$

Thus, $\mathrm{Im}(H(z,\omega))$ is an integer for all $z,\omega \in \Lambda$,
Now, theorem 5.1.6 gives us the claim. □

Remark 5.1.8

1. Note that to make idea about period matrices clear, the period matrix related to lattice Λ is a matrix whose columns correspond to a basis Λ of the lattice Λ expanded out using a basis of V. The two relations in Theorem 5.2.3 are called Riemann's bilinear period relations. Some texts call this a definition for a complex torus to be an abelian variety. If the dimension of the lattice is greater than one, then many tori do not admit a nonzero Riemann form. For example, let e_1, e_2, e_3, e_4 be vectors in $\mathbb{C}^2$ whose coordinates are all algebraically independent over $\mathbb{Q}$, and let Λ be the lattice that they span. Then the torus $\mathbb{C}^2/\Lambda$ is not an abelian variety.

2. Let τ be a $g \times g$ symmetric matrix whose imaginary part is positive definite. The torus $\mathbb{C}^g/\left(\mathbb{Z}^g + \tau\mathbb{Z}^g\right)$ is an abelian variety. We will use this in the next section to show that the Jacobian variety of a curve is an abelian variety.

Proposition 5.1.9

Let $A = V/\Lambda$ be an abelian variety, and let B be an abelian subvariety of A. Then another abelian subvariety exists C such that $B + C = A$ and $B \cap C$ is finite. In other words, the map

$$B \times C \longrightarrow A, \quad (b,c) \longmapsto b + c$$

is an isogeny.

Proof:
[HS00] Consider a nondegenerate Riemann form, H, for the given abelian variety A, and E be its imaginary part. We naturally identify the tangent space of A with V. This identification is made using the fact that the tangent space of an abelian variety at a point is isomorphic to the dual of the cotangent space at that point, which is in turn isomorphic to V itself. The isomorphism between the tangent space and the dual cotangent space is given by the Riemann form H.

We define another vector space V_1 as the tangent space of another abelian variety B. Additionally, we set Λ_1 to be the intersection of V_1 with the lattice Λ so that B can be expressed as the quotient V_1/Λ_1. Now, let's consider the vector space V_2, which is the orthogonal complement of V_1 with respect to the Riemann form H. In other words, it consists of all vectors in V that satisfy $H(v, w) = 0$ for all w in V_1. It's important to note that if w belongs to V_1, then iw also belongs to V_1 since V_1 is a complex vector space. Using the definition of the Riemann form, we can express V_2 alternatively as:

$$V_2 = \{v \in V \mid E(v, w) = 0 \text{ for all } w \in V_1\}$$

Now, consider the intersection Λ_2, which is defined as the intersection of the lattice Λ with V_2. In other words, Λ_2 consists of all lattice points x in Λ such that $E(x, y) = 0$ for all y in Λ_1. The nondegeneracy of E and the fact that Λ_1 is a lattice in V_1 imply that Λ_2 has a certain rank given by:

$$\operatorname{rank} \Lambda_2 = \operatorname{rank} \Lambda - \operatorname{rank} \Lambda_1 = 2 \dim_{\mathbb{C}} V_2$$

This implies that Λ_2 is a lattice in V_2, and as a result, we can define an abelian subvariety C of A as $C = V_2/\Lambda_2$. Using the decomposition $V = V_1 \oplus V_2$, we can deduce that A can be expressed as the sum of B and C such that their intersection, $B \cap C$, consists of only finitely many points. □

Definition 5.1.10
A torus is said to be simple if it does not contain any nontrivial subtori.

Corollary 5.1.11
Any abelian variety A is isogenous to a product of the form

$$A_1^{n_1} \times \cdots \times A_s^{n_s},$$

where the A_i 's are simple, pairwise nonisogenous abelian varieties.

Proof:
The follows by induction. It is clear that an abelian variety of dimension 1 is simple. Now, for an abelian variety A of dimension n, there are two cases to consider. If A is simple itself, there is nothing to prove. But if not, by 5.1.9 we get an isogenous decomposition of A. Now, by applying the induction hypothesis, the claim follows. □

5.2 Jacobians over $\mathbb{C}$

In this section, we will sketch the construction of the Jacobian of a compact Riemann surface. The Jacobian will be a complex torus carrying a nondegenerate Riemann form, an abelian variety.The Jacobian is

one of the central tools in studying curves and is why abelian varieties enter the picture. If we combine the theory of Jacobians with Theorem 5.1.6, we get a nice proof that all compact Riemann surfaces can be embedded into projective space.

Historical Remark:
The concept of abelian varieties originated in the 19th century when attempting to compute or describe integrals of the form $\int R(t, \sqrt{P(t)})dt$, where R is a rational function and P is a polynomial. More broadly, integrals $\int R(t, s)dt$ were investigated, subject to an algebraic relation $P(s, t) = 0$. These integrals were eventually termed abelian integrals.

For instance, let's consider the integral $u = \int_0^x 1/\sqrt{1-t^2}dt$. It is widely known from calculus that $u = \sin^{-1}(x)$, making it more convenient to examine the inverse function of u. In simpler terms, we work with the function S that satisfies $x = S(u)$, revealing that S is, in fact, the sine function. Remarkably, it has a periodic nature, $S(u + 2\pi) = S(u)$, and adheres to a differential equation $S(u)^2 + S'(u)^2 = 1$. To be precise, the mapping $u \to (S(u), S'(u))$ parameterizes the curve $x^2 + y^2 = 1$.
A significant breakthrough by Abel unveiled that when delving into integrals of the form $u = \int_0^x 1/\sqrt{Q(t)}dt$ where $\deg(Q) \geq 5$, an expansion into additional variables becomes necessary.

Consider a polynomial P of degree $2g+2$ with distinct roots. We construct a Riemann surface X by gluing two affine curves: $y^2 = P(x)$ and $v^2 = P^*(u) = u^{2g+2}P(u^{-1})$, via the mapping $(u, v) \mapsto (x^{-1}, yx^{-1-g})$. On X, the set $\{dx/y, xdx/y, \ldots, x^{g-1}dx/y\}$ forms a basis for the space of regular differentials. See, [HS00], exercise A.4.2.

Consider a differential form $\omega = dx/y$. Let γ be a path on X connecting the points $(a, \sqrt{P(a)})$ and $(b, \sqrt{P(b)})$. The line integral $\int_\gamma \omega$ on X then defines the concept of the multi-valued integral $\int_a^b 1/\sqrt{P(t)}dt$. The choice of γ resolves any ambiguity inherent in the integral's definition. This explains the motivation behind considering differentials, paths and integrals.

We build on this and elaborate on how the integral depends on the chosen path, we turn to the theory of homology. Suppose $\gamma_1, \ldots, \gamma_{2g}$ form a homology basis for $H_1(X, \mathbb{Z})$, the first homology group of X.

Given two paths γ and γ' that connect the same points A and B on X, the path γ followed by γ' in reverse defines a closed path. As such, this loop is homologous to a sum of the form $\sum m_i\gamma_i$, where m_i are integers. Hence, for any regular differential 1-form on X, the integral along γ minus the integral along γ' equals the sum of integrals along these homology basis elements, each multiplied by its respective coefficient m_i.

Definition 5.2.1
Let X be a Riemann surface. Let $\gamma_1, \ldots, \gamma_{2g}$ form a homology basis for $H_1(X, \mathbb{Z})$, the first homology group of X. Let $\omega_1, \ldots, \omega_g$ be a basis of the vector space of regular 1-forms, and let Λ be the $g \times 2g$ matrix with entries

$$\Lambda = \left(\Lambda_i^j\right)_{\substack{1\leq i\leq 2g \\ 1\leq j\leq g}} = \left(\int_{\gamma_i} \omega_j\right)_{\substack{1\leq i\leq 2g \\ 1\leq j\leq g}}.$$

We call Λ a period matrix of X, and we let L_Λ be the $\mathbb{Z}$-module generated by the columns of Λ.

It is important to note that, choosing a different basis for the homology and the space of 1-forms will give another period matrix Λ'.

Theorem 5.2.2 (Riemann's period relations)
Let $\gamma_1, \ldots, \gamma_{2g}$ be a basis for the homology group $H_1(X, \mathbb{Z})$, chosen to satisfy the following property: For each $1 \leq i \leq g$,

$$\gamma_i \cdot \gamma_j = \begin{cases} 1 & \text{if } j = i + g \\ 0 & \text{otherwise.} \end{cases}$$

Then for any nonzero regular 1 -forms ω and ω',

$$\sum_{k=1}^{g} \left(\int_{\gamma_k} \omega \int_{\gamma_{g+k}} \omega' - \int_{\gamma_k} \omega' \int_{\gamma_{g+k}} \omega \right) = 0.$$
$$\sqrt{-1} \sum_{k=1}^{g} \left(\overline{\int_{\gamma_{g+k}} \omega} \int_{\gamma_k} \omega - \int_{\gamma_{g+k}} \omega \overline{\int_{\gamma_k} \omega} \right) > 0.$$

Proof:
See, [GH14],Page, 231. □

We also have another criterion/definition for a complex torus to be an Abelian Variety.

Theorem 5.2.3
Let V be a complex vector space of dimension g, and Λ a lattice of full rank. Fix a basis $(e_1, \ldots, e_g)$ for V, and a basis $(\lambda_1, \ldots, \lambda_{2g})$ for Λ. Let P be the period matrix of Λ, i.e. the $g \times 2g$ matrix such that $X \cong \mathbb{C}^g / P\mathbb{Z}^{2g}$. Then, X is projective if and only if there is a non-degenerate alternating matrix $E \in M_{2g}(\mathbb{Z})$ such that
1)$PE^{-1}P^T = 0$;
2) $iPE^{-1}\bar{P}^T > 0$.

Remark 5.2.4
Let us address the decomposition of the lattice Λ into two components Λ_1 and Λ_2, each being a $g \times g$ matrix, we can approach Riemann's relations from a perspective of matrices:

$$\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1 \quad \text{and} \quad -i(\overline{\Lambda_1}\Lambda_2 - \overline{\Lambda_2}\Lambda_1) \text{is positive definite.}$$

Assume that $\Lambda_1 Y = 0$. This assumption leads us to the following equation:

$$\bar{Y}(-i(\overline{\Lambda_1}\Lambda_2 - \overline{\Lambda_2}\Lambda_1))Y = 0.$$

Given the positive definiteness of the matrix involved, the only solution to this equation is $Y = 0$. Therefore, Λ_1 must be invertible. This insight allows for a transformation from differential forms into Matrices, by considering Λ_1 as the identity matrix, and by redefining Λ_2 as a new matrix $\tau = \Lambda_1^{-1}\Lambda_2$. Under this new framework, the period matrix Λ becomes (I, τ), and Riemann's relations confirm that τ is symmetric, with its imaginary part, $\mathrm{Im}(\tau)$, being positively definite.

Corollary 5.2.5

The column vectors of Λ generate a lattice L_Λ inside $\mathbb{C}^g$.

Proof:
We use the previous discussion to get the new basis, we have the lattice Λ of the form, $\mathbb{Z}^g + \tau\mathbb{Z}^g$. □

Finally, we are prepared to define the Jacobian of a Riemann surface.

Definition 5.2.6
Let X be a be a Riemann surface. Let V^* denote the dual vector space of a complex vector space V, let $H^0\left(X, \Lambda^1_X\right)$ be the vector space of regular differentials on X, and let $H_1(X, \mathbb{Z})$ be the homology group of X. We can identify $H_1(X, \mathbb{Z})$ as a lattice in $H^0\left(X, \Lambda^1_X\right)^*$ via the map

$$H_1(X, \mathbb{Z}) \longrightarrow H^0\left(X, \Lambda^1_X\right)^*, \quad \gamma \longmapsto \left(\omega \mapsto \int_\gamma \omega\right)$$

Then the Jacobian of X is equal to

$$\mathrm{Jac}(X) = H^0\left(X, \Lambda^1_X\right)^* / H_1(X, \mathbb{Z}).$$

Remark 5.2.7
One can explicitly construct a Riemann form with respect to the lattice L_Λ. We may assume the lattice to be normalized, $L_\Lambda = \mathbb{Z}^g + \tau\mathbb{Z}^g$.
Consider, $H(z, w) = z^t \operatorname{Im}(\tau)^{-1}\bar{w}$. This form is positive definite from Riemann's relations, and if k, ℓ, m, n are vectors with integer coordinates, then $\operatorname{Im} H(m + \tau n, k + \tau\ell) = m^t\ell - n^t k$ is an integer.
Hence H is a Riemann form. By Theorem 5.1.6, Jacobian is a projective Variety.
More precisely, for each fixed basepoint $a \in X$ we define a holomorphic map

$$\Phi_a : X \longrightarrow \mathrm{Jac}(X) = \mathbb{C}^g/L_\Lambda, \quad b \longrightarrow \left(\int_a^b \omega_1, \ldots, \int_a^b \omega_g\right) \bmod L_\Lambda.$$

Definition 5.2.8
The map Φ_a is called the Jacobian embedding of X.

Remark 5.2.9
Moreover, we observe that up to translation, the map Φ_a is independent of a. Thus $\Phi_{a'}(b) = \Phi_a(b) - \Phi_a(a')$. So if we extend Φ_a linearly to the divisor group, then it will be completely independent of a on the group of divisors of degree zero. We denote this map by Φ,

$$\Phi : \mathrm{Div}^0(X) \longrightarrow \mathrm{Jac}(X), \quad \sum n_i (b_i) \longmapsto \sum n_i \Phi_a (b_i).$$

The map Φ is very important and the same can also be seen from the following celebrated theorem due to Abel and Jacobi:

Theorem 5.2.10
The map $\Phi : \mathrm{Div}^0(X) \longrightarrow \mathrm{Jac}(X)$ is surjective, and its kernel is exactly the subgroup of principal divisors.

Proof:
See, [HS00],section A.6. □

Corollary 5.2.11
Assume that X has genus $g \geq 1$. Then the map $\Phi_a : X \to \mathrm{Jac}(X)$ is an embedding.
In particular, if X has genus one, then X is isomorphic to its Jacobian. Further, a divisor $\sum n_i (P_i)$ will be principal if and only if $\sum n_i = 0$ and $\sum n_i P_i = 0$.

Proof:
See, [HS00],section A.6. □

Theorem 5.2.12 (Abel's theorem)
The map Φ descends to an isomorphism of groups, $\Phi : \mathrm{Pic}^0(X) \longrightarrow \mathrm{Jac}(X)$

Proof:
See, [HS00],section A.6. □

5.3 Modular Jacobians and Hecke operators

The main idea of this section is to explore the relationships between Jacobians of modular curves and the action of double coset operators.

Let's recall the double coset operator. Assume we have two congruence subgroups, Γ_1 and Γ_2, belonging to $\mathrm{SL}_2(\mathbb{Z})$. Additionally, suppose we have an element α from $\mathrm{GL}_2^+(\mathbb{Q})$. Defining Γ_3 as $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2$, we can identify representatives, denoted as $\{\gamma_{2,j}\}$, for which $\Gamma_3\backslash\Gamma_2$ can be represented as $\bigcup_j \Gamma_3\gamma_{2,j}$. Using these, we can determine another set of representatives, $\{\beta_j\} = \{\alpha\gamma_{2,j}\}$, satisfying $\Gamma_1\alpha\Gamma_2 = \bigcup_j \Gamma_1\beta_j$. Introducing Γ_3' as $\Gamma_3' = \alpha\Gamma_3\alpha^{-1} = \Gamma_1 \cap \alpha\Gamma_2\alpha^{-1}$, we have the following:

$$\Gamma_2 \longleftarrow \Gamma_3 \xrightarrow{\sim} \Gamma_3' \longrightarrow \Gamma_1$$

Here, the isomorphism between groups is given by $\gamma \mapsto \alpha\gamma\alpha^{-1}$, with the remaining arrows are simply inclusions.
Now, these groups correspond to modular curves denoted by X_1, X_2, X_3, and X_3', giving us:

$$X_2 \xrightarrow{\pi_2} X_3 \xrightarrow{\sim} X_3' \xrightarrow{\pi_1} X_1$$

The map between these modular curves is described by $\Gamma_3\tau \mapsto \Gamma_3'\alpha(\tau)$, denoted as α. Consider a point on X_2. When mapped through $\pi_1 \circ \alpha \circ \pi_2^{-1}$, this point corresponds to multiple points on X_1, shown as:

$$\Gamma_2\tau \xrightarrow{\pi_2^{-1}} \{\Gamma_3\gamma_{2,j}(\tau)\} \xrightarrow{\alpha} \left\{\Gamma_3'\beta_j(\tau)\right\} \xrightarrow{\pi_1} \{\Gamma_1\beta_j(\tau)\}$$

In this mapping, π_2^{-1} associates a point with its corresponding points on the layer above. The multiplicity of each associated point depends on its ramification degree, given by

$$\pi_2^{-1}(x) = \{e_y \cdot y : y \in X_3, \pi_2(y) = x\}.$$

This essentially gives an interpretation of the double coset operator as a reverse map of divisor groups:

$$[\Gamma_1\alpha\Gamma_2]_2 : \mathrm{Div}(X_2) \longrightarrow \mathrm{Div}(X_1),$$

We now consider the linear extension of the map given by $\Gamma_2\tau \mapsto \sum_j \Gamma_1\beta_j(\tau)$ that takes elements from X_2 into the divisor group $\mathrm{Div}(X_1)$. We can thus interpret the operation on divisor groups involving double cosets as a sequentialapplication of both forward and reverse maps.

$$[\Gamma_1\alpha\Gamma_2]_2 = (\pi_1)_D \circ \alpha_D \circ \pi_2^D.$$

This implies that the operation can be carried over to the level of Picard groups,

$$[\Gamma_1\alpha\Gamma_2]_2 = (\pi_1)_P \circ \alpha_P \circ \pi_2^P : \mathrm{Pic}^0(X_2) \longrightarrow \mathrm{Pic}^0(X_1)$$

The effect of this map on an element represented by a formal sum of τ's each scaled by an integer n_τ within Γ_2 is expressed as:

$$[\Gamma_1\alpha\Gamma_2]_2 \left[\sum_\tau n_\tau \Gamma_2\tau\right] = \left[\sum_\tau n_\tau \sum_j \Gamma_1\beta_j(\tau)\right].$$

Let us discuss this in terms of Jacobians and modular forms, let us denote Γ as a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$. We know from our brief exposition into the differential forms, there is a one-to-one correspondence between holomorphic differentials $\Omega^1_{\mathrm{hol}}(X(\Gamma))$ and weight 2 cusp forms $\mathcal{S}_2(\Gamma)$. In essence, each cusp form f is associated with a unique holomorphic differential $\omega(f)$, $f(\tau)d\tau$ defined $X(\Gamma)$.

Consequently, the map $\omega : \mathcal{S}_2(\Gamma) \to \Omega^1_{\mathrm{hol}}(X(\Gamma))$ is a linear isomorphism. This identification extends to their dual spaces through the mapping $\omega^\wedge$ and we have:

$$\mathcal{S}_2(\Gamma)^\wedge = \omega^\wedge(\Omega^1_{\mathrm{hol}}(X(\Gamma))^\wedge).$$

Within the framework of modular forms, the subgroup $\mathrm{H}_1(X(\Gamma), \mathbb{Z})$ is associated with $\mathcal{S}_2(\Gamma)^\wedge$, and would be denoted as $\omega^\wedge(\mathrm{H}_1(X(\Gamma), \mathbb{Z}))$ in differential terms.

In this setting, the Jacobian of $X(\Gamma)$ is appropriately redefined as a quotient of the dual space of the weight 2 cusp forms. This gives us the following definition.

Definition 5.3.1
For Γ a congruence subgroup of the full Modular group, we have that

$$\mathrm{Jac}(X(\Gamma)) = \mathcal{S}_2(\Gamma)^\wedge / \mathrm{H}_1(X(\Gamma), \mathbb{Z})$$

5.4 Abelian Varieties and Modularity

Let us recall the weight-2 Hecke operators, $T = T_p$ and $T = \langle d \rangle$, act on the dual space via:

$$T : \mathcal{S}_2\left(\Gamma_1(N)\right)^\wedge \longrightarrow \mathcal{S}_2\left(\Gamma_1(N)\right)^\wedge, \quad \varphi \mapsto \varphi \circ T.$$

This action then descends to the quotient, $\mathrm{J}_1(N)$. As such, these operators serve as endomorphisms on the kernel $\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)$, which is a finitely generated Abelian group. Consequently we have that the characteristic polynomial, $f(x)$, for T_p when acting on $\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)$ has integer coefficients and is monic. Furtheremore, We know from linear algebra that operators satisfy $f\left(T_p\right) = 0$ for $\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)$.

Besides, due to the C-linear nature of T_p, this equality holds true for both $\mathcal{S}_2(\Gamma_1(N))^\wedge$ and $\mathcal{S}_2(\Gamma_1(N))$. As a consequence, the eigenvalues of T_p must satisfy $f(x)$, thus being algebraic integers. This essentially proves:

Theorem 5.4.1
If $f \in \mathcal{S}_2(\Gamma_1(N))$ is a normalised eigenform for the Hecke operators T_p, then the corresponding eigenvalues $a_n(f)$ are algebraic integers.

To be a bit more general, an idea would be to see the Hecke operators within an algebraic structure rather than inside just a set. This hints us towards the definition of Hecke Algebras. When we talk about the Hecke algebra over $\mathbb{Z}$, we're essentially referring to the endomorphisms of $\mathcal{S}_2(\Gamma_1(N))$, defined over $\mathbb{Z}$ by the inclusion of Hecke operators. To define precisely, we have the following definition.

Definition 5.4.2
The Hecke Algebra over $\mathbb{Z}$ is the algebra of endomorphisms generated over $\mathbb{Z}$ by the Hecke operators. Set theoretically, it is given by

$$\mathbb{T}_{\mathbb{Z}} = \mathbb{Z}\left[\left\{T_n, \langle n\rangle : n \in \mathbb{Z}^+\right\}\right].$$

Similarly, The Hecke Algebra $\mathbb{T}_{\mathbb{C}}$ can be defined over $\mathbb{C}$.
Each level N has a distinct Hecke algebra. We often leave out N from our notation since it's implied within the context. From this point onwards, the text will be mainly algebraic instead of focusing particularly on analytic aspects of objects like modular forms or cusp forms.

If we consider $f(\tau) = \sum_{n=1}^{\infty} a_n(f)q^n$ as an eigenform, the mapping:

$$\lambda_f : \mathbb{T}_{\mathbb{Z}} \longrightarrow \mathbb{C}, \quad Tf = \lambda_f(T)f$$

has its image as a finitely generated $\mathbb{Z}$-module. This is because if we view $\mathbb{Z}$-module $\mathbb{T}_{\mathbb{Z}}$ as a ring of endomorphisms of the finitely generated free $\mathbb{Z}$-module $\mathrm{H}_1(X_1(N), \mathbb{Z})$ then it is finitely generated as well. This follows from the following lemma.

Lemma 5.4.3
Let M be a free $\mathbb{Z}$-module of rank r. Show that the ring of endomorphisms of M is a free $\mathbb{Z}$-module of rank r^2, and so any subring is a free $\mathbb{Z}$-module of finite rank.

Proof:
Observe as in the case of a finite-dimensional vector space, Given a free $\mathbb{Z}$-module M of rank r with a basis $\{e_1, e_2, ..., e_r\}$, any endomorphism $f \in \mathrm{End}_{\mathbb{Z}}(M)$ can be uniquely determined by its action on the basis elements. That is, for each basis element e_i, the image $f(e_i)$ can be expressed as a $\mathbb{Z}$-linear combination of the basis elements:

$$f(e_i) = a_{1i}e_1 + a_{2i}e_2 + ... + a_{ri}e_r$$

where each $a_{ji} \in \mathbb{Z}$. The above expression for $f(e_i)$ essentially gives us the columns of a matrix representation of the endomorphism f with respect to the basis $\{e_1, ..., e_r\}$. There will be r such columns, each with r entries from $\mathbb{Z}$, making up a total of r^2 entries. Consider the set of all such matrices that is a $\mathbb{Z}$ module isomorphic to $\mathbb{Z}^{r^2}$.

The isomorphism between $\mathrm{End}_{\mathbb{Z}}(M)$ and $\mathbb{Z}^{r^2}$ is given by the correspondence between each endomorphism

and its $r \times r$ matrix representation with respect to the basis of M. Every such matrix corresponds to a unique endomorphism, and every endomorphism can be uniquely represented by such a matrix.
To precisely state a basis of the ring of endomorphisms of M as $\mathbb{Z}-$ module, consider the r^2 endomorphisms that correspond to the elementary matrices where each matrix has a single entry of 1 in a unique position and 0s elsewhere. These elementary matrices E_{ij} are $r \times r$ matrices where the entry in the i-th row and j-th column is 1 and all other entries are 0. The corresponding endomorphism f_{ij} maps the basis vector e_j to the basis vector e_i and all other basis vectors to 0:

$$f_{ij}(e_k) = \begin{cases} e_i & \text{if } j = k, \\ 0 & \text{otherwise.} \end{cases}$$

Since any $r \times r$ matrix over $\mathbb{Z}$ is can expressed uniquely as a $\mathbb{Z}-$linear combination of these matrices, we have that the matrices that these elementary matrices correspond to form a basis of $\mathrm{End}_{\mathbb{Z}}(M)$. Therefore, $\mathrm{End}_{\mathbb{Z}}(M)$ is free of rank r^2, since its elements can be put in a one-to-one correspondence with the elements of a free $\mathbb{Z}$-module of rank r^2, which is $\mathbb{Z}^{r^2}$. □

Given the image $\mathbb{Z}[\{a_n(f) : n \in \mathbb{Z}^+\}]$, we have that despite of infinitely many eigenvalues, it is still a $\mathbb{Z}$-module of finite rank. More precisey we have the following proposition.

Proposition 5.4.4
Consider $f(\tau) = \sum_{n=1}^{\infty} a_n(f)q^n$ as an eigenform, consider the map:

$$\lambda_f : \mathbb{T}_{\mathbb{Z}} \longrightarrow \mathbb{C}, \quad Tf = \lambda_f(T)f.$$

Let

$$I_f = \ker(\lambda_f) = \{T \in \mathbb{T}_F : Tf = 0\}$$

be the kernel of the given mapping. Then we have a $\mathbb{Z}$-module isomorphism

$$\mathbb{T}_{\mathbb{Z}}/I_f \xrightarrow{\sim} \mathbb{Z}[\{a_n(f)\}]$$

Proof (Rough sketch)*:*
Let $f \in \mathcal{S}_2(\Gamma_1(N))$ be a normalized eigenform. Thus $f \in \mathcal{S}_2(N, \chi)$ for some Dirichlet character $\chi : (\mathbb{Z}/N\mathbb{Z})^* \longrightarrow \mathbb{C}^*$ and $\lambda_f(\langle d \rangle) = \chi(d)$ for all $d \in (\mathbb{Z}/N\mathbb{Z})^*$. The map λ_f gives a $\mathbb{Z}$-module surjective map onto its image $\mathbb{Z}[\{a_n(f), \chi(d)\}]$. The surjectivity is clear due to the fact that the hecke algebra is generated by diamond operators and T_p operators together with the fact that eigenvalues of T_p operators are given by the fourier coefficients of f, and $\lambda_f(\langle d \rangle) = \chi(d)$. Modding out by the kernel gives the isomorphism by the third isomorphism theorem. $\mathbb{T}_{\mathbb{Z}}/I_f \xrightarrow{\sim} \mathbb{Z}[\{a_n(f), \chi(d)\}]$. Furthermore, For each $d \in (\mathbb{Z}/N\mathbb{Z})^*$ take two primes p and p' both congruent to d modulo N. Generalising on 3.1.10, one can define

$$T_{p^r} = T_p T_{p^{r-1}} - p^{k-1} < p > T_{p^{r-2}}.$$

In particular, we can use express $\chi(d)$ in terms of $a_p(f)$, $a_{p^2}(f), a_{p'}(f)$, and $a_{p'^2}(f)$. This shows, that in fact $\chi(d)$'s are redundant in, $\mathbb{Z}[\{a_n(f), \chi(d)\}]$. Thus we have that,

$$\mathbb{T}_{\mathbb{Z}}/I_f \xrightarrow{\sim} \mathbb{Z}[\{a_n(f)\}]$$

.

□

Definition 5.4.5
Let $f \in \mathcal{S}_2(\Gamma_1(N))$ be a normalized eigenform, $f(\tau) = \sum_{n=1}^{\infty} a_n q^n$. The field $\mathbb{K}_f = \mathbb{Q}(\{a_n\})$ generated by the Fourier coefficients of f is called the number field of f.

Note: This is a number field because of the finite rank of the $\mathbb{Z}$-module, $\mathbb{Z}[\{a_n(f)\}]$ implying the finite degree of the extension over $\mathbb{Q}$.
Consider an arbitrary embedding $\sigma : \mathbb{K}_f \hookrightarrow \mathbb{C}$. More precisely, if $f(\tau) = \sum_{n=1}^{\infty} a_n q^n$, with τ with a_n being their Fourier coefficients, then, we denote the action with a superscript, and the conjugate of f $f^\sigma(\tau) = \sum_{n=1}^{\infty} a_n^\sigma q^n$ is obtained.
It is also noteworthy that the action gives rise to yet another eigenform. More precisely,

Theorem 5.4.6
Let f be a weight 2 normalized eigenform of the Hecke operators, so that $f \in \mathcal{S}_2(N, \chi)$ for some N and χ. Let $\mathbb{K}_f$ be its number field. For any embedding $\sigma : \mathbb{K}_f \hookrightarrow \mathbb{C}$ the conjugated f^σ is also a normalized eigenform in $\mathcal{S}_2(N, \chi^\sigma)$ where $\chi^\sigma(n) = \chi(n)^\sigma$. If f is a newform then so is f^σ.

Proof:
See, [DS05], Theorem 6.5.4. □

Corollary 5.4.7
The space $\mathcal{S}_2(\Gamma_1(N))$ has a basis of forms with rational integer coefficients.

Proof:
Consider a newform f of level M, where M divides N. Let $\mathbb{K} = \mathbb{K}_f$ represent the number field associated with f. Here, $\mathcal{O}_\mathbb{K}$ denotes the ring of integers of $\mathbb{K}$, treated as a $\mathbb{Z}$-module. Let $\{\alpha_1, \ldots, \alpha_d\}$ be a basis for $\mathcal{O}_\mathbb{K}$ as a $\mathbb{Z}$-module, and let $\{\sigma_1, \ldots, \sigma_d\}$ be the embeddings of $\mathbb{K}$ into $\mathbb{C}$.
Consider the matrix A formed by the basis elements and their corresponding embeddings:

$$A = \begin{bmatrix} \alpha_1^{\sigma_1} & \cdots & \alpha_1^{\sigma_d} \\ \vdots & \ddots & \vdots \\ \alpha_d^{\sigma_1} & \cdots & \alpha_d^{\sigma_d} \end{bmatrix}$$

Additionally, define vectors $\vec{f}$ and $\vec{g}$ as follows:

$$\vec{f} = \begin{bmatrix} f^{\sigma_1} \\ \vdots \\ f^{\sigma_d} \end{bmatrix}, \quad \vec{g} = A\vec{f}$$

In explicit terms, $g_i = \sum_{j=1}^{d} \alpha_i^{\sigma_j} f^{\sigma_j}$ for $i = 1, \ldots, d$. The linear independence of the basis $\{\alpha_1, \ldots, \alpha_d\}$ implies that $\text{span}(\{g_1, \ldots, g_d\}) = \text{span}(\{f^{\sigma_1}, \ldots, f^{\sigma_d}\})$ due to the invertibility of matrix A.
Each g_i can be expressed as $g_i(\tau) = \sum_n a_n(g_i) q^n$, where $a_n(g_i)$ are algebraic integers. For any automorphism $\sigma : \mathbb{C} \longrightarrow \mathbb{C}$, the action on the embeddings σ_j extends to $\sigma_j\sigma$, where composition is performed from left to right.
This leads to the crucial result that $g_i^\sigma = \sum_{j=1}^{d} \alpha_j^{\sigma_j\sigma} f^{\sigma_j\sigma} = g$. In simpler terms, each coefficient $a_n(g_i)$ remains fixed under all automorphisms of $\mathbb{C}$, proving that $a_n(g_i)$ lies in the intersection $\overline{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$.

Iterating this argument for each newform f whose level divides N establishes the desired conclusion. □

Since $\mathbb{T}_F$ acts on the Jacobian $\mathrm{J}_1(M_f)$, the subgroup $I_f\ \mathrm{J}_1(M_f)$ of $\mathrm{J}_1(M_f)$ makes sense.

Definition 5.4.8
The Abelian variety associated to f is defined as the quotient

$$A_f = \mathrm{J}_1(M_f)/I_f\ \mathrm{J}_1(M_f).$$

By this definition $\mathbb{T}_{\mathbb{Z}}/I_f$ acts on A_f and hence so does its isomorphic image $\mathbb{Z}[\{a_n\}]$.

To study how **Jacobians decompose** into **Abelian Varieties** A_f let us now, introduce an equivalence relation on newforms denoted as $\tilde{f} \sim f$, defined by $\tilde{f} = f^\sigma$ for some automorphism $\sigma : \mathbb{C} \longrightarrow \mathbb{C}$. Symbolically,

$$\tilde{f} \sim f \iff \tilde{f} = f^\sigma \quad \text{for some automorphism } \sigma : \mathbb{C} \longrightarrow \mathbb{C}.$$

Let $[f]$ represent the equivalence class of f, defined as

$$[f] = \{f^\sigma : \sigma \text{ is an automorphism of } \mathbb{C}\}.$$

The cardinality of $[f]$ corresponds to the number of embeddings $\sigma : \mathbb{K}_f \hookrightarrow \mathbb{C}$. According to 5.4.6 each $f^\sigma \in [f]$ is a newform at level M_f.
Define the subspace V_f of $\mathcal{S}_2(\Gamma_1(M_f))$ associated with f and its equivalence class as

$$V_f = \mathrm{span}([f]) \subset \mathcal{S}_2(\Gamma_1(M_f)),$$

which has a dimension of $[\mathbb{K}_f : \mathbb{Q}]$, representing the number of embeddings.
By restricting the subgroup $\mathrm{H}_1(X_1(M_f), \mathbb{Z})$ of $\mathcal{S}_2(\Gamma_1(M_f))^\wedge$ to functions on V_f, a subgroup of the dual space $V_f^\wedge$ is obtained:

$$\Lambda_f = \mathrm{H}_1(X_1(M_f), \mathbb{Z})|_{V_f}.$$

With this, a well-defined homomorphism is established by restricting to V_f:

$$\mathrm{J}_1(M_f) \longrightarrow V_f^\wedge/\Lambda_f, \quad [\varphi] \mapsto \varphi|_{V_f} + \Lambda_f \text{ for } \varphi \in \mathcal{S}_2(\Gamma_1(M_f))^\wedge.$$

This homomorphism provides mapping from the Jacobian $\mathrm{J}_1(M_f)$ to the dual space quotient $V_f^\wedge/\Lambda_f$, linking the structure of the Jacobian to the properties of the associated newform f and its equivalence class.

Proposition 5.4.9
Let $f \in \mathcal{S}_2(\Gamma_1(M_f))$ be a newform with number field $\mathbb{K}_f$. Then restricting to V_f induces an isomorphism

$$A_f \xrightarrow{\sim} V_f^\wedge/\Lambda_f, \quad [\varphi] + I_f\ \mathrm{J}_1(M_f) \mapsto \varphi|_{V_f} + \Lambda_f \text{ for } \varphi \in \mathcal{S}_2(\Gamma_1(M_f))^\wedge,$$

and the right side is a complex torus of dimension $[\mathbb{K}_f : \mathbb{Q}]$.

Proof:
See [DS05], Proposition, 6.6.4.

□

Let us now define isogenies in higher dimension.

Definition 5.4.10
An isogeny is a holomorphic homomorphism between complex tori that surjects and has finite kernel.

Lemma 5.4.11
For any $\gamma \in \Gamma_1(N)$ and any positive integer $n \mid N/M_f$, we have that $\begin{bmatrix} n & 0 \\ 0 & 1 \end{bmatrix} \gamma \begin{bmatrix} n & 0 \\ 0 & 1 \end{bmatrix}^{-1} \in \Gamma_1(M_f)$. Furthermore, suppose that the path $\alpha : [0,1] \longrightarrow \mathcal{H}$ is the lift of a loop in $X_1(N)$. Then the path $\tilde{\alpha}(t) = n\alpha(t)$ is the lift of a loop in $X_1(M_f)$.

Proof:
Let $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_1(N)$. The action of a matrix $A = \begin{bmatrix} n & 0 \\ 0 & 1 \end{bmatrix}$ via conjugation on $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma_1(N)$ is given by $A\gamma A^{-1}$. Let's compute this explicitly. The inverse of A is $A^{-1} = \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & 1 \end{bmatrix}$.

Now, the product $A\gamma A^{-1}$ is given by:

$$A\gamma A^{-1} = \begin{bmatrix} n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & 1 \end{bmatrix}$$

Multiplying these matrices, we get:

$$B = A\gamma A^{-1} = \begin{bmatrix} n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \frac{1}{n} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & bn \\ \frac{c}{n} & d \end{bmatrix}$$

.
We must show that the matrix $B \in \Gamma_1(M_f)$. To begin, note that, matrix B has integer entries because $\gamma \in \Gamma_1(N)$ which means that N divides c. But, $n \mid N/M_f$. Thus we have that $\frac{c}{n} = \frac{c}{N}.M_f.k$, for some integer k. Thus we get that $B \in M_2(\mathbb{Z})$. Furthermore, Matrices B, γ have the same determinant. It remains to show the congruence condition for $\Gamma_1(M_f)$. But this is also clear because the congruence conditions for N automatically imply congruence conditions for any divisor of N, particularly for M_f. Lastly, from the equation, $\frac{c}{n} = \frac{c}{N}.M_f.k$, we also have that $M_f \mid \frac{c}{n}$, which finishes the first claim.

Now, we prove the latter part, which follows from the first part.
Consider a path α, which is the lift of a loop in $H\ \Gamma_1(N)$. Thus, this just means that under the action of some, $\gamma \in \Gamma_1(N)$, $\alpha(0)$ is mapped to $\alpha(1)$, so under quotient we have a loop. Thus to finish the proof we need a matrix $\gamma' \in \Gamma_1(M_f)$ which essentially sends $\tilde{(}\alpha(0))$ to $\tilde{(}\alpha(1))$. Given a positive integer n and given a γ take $A = \begin{bmatrix} n & 0 \\ 0 & 1 \end{bmatrix}$. Taking $\gamma' = A\gamma A^{-1}$ finishes the claim. □

Note about the cusps: We will initially have a lift to $\mathcal{H}^*$. But what the matrix B does is nothing but the scaling of the original path by n. So if we are avoiding cusps in the earlier case by having a lift to $\mathcal{H}$, we will avoid them now as well, because one suddenly cannot get a rational number(or ∞) because we are just scaling by n.

Theorem 5.4.12
The Jacobian associated with the modular group $\Gamma_1(N)$ is isogeneous to a direct sum of Abelian varieties linked to equivalence classes of newforms:

$$\mathrm{J}_1(N) \longrightarrow \bigoplus_f A_f^{m_f}$$

In this expression, the summation is taken over a set of representatives $f \in \mathcal{S}_2\left(\Gamma_1\left(M_f\right)\right)$ at levels M_f that divide N. Each m_f represents the number of divisors of N/M_f.

Proof:
[DS05] By, 3.3.9 the space $\mathcal{S}_2\left(\Gamma_1(N)\right)$ has a basis $\mathcal{B}_2(N)$, defined as the union over equivalence class representatives f, divisors of N/M_f, and embeddings of $\mathbb{K}_f$ in $\mathbb{C}$:

$$\mathcal{B}_2(N) = \bigcup_f \bigcup_n \bigcup_\sigma f^\sigma(n\tau)$$

Here, the first union is taken over equivalence class representatives, the second over divisors of N/M_f, and the third over embeddings of $\mathbb{K}_f$ in $\mathbb{C}$. Each m_f is at most the number of divisors of N/M_f.
For each pair (f, n), let $d = [\mathbb{K}_f : \mathbb{Q}]$, and let $\sigma_1, \ldots, \sigma_d$ be the embeddings of $\mathbb{K}_f$ in $\mathbb{C}$. Define the map $\Psi_{f,n} : \mathcal{S}_2\left(\Gamma_1(N)\right)^\wedge \longrightarrow V_f^\wedge$ as follows:

$$\psi\left(\sum_{j=1}^{d} z_j f^{\sigma_j}(\tau)\right) = \sum_{j=1}^{d} z_j n\varphi\left(f^{\sigma_j}(n\tau)\right)$$

This map takes $\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)$ into $\Lambda_f = \mathrm{H}_1\left(X_1\left(M_f\right), \mathbb{Z}\right)|_{V_f}$. To see this, consider $\varphi = \int_\alpha$ for some loop α in $X_1(N)$. Consequently we have ,

$$\psi\left(f^\sigma(\tau)\right) = n\int_\alpha f^\sigma(n\tau)d\tau = \int_{\tilde{\alpha}} f^\sigma(\tau)d\tau$$

where $\tilde{\alpha}(t) = n\alpha(t)$. Note that the holomorphic differential $\omega\left(f^\sigma(n\tau)\right)$ on $X_1(N)$ is identified with its pullback to $\mathcal{H}$, and α is identified with some lift in $\mathcal{H}$. Consequently, $\tilde{\alpha}$ is the lift of a loop in $X_1\left(M_f\right)$ which follows from our previous lemma 5.4.11.
Furthermore, $\Psi_{f,n}$ takes $\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)$ to Λ_f as claimed.
Taking the product map over all pairs (f, n) gives:

$$\Psi = \prod_{f,n} \Psi_{f,n} : \mathcal{S}_2\left(\Gamma_1(N)\right)^\wedge \longrightarrow \bigoplus_{f,n} V_f^\wedge = \bigoplus_f \left(V_f^\wedge\right)^{m_f}$$

A Counting of dimensions argument shows that Ψ is a vector space isomorphism.
This isomorphism descends to an isomorphism of quotients:

$$\bar{\Psi} : \mathrm{J}_1(N) \xrightarrow{\sim} \bigoplus_f \left(V_f^\wedge\right)^{m_f} / \Psi\left(\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)\right)$$

Since $\Psi\left(\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)\right) \subset \bigoplus_f \Lambda_f^{m_f}$ is a containment of Abelian groups of the same rank, the natural surjection:

$$\pi : \bigoplus_f \left(V_f^\wedge\right)^{m_f} / \Psi\left(\mathrm{H}_1\left(X_1(N), \mathbb{Z}\right)\right) \longrightarrow \bigoplus_f \left(V_f^\wedge / \Lambda_f\right)^{m_f}$$

has a finite kernel. By 5.4.9 , there is an isomorphism:

$$i : \bigoplus_f \left(V_f^\wedge / \Lambda_f\right)^{m_f} \xrightarrow{\sim} \bigoplus_f A_f^{m_f}$$

Hence, $i \circ \pi \circ \bar{\Psi} : \mathrm{J}_1(N) \longrightarrow \bigoplus_f A_f^{m_f}$ is the desired isogeny. □

To conclude this chapter, we present a version of the Modularity Theorem involving Abelian varieties. While we develop results in the larger context of $\Gamma_1(N)$ for generality, we state versions of the Modularity Theorem in the restricted context of $\Gamma_0(N)$, making them slightly sharper.
Specifically, for each newform $f \in \mathcal{S}_2\left(\Gamma_0\left(M_f\right)\right)$, let

$$A'_f = \mathrm{J}_0\left(M_f\right) / I_f \ \mathrm{J}_0\left(M_f\right).$$

This is another Abelian variety associated with f. The proofs in the earlier case transfer to $\Gamma_0(N)$, showing,

Theorem 5.4.13 (Isogeneous decomposition)
There is an isomorphism

$$A'_f \xrightarrow{\sim} V_f^\wedge / \Lambda'_f \quad \text{where } \Lambda'_f = \mathrm{H}_1\left(X_0\left(M_f\right), \mathbb{Z}\right)|_{V_f}$$

and an isogeny

$$\mathrm{J}_0(N) \longrightarrow \bigoplus_f \left(A'_f\right)^{m_f}$$

where now the sum is taken over the equivalence classes of newforms $f \in \mathcal{S}_2\left(\Gamma_0\left(M_f\right)\right)$.

The introduction of A'_f serves to phrase the following version of the Modularity Theorem entirely in terms of $\Gamma_0(N)$. By incorporating Abelian varieties, this version associates a modular form, in fact a newform, to an elliptic curve.

Theorem 5.4.14 (**Modularity theorem**)
Let E be a complex elliptic curve with $j(E) \in \mathbb{Q}$. Then for some positive integer N and some newform $f \in \mathcal{S}_2\left(\Gamma_0(N)\right)$ there exists a surjective holomorphic homomorphism of complex tori

$$A'_f \longrightarrow E.$$

Let us now shift our attention back to the example of the Modular curve, $X_0(38)$ and demonstrate some of the things that we discussed in this chapter pertaining to this curve. We recall from our previous computations that this is a genus 4 curve.

The Jacobian $\mathrm{J}_0(N)$ of a modular curve $X_0(N)$ is an abelian variety whose dimension is equal to the genus g of the curve. The Jacobian is a g-dimensional complex torus. Theorem 5.4.13 states that the

Jacobian $J_0(N)$ can be decomposed into a direct sum of abelian varieties associated with newforms. Each of these abelian varieties, say A'_f, corresponds to a newform f and is an abelian variety of some dimension. The dimension of each abelian variety A'_f in the decomposition contributes to the total dimension of the Jacobian. The sum of the dimensions of these individual abelian varieties must equal the genus of the modular curve. The modular curve's genus g gives the total number of 'independent directions' or 'degrees of freedom' in the Jacobian. In simpler terms, each abelian variety A'_f takes up some degrees of freedom. The dimension of each A'_f tells us how many degrees of freedom it occupies. For example let us suppose the genus of $X_0(N)$ is 3 for some N. If the Jacobian decomposes into two Abelian varieties, one could be 1-dimensional and the other 2-dimensional, totalling 3.

Thus, we need to calculate the newforms associated to the Modular curve $X_0(38)$. We will intensively use Magma for some of the computations. As it turns out, an efficient way to do so is via Modular symbols. We encourage the readers to go through [Ste07].
Since, this is only an example and the modular symbols are not as much relevant for the further discussion, we will not develop the theory of Modular symbols. We chose to instead emphasize on the fact that Modular symbols lie at the heart of these computations and it should not be forgiven. In [Ste07], Stein describes how to use modular symbols to construct a basis of $S_2(\Gamma_0(N))$ consisting of modular forms that are eigenvectors for every element of the ring $\mathbb{T}'$ generated by the Hecke operator T_p, with $p \nmid N$. Recall that such eigenvectors are called eigenforms.

Consider M as a positive integer which is a divisor of N. In [Lan95], Serge lang describes a way to link the space of modular forms of level M, $S_2(M)$, to the space $S_2(\Gamma_0(N))$. For every divisor d of N/M, there is a map, the so-called degeneracy map $\beta_{M,d}$, which takes a modular form f in $S_2(M)$ and maps it to $S_2(\Gamma_0(N))$ by the rule $\beta_{M,d}(f(q)) = f(q^d)$. This map essentially adjusts the level of the modular form from M to N, accommodating the change by raising the modular argument q to the power d.

Let us recall the Newspace $S_2(\Gamma_0(N))_{\text{new}}$. This subspace can be understood as the orthogonal complement, in the sense of Hecke operators, of the space spanned by all images of the degeneracy maps $\beta_{M,d}$ for all choices of M and d.

Drawing on the foundational work of Atkin and Lehner discussed in Chapter 3, we understand that the space $S_2(\Gamma_0(N))$ can be decomposed as a direct sum of the images of these new spaces under the degeneracy maps. More formally, this is expressed as:

$$S_2(\Gamma_0(N)) = \bigoplus_{\substack{M|N \\ d|N/M}} \beta_{M,d}(S_2(M)_{\text{new}}).$$

Consequently, to calculate $S_2(\Gamma_0(N))$, it is sufficient to compute $S_2(M)_{\text{new}}$ for each positive divisor M of N and then consider the images of these spaces under the respective degeneracy maps.
As explained before, we follow [Ste07]. We first define the space of modular symbols of level 38 and find out its cuspidal subspace. We then find out the new form decomposition of the cuspidal subspace of M. We do all of this by applying the following series of commands in Magma:

```
M := ModularSymbols(38);
```

```
M_cusp := CuspidalSubspace(M);
M_dec := NewformDecomposition(M_cusp);
Eigenform(M_dec[1],10);
Eigenform(M_dec[2],10);
Eigenform(M_dec[3],10);
```

We get the output as follows:

```
q - q^2 + q^3 + q^4 - q^6 - q^7 - q^8 - 2*q^9 + O(q^10)
q + q^2 - q^3 + q^4 - 4*q^5 - q^6 + 3*q^7 + q^8 - 2*q^9 + O(q^10)
q - 2*q^3 - 2*q^4 + 3*q^5 - q^7 + q^9 + O(q^10)
```

Let us fix some notations.

$$f_1 = q - q^2 + q^3 + q^4 - q^6 - q^7 - q^8 - 2q^9 + O(q^{10}) \tag{5}$$
$$f_2 = q + q^2 - q^3 + q^4 - 4q^5 - q^6 + 3q^7 + q^8 - 2q^9 + O(q^{10}) \tag{6}$$
$$g_1 = q - 2q^3 - 2q^4 + 3q^5 - q^7 + q^9 + O(q^{10}) \tag{7}$$

Note that, f_1, f_2 are newforms of level 38 with trivial character and g_1 is a newform of level 19 with trivial character. This coincides with the information that we have on the LMFDB Page for the Modular curve $X_0(38)$.
Consider g_1 and g_2, where $g_2(z) = g_1(2z)$. Our goal is to show that they are linearly independent. To demonstrate this, we investigate the q-expansions of both functions. Recall that q-expansions provide a unique power series representation for modular forms.
Assume, for argument's sake, there exist constants c_1 and c_2 such that $c_1 g_1 + c_2 g_2 = 0$. Mathematically, this can be expressed in terms of their q-expansions as:

$$c_1 \cdot \sum_n a_n q^n + c_2 \cdot \sum_n b_n q^n = 0.$$

Here, a_n and b_n are the coefficients in the q-expansions of g_1 and g_2, respectively. The crucial observation is that the odd-powered terms in this expansion are exclusively contributed by g_1. The absence of odd powers in g_2's expansion is because it is defined as $g_1(2z)$. Note here that, since $q = e^{2\pi i z}$, replacing z by $2z$ amounts to replacing q by q^2, so that the q-series of g_2 is given by $q^2 - 2q^6 - 2q^8 + O(q^{10})$. Specifically, the cancellation of odd terms in the series shows that c_1 must be zero. Once $c_1 = 0$ is established, c_2 must also be zero to nullify the even terms, due to the structure of g_2.
Thus, $c_1 = c_2 = 0$ is the only solution, proving that g_1 and g_2 are linearly independent.
Since the dimension of the newspace is 2 since by definition f_1, f_2 generate the Newspace., and we have found two linearly independent functions within it, g_1 and g_2 generate the old space in $S_2(38)$.
Thus, f_1, f_2, g_1, g_2 form a basis of $S_2(38)$. Note that, this is coherent with our previous discussion since g_2 is given by q with q^d for $d = 2$ which is a divisor of $\frac{38}{19}$. This amounts to saying that g_2 was obtained by taking the image of degeneracy map with respect to $M = 19, d = 2$ of g_1.
Thus, everything connects so beautifully.
Let us now end our discussion by stating the isogenous decomposition as stated in 5.4.13. Note that, since 2 divides 38, only relevant newforms are from level 2, level 19 and level 38. The space $S_2(2)$ is trivial so only relevant newforms come from level 19 and 38.

In conclusion, we get that,

$$J_0(38) \cong A'_{f_1} \bigoplus A'_{f_2} \bigoplus (A'_{g_1})^2$$

where the congruence is upto an isogeny and the abelian variety (A'_{g_1}) appears with multiplicity 2 because the number of divisors of 19 are 2.

6 Galois Representations

In this section, we briefly discuss Galois representations of the absolute Galois group of the field of rational numbers $\mathbb{Q}$. We will start with a brief background and introduce the required terminology. This section aims to give motivation for studying 2-dimensional representations by taking inspiration from the classical result in class field theory, the Kronecker-Weber theorem, which essentially describes one-dimensional representations of $G_{\mathbb{Q}}$. We will then describe some of the primary sources of Galois representations, namely, Elliptic curves and Modular forms. This will already give us enough background to appreciate the landmark results that stood as the stepping stones in the proof of Fermat's last theorem. We will closely follow the book [DS05] and [DDT95]. We will sometimes closely follow proofs from either of these sources for the sake of completeness. Due to the extreme complexity of the overall topic and to demonstrate all of this within 6 months, it was inevitable to closely follow some of the proofs but written in my own words. Having said that, I must add that in many places, I have expanded on my own, giving more details.

6.1 Motivation and Basics

To state the motivation behind Galois representations in one line would be to study the absolute Galois Group of $\mathbb{Q}$, $G_{\mathbb{Q}} = Gal(\overline{\mathbb{Q}}/\mathbb{Q})$. One key idea is to look at

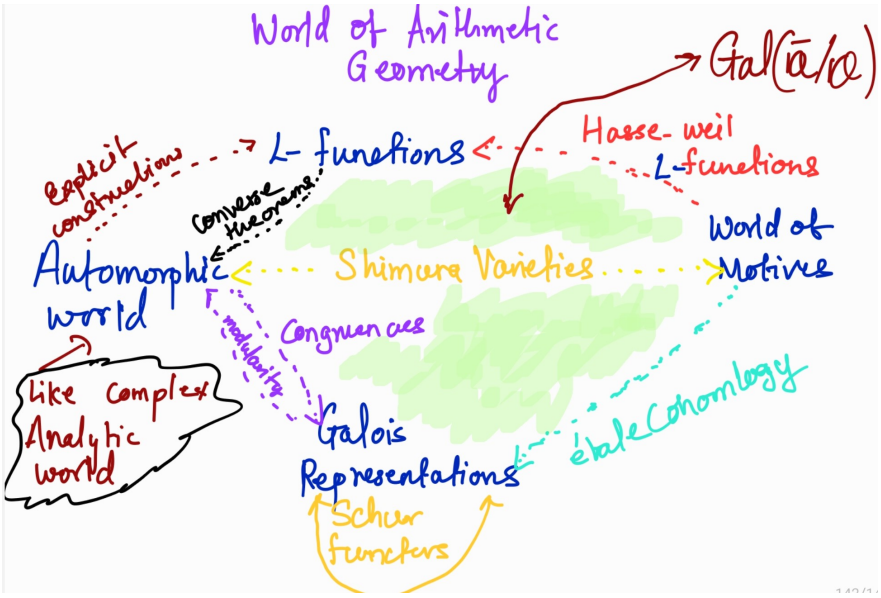
$$\rho : G_Q \longrightarrow G,$$

where G is some group that we nicely understand. This gives us our first definition of this chapter.

Definition 6.1.1
A d-dimensional representation of $G_{\mathbb{Q}}$ is a continous homomorphism

$$G_{\mathbb{Q}} \longrightarrow GL_d(K)$$

where K is a field.
Let us look at the following picture which depicts the current world of Arithmetic Geometry.



The main idea usually behind the representations is that we understand good chunk of information about the domain group especially when it is mapped to a group that we fully or almost completely understand. The motivation behind studying Galois representations is to create a bridge between field theory, which studies the algebraic structure of fields, and group theory, which studies the algebraic structure of symmetries. Galois representations provide a way to encode the action of Galois groups on various mathematical structures into linear algebraic terms, often as matrices acting on vector spaces. This encoding allows complex symmetries to be understood and manipulated using the well-developed tools of linear algebra.

To briefly current scheme of Arithmetic geometry without going into too many details, **Shimura Varieties** are a class of higher-dimensional spaces that generalize elliptic curves and provide a geometric setting for studying arithmetic properties of automorphic forms. They serve as a bridge between complex analysis and number theory. **Automorphic Forms** are complex functions that are symmetric with respect to the action of a discrete group. They are central to the Langlands program, which predicts a correspondence between automorphic forms and Galois representations. **Galois Representations** provide a way to represent Galois groups, which are fundamental in understanding the symmetries of the roots of polynomials. Galois representations can be associated with automorphic forms, which is a key aspect of the Langlands correspondence. **L-Functions** are complex functions that encode number-theoretic information in their coefficients and zeros. Hasse-Weil L-functions, for instance, are associated with algebraic varieties and are predicted to correspond to automorphic L-functions. **Motives** are an abstract formulation of various cohomology theories. The Langlands program can be seen as an attempt to realize a correspondence between motives (arising from algebraic geometry) and automorphic forms. **The Langlands program** suggests that there's a deep connection between Galois groups and automorphic forms (as represented by Galois representations and automorphic L-functions), mediated through geometric objects like Shimura varieties and abstract objects like motives.

As one could, see from this how important this is for us to study Galois representations and thus we shift our focus to basic definitions and understanding more about Galois representations.

Let us proceed with recalling that the absolute Galois group of a field $\mathbb{F}$, G_F is profinite. Specifically, it is determined by $G_{\mathbb{F}} = \lim \mathrm{Gal}(L/\mathbb{F})$, where L iterates over the finite Galois extensions of F that are contained in $\overline{\mathbb{F}}$. This gives $G_{\mathbb{F}}$ an inherent topology known as the so-called Krull topology.
Considering a prime ℓ that is different than that of characteristic of $\mathbb{F}$, we consider $\varepsilon_\ell : G_{\mathbb{F}} \to \mathbb{Z}_\ell^\times$. It represents the ℓ-adic cyclotomic character. Put in simpler terms: for every $\sigma \in G_{\mathbb{F}}$, if ζ is an ℓ-power root of unity in $\bar{F}$, then $\sigma(\zeta) = \zeta^{\varepsilon_\ell(\sigma)}$. When the context indicates the choice of ℓ, we can simply use ε.

Definition 6.1.2
Consider a Galois number field $\mathbb{F}$ and a rational prime p. For each maximal ideal $\mathfrak{p}$ of the ring of integers $\mathcal{O}_{\mathbb{F}}$ lying over p, the decomposition group $D_{\mathfrak{p}}$ is defined as the subgroup of the Galois group $\mathrm{Gal}(\mathbb{F}/\mathbb{Q})$ that fixes $\mathfrak{p}$ as a set:

$$D_{\mathfrak{p}} = \{\sigma \in \mathrm{Gal}(\mathbb{F}/\mathbb{Q}) : \mathfrak{p}^\sigma = \mathfrak{p}\}.$$

The order of the decomposition group is ef, where e is the ramification index and f is the residue class degree. Consequently, its index in $\mathrm{Gal}(\mathbb{F}/\mathbb{Q})$ is the decomposition index g. The action of $D_{\mathfrak{p}}$ on the residue field $\mathfrak{O}_{\mathfrak{p}} = \mathcal{O}_{\mathbb{F}}/\mathfrak{p}$ is given by:

$$(x+\mathfrak{p})^{\sigma} = x^{\sigma} + \mathfrak{p}, \quad x \in \mathcal{O}_{\mathbb{F}}, \sigma \in D_{\mathfrak{p}}.$$

Definition 6.1.3
The inertia group $I_{\mathfrak{p}}$ of $\mathfrak{p}$ is defined as the kernel of this action:

$$I_{\mathfrak{p}} = \{\sigma \in D_{\mathfrak{p}} : x^{\sigma} \equiv x (\mathrm{mod}\mathfrak{p}) \text{ for all } x \in \mathcal{O}_{\mathbb{F}}\}.$$

The order of the inertia group is e, and it is trivial for all $\mathfrak{p}$ lying over any unramified prime p.

Definition 6.1.4
Let $\mathbb{F}/\mathbb{Q}$ be a Galois extension. Let p be a rational prime and let $\mathfrak{p}$ be a maximal ideal of $\mathcal{O}_{\mathbb{F}}$ lying over p. A Frobenius element of $\mathrm{Gal}(\mathbb{F}/\mathbb{Q})$ is any element $\mathrm{Frob}_{\mathfrak{p}}$ satisfying the condition

$$x^{\mathrm{Frob}_{\mathfrak{p}}} \equiv x^{p} (\mathrm{mod}\mathfrak{p}) \quad \text{for all } x \in \mathcal{O}_{\mathbb{F}}.$$

Note that when p is unramified, we have that order of the Inertia group 1 and thus giving the uniqueness of Frobenius element.
We can also take motivation from the above defintions and define these groups also over a fixed algebraic closure of $\mathbb{Q}$.

Definition 6.1.5
For a family of elements, let $p \in \mathbb{Z}$ be any prime and let $\mathfrak{p} \subset \overline{\mathbb{Z}}$ be any maximal ideal over p. Let $\mathfrak{p}$ be the kernel of the reduction map $\overline{\mathbb{Z}} \longrightarrow \overline{\mathbb{F}}_p$.

The decomposition group of $\mathfrak{p}$ is

$$D_{\mathfrak{p}} = \{\sigma \in G_{\mathbb{Q}} : \mathfrak{p}^{\sigma} = \mathfrak{p}\}.$$

The reduction map

$$D_{\mathfrak{p}} \longrightarrow G_{\mathbb{F}_p}$$

is surjective. An absolute Frobenius element over p is any preimage $\mathrm{Frob}_{\mathfrak{p}} \in D_{\mathfrak{p}}$ of the Frobenius automorphism $\sigma_p \in G_{\mathbb{F}_p}$. Thus $\mathrm{Frob}_{\mathfrak{p}}$ is defined only up to the kernel of the reduction map which brings us to the next definition.

Definition 6.1.6
The inertia group of $\mathfrak{p}$ is defined as,

$$I_{\mathfrak{p}} = \left\{\sigma \in D_{\mathfrak{p}} : x^{\sigma} \equiv x (\mathrm{mod}\mathfrak{p}) \text{ for all } x \in \overline{\mathbb{Z}}\right\}.$$

Definition 6.1.7
Let d be a positive integer. A d-dimensional ℓ-adic Galois representation is a continuous homomorphism

$$\rho : G_{\mathbb{Q}} \longrightarrow \mathrm{GL}_d(\mathbb{L})$$

where $\mathbb{L}$ is a finite extension field of $\mathbb{Q}_\ell$. If $\rho' : G_{\mathbb{Q}} \longrightarrow \mathrm{GL}_d(\mathbb{L})$ is another such representation and there is a matrix $m \in \mathrm{GL}_d(\mathbb{L})$ such that $\rho'(\sigma) = m^{-1}\rho(\sigma)m$ for all $\sigma \in G_{\mathbb{Q}}$ then ρ and ρ' are equivalent. Equivalence is denoted by $\rho \sim \rho'$.

Remark 6.1.8
This defintion comes from the motivation that dirichlet characters lead to homomorphisms from the Absolute galois group of $\mathbb{Q}$ into $\mathbb{C}*$ of certain type and all such homomorphism come from Dirichlet characters. One can more generally show that any continuous homomorphism ρ from the absolute Galois group $G_{\mathbb{Q}}$ to the general linear group of degree d over the complex numbers, denoted $\mathrm{GL}_d(\mathbb{C})$ has finite image. Extending this notion, the image of a Dirichlet character χ is situated within a number field $\mathbb{K}$, and consequently, it can be associated with a field $\mathbb{K}_\lambda$, in which λ is an ideal above a specified rational prime ℓ. Thus, the multiplicative group of the complex numbers, $\mathbb{C}^*$, can be replaced by $\mathbb{K}_\lambda^*$ within the framework of diagram (9.11), [DS05]. Following this substitution, the continuity of the representation ρ_χ: $G_{\mathbb{Q}} \to \mathbb{K}_\lambda^*$ is preserved.

Before going to the next section, we define the last important defintion in this section, which will facilitate the discussions from this point onwards.

Definition 6.1.9
Let ρ be a Galois representation and let p be prime. Then ρ is unramified at p if $I_{\mathfrak{p}} \subset \ker \rho$ for any maximal ideal $\mathfrak{p} \subset \overline{\mathbb{Z}}$ lying over p.

Definition 6.1.10
An algebraic extension $F/\mathbb{Q}$ is termed unramified at p if all conjugates of I_p lie within $G_F \subset G_{\mathbb{Q}}$. Conversely, it's termed ramified otherwise.

If $F/\mathbb{Q}$ is Galois and unramified at p, a distinct conjugacy class $[\mathrm{Frob}_p] \subset \mathrm{Gal}(F/\mathbb{Q})$ arises.
Replacing the p-adic completion with an Archimedean one, there's a unique conjugacy class $[c]$ in $G_{\mathbb{Q}}$ corresponding to complex conjugation for some embeddings of $\overline{\mathbb{Q}}$ into $\mathbb{C}$. For a chosen c, its stabiliser is G_∞.
We state a couple of important results without proofs which help us in the understanding of the absolute Galois group of Q. See, [DDT95] for more details.

Theorem 6.1.11
A finite extension $F/\mathbb{Q}$ is ramified only at finitely many primes.

Theorem 6.1.12 (Chebotarev Density theorem)
For a Galois extension $F/\mathbb{Q}$ unramified outside a finite set of primes S, the set $\bigcup_{p \notin S} [\mathrm{Frob}_p]$ is dense in $\mathrm{Gal}(F/\mathbb{Q})$.

:

Theorem 6.1.13 (Kronecker-Weber)
There is an isomorphism:

$$\prod_p \varepsilon_p : G_{\mathbb{Q}}^{\mathrm{ab}} \xrightarrow{\sim} \prod_p \mathbb{Z}_p^\times$$

The Local Kronecker-Weber theorem, state that the following map is an isomorphism:

$$\varrho \times \varepsilon_p : G_{\mathbb{Q}_p}^{\mathrm{ab}} \longrightarrow G_{\mathbb{F}_p} \times \mathbb{Z}_p^{\times}$$

Where, G^{ab} denotes the abelianisation of a profinite group G. Under this mapping, I_p is mapped to $\mathbb{Z}_p^{\times}$, and for positive u, I_p^u is mapped to $\left(1 + p^{\lceil u \rceil}\mathbb{Z}_p\right) \subset \mathbb{Z}_p^{\times}$, where $\lceil u \rceil$ denotes the ceiling function. For cases where $\ell \neq p$, the map ε_ℓ is trivial on I_p and sends Frob $_p$ to $p \in \mathbb{Z}_\ell^{\times}$.
Note: We're primarily interested in representations into fields with a natural topology, emphasizing continuous maps. A one-dimensional representation inherently has an abelian image. Kronecker-Weber theorem essentially characterizes these representations and their behaviour on decomposition groups at all primes. The goal of this chapter is to study some of the things which were done to expand upon this for two-dimensional representations of $G_{\mathbb{Q}}$.

6.2 Elliptic curves, Modular forms and Galois representations

The idea behind this section is to introduce the connection of elliptic curves to Galois representations. We start by introducing ℓ-adic module associated with an elliptic curve E. Recall, that $E[n](\overline{\mathbb{Q}})$ denote the group of n-torsion points on $E(\overline{\mathbb{Q}})$. For simplicity, we will use the notation $E[n]$.

Definition 6.2.1
Let ℓ be a prime number. Let E be an elliptic curve. Then, the ℓ-adic Tate module of E is given by,

$$\mathcal{T}_\ell E = \lim_{\leftarrow} E\left[\ell^n\right],$$

where the inverse limit runs over natural numbers.

We recall that, for each n, we have the isomorphism between $E\left[\ell^n\right]$ and $\mathbb{Z}/\ell^n\mathbb{Z}$. This, gives

$$\mathcal{T}_\ell(E) \cong \mathbb{Z}_\ell^2.$$

Furthermore, one can easily see that Tate module is indeed a $G_{\mathbb{Q}}$-module. To see this quickly, note that if attach the ℓ^n-torsion points to $\mathbb{Q}$, one gets a Galois number field. Thus, one can restrict from the absolute Galois group of $\mathbb{Q}$ to $\mathrm{Gal}(\mathbb{Q}(E[\ell^n])/\mathbb{Q})$. Furthermore, the latter Galois group injects into $\mathrm{Aut}(E[\ell^n])$. Combining and Checking compatibility, of restriction map into $Aut[E[\ell^n]], Aut[E[\ell^{n+1}]]$, one realises that the ℓ-adic Tate module asosciated to E is indeed a $G_{\mathbb{Q}}$-module.
One has the natural isomorphism, $\mathrm{Aut}\left(E[\ell^n]\right) \xrightarrow{\sim} \mathrm{GL}_2\left(\mathbb{Z}/\ell^n\mathbb{Z}\right)$, and these isomorphisms combine to give $\mathrm{Aut}\left(\mathrm{Ta}_\ell(E)\right) \xrightarrow{\sim} \mathrm{GL}_2\left(\mathbb{Z}_\ell\right)$. Since $G_{\mathbb{Q}}$ acts on $\mathrm{Ta}_\ell(E)$, the cumulative result is a homomorphism

$$\rho_{E,\ell} : G_{\mathbb{Q}} \longrightarrow \mathrm{GL}_2\left(\mathbb{Z}_\ell\right) \subset \mathrm{GL}_2\left(\mathbb{Q}_\ell\right).$$

.
The continuity of this homomorphism follows from the fact that the Tate module has a natural continuous action of $G_{\mathbb{Q}}$.
Thus, we get the following definition.

Definition 6.2.2

$\rho_{E,\ell}$ is a Galois representation, the 2-dimensional Galois representation associated to E.

Similarly, due to the action of G_Q on the group of n-torsion points of E, one gets a representation defined below.

Definition 6.2.3

$$\bar{\rho}_{E,n} : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}/n\mathbb{Z}).$$

is representation assosciated with E, for n, a natural number.

We have the following global properties for these representations.

Proposition 6.2.4 1. The determinant of $\rho_{E,\ell}$ is ε_ℓ.

2. The representation $\rho_{E,\ell}$ is absolutely irreducible for all ℓ and for fixed E, $\bar{\rho}_{E,\ell}$ is absolutely irreducible for all but finitely many ℓ.

3. Let E be an elliptic curve over $\mathbb{Q}$ with conductor N. The Galois representation $\rho_{E,\ell}$ is unramified at every prime $p \nmid \ell N$. For any such p let $\mathfrak{p} \subset \overline{\mathbb{Z}}$ be any maximal ideal over p. Then the characteristic equation of $\rho_{E,\ell}(\text{Frob}_{\mathfrak{p}})$ is

$$x^2 - a_p(E)x + p = 0.$$

Proof:
First result follows from the existence of the non-degenerate alternating Galois-equivariant Weil pairing

$$\mathcal{T}_\ell E \times \mathcal{T}_\ell E \longrightarrow \mathbb{Z}_\ell(1) := \varprojlim \mu_{\ell^n}.$$

The second result is the main result of [Ser68]. The last part is proved in detail in [DS05], see for example, proposition 9.4.1

□

Remark 6.2.5
There is a stronger version than 2nd part of the previous proposition due to Barry Mazur which states that, for $E/\mathbb{Q}$ an elliptic curve, we have following:

1. If $\ell > 163$ is a prime then $\bar{\rho}_{E,\ell}$ is irreducible.

2. If E is semistable then $\bar{\rho}_{E,\ell}$ is irreducible for $\ell > 7$.

3. If E is semistable and $\bar{\rho}_{E,2}$ is trivial then $\bar{\rho}_{E,\ell}$ is irreducible for $\ell > 3$.

Note that the irreducibility result in the second part is stronger than in the general case and requires E to have semistable reduction everywhere. The condition that $\bar{\rho}_{E,2}$ is trivial adds additional constraints on the curve and its torsion structure, which can be used to deduce the irreducibility of $\bar{\rho}_{E,\ell}$ for the remaining primes ℓ greater than 3. Furthermore, once we look at this result and the results due to Serre [Ser68], the cumulative result is that if E is semistable everywhere, then $\bar{\rho}_{E,\ell}$ is surjective for $\ell > 7$(See, [Maz78], Theorem 4).

Now, we briefly discuss the local behaviour of these representations before going further to discuss the connection of Galois representations with Modular curves and in turn with Modular forms.

Theorem 6.2.6
Suppose E has good reduction at a prime p. If $\ell \neq p$, then $\rho_{E,\ell}$ is unramified at p, and we have the formula

$$\operatorname{tr} \rho_{E,\ell}\left(\mathrm{Frob}_p\right) = p + 1 - \#\bar{E}_p\left(\mathbb{F}_p\right).$$

In particular $\operatorname{tr} \rho_{E,\ell}(\ \mathrm{Frob}\ _p)$ belongs to $\mathbb{Z}$ and is independent of $\ell \neq p$.

Let us now attempt to explain the connection of Galois representations with modular curves and how these representations decompose into 2-dimensional representations associated to modular forms.
Let N be a positive integer and let ℓ be prime. The modular curve $X_1(N)$ is a projective nonsingular algebraic curve over $\mathbb{Q}$. Let g denote its genus. The curve $X_1(N)_{\mathbb{C}}$ over $\mathbb{C}$ defined by the same equations can also be viewed as a compact Riemann surface.

Definition 6.2.7
The Picard group of the modular curve is the Abelian group of divisor classes on the points of $X_1(N)$,

$$\mathrm{Pic}^0\left(X_1(N)\right) = \mathrm{Div}^0\left(X_1(N)\right) / \mathrm{Div}^{\ell}\left(X_1(N)\right).$$

Definition 6.2.8
The ℓ-adic Tate module of $X_1(N)$ is

$$\mathrm{Ta}_{\ell}\left(\mathrm{Pic}^0\left(X_1(N)\right)\right) = \lim_{n} \left\{\mathrm{Pic}^0\left(X_1(N)\right)\left[\ell^n\right]\right\}.$$

Section 7.9 from [DS05] discusses the identification of Picard group with the complex analytic Picard group. After following a series of observations which follow from Abel's theorem and Igusa's theorem discussed in the last chapter, we get the following isomorphisms.

$$i_n : \mathrm{Pic}^0\left(X_1(N)\right)\left[\ell^n\right] \longrightarrow \mathrm{Pic}^0\left(X_1(N)_{\mathbb{C}}\right)\left[\ell^n\right] \cong \left(\mathbb{Z}/\ell^n\mathbb{Z}\right)^{2g}$$

$$\pi_n : \mathrm{Pic}^0\left(X_1(N)\right)\left[\ell^n\right] \longrightarrow \mathrm{Pic}^0\left(\widetilde{X}_1(N)\right)\left[\ell^n\right], p \nmid \ell N.$$

Analogous to the preceding discussion, if we select bases for the torsion subgroup $\mathrm{Pic}^0(X_1(N))[\ell^n]$ compatibly for each n, we have that:

$$\mathrm{Ta}_{\ell}(\mathrm{Pic}^0(X_1(N))) \simeq \mathbb{Z}_{\ell}^{2g}.$$

Here, the isomorphism indicates that the Tate module Ta_{ℓ} of the Picard group of degree zero Pic^0 of the modular curve $X_1(N)$ is isomorphic to a free $\mathbb{Z}_{\ell}$-module of rank $2g$, where g denotes the genus of the modular curve $X_1(N)$. Following on the footsteps of the previous discussion as in the case of Elliptic curves, one gets the Galois representation associated with $X_1(N)$.

$$\rho_{X_1(N),\ell} : G_{\mathbb{Q}} \longrightarrow \mathrm{GL}_{2g}\left(\mathbb{Z}_{\ell}\right) \subset \mathrm{GL}_{2g}\left(\mathbb{Q}_{\ell}\right).$$

Theorem 6.2.9

Let ℓ be prime and let N be a positive integer. The Galois representation $\rho_{X_1(N),\ell}$ is unramified at every prime $p \nmid \ell N$. For any such p let $\mathfrak{p} \subset \overline{\mathbb{Z}}$ be any maximal ideal over p. Then $\rho_{X_1(N),\ell}(\text{ Frob }_{\mathfrak{p}})$ satisfies the polynomial equation

$$x^2 - T_p x + \langle p \rangle p = 0$$

Proof (sketch)*:*
Assuming p is not a divisor of ℓN and let $\mathfrak{p}$ be an ideal above p, we have the following diagram, which gives the first claim due to the fact that the mapping on the right is an isomorphism:

$$\begin{array}{ccc} D_{\mathfrak{p}} & \longrightarrow & \mathrm{Aut}(\mathrm{Pic}^0(X_1(N))[\ell^n]) \\ \downarrow & & \downarrow \\ G_{F_p} & \longrightarrow & \mathrm{Aut}(\mathrm{Pic}^0(\tilde{X}_1(N))[\ell^n]) \end{array}$$

For the subsequent part, Eichler-Shimura relation when restricted to ℓ^n-torsion, gives another commutative diagram:

$$\begin{array}{ccc} \mathrm{Pic}^0(X_1(N))[\ell^n] & \xrightarrow{T_p} & \mathrm{Pic}^0(X_1(N))[\ell^n] \\ \downarrow & & \downarrow \\ \mathrm{Pic}^0(\widetilde{X}_1(N))[\ell^n] & \xrightarrow{\sigma_{p,*}+(\langle p\rangle\cdot\sigma_p)^*} & \mathrm{Pic}^0(\widetilde{X}_1(N))[\ell^n] \end{array}$$

The diagram, when modified with $\mathrm{Frob}_{\mathfrak{p}} + \langle p \rangle p \mathrm{Frob}_{\mathfrak{p}}^{-1}$ across the top row, preserves commutativity. Given that the vertical mappings are isomorphisms, we infer that $T_p = \mathrm{Frob}_{\mathfrak{p}} + \langle p \rangle p \mathrm{Frob}_{\mathfrak{p}}^{-1}$ on $\mathrm{Pic}^0(X_1(N))[\ell^n]$. This holds for all n, and thus, the relationship extends to $\mathrm{Ta}_\ell(\mathrm{Pic}^0(\widetilde{X}_1(N)))$, from which the result follows.

□

6.3 Modularity

In this brief section, we attempt to briefly explain the Modularity theorem in terms of Galois representations.

Definition 6.3.1
Let $\rho : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Q}_\ell)$ be an irreducible Galois representation where $G_{\mathbb{Q}}$ is the absolute Galois group of $\mathbb{Q}$, and $\mathrm{GL}_2(\mathbb{Q}_\ell)$ denotes the group of 2x2 invertible matrices over the ℓ-adic numbers $\mathbb{Q}_\ell$. Furthermore, let the determinant of ρ be the ℓ-adic cyclotomic character χ_ℓ.

The representation ρ is said to be modular if the following conditions are satisfied:

1. There exists a newform f in the space of cusp forms of weight 2 for the congruence subgroup $\Gamma_0(M_f)$, symbolized as $f \in \mathcal{S}_2(\Gamma_0(M_f))$.
2. The field $\mathbb{K}_{f,\lambda}$, which is the completion of the number field generated by the Fourier coefficients of f at a maximal ideal λ, is identical to $\mathbb{Q}_\ell$, where λ is an ideal in the ring of integers $\mathcal{O}_{\mathbb{K}_f}$ that lies above ℓ.
3. The representation $\rho_{f,\lambda}$, associated with the newform f at the ideal λ, is isomorphic to ρ, denoted as $\rho_{f,\lambda} \sim \rho$.

Now, we are ready to the state the version of Modularity theorem that was finally proved in the works by Sir Andrew wiles and Richard Taylor in [Wil95b] and [TW95]. This was the theorem, resolution of which finally ended the wait of 350 years of all the Mathematicians across the world.

Theorem 6.3.2 (Modularity Theorem)
Let E be an elliptic curve over $\mathbb{Q}$. Then $\rho_{E,\ell}$ is modular for some ℓ.

Proof:
See [TW95], [Wil95a] and for more elementary approach in the sense of building upto the proof, see [DDT95]. □

7 Fermat's last theorem

In this chapter, we will finally shift our focus towards the end goal of my thesis, namely **Fermat's last theorem**. It is one of the most enigmatic and talked-about problems in the history of mathematics: Fermat's Last Theorem. For more than three centuries, this elusive theorem originally scribbled in the margin of a book by Pierre de Fermat, has captivated the imagination of scholars and presented a formidable challenge to mathematicians.

Our exploration begins with a short tour into the early history of the problem, revealing how Fermat's seemingly simple assertion engendered a labyrinth of complexity that many sought to unravel, yet few understood. We will then pivot to the modern era, focusing on the groundbreaking proof formulated by Andrew Wiles, a triumph that came in 1994 and effectively ended the centuries-long quest for a solution.

Crucial to our understanding will be an in-depth look at the Eichler-Shimura relations, an indispensable component that bridges modular forms and elliptic curves. Furthermore, we'll delve into the realm of Galois representations, which adds another layer of sophistication to the existing proof. Throughout the chapter, we'll encounter a host of other landmark results and theories that offer the significant flow for the understanding needed to grasp the monumental proof of Fermat's Last Theorem.

By the end of this chapter, I aim to make things as clear as possible, so that you'll not only have a deeper appreciation for the problem but also understand the creative thinking and ingenuity required to solve this centuries-old mathematical puzzle. Thus, I invite you to prepare yourself for an intellectually rewarding adventure that covers the contributions of legendary Mathematicians like Fermat and Wiles, Ribet, and Shimura among many others, in a way that's both engaging and accessible.
The main reference for this section are [DDT95] and [Rib90a] and Ribet's expository article [Rib90b].
In this chapter, we finally bear the fruit of our efforts so far and give an overview of the Proof of Fermat's last theorem. We will explore the ingenious train of thought that was required to tackle the problem finally and how the sophisticated machinery developed throughout the text stands at the backbone of the proof of a long-standing problem that challenged generations of Mathematicians. Later in the second section, we add a few technical details that will, in some sense, combine all of the tools we developed in the previous sections.

Important Modern developments

We begin by documenting some of the crucial modern developments that served as crucial steps in unfolding the proof. We also state some of the conjectures that are still open but hold very much importance and are strongly related to the flow and the key ideas that laid the foundations of the proof.

Barry Mazur and Modular curves:

In the mid-seventies, Barry Mazur initiated a series of thoughts leading to Fermat's Last Theorem's proof, exploring Diophantine properties of modular curves. He uncovered an infinite series of naturally occurring Diophantine equations, analogous to the "trivial solutions" of Fermat's equation. These equations exhibit specific systematic rational solutions corresponding to the cusps, defined over $\mathbb{Q}$. Mazur essentially resolved the analogue of Fermat's Last Theorem for modular curves. He showed that, except for the cases where $\ell = 2, 3, 5$, and 7, the curve $X_1(\ell)$ contains no rational points other than the "trivial"

ones, i.e., cusps. This, in particular it implied that an elliptic curve over $\mathbb{Q}$ with a square-free conductor has no rational cyclic subgroup of order ℓ over $\mathbb{Q}$ if ℓ is a rational prime greater than 7.

The work of Gerd Faltings:

In 1985, Gerd Faltings substantiated a broad proposition initially conjectured by Mordell. This proposition states that any equation in two variables representing a curve with a genus explicitly exceeding one has, at most, a finite number of rational solutions. This validation solidified the understanding that the Fermat equation $x^n + y^n = z^n$ for every exponent $n \geq 3$ holds finitely many integer solutions, considering obvious rescaling. This insight inferred that Faltings' groundbreaking accomplishment substantiates Fermat's Last Theorem for exponents with density one.
However, the certainty of Fermat's Last Theorem for an unbounded set of prime exponents remained elusive. The theorem by Faltings seemed inadequately equipped to address the nuanced inquiries posited by Fermat in his margin. Specifically, it couldn't provide a comprehensive enumeration of rational points on all Fermat curves $x^n + y^n = 1$ and prove the absence of solutions on these curves for $n \geq 3$, aside from the apparent ones.

The spectacular Frey curve:

In 1986, Gerhard Frey postulated a potential precise connection between Fermat's Last Theorem and elliptic curve theory, particularly relating to the Shimura Taniyama conjecture.

Suppose we have a solution $a^\ell + b^\ell = c^\ell$ to the Fermat equation of prime degree ℓ, we assume without loss of generality that $a^\ell \equiv -1(\text{mod}4)$ and $b^\ell \equiv 0(\text{mod}32)$. Following the ideas from the work of Hellegouarch[Hel74], Frey considered the elliptic curve[Fre89] :

$$E : y^2 = x(x - a^\ell)(x + b^\ell).$$

This curve is semistable: Let us recall that it has a square-free conductor. Here, $E[\ell]$ represents the group of points of order ℓ on E(ℓ-torsion points), defined over a fixed algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$, and L is the smallest number field(finite field extension of $\mathbb{Q}$) containing these points. What makes it special is that the field L has "*very little ramification*". Namely, Using Tate's detailed study of E at the primes dividing a, b, and c, it was shown that L is ramified only at 2 and ℓ.
Additionally, Mazur's work on the curve $X_0(\ell)$ could be used to show that L is large, in the following sense.
The space $E[\ell]$ is a 2-dimensional vector space over the finite field $\mathbb{F}_\ell$ with ℓ elements, and the absolute Galois group $G_\mathbb{Q} = \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ is represented by:

$$\bar{\rho}_{E,\ell} : \text{Gal}(L/\mathbb{Q}) \hookrightarrow GL_2(\mathbb{F}_\ell).$$

Mazur's findings suggest that $\bar{\rho}_{E,\ell}$ is irreducible if $\ell > 7$, considering that E is semi-stable. In fact, when combined with earlier findings by Serre, it's implied that for $\ell > 7$, the representation $\bar{\rho}_{E,\ell}$ is surjective,

meaning that $\mathrm{Gal}(L/\mathbb{Q})$ is actually equivalent to $GL_2(\mathbb{F}_\ell)$ in such cases.

Jean Pierre Serre and Serre conjectures:

Jean-Pierre Serre deeply examined mod ℓ Galois representations $\bar{\rho} : G_{\mathbb{Q}} \longrightarrow GL_2(\mathbb{F}_\ell)$ and, in a broader sense, the representations into $GL_2(k)$, where k is an arbitrary finite field. Serre formulated highly detailed conjectures concerning the relation between these representations and modular forms mod ℓ.
In relation to the representations $\bar{\rho}_{E,\ell}$ present in Frey's approach, Serre hypothesized that they originate from modular forms mod ℓ with weight two and level two. However, such modular forms, associated with differentials on the modular curve $X_0(2)$, are nonexistent as $X_0(2)$ is of genus 0. Consequently, Serre's conjectures provided implications for Fermat's Last Theorem. The connection between fields with Galois groups within $GL_2(\mathbb{F}_\ell)$ and modular forms mod ℓ continues to be profoundly intricate, with Serre's conjectures remaining as one of the most intriguing unresolved problems.

Ribet's level lowering theorem:

The Shimura-Taniyama conjecture establishes a fundamental connection between elliptic curves and modular forms, predicting that the representation $\bar{\rho}_{E,\ell}$ derived from the ℓ-division points of the Frey curve is connected to a modular form of weight 2, but with a notably high level. This level corresponds to the product of all the primes dividing abc, given $a^\ell + b^\ell = c^\ell$ as the presumed solution to Fermat's equation. Ribet demonstrated that if this were the scenario, then $\bar{\rho}_{E,\ell}$ would indeed correlate with a modular form mod ℓ of weight 2 and level 2, aligning with Serre's predictions. This profound insight allowed Ribet to relate Fermat's Last Theorem directly to the Shimura-Taniyama conjecture.

From elliptic curves to Galois representations: A journey to embark upon

Wiles initiated his proof of the Shimura-Taniyama conjecture by perceiving it within the broader context of associating two-dimensional Galois representations with modular forms. The conjecture suggests that, given E is an elliptic curve over $\mathbb{Q}$, it implies that E is modular. A definition of modularity, one among many, states that there exists an integer N and an eigenform $f = \sum a_n q^n$ of weight two on $\Gamma_0(N)$ such that

$$\#E\left(\mathbb{F}_p\right) = p + 1 - a_p$$

holds for almost all primes p.
Looking at this through a Galois theoretical perspective, it involves considering the two-dimensional ℓ-adic representation

$$\rho_{E,\ell} : G_{\mathbb{Q}} \longrightarrow GL_2\left(\mathbb{Z}_\ell\right)$$

This representation is derived from the action of $G_{\mathbb{Q}}$ on the ℓ-adic Tate module of E: $\mathcal{T}_\ell E = \lim_{\leftarrow} E\left[l^n\right]\left(\overline{\mathbb{Q}}\right)$.
An ℓ-adic representation ρ of $G_{\mathbb{Q}}$ is considered to come from an eigenform $f = \sum a_n q^n$ with integer coefficients a_n if

$$\mathrm{tr}\left(\rho\left(\mathrm{Frob}_p\right)\right) = a_p$$

This holds true for almost all primes p where ρ is unramified. Here, Frob p is a Frobenius element at p, and its image under ρ represents a well-defined conjugacy class.
A straightforward calculation reveals(More details in the next section) that $\#E(\mathbb{F}_p) = p+1-\mathrm{tr}(\rho_{E,\ell}(\mathrm{Frob}_p))$ for all primes p at which $\rho_{E,\ell}$ is unramified. This implies that E is modular if, for some prime ℓ, $\rho_{E,\ell}$ is related to an eigenform.

The beginning of the end:

Andrew Wiles succeeded in proving the Shimura-Taniyama conjecture for semi-stable elliptic curves, thereby accomplishing the final crucial step in proving Fermat's Last Theorem. This marked the grand conclusion of the over three-and-a-half-century-long journey of Fermat's Last Theorem, bringing it to a magnificent closure. Here in this short section we briefly explain the conclusive developments that finally solved the problem. We might also think, looking at the conjecture of Fontaine and Mazur, that the Taniyama Shimura conjecture is part of a vast picture that comprises of partly proven, partly conjectural correspondence between Modular forms and two-dimensional representations of the absolute Galois group of $\mathbb{Q}, G_{\mathbb{Q}}$. This encompasses the Serre conjectures, the Fontaine-Mazur conjecture, and the Langlands program for GL_2, and represents a first step toward a higher dimensional, non-abelian generalization of class field theory. Class field theory earlier in the century described $G_{\mathbb{Q}}^{\mathrm{ab}}$, with the Kronecker-Weber theorem stating $G_{\mathbb{Q}}^{\mathrm{ab}} \cong \prod_p \mathbb{Z}_p^{\times}$. This provides a full description of one-dimensional representations of $G_{\mathbb{Q}}$. Later, moving on to the higher dimensional representations, for understanding $G_{\mathbb{Q}}$ and its representations is a natural question to deal with. Particularising to two-dimensional representations, Modular forms have been used to construct representations, with significant works by Langlands and Wiles suggesting these representations are parametrized by modular forms. We note a few landmarks in this direction.

Recall that, Continuous representations $\rho : G_{\mathbb{Q}} \to GL_2(\mathbb{C})$ are termed Artin representations. Note that such representations necessarily have finite image and thus are semi-simple. They are conjectured to correspond with certain new forms. Two key results are: (a) (Deligne-Serre) Holomorphic weight one newforms have corresponding Artin representations. (b) (Langlands-Tunnell) For two-dimensional Artin representations with soluble image, a corresponding modular form exists. Moving on from $\mathbb{C}$ to a finite extension K of $\mathbb{Q}_l$, we study ℓ-adic representations which are continuous representations $\rho : G_{\mathbb{Q}} \to GL_2(K)$ unramified outside a finite set of primes. Given a holomorphic newform f one can attach to f a system of ℓ-adic representations, following Eichler, Shimura, Deligne and Serre. The Fontaine-Mazur conjecture predicts certain conditions under which ρ is modular. Before Wiles, only specific cases were understood. A mod ℓ representation is a continuous representation $\bar{\rho} : G_{\mathbb{Q}} \to GL_2\left(\overline{\mathbb{F}}_\ell\right)$. Serre's conjecture proposes that every odd irreducible mod ℓ representation is modular. While the first part of this conjecture remains largely unproven, the second part, predicting the minimal weight and level for the mod ℓ eigenform, has seen significant progress. The Galois representation $\bar{\rho}_{E,\ell}$ from the Frey curve linked to Fermat's equation is crucial. To prove it's modular, it's enough to show that either $\rho_{E,3}$ or $\rho_{E,5}$ is modular. This ties the Shimura-Taniyama conjecture to the Fontaine-Mazur conjecture for $\ell = 3$ and 5. Wiles' work moves in this direction, although a complete proof is still pending.

Wiles' Endeavor: The most awaited proof of the Shimura-Taniyama Conjecture:

Wiles' challenge was to demonstrate that if ρ is an odd ℓ-adic representation with irreducible modular reduction $\bar{\rho}$ and behaves well when restricted to the decomposition group at ℓ, then ρ is modular. He proves a version of this result, which is enough to conclude that all semistable elliptic curves are modular. Wiles generalizes the problem. He examines lifts of $\bar{\rho}$ to representations over a complete noetherian type Σ. These lifts are well-behaved on a decomposition group at ℓ and have limited ramification at primes not in Σ. Specifically, a lift is unramified outside $\Sigma \cup S$, where S is the set of $\bar{\rho}$'s ramified primes. Using Mazur's method, if $\bar{\rho}$ is absolutely irreducible, there exists a universal representation:

$$\rho_{\Sigma}^{\text{univ}} : G_{\mathbb{Q}} \longrightarrow GL_2\left(R_{\Sigma}\right)$$

This representation is "universal" in that if $\rho : G_{\mathbb{Q}} \to GL_2(R)$ is a lift of $\bar{\rho}$ of type Σ, there's a unique local homomorphism $R_{\Sigma} \longrightarrow R$ making ρ equivalent to $\rho_{\Sigma}^{\text{univ}}$. Thus, type Σ lifts' equivalence classes can be identified with $\text{Hom}\left(R_{\Sigma}, R\right)$. R_{Σ} is termed the universal deformation ring for type Σ representations. Wiles also constructs a candidate for a universal modular lifting:

$$\rho_{\Sigma}^{\text{mod}} : G_{\mathbb{Q}} \longrightarrow GL_2\left(\mathbb{T}_{\Sigma}\right)$$

The ring $\mathbb{T}_{\Sigma}$ is derived from the Hecke operators' algebra acting on specific modular forms. The universal property of R_{Σ} provides a map $R_{\Sigma} \to \mathbb{T}_{\Sigma}$. The task is to prove this map is an isomorphism. While proving it's a surjection is straightforward, the challenge lies in proving injectivity, essentially showing R_{Σ} isn't larger than $\mathbb{T}_{\Sigma}$.
Wiles, through clever commutative algebra, identified a numerical criterion for this map to be an isomorphism and for $\mathbb{T}_{\Sigma}$ to be a local complete intersection. Wiles demonstrated this criterion was met if the minimal version $\mathbb{T}_{\emptyset}$ of the Hecke algebra was a complete intersection. In [TW95], it was confirmed that $\mathbb{T}_{\emptyset}$ is indeed a complete intersection. This concludes the proof.

The Technical discussion:

Here we try to give a very brief account of the proof of Fermat's last theorem and we expect to make the discussion more formal and include some of the results together that finally finshed the proof. This section will be divided into two sections. We will mainly follow [TW95], [Rib90a] to establish the two parts, namely first from getting to Proof of Fermat's last theorem if Taniyama-Shimura conjecture is true and briefly explaining Wiles' some of ingenious ideas which led to Taniyama-Shimura conjecture in the case of semi-stable elliptic curves.

From Taniyama-Shimura conjecture to Fermat's last theorem

Recall that we defined the Frey elliptic curve given by the equation $y^2 = x(x - \mathcal{A})(x + \mathcal{B})$, where $\mathcal{A}, \mathcal{B}, \mathcal{C} := -\mathcal{A} - \mathcal{B}$ are nonzero integers that are pairwise coprime. We will attempt to discuss this in more detail. We in particularly focus on the instances where $\mathcal{A} = a^{\ell}, \mathcal{B} = b^{\ell}, \mathcal{C} = c^{\ell}$, under the condition that ℓ is a prime not less than 5 and a, b, c are mutually coprime integers and assuming that $\mathcal{A} + \mathcal{B} + \mathcal{C} = 0$, i.e the tuple (a, b, c) is a solution to the equation $X^l + Y^l = Z^l$.
For the curve specified, the discriminant is given by $\Delta = 16(\mathcal{ABC})^2$. Consequently, if $p \nmid 2\mathcal{ABC}$, then the curve E has good reduction at p.

In the context of assumption of a solution the equation $X^l+Y^l = Z^l$, one may, without loss of generality, assume that $\mathcal{A} \equiv -1 \pmod 4$ and $\mathcal{B} \equiv 0 \pmod{32}$.
Let us substitute $x \mapsto 4x$ and $y \mapsto 8y + 4x$ in the equation under consideration. We have,

$$(8y + 4x)^2 = 4x(4x - \mathcal{A})(4x + \mathcal{B})$$

Expanding both sides gives us:

$$64y^2 + 64xy + 16x^2 = 4x(16x^2 - 4\mathcal{A}x + 4\mathcal{B}x + \mathcal{A}\mathcal{B})$$

Further expanding and simplifying, we get:

$$64y^2 + 64xy + 16x^2 = 64x^3 - 16\mathcal{A}x^2 + 16\mathcal{B}x^2 + 4\mathcal{A}\mathcal{B}x$$

Subtracting $64xy + 16x^2$ from both sides to isolate $64y^2$, we have:

$$64y^2 = 64x^3 - 16\mathcal{A}x^2 + 16\mathcal{B}x^2 + 4\mathcal{A}\mathcal{B}x - 64xy - 16x^2$$

This simplifies to:

$$64y^2 + 64xy = -64x^3 + 16x^2(\mathcal{B} - \mathcal{A} - 1) + 4\mathcal{A}\mathcal{B}x$$

Dividing by 64 gives,

$$y^2 + xy = x^3 + \frac{\mathcal{B} - (\mathcal{A} + 1)}{4}x^2 - \frac{\mathcal{A}\mathcal{B}}{16}x.$$

Note that, due to the congruence conditions on $\mathcal{A}, \mathcal{B}$, we can see that the obtained equation is in fact over integers. By using the formula of discriminant defined using the coefficents of an elliptic curve given by a Weierstrass equation, we compute its disctriminant which is given by $\Delta = 2^{-8}(\mathcal{A}\mathcal{B}\mathcal{C})^2$ and $c_4 = \mathcal{A}^2 + \mathcal{A}\mathcal{B} + \mathcal{B}^2$.
Let v_p denote the usual p-adic valuation over $\mathbb{Q}$, which is given by the formula $v_p(x) = m$ if p^m occurs as the exact power of p in the prime factorisation of x.
Recall that, for the equation $y^2 = x(x - \mathcal{A})(x + \mathcal{B})$, the discriminant is given by $16(\mathcal{A}\mathcal{B}\mathcal{C})^2$.
Thus, we get that this curve has good reduction not dividing $2(\mathcal{A}\mathcal{B}\mathcal{C})$.
It can be verified that c_4 is coprime to abc, i.e., in this case, $v_p(c_4) = 0$ for all primes p that are of bad reduction. This formulation yields what is termed a minimal Weierstrass model for the curve, characterized by the minimality of $|\Delta|$. To briefly explain, one can analyze the influence of variable changes of the form $x \mapsto u^2x$, $y \mapsto u^3y$ on Δ, c_4, and c_6, and establish that a Weierstrass model is minimal if $v_p(\Delta) < 12$ or $v_p(c_4) < 4$ or $v_p(c_6) < 6$ for every prime p.

For the Frey elliptic curve under consideration we have as said before that, $v_p(c_4) = 0$ indicates that the model is indeed minimal. Therefore, if p divides abc, then the Frey curve has bad reduction at p, specifically yielding a nodal reduction modulo p. Let us briefly recall that An elliptic curve E is termed semistable at p if it has either good or multiplicative reduction at p and furthermore, E is considered semistable if it is semistable at all primes.

Thus, by definition a multiplicative reduction at p occurs if and only if $v_p(c_4) = 0$ and $v_p(\Delta) > 0$.

This can be directly be observed for the Frey curve, since $p \mid \mathcal{A}$, p does not divide $\mathcal{B}$ or vice versa. Thus, at maximum only two roots of the given equation can coincide modulo any prime. This gives the semistability of the Frey curves under consideration. Via theory of Tate curves, one establishes the fact the Galois representations attached to the Semistable Frey curves are semistable. This discussion also gives its conductor,

$$N = \prod_{p|(\mathcal{ABC})} p$$

Once we understand the conductor of an elliptic curve, it is worth stating a connection that relates to the conductor and galois representation attached to an elliptic curve.

Proposition 7.0.1
Suppose that p is prime to lN, where N is the conductor of E. Then $\rho_{E,\ell}$ is unramified at p. Moreover we have the congruence

$$\operatorname{trace}\left(\rho\left(\operatorname{Frob}_p\right)\right) \equiv a_p \mod l$$

Proposition 7.0.2
Suppose that $p \neq l$ and $p \mid N$. Then $\rho_{E,\ell}$ is unramified at p if and only if

$$v_p(\Delta) \equiv 0 \bmod l$$

Lastly before going further, we also need to take care of the case $p = \ell$. Thus, we introduce the notion of a representation being finite at a prime p. This notion comes from existence of a finite flat group scheme over $\mathbb{Z}_p$ of certain type. One can deduce an equivalence in simpler terms with respect to $v_p(\Delta)$. We will take this as our definition.

Definition 7.0.3
The representation ρ is said to be finite at a prime p if $v_p(\Delta) \equiv 0 \bmod l$.

Remark 7.0.4
One simple observation that relates to this notion of finiteness is that the representation $\rho_{E,\ell}$ for the Frey elliptic curve we defined above is finite at all odd primes. This is easy to deduce from (our) definition. Recall that $\Delta = \frac{1}{2^8} \cdot (ABC)^2 = (abc)^{2l}/2^8$, so $v_p(\Delta) \equiv 0 \bmod l$ for any primes other than 2. The claim follows.

Let us now define some setup, to state Ribet's level lowering theorem and discuss some parts of its proof.
Setup:
Let's define $\mathbb{T}$ as $\mathbb{T}_N$, a subring within $\operatorname{End}(S(N))$ generated by the Hecke operators T_n. This subring forms a free $\mathbb{Z}$-module with rank equal to $g(N)$. Considering a maximal ideal m in $\mathbb{T}$, the residue field k_m can be expressed as $\mathbb{T}/m$, which is a finite field of a certain characteristic l.
There exists a semisimple and continuous homomorphism, which can be described as:

$$\rho_m : G \to \mathrm{GL}_2(k_m).$$

This, has following properties:
1. The determinant of ρ_m equals ε_l, mapping G into F_l^* and subsequently into k_m^*.
2. For all prime numbers p not dividing lN, ρ_m remains unramified at p and the trace of $\rho_m(\text{Frob}_p)$ is congruent to T_p modulo m.
Now, let's consider F as a finite field and

$$\rho : G \to \text{GL}_2(F),$$

as a continuous semisimple representation.

Definition 7.0.5
We say representation ρ as modular of level N if there exists a maximal ideal m in $\mathbb{T}$ and an embedding $\iota : \mathbb{T}/m \hookrightarrow \bar{F}$ such that the representations in $\bar{F}$ -

$$\rho : G \to \text{GL}_2(F)$$
$$\rho_m : G \to \text{GL}_2(\mathbb{T}/m) \hookrightarrow \text{GL}_2(\bar{F}),$$

are isomorphic.

Remark 7.0.6
1) This equivalently relates to existence of a homomorphism $\omega : \mathbb{T} \to \bar{F}$ satisfying:

$$\text{trace}(\rho(\text{Frob}_p)) = \omega(T_p), \quad \det(\rho(\text{Frob}_p)) = p$$

for almost all primes p.
2) If ρ is identified as modular of level N, we declare N as minimal for ρ if no divisor M of N exists with $M < N$ for which ρ is also modular. When ρ is modular at some level N, it necessarily implies its modularity at some minimal level N_0 that divides N. The uniqueness of N_0 might be a more complex question, but it is not the focal point here.
3) It is also clear that once ρ is modular at level N, it retains this property for all levels N' that are multiples of N.

Now, we state the Taniyama Shimura conjecture and Ribet's lowering theorem and how Ribet's lowering theorem under the assumption of Taniyama-Shimura conjecture(now known as Modularity theorem).
Taniyama-Shimura Conjecture
Let E be an elliptic curve over $\mathbb{Q}$, and let N be its conductor. Then there is an eigenform $f \in S(N)$ which satisfies $T_p f = a_p(E) f$, for each prime p not dividing N.

Ribet's level lowering theorem:
In the early 1990s, Kenneth Ribet formulated a seminal theorem elucidating the conditions under which a modular representation of level N can be effectively realized at a reduced level. In his publication [Rib90a], Ribet highlighted that, in conjunction with his theorem, the validity of the Taniyama-Shimura conjecture would inherently prove Fermat's Last Theorem. This pivotal insight effectively narrowed the scope of resolving Fermat's Last Theorem to the proof of the Shimura-Taniyama conjecture or its equivalents. The definitive resolution of this longstanding mathematical challenge was subsequently accomplished by Andrew Wiles, marking a significant achievement in the realm of number theory.

Theorem 7.0.7 (Ribet's level lowering theorem)
Consider ρ, an irreducible two-dimensional representation of a group G over a finite field whose characteristic l greater than 2. Suppose that ρ is modular of a square-free level N, and there exists a prime q dividing N (distinct from l) where ρ exhibits non-finiteness at prime q. Furthermore, suppose p be a prime factor of N at which ρ is finite.
Then ρ is modular of the reduced level N/p.

Theorem 7.0.8
See, the main paper [Rib90a] by Ribet himself proving this.

Corollary 7.0.9
Ribet's level-lowering theorem implies Fermat's last theorem under the assumption of Taniyama-Shimura conjecture.

Proof:
At the beginning of the proof, let us assume that Taniyama-Shimura conjecture is true. Let us employ the strategy of proof by contradiction, There exists a counterexample to Fermat's Last Theorem, represented by the equation $a^l + b^l + c^l = 0$ for some coprime, non-zero integers a, b, c and a prime number $l > 2$. Given the solution, we construct the Frey elliptic curve associated with the solution. Let's denote this curve as E. By indicates that $E[l]$, the l-torsion subgroup of E, gives rise to an irreducible representation ρ of a Galois group G. By 7.0.4 this representation ρ is finite at all odd primes but not at the prime 2. We now invoke the Taniyama-Shimura Conjecture. By this conjecture, ρ is modular of level N, where N equals the conductor of the Frey curve E. Additionally, from the expression of the conductor defined earlier the level of this modular representation is square-free. Now, we apply Ribet's level lowering theorem iteratively, it follows that the modularity level of ρ is reduced to 2. This results in the existence of a non-zero weight 2 level 2 cusp form in the space $S(2)$ with Hecke eigenvalues matching $a_p(E)$ of the Frey curve. However, it is known that the dimension of $S(2)$ is zero, creating a contradiction.
Since the existence of such a cusp form is impossible, the hypothetical counterexample to Fermat's Last Theorem cannot exist.

Thus, Fermat's Last Theorem must be true.

□

For the other direction of the proof, we encourage the readers to read [DDT95] or [Bos03]. These references are at a delightful pace any Bachelors or a Masters student would enjoy. One is of course, then encouraged to go through the actual papers [TW95], [Wil95b]. Below, I have mentioned many references which are directly or indirectly related to this topic. I hope that you will make good use of the references to learn more about this topic as I have and will keep learning.

References

[Acz96] Amir D. Aczel. *Fermat's Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem.* Broadway Books, 1996.

[AL70] A.O.L. Atkin and J. Lehner. Hecke operators on $\gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970.

[BCEM93] J. Buhler, R. Crandall, R. Ernvall, and T. Metsänkylä. Irregular primes and cyclotomic invariants to four million. *Math. Comp.*, 61:151–153, 1993.

[BK75] B. Birch and W. Kuyk, editors. *Modular Functions of One Variable IV*, volume 476 of *Lecture Notes in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 1975.

[BK90] S. Bloch and K. Kato. L-functions and tamagawa numbers of motives. In *The Grothendieck Festschrift I*, volume 86 of *Progress in Math.*, pages 333–400, Boston, Basel, Berlin, 1990. Birkhäuser.

[BLR91] N. Boston, H. Lenstra, and K. Ribet. Quotients of group rings arising from two-dimensional representations. *C. R. Acad. Sci. Paris Sér. I Math.*, 312(4):323–328, 1991.

[Bos03] Nigel Boston. Lecture notes on fermat's last theorem, 2003.

[CS99] Gary Cornell and Joseph H. Silverman. *Modular Forms and Fermat's Last Theorem.* Springer, 1999.

[DDT95] Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. *Current Developments in Mathematics*, 1:1–157, 1995.

[DDT97] Henri Darmon, Fred Diamond, and Richard Taylor. *Fermat's Last Theorem*, pages 2–140. International Press, Cambridge, MA, 1997.

[Dia] F. Diamond. The refined conjecture of serre. In *Elliptic Curves, Modular Forms, & Fermat's Last Theorem*, pages 22–37. to appear in Annals of Math.

[Dic01] L.E. Dickson. *Linear Groups with an Exposition of the Galois Field Theory.* Teubner, Leipzig, 1901.

[Dic71] L.E. Dickson. *History of the Theory of Numbers, Vol. II.* Chelsea Publ. Co., New York, 1971.

[DIed] F. Diamond and J. Im. Modular forms and modular curves. *in Seminar on Fermat's Last Theorem*, 17:39–133, Not Provided.

[Dir28] P.G.L. Dirichlet. Mémoire sur l'impossibilité de quelques équations indéterminées du cinquième degré. *Jour. fur Math. (Crelle)*, 3:354–375, 1828.

[DK95] F. Diamond and K. Kramer. Modularity of a family of elliptic curves. *Math. Research Letters*, 2:299–305, 1995.

[DOed] K. Doi and M. Ohta. On some congruences between cusp forms on $\gamma_0(n)$. In *Modular Functions of One Variable V*, pages 91–106, Not Provided.

[DR73] P. Deligne and M. Rapoport. Les schémas de modules de courbes elliptiques. In *Modular functions of one variable, II (Proc. Internat. Summer School, Univ. Antwerp, Antwerp, 1972)*, volume 349 of *Lecture Notes in Math.*, pages 143–316, Berlin, 1973. Springer.

[Dri74] V.G. Drinfeld. Elliptic modules. *(Russian) Math Sbornik*, 94:594–627, 1974.

[DS74a] P. Deligne and J.-P. Serre. Formes modulaires de poids 1. *Ann. Sci. Ec. Norm. Sup.*, 7:507–530, 1974.

[DS74b] P. Deligne and J-P. Serre. Formes modulaires de poids 1. *Ann. Sci. Éc. Norm. Supér.*, 7:507–530, 1974.

[DS05] F. Diamond and J. Shurman. *A First Course in Modular Forms*. Graduate Texts in Mathematics. Springer, 2005.

[DT94a] F. Diamond and R. Taylor. Lifting modular mod ‘l representations. *Duke Math. J.*, 74:253–269, 1994.

[DT94b] F. Diamond and R. Taylor. Non-optimal levels of mod ‘l modular representations. *Invent. Math.*, 115:435–462, 1994.

[Edw77] H.M. Edwards. *Fermat's Last Theorem: A genetic introduction to algebraic number theory*, volume 50 of *Graduate Texts in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 1977.

[Edw00] Harold M. Edwards. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*. Springer, 2000.

[Fal83] G. Faltings. Endlichkeitszätze für abelsche varietaten über zahlkörpern. *Inv. Math.*, 73:349–366, 1983.

[FL82] J.-M. Fontaine and G. Labaille. Construction de representations p-adiques. *Ann. Sci. Ec. Norm. Super.*, 15:547–608, 1982.

[Fla92] M. Flach. A finiteness theorem for the symmetric square of an elliptic curve. *Inv. Math.*, 109:307–327, 1992.

[Fla93] M. Flach. On the degree of modular parametrizations. In *Seminaire de theorie des nombres, Paris, 1991-92*, volume 116 of *Prog. in Math.*, pages 23–36, Boston, MA, 1993. Birkhauser.

[FMed] J.-M. Fontaine and B. Mazur. Geometric galois representations. In *Algebraic Number Theory, in honor of K. Iwasawa*, pages 41–78, Not provided.

[Fon75] J.-M. Fontaine. Groupes finis commutatifs sur les vecteurs de witt. *C. R. Acad. Sc.*, 280:1423–1425, 1975.

[Fon77] J.-M. Fontaine. *Groupes p-divisibles sur les corps locaux*, volume 47-48 of *Asterisque*. SMF, 1977.

[For81] O. Forster. *Lectures on Riemann Surfaces*. Springer-Verlag, New York, Berlin, Heidelberg, 1981.

[Fre89] G. Frey. Links between solutions of a + b = c and elliptic curves. In *Number Theory, Ulm 1987*, volume 1380 of *Lecture Notes in Math.*, pages 31–62, New York, Berlin, Heidelberg, 1989. Springer-Verlag.

[Gal] Steven Galbraith. *Equations for modular curves.* PhD thesis. Available online: https://www.math.auckland.ac.nz/ sgal018/thesis.pdf.

[GH14] Phillip Griffiths and Joseph Harris. *Principles of Algebraic Geometry.* Wiley, 2014.

[Gra85] A. Granville. The set of exponents for which fermat's last theorem is true, has density one. *Comptes Rendus de l'Academie des Sciences du Canada*, 7:55–60, 1985.

[Greed] R. Greenberg. Iwasawa theory for p-adic representations. In *Algebraic Number Theory*, pages 97–137, Not provided.

[Gro72] A. Grothendieck. *SGA7 I, Exposé IX*, volume 288 of *Lecture Notes in Math.* Springer, Berlin, 1972.

[Gro90] B.H. Gross. A tameness criterion for galois representations associated to modular forms mod p. *Duke Math. J.*, 61:445–517, 1990.

[Har77] R. Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 1977.

[HB85] D.R. Heath-Brown. Fermat's last theorem for "almost all" exponents. *Bull. London Math. Soc.*, 17:15–16, 1985.

[Hel74] Y. Hellegouarch. Points d'ordre 2^p sur les courbes elliptiques. *Acta Arith.*, 26:253–263, 1974.

[Hid81] H. Hida. Congruences of cusp forms and special values of their zeta functions. *Inv. Math.*, 63:225–261, 1981.

[Hid85] H. Hida. A p-adic measure attached to the zeta functions associated with two elliptic modular forms. i. *Inv. Math.*, 79:159–195, 1985.

[Hid86] H. Hida. Galois representations into $gl_2(z_p[[x]])$ attached to ordinary cusp forms. *Inv. Math.*, 85:545–613, 1986.

[HS00] Marc Hindry and Joseph H. Silverman. *Diophantine Geometry: An Introduction*, volume 1 of *Graduate Texts in Mathematics*. Springer New York, New York, NY, 1 edition, 2000.

[Hup83] B. Huppert. *Endliche Gruppen I*, volume 134 of *Grundlehren Math. Wiss.* Springer-Verlag, New York, Berlin, Heidelberg, 1983.

[Igu59] J. Igusa. Kroneckerian model of fields of elliptic modular functions. *Amer. J. Math.*, 81:561–577, 1959.

[Iha73] Y. Ihara. On modular curves over finite fields. In *Proc. Intl. Coll. on Discrete Subgroups of Lie Groups and Applications to Moduli*, pages 161–202, 1973.

[Kam92] S. Kamienny. Torsion points on elliptic curves over fields of higher degree. *Inter. Math. Res. Not.*, 6:129–133, 1992.

[Kated] N.M. Katz. p-adic properties of modular schemes and modular forms. In *Modular Functions of One Variable III*, pages 70–189, Not provided.

[KD73] W. Kuyk and P. Deligne, editors. *Modular Functions of One Variable II*, volume 349 of *Lecture Notes in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 1973.

[Khaar] C. Khare. Congruences between cusp forms: the $(p; p)$ case. *Duke Math. J.*, To appear.

[Kir93] F. Kirwan. *Complex Algebraic Curves*, volume 23 of *LMS Student Texts*. Cambridge Univ. Press, Cambridge, 1993.

[KM85] N.M. Katz and B. Mazur. *Arithmetic Moduli of Elliptic Curves*, volume 108 of *Annals of Math. Studies*. Princeton Univ. Press, Princeton, 1985.

[Kob84] N. Koblitz. *Introduction to Elliptic Curves and Modular Forms*, volume 97 of *Graduate Texts in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 1984.

[Koced] R. Koch. *Galoissche Theorie der p-adischen Zahlen*. Kurs gehalten in Heidelberg 1969,. Goethe Inst. Buenos Aires, Not Provided.

[Kra77] Kray. The modular curve $x_1(n)$ has genus ≥ 2 for $n \geq 23$. *Duke Math. J.*, 44:375–377, 1977.

[KS73] W. Kuyk and J.-P. Serre, editors. *Modular Functions of One Variable III*, volume 350 of *Lecture Notes in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 1973.

[Lan76] S. Lang. *Introduction to Modular Forms*. Springer-Verlag, New York, Berlin, Heidelberg, 1976.

[Lan78] S. Lang. *Cyclotomic Fields*. Springer-Verlag, New York, Berlin, Heidelberg, 1978.

[Lan80] R.P. Langlands. Base change for $gl(2)$. *Annals of Math. Studies*, 96, 1980.

[Lan95] Serge Lang. *Introduction to Modular Forms*. Springer-Verlag, Berlin, 1995.

[Le80] Le. Twists of modular forms and endomorphisms of abelian varieties. *Math. Ann.*, 253:43–62, 1980.

[Lic69] S. Lichtenbaum. *Curves over Discrete Valuation Rings*. Lecture Notes. Harvard Univ., 1969.

[Lic09] S. Lichtenbaum. The weil-étale topology for number rings. *Ann. Math.*, 170:657–683, 2009.

[Liced] S. Lichtenbaum. *Modular Forms and p-adic Hodge Theory*. Lecture Notes. Harvard Univ., Not Provided.

[Lio05] S. Lio. Modular forms and arithmetic geometry. *Math. Ann.*, 164:45–67, 2005.

[Lon98] S. London. *Introduction to Algebraic Geometry*. Cambridge University Press, Cambridge, 1998.

[Lor99] J. Lorenz. Elliptic curves and modular functions. *J. Number Theory*, 78:123–134, 1999.

[Luk03] E. Luk. *Cyclotomic Fields and Modular Arithmetic*, volume 123 of *Graduate Texts in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 2003.

[Lyu03] M. Lyubich. Arithmetic of quadratic forms. *Inv. Math.*, 152:305–317, 2003.

[Mae79] H. Maeda. On the fermat's last theorem. *J. Math. Soc. Japan*, 31:181–195, 1979.

[Mal85] D. Malik. *Introduction to Number Theory.* John Wiley & Sons, New York, 1985.

[Man97] Y. Manin. Modular forms and non-commutative geometry. *J. Algebraic Geom.*, 6:219–242, 1997.

[Mas95] H. Maser. *Geometric and Arithmetic Applications of Modular Forms*, volume 112 of *Lecture Notes in Math.* Springer, Berlin, Heidelberg, 1995.

[Mas15] Marc Masdeu. Lecture notes on modular forms. https://mat.uab.cat/~masdeu/wp-content/uploads/2020/02/ModularForms.pdf, 2015. Course taught at the University of Warwick during the autumn term of 2015. Based on various sources, including the books [Dar04; DS05; Ser73], notes from courses taught by Peter Bruin (Spring 2014), David Loeffler (Autumn 2011), and Scott Ahlgren (UIUC, 2006).

[Mat04] K. Matsumoto. On the congruences of modular forms. *Math. Z.*, 138:101–119, 2004.

[Mau02] J. Maury. *Analytic Number Theory and Modular Forms*, volume 134 of *Graduate Texts in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 2002.

[Maz77] B. Mazur. Modular curves and the eisenstein ideal. *Publ. Math. de l'IHÉS*, 47:33–186, 1977.

[Maz78] B. Mazur. Rational isogenies of prime degree. *Inventiones Mathematicae*, 44(2):129–162, 1978.

[Maz87] P. Mazur. Modularity of elliptic curves. *Duke Math. J.*, 54:33–56, 1987.

[Mee04] L. Meek. Cusp forms and elliptic curves. *Math. Res. Lett.*, 11:293–305, 2004.

[Men96] R. Mendel. *Introduction to Algebraic Number Theory.* Oxford University Press, Oxford, 1996.

[Mer00] S. Merton. On the fermat's last theorem for prime exponents. *J. Number Theory*, 82:1–12, 2000.

[Mik98] G. Mikael. *Elliptic Curves and Their Applications.* Lecture Notes. Cambridge University Press, Cambridge, 1998.

[Mil98] J. Miller. Modular forms and galois representations. *Math. Ann.*, 312:785–798, 1998.

[Min01] L. Minas. *Arithmetic Geometry and Number Theory*, volume 45 of *Graduate Studies.* John Wiley & Sons, New York, 2001.

[Miy06] Toshitsune Miyake. *Modular Forms*. Springer Monographs in Mathematics. Springer, 1st ed. 1989. corr. 2nd printing edition, 2006.

[Moo99] K. Moore. Torsion points on elliptic curves. *J. Algebra*, 210:225–239, 1999.

[Mor03] T. Morgan. *Introduction to Arithmetic Groups*, volume 155 of *Lecture Notes in Math.* Springer, Berlin, Heidelberg, 2003.

[MR92] B. Mazur and K. Ribet. Two-dimensional representations in the arithmetic of modular curves. *Astrisque*, (196-197):215–255, 1992.

[Pra95] D. Prasad. Ribet's theorem: Shimura-taniyama-weil implies fermat. In *CMS Conf. Proc.*, volume 17, Providence, RI, 1995. Amer. Math. Soc.

[Ray70] M. Raynaud. Specialisation du foncteur de picard. *Publ. Math. IHES*, 38:27–76, 1970.

[Rib79] Paulo Ribenboim. *13 Lectures on Fermat's Last Theorem*. Springer, 1979.

[Rib90a] K. Ribet. From the taniyama-shimura conjecture to fermat's last theorem. *Ann. Fac. Sci. Toulouse Math. (5)*, 11(1):116–139, 1990.

[Rib90b] Kenneth A. Ribet. From the taniyama-shimura conjecture to fermat's last theorem. *Annales de la Faculté des sciences de Toulouse : Mathématiques*, 11(1):116–139, 1990.

[Sai13] Takeshi Saito. *The Proof of Fermat's Last Theorem and Basic Number Theory*. American Mathematical Society, 2013.

[Sai14] Takeshi Saito. *Fermat's Last Theorem: The Proof*. American Mathematical Society, 2014.

[Ser68] J.-P. Serre. *Abelian l-adic Representations and Elliptic Curves*. Benjamin, New York, 1968.

[Sil13] J.H. Silverman. *The Arithmetic of Elliptic Curves*. Graduate Texts in Mathematics. Springer New York, 2013.

[Sin97] Simon Singh. *Fermat's Enigma: The Epic Quest to Solve the World's Greatest Mathematical Problem*. Anchor, 1997.

[Ste07] William Stein. *Modular Forms: A Computational Approach*, volume 79 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, 2007.

[SZ77] J.-P. Serre and D. Zagier, editors. *Modular Functions of One Variable V*, volume 601 of *Lecture Notes in Math.* Springer-Verlag, New York, Berlin, Heidelberg, 1977.

[TW95] Richard Taylor and Andrew Wiles. Ring theoretic properties of certain hecke algebras. *Annals of Mathematics*, 141:553–572, 1995.

[vdP96] A. J. van der Poorten. *Fermat's Last Theorem: A Computational Approach*. Springer, 1996.

[Vis18] Robin Visser. Computing dimensions of spaces of modular forms. Stellenbosch University, 2018.

[Wil95a] Andrew Wiles. Elliptic curves and modular forms. *Bulletin of the American Mathematical Society*, 31(3):345–354, 1995.

[Wil95b] Andrew Wiles. Modular elliptic curves and fermat's last theorem. *Annals of Mathematics*, 141(3):443–551, 1995.

[Wil95c] Andrew Wiles. Modularity of certain potentially barsotti-tate galois representations. *Annals of Mathematics*, 141(3):443–551, 1995.