The Theory of Modular Forms and the Converse Theorem of Weil

Tejasi Bhatnagar

A dissertation submitted for the partial fulfilment of BS-MS dual degree in Science

INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH MOHALI

April 2019

Certificate of Examination

This is to certify that the dissertation titled "The theory of modular forms and the Converse Theorem of Weil" submitted by Ms. Tejasi Bhatnagar (Reg. number –MS14071) for the partial fulfillment of BS-MS dual degree program of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Abhik Ganguli (Supervisor)

Dr. Chetan Balwe

Dr. Varadharaj Srinivasan

Date: April 2019

Declaration of Authorship

The work presented in this dissertation has been carried out by me under the guidance of Dr. Abhik Ganguli at the Indian Institute of Science Education and Research, Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Tejasi Bhatnagar
(Candidate)

Date: April 2019

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Abhik Ganguli
(Supervisor)

Date: April 2019

"The universe is infinite and so are we!"

-Anonymous

Acknowledgements

There are many people I want to thank. I'll start with my parents for giving me the freedom of mind and motivating me to pursue excellence in whatever I choose to do.

I am grateful to IISER Mohali, where I have grown both personally and academically. A special mention to the library of IISER, where I spent ninety percent of my time figuring out all the math they taught us.

I want to express my gratitude to all the professors who guided or taught me. Part of the reason I am pursuing math is because of all of you!

A big thank you to all my friends at IISER. Special shoutout to Avneet, Diksha, and Nitin for always being there for me. I will really miss you all.

Above all, to Dr. Abhik Ganguli, for being a fine mentor, catching the details I missed while understanding the area during our discussions, helping me with complex analysis in the later part of my thesis and finally, for giving a wonderful start to my mathematical career. Thank you Sir!

Contents

Abstract

In this thesis we provide an introduction to the theory of modular forms. We aim to show two major results. The first one shows that the space of modular forms is finite dimensional. The second result is the "Converse Theorem of Weil" on L-functions associated to modular forms. As we go on we will encounter some very interesting mathematical objects such as modular curves, Hecke operators and L-functions.

Introduction

A *modular form of weight k* is a holomorphic function on the upper half complex plane $\mathbb{H}$, such that it is invariant up to an "automorphy factor" with respect to the action of congruence subgroups of $SL_2(\mathbb{Z})$ on $\mathbb{H}$. This transformation law is known as the "modularity condition"(See Section 1.1). There are two parameters associated to a modular form – its weight and the level. The weight is determined by the automorphy factor and the level depends on the finite index congruence subgroup with respect to which it satisfies the modularity condition. Modular forms are "holomorphic at the points at infinity" known as "cusps". This means that they have a power series expansion about 0, when viewed as a function on the unit disc (See Section 1.2). Moreover, the modularity condition implies that these functions are periodic. This gives us a Fourier series expansion for the function which is exactly the above power series expansion.

Modular forms are one of the central objects of interest in the area of number theory. They are studied extensively because of the much celebrated Taniyama–Shimura conjecture which highlights a deep relation between these functions and special cubic curves known as elliptic curves. This conjecture was the key to proving Fermat's Last Theorem. However, the classical theory of modular forms can be traced back to the works of Jacobi and decades later when Ramanujan introduced the famous τ–function. He studied the product $\prod_{n=1}^{\infty} q(1-q^n)^{24}$ which can be expanded as a q–series given by $\sum_{n=1}^{\infty} \tau(n)q^n$. Ramanujan conjectured that the τ–function is multiplicative. Moreover he discovered that for a prime p, the coefficients satisfy the property: $\tau(p^{n+2}) = \tau(p)\tau(p^{n+1}) - p^{11}\tau(p^n)$. In addition to the above properties, he considered the Dirichlet series $\sum_{n=1}^{\infty} \tau(n)n^{-s}$ and believed that it has an Euler product representation.

$$\sum_{n=1}^{\infty} \tau(n)n^{-s} = \prod_{p} (1 - \tau(p)p^{-s} + p^{11-2s})^{-1}$$

The answers to these remarkable results lie in the fact that the sum defined above with $\tau(n)$ as its coefficients is a modular form. In this thesis, we aim to provide the general theory behind such results.

To begin with, the first three chapters of the thesis is dedicated to proving that the space of modular forms is finite dimensional. To that end, we will study modular curves. Suppose Γ is a subgroup of $SL_2(\mathbb{Z})$ of finite index. A *modular curve* with respect to Γ is the quotient space of orbits under the action of Γ on $\mathbb{H}$. We will see that every modular curve is in fact a Riemann surface. We will add the points at infinity called "cusps" to the modular curve so as to make it into a compact Riemann surface. We will see that a weight $2k$ modular form with respect to Γ can be viewed as a k-fold differential form on the associated compact modular curve $X(\Gamma)$. The dimension is then calculated using the Riemann Roch Theorem in terms of the known data of the modular curve (See Theorem 2.13.5).

The answer to the first two results related to the τ–function is dealt by the Hecke theory of modular forms. This is developed in Chapter 3. We study the space of modular forms via a family of normal operators called "Hecke operators" acting on it. In particular for the space of cusp forms we find a suitable basis of orthogonal "eigenforms" for the family of Hecke operators.

The central object of study in chapter 4 is the *L-function* associated to a modular form. Here we will look for an answer to the third property of the τ–function dealing with the Euler product expansion of the Dirichlet series, in a much more general setting. Suppose the q–expansion of a modular form f is given by $\sum_{n=0}^{\infty} a_n q^n$. We associate the Dirichlet series to f given by the sum $\sum_{n=1}^{\infty} a_n n^{-s}$ known as its L-function. Furthermore, we will show that the modular forms whose L-function have an Euler product expansion are precisely the "normalized eigenforms". (See Theorem 4.2.11).

The L-function of a cusp form can be analytically continued to the whole complex plane. It also satisfies a functional equation. Hecke conjectured and later Weil proved the Converse Theorem for L-functions associated to cusp forms of level N. Here we will work with a family of "twisted L-functions." More precisely these are Dirichlet series of the form $\sum_{n=1}^{\infty} \chi(n) a_n n^{-s}$ where χ is a suitable primitive Dirichlet charcter. Given sufficiently many such L-series with analytic continuation and a functional equation, along with appropriate growth conditions on the coefficients a_n, the Converse Theorem guarantees that the coefficients come from a cusp form $\sum_{n=1}^{\infty} a_n q^n$ of level N. (See Theorem 4.4.6).

Our main references are the books *A first course in modular forms* by Diamond and Shurman[2] and *Automorphic forms and representations* by Daniel Bump[1].

Chapter 1

The theory of modular forms

Definition 1.0.1 (The modular group)*.* The modular group is the group of all invertible 2 by 2 matrices with integer entries having determinant 1. We denote it by $SL_2(\mathbb{Z})$.

The modular group acts on the upper half complex plane, $\mathbb{H}$ as follows: For any $\gamma \in SL_2(\mathbb{Z})$ such that $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ define:

$$\gamma\tau = \frac{a\tau + b}{c\tau + d}.$$

Lemma 1.0.2*.* The modular group maps the upper half plane to itself.

Proof. Easy calculations show that $\mathrm{Im}(\gamma\tau) = \frac{(ad-bc)\,\mathrm{Im}(\tau)}{|c\tau+d|^2}$. Since $ad - bc = 1, \tau \in \mathbb{H}$ implies that $\mathrm{Im}(\tau) > 0$. We therefore have that $\mathrm{Im}(\gamma\tau) = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2} > 0$. □

One can easily check the following to confirm that the action defined above is in fact a group action.

1. $I(\tau) = \tau$ where I denotes the identity matrix.
2. $(\gamma\gamma')\tau = \gamma(\gamma'\tau)$ for some $\gamma, \gamma' \in SL_2(\mathbb{Z})$

Lemma 1.0.3*.* The modular group is generated by the matrices:

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Proof. Notice that $T^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$ and $S^2 = -I$. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. If $c = 0$, then $a = d = \pm 1$ so that $\gamma = \pm T^{b'}$ where $b' = \pm b$. In this case, $\gamma = T^{b'}$ or $S^2 T^{b'}$. Next, assume that $c \neq 0$. Observe that multiplying S with γ switches the rows with a sign change. Therefore without loss of generality we can assume that $|a| \geq |c|$, otherwise we can multiply by S to interchange the rows. By the division algorithm, there exists integers q and r such that $a = cq + r$ with $0 \leq r < c$. Multiplying T^{-q} with γ, we see that

$$T^{-q}\gamma = \begin{pmatrix} a - cq & b - qd \\ c & d \end{pmatrix} = \begin{pmatrix} r & b - qd \\ c & d \end{pmatrix}$$

Since $r < |c|$, we again switch the rows by multiplying by S to get

$$ST^{-q}\gamma = \begin{pmatrix} -c & -d \\ r & b - qd \end{pmatrix}$$

Applying the division algorithm and repeating this process, at some point the lower left entry will become 0 and we'll be done by the previous case. It follows that, any arbitrary matrix in $SL_2(\mathbb{Z})$ is obtained by multiplying suitable powers of S and T. □

The matrix T maps τ to $\tau + 1$. In other words, it acts via translation by 1 on τ. The second matrix takes τ to $-1/\tau$. Under this transformation the points inside the unit disc in the upper half plane are mapped to points outside the unit disc and vice versa. These transformations will help us determine the fundamental domain for the action of $SL_2(\mathbb{Z})$ on $\mathbb{H}$ in section 2.1.

1.1 Modular forms

Definition 1.1.1 (Modular form of weight k). Let k be an integer. A modular form of weight k, for the group $SL_2(\mathbb{Z})$, is a function satisfying the following:

1. f is holomorphic on $\mathbb{H}$.
2. (Modularity condition) For all $\gamma \in SL_2(\mathbb{Z}), \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$f(\gamma\tau) = (c\tau + d)^k f(\tau) \text{ for all } \tau \in \mathbb{H}.$$

3. The function f is holomorphic at ∞. We will make this precise in section 1.2.

The first example of a modular form of weight k is the zero function. In fact one can check that for odd weights, the zero function is the only modular form.

To check the modularity condition for all the matrices of $SL_2(\mathbb{Z})$ is a very tedious job. However the modularity condition reduces to the following when we check it on the generators.

$$f(\tau+1) = f(\tau) \text{ and } f(-1/\tau) = \tau^k f(\tau) \text{ for all } \tau \in SL_2(\mathbb{Z}). \tag{1.1.2}$$

The next proposition helps us to conclude that checking the modularity condition for all the matrices in $SL_2(\mathbb{Z})$ is equivalent to checking the condition on the generators. We introduce some notation first. Suppose $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. Let $\jmath(\gamma, \tau) = (c\tau + d)^k$. Denote $(c\tau+d)^{-k} f(\gamma\tau)$ by $f[\gamma]_k$. The term $\jmath(\gamma,\tau)$ is called the automorphy factor, while $[\gamma]_k$ is called the weight k operator. The modularity condition for a modular form f is now equivalent to the following statement: $f[\gamma]_k = f$.

Proposition 1.1.3. Suppose that f is a modular form of weight k and $\gamma_1, \gamma_2 \in SL_2(\mathbb{Z})$. Then,

1. $\jmath(\gamma_1\gamma_2, \tau) = \jmath(\gamma_1, \gamma_2\tau)\jmath(\gamma_2, \tau)$.
2. $(f[\gamma_1]_k)[\gamma_2]_k = f[\gamma_1\gamma_2]_k$.

The second part, tells us that, if the modularity condition holds for two matrices then it holds for their product as well. Therefore it is now sufficient to look at the conditions in (1.1.2).

Proof. Part 4.4.11 is a direct calculation. To see the second part, compute that

$$\begin{aligned} (f[\gamma_1]_k)[\gamma_2]_k(z) &= \jmath(\gamma_1, \gamma_2 z)^{-k} \jmath(\gamma_2, z)^{-k} f(\gamma_1\gamma_2 c) \\ &= \jmath(\gamma_1\gamma_2, z) f(\gamma_1\gamma_2 z) \\ &= f[\gamma_1\gamma_2]_k \end{aligned}$$

□

It is easy to observe that the set of modular forms of weight k form a vector space over $\mathbb{C}$ with the addition defined by the usual addition of functions.

1.2 The q-expansion of a modular form

The map $\tau \mapsto e^{2\pi i\tau}$ is a surjective holomorphic map of the upper half plane to the punctured unit disc. This transformation will help us view modular forms as functions on the unit disc instead of the upper half plane. Let $q = e^{2\pi i \tau}$. Since modular forms are periodic, it is well defined to set $\widetilde{f}(q) = f(\log(q)/2\pi i) = f(\tau)$. The function f is now defined on the open punctured unit disc via $\widetilde{f}$. Note that, as $\tau \to i\infty, q \to 0$. We say that the function f is holomorphic at infinity if $\widetilde{f}$ can be analytically extended to the whole unit disc, so that it is analytic at 0. The function therefore can be written as a power series around 0. By abuse of notation we write:

$$\widetilde{f}(q) = f(q) = \sum_{n \geq 0} a_n q^n.$$

1.3 First examples

For the matrix $-I$, the modularity condition implies that $f(\tau) = (-1)^k f(\tau)$. Therefore, for odd weights the only modular form is the 0 function. For a non-trivial example of a modular form of weight k, consider the double sum:

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

The series converges absolutely when $k > 2$. We see that when $k > 2$,

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

The last equality comes from the fact that the series is absolutely convergent, thus we can change the order of the sum over (m, n) to $(m, m + n)$. One can similarly check the second modularity condition. The series $G_k(\tau)$ is known as the Eisenstien series of weight k. A specific example of a modular form of weight 12 is the Δ function given by:

$$\Delta(\tau) = (60G_4(\tau))^3 - 27(140(G_6))^2.$$

The q-expansion of some Eisenstien series are given as follows:

$$G_4(q) = 1/240 + q + 9q^2 + 28q^3 + 73q^4 + O(q^5)$$

$$G_6(q) = -1/504 + q + 33q^2 + 244q^3 + 1057q^4 + O(q^5)$$

The q-expansion of the Δ function is:

$$\Delta(q) = q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 + O(q^6)$$

Notice that the constant term a_0 of the q-series for the Δ function is 0. Such forms are called cusp forms. These form a subspace of $\mathcal{M}_k(SL_2(\mathbb{Z}))$, which we denote by $\mathcal{S}_k(SL_2(\mathbb{Z}))$.

1.4 Congruence subgroups of $SL_2(\mathbb{Z})$

We saw that there are no modular forms of odd weights with respect to $SL_2(\mathbb{Z})$. There are also some modular forms which satisfy the modularity condition not for the whole group but for a particular subgroup of $SL_2(\mathbb{Z})$. For example, the theta function

$$\theta(\tau) = \sum_{n=0}^{\infty} r(n,4)e^{2\pi i n\tau}$$

is a modular form of weight 2 with respect to the subgroup generated by the matrices:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 \\ 4 & 1 \end{pmatrix}$$

The coefficient $r(n,4)$ denotes the number of ways n can be written as a sum of four squares. This is a non trivial result. The above fact and the fact that the space of modular forms is finite dimensional is used to derive an explicit formula for $r(n,4)$. We won't go into the details, but this gives us enough motivation to describe some subgroups of $SL_2(\mathbb{Z})$ of particular interest to us. The subgroup $\Gamma_0(N)$ is described by the following set:

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \mod N \right\}$$

We also have the subgroup:

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \mod N \right\}$$

The principle congruence subgroup $\Gamma(N)$ is defined as,

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mod N \right\}$$

We will work with a much generalized class of subgroups. These are called the congruence subgroups.

Definition 1.4.1 (Congruence subgroup). A subgroup Γ of $SL_2(\mathbb{Z})$ is a congruence subgroup of level N if $\Gamma(N) \subseteq \Gamma$ for some positive integer N.

When $-I \notin \Gamma$, modular forms of odd weights may exist with respect to Γ, unlike in the case of $SL_2(\mathbb{Z})$.
Next, we move onto proving certain properties of the subgroups described above.

Proposition 1.4.2. The group $SL_2(\mathbb{Z}/N\mathbb{Z})$ isomorphic to $SL_2(\mathbb{Z})/\Gamma(N)$.

This requires a small result to be proven first:

Lemma 1.4.3. Suppose $(c, d, N) = 1$, where (c, d, N) denotes the greatest common divisor of c, d and N. Then there exists $c' = c + tN$ and $d' = d + tN$ for some integers s and t such that $(c', d') = 1$.

Proof. Let $c_1 = (c, N)$ so that $c = c_1c_2$ for some c_2. Now, $(c_2, N) = 1$. This implies that there exists u and $v \in \mathbb{Z}$ such that $1 = c_2u + Nv$. Let $m = v - dv$. Then observe the following:

$$\begin{aligned} d + mN &= d + (v - dv)N \\ &\equiv 1 \bmod c_2 \end{aligned}$$

using the fact that $Nv \equiv 1 \bmod c_2$. Let $c = c'$ and $d + mN = d'$.

Claim 1.4.4. $(c', d') = 1$.

Suppose $p \mid c'$ and $p \mid d'$ for some prime p. Then $p \mid c_1c_2$ and $p \mid (d+mN)$. Thus $p \mid c_1$ or $p \mid c_2$ and $p \mid (d + mN)$. Discard the case when $p \mid c_2$ and $p \mid (d + mN)$. This is because $c_2 \mid (d + mN - 1)$, and so $p \nmid (d + mM)$. Therefore, $p \mid c_1 = (c, N)$ and $p \mid (d + mN)$. Thus, $p \mid (c, d, N) = 1$ proving that $p = 1$. □

Proof of Proposition 2.8. Consider the following homomorphism:

$$\varphi : SL_2(\mathbb{Z}) \to SL_2(\mathbb{Z}/N\mathbb{Z})$$

such that $\varphi(\gamma) = \gamma \bmod N$. Using the lemma we will show that this map is a surjection. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}/N\mathbb{Z})$. Since c and $d \in SL_2(\mathbb{Z}/N\mathbb{Z})$, $(c, d, N) = 1$. Using Lemma 1.4.3, we can find integers c' and d' such that $c' \equiv c \bmod N$ and $d' \equiv d \mod N$ and $(c', d') = 1$. We need to find k and l such that there exists a lift of γ given by $\begin{pmatrix} a + kN & b + lN \\ c' & d' \end{pmatrix} \in SL_2(\mathbb{Z})$. That is, we want integers k and l such that the following condition is satisfied: $(a + kN)d' - (b + lN)c' = 1$. This is equivalent to saying that

$$ad' + bc' + (kd' - lc')N = 1 \tag{1.4.5}$$

Now $ad' - bc' = 1 + jN$ for some integer j. Substituting in (1.4.5), we see that, we need integers k and l which satisfy that $j = lc' - kd'$. Since c' and d' are co-prime, such integers clearly exist. The kernel of the map is clearly $\Gamma(N)$. □

Corollary 1.4.6. The subgroup $\Gamma(N)$ is normal in $SL_2(\mathbb{Z})$.

Proof. Being the kernel of a surjective homomorphism, this follows directly from the proof of Proposition 1.4.2. □

Proposition 1.4.7. The group $(\mathbb{Z}/N\mathbb{Z})^*$ is isomorphic to $\Gamma_0(N)/\Gamma_1(N)$.

Proof. Consider the homomorphism

$$\psi : \Gamma_0(N) \to (\mathbb{Z}/N\mathbb{Z})^*$$

$$\begin{pmatrix} a' & b' \\ c'N & d' \end{pmatrix} \mapsto d' \bmod N$$

To prove that this map is a surjection, let $d \in (\mathbb{Z}/N\mathbb{Z})^*$ with $(N, d) = 1$. This implies that there exist integers a and b such that $bd - aN = 1$. Take $\gamma = \begin{pmatrix} b & a \\ N & d \end{pmatrix}$. This clearly belongs to $\Gamma_0(N)$ and $\psi(\gamma) = d$. The kernel of ψ is clearly $\Gamma_1(N)$ and so the result follows. □

The above proof also shows that $\Gamma_1(N)$ is normal in $\Gamma_0(N)$.

Proposition 1.4.8. Any congruence subgroup Γ is of finite index in $SL_2(\mathbb{Z})$.

Proof. Since $\Gamma(N) \subseteq \Gamma$ for some N and $[SL_2(\mathbb{Z}) : \Gamma(N)]$ is finite. Therefore, Γ has finite index. □

We further develop the definition of a modular form with respect to a congruence subgroup. In the usual definition of a modular form, the only thing which needs more work is the "holomorphy at ∞" condition. We develop this definition keeping in mind the following points:

1. Since $\Gamma(N) \subseteq \Gamma$ for some N, each congruence subgroup Γ of $SL_2(\mathbb{Z})$ contains a translation matrix $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$ for some minimal $h \in \mathbb{Z}^+$. Therefore any modular form with respect to Γ is h-periodic and therefore by the same argument in section 1.2, there exists a corresponding function $\widetilde{f} : D' \to \mathbb{C}$ where D' is a punctured disc around 0. But now, $f(\tau) = \widetilde{q_h}$ where $q_h = e^{2\pi i\tau/h}$. As argued before f is

holomorphic on $\mathbb{H}$ and so it is analytic on D'. We define f to be *holomorphic at* ∞ if $\widetilde{f}$ can be analytically extended to the whole disc. Again, we will get a Fourier series expansion on D' given by the following expression:

$$f(q) = \sum_{n \geq 0} a_n q_h^n.$$

2. Till now we have only defined the action of $SL_2(\mathbb{Z})$ on the points of $\mathbb{H}$. However, we will later see that in order to make the space of modular forms finite dimensional, we need to define the action on the point "∞" as well and adjoin its "$SL_2(\mathbb{Z})$ translates" to $\mathbb{H}$. These are precisely the rational numbers $\mathbb{Q}$ as we will see in section 2.5. We call the set $\{\mathbb{Q}\} \cup \infty$ as *cusps*.

3. The crucial point to note is that we need the modular forms to be holomorphic at cusps as well. We will prove in section 2.5 that each $s \in \mathbb{Q}$ is of the form $\alpha(\infty)$ for some $\alpha \in SL_2(\mathbb{Z})$. Thus, holomorphy of f at the cusps is naturally defined in terms of holomorphy of the function $f[\alpha]_k$ at ∞.

4. In order to make sense of the last sentence of the previous point note that for any $\delta \in SL_2(\mathbb{Z})$, $\delta^{-1}\Gamma(N)\delta = \Gamma(N)$ by Corollary 1.4.6. This implies that $\Gamma(N) \subseteq \alpha^{-1}\Gamma\alpha$ for some N. Thus, $\alpha^{-1}\Gamma\alpha$ is a congruence subgroup. Notice that $f[\alpha]_k$ satisfies the modularity condition with respect to the subgroup $\alpha^{-1}\Gamma\alpha$ since for any $\gamma \in \Gamma$, $(f[\alpha]_k)[\alpha^{-1}\gamma\alpha]_k = f[\alpha]_k$ By the argument in 1, holomorphy at ∞ of the function $f[\alpha]_k$ is precisely defined.

Definition 1.4.9. (Modular form with respect to Γ) Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$ and let k be an integer. A function $f : \mathbb{H} \to \mathbb{C}$ is a modular form of weight k with respect to Γ if

1. f is holomorphic on $\mathbb{H}$.

2. f satisfies the modularity condition for all matrices in Γ.

3. The function $f[\alpha]_k$ is holomorphic at ∞ for all $\alpha \in SL_2(\mathbb{Z})$.

In addition, if $a_0 = 0$ in the Fourier expansion of $f[\alpha]_k$ for all $\alpha \in SL_2(\mathbb{Z})$, then f is a cusp form of weight k with respect to Γ.

For a modular form f with respect to $SL_2(\mathbb{Z})$, condition 3 in the above definition is reduced to the usual definition of "holomorphy of f" at ∞. We denote the set of modular forms with respect to Γ as $\mathcal{M}_k(\Gamma)$ and similarly, the cusp forms as $\mathcal{S}_k(\Gamma)$.

The following proposition is useful in checking the third condition mentioned in the above definition of modular forms.

Proposition 1.4.10. Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$ of level N, and let $q_N = e^{2\pi i\tau/N}$ for $\tau \in \mathbb{H}$. Suppose that the function $f : \mathbb{H} \to \mathbb{C}$ satisfies conditions 1 and 2 in definition 1.4.9. In the Fourier expansion $f(\tau) = \sum_{n=0}^{\infty} a_n q_N^n$, let the coefficients for $n > 0$ be such that $|a_n| \leq Cn^r$ for some positive constants C and r. Then f also satisfies condition 3 in definition 1.4.9. Therefore, $f \in \mathcal{M}_k(\Gamma)$.

Proof. For every $\alpha \in SL_2(\mathbb{Z})$, the function $(f[\alpha]_k)(\tau)$ is holomorphic and weight k invariant with respect to the subgroup $\alpha^{-1}\Gamma\alpha$. As pointed out in the above discussion, we get a Laurant series expansion about 0.

$$f([\alpha]_k)(\tau) = \sum_{n\in\mathbb{Z}}^{\infty} a'_n q_N^n$$

In order to show that $f[\alpha]_k$ is holomorphic for all $\alpha \in SL_2(\mathbb{Z})$, we show that the Laurant series truncates from the left to give a power series. This amounts to showing that $\lim_{q_N \to 0} |f[\alpha]_k(\tau) q_N| = 0$. To that end, we begin by estimating the function f. Since $f = \sum_{n\geq 0}^{\infty} a_n q_N^n$, and $|a_n| \leq Cn^r$ writing $\tau = x + iy$, observe that,

$$|f(\tau)| \leq |a_0| + C \sum_{n=1}^{\infty} n^r e^{-2\pi ny/N}. \tag{1.4.11}$$

Consider the function $g(t) = t^r e^{-2\pi ty/N}$. A little bit of calculus shows that g has only one maximum at $t = rN/2\pi y$. Also, $g(0) = 0$. So the graph of g increases from 0 to $rN/2\pi y$ and then decays to 0. We estimate the sum in (1.4.11) using the area under the curve of the function g. Suppose $g(t_0) = rN/2\pi y$ such that t_0 lies between the integers d and $d+1$. Then the graph of the function g helps us conclude that,

$$\begin{aligned} |f(\tau)| &\leq |a_0| + C(e^{-2\pi y/N} + \cdots + d^r e^{-2\pi dy/N} + (d+1)^r e^{-2\pi(d+1)y/N} + \dots) \\ &\leq |a_0| + C\left(\int_0^{\infty} g(t)dt + (d+1)^r e^{-2\pi i(d+1)y/N} \right) \\ &= |a_0| + C\left(\int_0^{\infty} g(t)dt + (rN/2\pi y)^r e^{-r} \right) \end{aligned}$$

This implies that $|f(\tau)| \leq |a_0| + C'(\int_0^{\infty} g(t)dt + (1/y)^r)$ for some constant C'. Next, we estimate the integral of $g(t)$ from 0 to ∞.

$$\int_0^{\infty} g(t) = \int_0^{\infty} t^r e^{-2\pi yt/N} dt$$

Writing $x = 2\pi yt/N$ and doing a change of variable we arrive at the following integral:

$$(N/2\pi y)^{r+1} \int_0^{\infty} x^r e^{-x} dx$$

The integral in the above expression is the Gamma function denoted by $\Gamma(r)$ and for a positive integer r it takes the value $(r+1)!$ We will talk more about the Gamma function in section 4.1.

It now follows that $\int_0^\infty g(t)dt = \mathcal{O}(1/y^{r+1})$. Combining all the observations we land up with the following inequality for some positive constants c_0 and c_1: $|f(\tau)| \leq c_0 + c_1/(\operatorname{Im}\tau)^r$. Moreover this implies that,

$$|f(\alpha\tau)| \leq c_0 + c_1/(\operatorname{Im}(\alpha\tau))^r \tag{1.4.12}$$

Using this and the formula for $\operatorname{Im}(\alpha\tau)$ in the proof of Lemma 1.0.2, we compute:

$$\begin{aligned} \lim_{q_N\to 0} |f[\alpha]_k(\tau)q_N| &= \lim_{q_N\to 0} |\jmath(\alpha,\tau)^{-k} f(\alpha\tau)q_N| \\ &= \lim \left| c_0\jmath(\alpha,\tau)^{-k} + c_1\jmath(\alpha,\tau)|^{2r-k} \right| e^{-2\pi y/N} \end{aligned}$$

Notice that $|\jmath(\alpha,\tau)| = \mathcal{O}(y^2)$. Using the fact that exponential decay dominates polynomial growth, and observing that as $q_N \to 0, y \to \infty$, we see $\lim_{q_N\to 0} |f[\alpha]_k(\tau)q_N| = 0$. The result now follows. $\square$

1.5 Meromorphic modular forms

As the name suggests these are meromorphic functions on the upper half plane satisfying the same properties as holomorphic modular forms. Meromorphy at ∞ means that the function is now allowed to have a pole at infinity and so f has a Laurent series expansion around 0. To make this precise: Let $h \in \mathbb{Z}^+$ be the smallest number such that $\begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \in \Gamma$. Since f satisfies the modularity condition with respect to Γ, it has period h. Suppose f has no poles in the region $\{\tau \in \mathbb{H} \mid \operatorname{Im}(\tau) > c\}$. Then f has a Laurent series expansion in the corresponding punctured disk about 0:

$$\widetilde{f}(q) = \sum_{n=-\infty}^{\infty} a_n q_h^n \ ; \ \operatorname{Im}(\tau) > c \tag{1.5.1}$$

The function f is defined to be *meromorphic at* ∞ if $\widetilde{f}$ is meromorphic at 0, that is, the series in 1.5.1 truncates from the left. Meromorphy at the cusps is defined in a similar fashion as done previously.

Definition 1.5.2. (Meromorphic modular forms) Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$ and let k be an integer. A function $f : \mathbb{H} \to \mathbb{C}$ is a modular form of weight k with respect to Γ if

1. f is meromorphic on $\mathbb{H}$.

2. f satisfies the modularity condition for all matrices in Γ.

3. The function $f[\alpha]_k$ is meromorphic at ∞ for all $\alpha \in SL_2(\mathbb{Z})$.

One of the first examples of a meromorphic modular form of weight 0 is the j function. It is given by the following expression:

$$j(\tau) = 1728\frac{(60G_4)^3}{\Delta}(\tau). \tag{1.5.3}$$

The first few terms of the q-expansion of the j are:

$$j(q) = q^{-1} + 744 + 196884q + 21493760q^2 + O(q^3) \tag{1.5.4}$$

Notice that the j function is of weight 0, that is, it is $SL_2(\mathbb{Z})$ invariant. It follows that, it is a well defined function on the space $SL_2(\mathbb{Z})/\mathbb{H}$. This space is called a modular curve of level 1. We denote it by $Y(1)$. We will elaborate more on modular curves in the next section.

The set of meromorphic modular forms of weight k with respect to the subgroup Γ also form a vector space. We denote it by $\mathcal{A}_k(\Gamma)$. It contains the subspace of holomorphic modular forms which we denote by $\mathcal{M}_k(\Gamma)$.

Chapter 2

Modular curves and the dimension theory

Throughout this chapter Γ dentotes a subgroup of $SL_2(\mathbb{Z})$ of finite index.

Definition 2.0.1 (Modular curves). Define a modular curve $Y(\Gamma)$ to be the quotient of the upper half plane under the action of Γ.

Two points τ and τ' are in the same orbit if and only if there exists some $\gamma \in \Gamma$ such that $\gamma\tau = \tau'$. We give $Y(\Gamma)$ the quotient topology via the map $\Pi : \mathbb{H} \to Y(\Gamma)$. This means that $U \subset Y(\Gamma)$ is open if and only if $\Pi^{-1}(U)$ is open in $\mathbb{H}$. Note that Π is an open map.

We next state a very important proposition and its consequences which we will use extensively later.

Proposition 2.0.2. Let τ_1 and $\tau_2 \in \mathbb{H}$ be given. Then there exist neighborhoods U_1 of τ_1 and U_2 of τ_2 in $\mathbb{H}$ with the property that for any $\gamma \in SL_2(\mathbb{Z})$, $\gamma(U_1) \cap U_2 \neq \varphi$ implies that $\gamma(\tau_1) = \tau_2$.

In simple words, when we take $\tau_1 = \tau_2$ and $U_1 = U_2$, the above proposition guarantees the existence of a neighborhood for every point $\tau \in \mathbb{H}$, in which it is the sole representative of its Γ-orbit. As a direct consequence we have that:

Corollary 2.0.3. Let $\tau \in \mathbb{H}$. For any $\gamma \in SL_2(\mathbb{Z})$, there exists a neighborhood U such that $\gamma U \cap U \neq \varphi$ implies that γ belongs to Γ_τ, where Γ_τ denotes the stabilizer of τ in $SL_2(\mathbb{Z})$.

In order to prove the proposition, we need a small lemma.

Lemma 2.0.4. For any compact sets A and $B \in \mathbb{H}$, the set $S = \{\gamma \in \Gamma \mid \gamma A \cap B \neq \varphi\}$ is finite.

Proof. Let A and B be compact subsets of $\mathbb{H}$. Then the set $\{\operatorname{Im}\tau \mid \tau \in B\} \subseteq [c_1, c_2]$ in $\mathbb{R}$ with $c_1 > 0$. Suppose that $\tau_A \in A$ and $\tau_B \in B$ such that $\gamma\tau_B = \tau_A$. Then $\operatorname{Im}(\tau_A)/|c\tau_A + d|^2 = \operatorname{Im}(\tau_B) \geq c_1$. We also have that,

$$\operatorname{Im}(\tau_A) \leq \operatorname{Im}(\tau_B)|c\tau_A + d|^2 \leq c_2|c\tau_A + d|^2 \tag{2.0.5}$$

Combining the two inequalities, we get

$$\operatorname{Im}(\tau_A)/c_2 \leq |c\tau_A + d|^2 \leq \operatorname{Im}(\tau_A)/c_1 \tag{2.0.6}$$

Since A and B are compact, the imaginary parts of τ_A and τ_B are bounded below and from above as well. Therefore, there are finitely many possibilities for tuples $(c, d) \in \mathbb{Z}^2$ which satisfy the inequality (2.0.6). We next show that we have only finitely many possibilities of matrices having the second row entries as (c, d) which satisfy (2.0.6). This follows from the following claim:

Claim 2.0.7. If two matrices $\gamma = \begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix}$ and $\delta = \begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix}$ have the same second row entries (c, d), then $\gamma\delta^{-1} = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$.

We need to show that for matrices in the above claim, $a_1 = a_2 + nc$ and $b_1 = b_2 + nc$ for some n. This is seen as follows: $a_1 d - b_1 c = a_2 d - b_2 c = 1$. This implies that $c/d = (a_1 - a_2)/(b_1 - b_2)$. Since $(c, d) = 1$, we conclude that $(a_1 - a_2)/n = c$ and $(b_1 - b_2)/n = d$ for some n. It follows that $a_1 = a_2 + nc$ and $b_1 = b_2 + nc$ so that the condition in the claim is satisfied.

The integer n is bounded since A and B are compact sets. This implies that there are only finitely many possibilities for the entries of matrices in S. This makes S finite. $\square$

Proof of Proposition 2.0.2. Suppose that τ_1 and $\tau_2 \in \mathbb{H}$ and C_1 and C_2 are closed neighborhoods about τ_1 and τ_2 respectively. Let $S(C_1, C_2) = \{\gamma \in \Gamma \mid \gamma C_1 \cap C_2 \neq \phi \text{ and } \tau_1 \neq \tau_2\}$. By Lemma 2.0.4, this set is finite. If the set $S(C_1, C_2)$ is empty, then we can take $U_1 \subset C_1$ and $U_2 \subset C_2$ both open to satisfy the condition stated in the proposition. For $\gamma U_1 \cap U_2 \neq \varphi$ will imply that $\gamma\tau_1 = \tau_2$. Now suppose that $S(C_1, C_2)$ is non empty. Let $\gamma \in S(C_1, C_2)$. As $\mathbb{H}$ is Hausdorff, we can find disjoint neighborhoods V_1 and V_2 of $\gamma\tau_1$ and τ_2 respectively. We see that $\gamma^{-1}V_1$ is an open neighborhood of τ_1 and it contains a closed disc $C_1' \subseteq C_1$ around τ_1. Similarly, we can find a closed disk $C_2' \subseteq V_2 \cap C_2$ about τ_2. This implies that $S(C_1', C_2') \subseteq S(C_1, C_2)$. Observe that $\gamma C_1' \subseteq V_1$. The intersection $V_1 \cap V_2 = \varphi$, implies that $\gamma C_1' \cap C_2' = \varphi$. Therefore $\gamma \notin S(C_1', C_2')$ and $S(C_1', C_2') \subsetneq S(C_1, C_2)$. This set is again finite, and so we can keep repeating the process to get $C_1^{(n)}$ and $C_2^{(n)}$ such that $S(C_1^{(n)}, C_2^{(n)}) = \varphi$. Now, by the argument in the beginning, we get open neighborhoods about τ_1 and τ_2 with the property in the proposition. $\square$

Corollary 2.0.8. The modular curve $Y(\Gamma)$ is Hausdorff.

Before going on to proving the corollary we will need a small result.

Claim 2.0.9. Let U_1 and U_2 be open neighborhoods. Then $\Pi(U_1) \cap \Pi(U_2) = \varphi$ in $Y(\Gamma)$ if and only if $\Gamma(U_1) \cap U_2 = \varphi$ in $\mathbb{H}$.

Suppose that $\gamma U_1 \cap U_2 = \varphi$ for all $\gamma \in \Gamma$. To the contrary assume that $\Pi(U_1) \cap \Pi(U_2) \neq \varphi$ Then there exists $x \in \Pi(U_1) \cap \Pi(U_2)$. Now, $x \in \Pi(U_1)$ implies that $\Pi^{-1}(x) = \bigcup_{\gamma \in \Gamma} \gamma U_1$, while $x \in \Pi(U_2)$ implies that there exists $\tau \in U_2$ such that $\Pi(\tau) = x$. Moreover this gives us that $\tau \in \Pi^{-1}(x)$. Therefore, $\tau \in \delta U_1$ for some $\delta \in \Gamma$. Finally, $\tau \in \delta U_1 \cap U_2$ which is a contradiction to our assumption. Conversely, suppose that $\Pi(U_1) \cap \Pi(U_2) = \varphi$. We want to show that $\Gamma(U_1) \cap U_2 = \varphi$. Suppose not. Then there exists $\tau \in U_2$ such that $y = \gamma\tau'$ for some $\tau' \in U_1$ and $\gamma \in \Gamma$. This implies that $\Pi(\tau) = \Pi(\tau')$. But then $\Pi(\tau) \in \Pi(U_1) \cap \Pi(U_2)$, a contradiction.

Proof of Corollary 2.0.8. Let $\Pi(\tau_1)$ and $\Pi(\tau_2)$ be two distinct points in $Y(\Gamma)$. Take neighborhoods U_1 and U_2, as in Proposition 2.0.2. Since $\Gamma\tau_1 \neq \tau_2$, by Proposition 2.0.2, $\Gamma U_1 \cap U_2 = \varphi$. By the previous claim 2.0.9, $\Pi(U_1) \cap \Pi(U_2) = \varphi$. Therefore, there exists disjoint neighborhoods $\Pi(U_1)$ and $\Pi(U_2)$ of $\Pi(\tau_1)$ and $\Pi(\tau_2)$ respectively, in $Y(\Gamma)$. This proves that $Y(\Gamma)$ is Hausdorff.

$\square$

2.1 The fundamental domain

A fundamental domain for the action of Γ on $\mathbb{H}$ is a region on the upper half plane such that for every point τ in $\mathbb{H}$, the region contains exactly one point which is in the same Γ-orbit as τ. In this section we will find out the fundamental domain for $SL_2(\mathbb{Z})$.

Proposition 2.1.1. The fundamental domain $\mathcal{D}$ for the action of $SL_2(\mathbb{Z})$ is the region $\mathcal{D} = \{\tau \in \mathbb{H} : |\operatorname{Re}(\tau)| \leq 1/2, |\tau| \geq 1\}$.

The region is shown in figure 2.1. Let S and T be matrices as given in 1.0.3.

Proof. We first show that for every $\tau \in \mathbb{Z}$ there exists $\tau' \in \mathcal{D}$ such that $\tau = \gamma\tau'$ for some $\gamma \in SL_2(\mathbb{Z})$. This will show that any point in $\mathbb{H}$ is $SL_2(\mathbb{Z})$-equivalent to a point in $\mathcal{D}$. Choose any $\tau \in \mathbb{H}$ and apply the matrix T^n for some $n \in \mathbb{Z}$ to translate τ by n to the

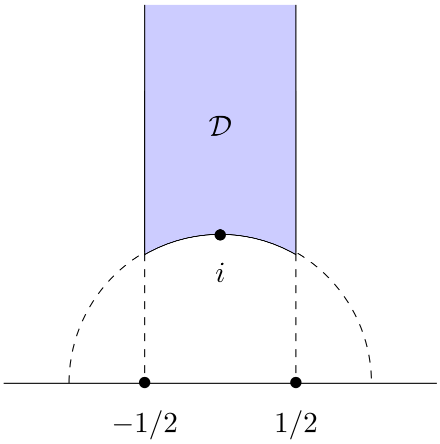

FIGURE 2.1: The fundamental domain

vertical strip $\{\tau : |\operatorname{Re}(\tau)| \leq 1/2\}$. If τ translates to $\mathcal{D}$, then we are done. If not, then $|\tau| < 1$. Now, notice that $\operatorname{Im}(-1/\tau) = \operatorname{Im}(\tau)/|\tau|$. Thus when $|\tau| < 1, \operatorname{Im}(-1/\tau) > \operatorname{Im}(\tau)$. Therefore, the next step would be to repeatedly apply the matrix S until $\operatorname{Im}(\tau) > 1$. We need to make sure the above process stops at a finite number of steps. To see this, notice that for any τ there are only finite number of integer pairs (c, d) such that $|c\tau + d| < 1$. Since $\operatorname{Im}(\gamma\tau) = \operatorname{Im}(\tau)/|c\tau + d|^2$, there are only finitely many transformations such that $\operatorname{Im}(\gamma\tau) > \operatorname{Im}(\tau)$. It follows that the above process terminates at some point. □

However, notice that the line $\operatorname{Re}(\tau) = 1/2$ is identified with the line $\operatorname{Re}(\tau) = -1/2$. The two halves of the boundary arc $\tau = 1$ are also identified via $\tau \mapsto -1/\tau$. We next show that these are the only two identifications.

Proposition 2.1.2. Suppose τ_1 and τ_2 are distinct points in $\mathcal{D}$ such that $\tau_2 = \gamma\tau_1$ for some $\gamma \in SL_2(\mathbb{Z})$. Then only the following cases are possible:

1. $\operatorname{Re}(\tau_1) = \pm 1/2$ and $\tau_2 = \tau_1 + 1$
2. $|\tau_1| = 1$ and $\tau_2 = -1/\tau_1$.

Proof. Without loss of generality, assume that $\operatorname{Im}(\tau_2) \geq \operatorname{Im}(\tau_1)$. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \neq \pm I$. Using the fact that $\operatorname{Im}(\tau_2) = \operatorname{Im}(\tau_1)/|c\tau + d|^2$, we see that $|c\tau + d| < 1$. Observe that $|c|\sqrt{3}/2 \leq |c| \operatorname{Im}(\tau_1) = \operatorname{Im}(c\tau_1 + d) \leq |c\tau_1 + d| \leq 1$ since $\tau_1 \in \mathcal{D}$. We finally get that $|c| \leq 2/\sqrt{3}$. Thus, the only possibility for c is that $|c| \in \{0, 1\}$. We now divide the argument into different cases as follows:

1. If $c = 0$, then $\gamma = \pm \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$. So, $\tau_2 = \tau_1 + b$ which implies that $\text{Re}(\tau_2) = \text{Re}(\tau_1) + b$. Further, we see $|\text{Re}(\tau_2) - \text{Re}(\tau_1)| = |b| \leq 1$ since τ_1 and $\tau_2 \in \mathcal{D}$ and so $b = 1$. It follows that $\tau_2 = \tau_1 + 1$ giving us that $\text{Re}(\tau_1) = \pm 1/2$. Thus 1 is satisfied.

2. If $|c| = 1$ then $|c\tau_1 + d| = |\tau \pm d| \leq 1$. This implies that

$$(\text{Re}(\tau_1) \pm d)^2 \leq 1 - \text{Im}(\tau_1)^2 \leq 1 - (\sqrt{3}/2)^2 = 1/4. \tag{2.1.3}$$

This gives us that $|\text{Re}(\tau_1 + d)| \leq 1/2$. Since $\text{Re}(\tau) \leq 1/2$ for all $\tau \in \mathcal{D}$, we get that $|d| = 1$ or 0. Considering two sub cases:

   (a) If $d = 1$ then $|\text{Re}(\tau_1) \pm 1| = 1/2$. This is because $|\text{Re}(\tau_1) \pm 1| < 1/2$ implies $\text{Re}(\tau_1) > 1/2$ which is not possible and so the equality holds. Therefore, in this case $\text{Re}(\tau_1) = \pm 1/2$. The inequality (2.1.3) implies that $\text{Im}(\tau_1) = \pm\sqrt{3}/2$. This implies that $|\tau_1|$ is one of $\pm 1/2 + \sqrt{3}/2$. We also have that $c\tau_1 + d = 1$ which gives us that $\text{Im}(\tau_1) = \text{Im}(\tau_2)$. Hence $\tau_2 = 1/2 + \sqrt{3}/2$ when $\tau_1 = -1/2 + \sqrt{3}/2$ and vice versa. So in this case, both 1 and 2 hold.

   (b) If $d = 0$, then $|c\tau + d| \leq 1$ implies that $|\tau_1| \leq 1$. Since $\tau_1 \in \mathcal{D}, |\tau_1| = 1$. As in the previous case we again have that $\text{Im}(\tau_1) = \text{Im}(\tau_2)$. By running the same argument as in case 2 for γ^{-1} instead of γ, we again land in 2 cases depending on the corner right entry of γ^{-1}. One sub case is already dealt with in point (a). The other sub case for γ_2^{-1} gives that $|\tau_2| = 1$ by the same reasoning done for τ_1 above. Therefore, τ_1 and τ_2 are lie on the arc in $\mathcal{D}$ with the same imaginary parts. This implies that $\text{Re}(\tau_1)^2 = \text{Re}(\tau_2)^2$. Since τ_1 and τ_2 are distinct,$\text{Re}(\tau_1) = -\text{Re}(\tau_2)$. Thus $\tau_1 = -\overline{\tau}_2$. The matrix γ in this case looks like $\pm \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}$. Finally what we have is the following:

$$\begin{aligned} \gamma\tau_1 = \tau_2 &= (a\tau_1 - 1)/\tau_1 \\ &= (-a\bar{\tau}_2 - 1)/(-\bar{\tau}_2) \end{aligned}$$

   which gives us that $-|\tau_2|^2 = -1 = (-a\bar{\tau}_2 - 1)$. Hence, $a = 0$ and in this case as well, 2 holds.

$\square$

The above discussion shows that up to some boundary identifications, the region $\mathcal{D}$ is homeomorphic to $Y(1)$. In the proposition to follow, we find a suitable fundamental domain for the action of any congruence subgroup on $\mathbb{H}$.

Proposition 2.1.4. Suppose Γ is a congruence subgroup and the set $\{\gamma_j\}_j$ are the coset representatives of $\{\pm I\}\Gamma$ in $SL_2(\mathbb{Z})$. That is, the group is decomposed as $SL_2(\mathbb{Z}) = \bigcup_j \{\pm I\}\Gamma\gamma_j$. Then, $\bigcup_j \{\pm I\}\gamma_j\mathcal{D}$ surjects to $Y(\Gamma)$.

The above proposition shows that up to some boundary identifications $\bigcup_j \{\pm I\}\gamma_j\mathcal{D}$ is homeomorphic onto $Y(\Gamma)$.

Proof. Taking inverse in the given decomposition of $SL_2(\mathbb{Z})$ we see that $SL_2(\mathbb{Z}) = \bigcup_j \gamma_j^{-1}\{\pm I\}\Gamma$. In order to show that we get a surjection onto $Y(\Gamma)$, we show that for any $\tau \in \mathbb{H}$, there exists some τ' in $\bigcup_j \{\pm I\}\gamma_j\mathcal{D}$ such that $\tau = \gamma\tau'$ for some $\gamma \in \Gamma$.

Let $\tau \in \mathbb{H}$, then there exists some matrix δ in $SL_2(\mathbb{Z})$ such that $\delta\tau \in \mathcal{D}$. The matrix δ lies in ${\gamma_j}^{-1}\{\pm\}\Gamma$ for some j. Writing $\delta = \{\pm I\}{\gamma_j}^{-1}\gamma'$ for some γ' in Γ we see that $\gamma'\tau \in \{\pm I\}\gamma_j\mathcal{D}$. This completes the proof of the proposition.

$\square$

2.2 Realising a modular curve as a Riemann surface

A Riemann surface is a Hausdorff, second countable topological space, which locally "looks like" the complex plane. In order to give modular curves a Riemann structure, we need to make sure that every point on the modular curve has a neighborhood homeomorphic to an open set in the complex plane. These are called local charts which define a local coordinate structure on the Riemann surface. More precisely, a Riemann surface X consists of the following:

1. A collection $(U_i, V_i, \varphi_i)_{i\in I}$ for some indexing set I, where U_i are open in $\mathbb{C}$, the neighborhood V_i are open in X and $\varphi_i : V_i \to U_i$ are homeomorphisms. We call V_i charts and the set $\{V_i\}_{i\in I}$, an atlas on X.
2. If the intersection of any two charts V_i and V_j is non-empty, then the map $\varphi_j^{-1} \circ \varphi_i : \varphi_i(U_i \cap U_j) \to \varphi_j(U_i \cap U_j)$ is holomorphic.

Condition 2 helps us to smoothly go from one chart to another via the intersection. The simplest example of a Riemann surface is the complex sphere. The torus is also a Riemann surface. In fact, any g holed surface is an example of a Riemann surface. We will later see that when we compactify the modular curve $Y(1)$, it is topologically a sphere.

It is simplest to define a coordinate chart around the points whose stabiliser consist only of $\pm I$. We define local charts around such points as follows: Take a neighborhood U of τ

which satisfies the condition in Corollary 2.0.3. Suppose V is its image in $Y(\Gamma)$. Define $\varphi = \Pi|_U$.

Claim 2.2.1. The map $\varphi : U \to V$ is a homeomorphism.

Proof. The map φ is clearly a surjection. Let τ_1 and τ_2 be two points in U such that $\Pi(\tau_1) = \Pi(\tau_2)$. Then $\tau_1 = \gamma\tau_2$ for some $\gamma \in \Gamma$. We know that the intersection $\gamma' U \cap U$ is empty for all $\gamma' \in \Gamma/\{\pm I\}$. Corollary 2.0.3 implies that γ is I or $-I$. Thus $\tau_1 = \tau_2$. The function φ is therefore an injection as well. Since Π is an open map, we see that φ is a homeomorphism from U to V. □

We have a slight problem when the stabilizer of a point τ in the upper half plane is a non trivial subgroup of the modular group. For example, consider the point $i \in \mathbb{H}$ and the transformation S given in 1.0.3. One can easily see that $Si = i$. If we take any neighborhood around i, the matrix S acts as a rotation by π around i. To see this, assume that U is any neighborhood around i. Notice that the transformation $\delta = \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$ takes i to 0. Thus $\delta(U)$ is a neighborhood around 0. Also, δ is a homeomorphism from U to $\delta(U)$. The transformation S fixes i if and only if $\delta S \delta^{-1}$ fixes 0. Therefore S acts on the point τ in U in the same way as $\delta S \delta^{-1}$ acts on the point $\delta\tau$ in $\delta(U)$. We can easily calculate that $\delta S \delta^{-1}$ maps z in $\delta(U)$ to $-z$, that is, it acts as a rotation by π about 0.

The above argument shows that we will always have Γ-equivalent points in any neighborhood around i and we cannot use the previous argument to define a local coordinate. In general, the points with a non-trivial stabilizer are called elliptic points.

Definition 2.2.2 (Elliptic point). We say $\tau \in \mathbb{H}$ is an elliptic point if $\Gamma_\tau/\{\pm I\}$ is not equal to $\{I\}$.

2.3 More on elliptic points

In this section we elaborate on elliptic points. We will give some idea on how to calculate elliptic points for the modular curve $Y(1)$.

2.3.0.1 Finding elliptic points for $Y(1)$

Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a non-trivial matrix in $\mathrm{SL}_2(\mathbb{Z})$ such that $\gamma\tau = \tau$ for some $\tau \in \mathbb{H}$. After solving the equation $(a\tau + b)/(c\tau + d) = \tau$, we get the quadratic equation $c\tau^2 + (d - a)\tau - b = 0$. If $c = 0$, then $a = d = \pm 1$ gives us that $\tau + b = \tau$. This implies that $b = 0$ and $\gamma = \pm I$. Since we assumed γ to be nontrivial, we have two distinct solutions to the above quadratic equation given by:

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

The term $(d-a)^2 + 4bc < 0$ as $\tau \in \mathbb{H}$. On simplifying we get that $|a+d| < 2$. We therefore have that $a + d = \pm 1$ or 0.

Next, notice that the characteristic polynomial of γ is $x^2 - (a+d)x + 1$. So the only possibilities for the characteristic polynomial are $x^2 + 1, x^2 - x + 1$ or $x^2 + x + 1$. The polynomial $x^2 + 1$ is a factor of $x^4 - 1$ while $x^2 - x + 1$ and $x^2 + x + 1$ are factors of $x^6 - 1$. By Caley Hamilton Theorem, γ satisfies its own characteristic polynomial. It follows that $\gamma^4 = I$ or $\gamma^6 = I$. Therefore the possible orders for γ are $1, 2, 3, 4$ or 6. One can easily check that $\gamma = \pm I$ when γ has order 1 or 2.

The next theorem characterizes the matrices of order $3, 4$ or 6 in $SL_2(\mathbb{Z})$.

Theorem 2.3.1. Let $\gamma \in SL_2(\mathbb{Z})$. Denote the following matrices as:

$$R = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}; S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}; W = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

1. if γ has order 3, then γ is conjugate to $R^{\pm 1}$.
2. if γ has order 4, then γ is conjugate to $S^{\pm 1}$.
3. if γ has order 6, then γ is conjugate to $W^{\pm 1}$.

An immediate consequence of the above theorem is the following corollary:

Corollary 2.3.2. The elliptic points for $SL_2(\mathbb{Z})$ are the points $SL_2(\mathbb{Z})i$ and $SL_2(\mathbb{Z})\mu_3$ where $\mu_3 = e^{2\pi i/3}$. The modular curve $Y(1)$ therefore has just two elliptic points.

Proof. The matrices R and W fix μ_3, while S fixes i. Recall that any matrix that fixes some point $\tau \in \mathbb{H}$ is conjugate to R, S or W in $SL_2(\mathbb{Z})$. Suppose δ is a transformation in $SL_2(\mathbb{Z})$ such that $\delta = \gamma S \gamma^{-1}$ for some $\gamma \in SL_2(\mathbb{Z})$. Then δ fixes γi. Similarly, we argue for W and R and conclude that the elliptic points for the modular group are $SL_2(\mathbb{Z})i$ and $SL_2(\mathbb{Z})\mu_3$.

From the discussion in section 2.1 we know that the orbits of i and μ_3 are not equal in $Y(1)$. Therefore, $Y(1)$ has two elliptic points. □

Proof of Proposition 2.3.1. We give a proof of part 3 first and the rest will follow similarly. Let γ be such that $\gamma^6 = I$. Define a lattice $L = \mathbb{Z}^2$ consisting of column vectors with integers. L can be made into a module over $\mathbb{Z}[\mu_6]$ where $\mu_6 = e^{2\pi i/6}$ by defining the action as follows: Let v be any vector in L. Then,

$$(a + b\mu_6) \cdot v = (aI + b\gamma)v$$

where the operation on the right hand side is the usual matrix multiplication. Now, $\mathbb{Z}[\mu_6]$ is a PID and L is finitely generated over $Z[\mu]_6$. By the structure theorem for finitely generated modules over a PID we have:

$$\mathbb{Z}^2 \cong \bigoplus_k \mathbb{Z}[\mu_6]/I_k$$

for some ideals I_k in $\mathbb{Z}[\mu_6]$. Observe that every non-zero ideal of $\mathbb{Z}[\mu_6]$ has rank 2 as an abelian group This is because if $I_k = \langle x \rangle$, then x and $\mu_6 x$ are linearly independent. This implies that $\mathbb{Z}[\mu_6]/I_k$ is a torsion abelian group. However, L as an abelian group is free. Thus, for each k, I_k is zero. It follows that $\mathbb{Z}^2 \cong \mathbb{Z}[\mu_6]$. Using this isomorphism, we will show that γ is conjugate to the matrix W. Let

$$\phi_\gamma : \mathbb{Z}[\mu_6] \to L$$

be the module isomorphism map such that $\phi_\gamma(1) = u$ and $\phi_\gamma(\mu_6) = v$. Let $[u\ v]$ denote a 2×2 matrix with first column u and second column v. This is the matrix of the module isomorphism. Therefore, $[u\ v]^{-1}$ exists and so $\det[u\ v] = \pm 1$. We make the following claim:

Claim 2.3.3. The matrix $\gamma = [u\ v]W[u, v]^{-1}$

Notice that $\gamma u = \mu_6 \cdot \phi_\gamma(1) = \phi_\gamma(\mu_6) = v$. Similarly, $\gamma v = \mu_6 \phi_\gamma(\mu_6) = \phi_\gamma(\mu_6 - 1) = -u + v$. This is because μ_6 satisfies the equation: $\mu_6^2 - \mu_6 + 1 = 0$. We therefore have that: $\gamma[u\ v] = [v\ \ -u + v]$. Writing $[u, v] = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}$, one can check that

$$\gamma[u\ v] = [v\ \ -u + v] = [u\ v]W$$

which proves the claim. We similarly see that $\gamma = [v\ u]W^{-1}[v\ u]^{-1}$. Note that if $[u\ v]$ has determinant -1, then $[v\ u]$ will have determinant 1. It follows that γ is conjugate to $W^{\pm 1}$ in $SL_2(\mathbb{Z})$.

Proof of part 1 follows from doing the same trick with the ring $\mathbb{Z}[i]$. Part 2 follows from the fact that if γ has order 3, then $-\gamma$ has order 6. By part 1, $-\gamma$ is conjugate to $W^{\pm 1}$ giving us that γ is conjugate to $R^{\pm 1}$ in $SL_2(\mathbb{Z})$. $\square$

It is easy to check the following corollary which explicitly calculates the stabilizers of the elliptic points i and μ_3.

Corollary 2.3.4. The stabilizer of i in $SL_2(\mathbb{Z})/\{\pm I\}$ is the cyclic subgroup of order 2 generated by the matrix S. The stabilizer of μ_3 in $SL_2(\mathbb{Z})/\{\pm I\}$ is the cyclic subgroup of order 3 generated by the matrix W.

Proof. The matrix S fixes i and so the subgroup generated by S also lies in the stabilizer of i. For the other way round inclusion, suppose $\gamma \neq \pm I$ fixes i. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that $\gamma i = (ai+b)/(ci+d) = i$. With a little bit of manipulation, we get $a = d$ and $b = -c$. Thus, the stabilizer of i is the set:

$$\left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \;\middle|\; a^2 + b^2 = 1, a \text{ and } b \in \mathbb{Z} \right\}$$

As a and $b \in \mathbb{Z}$, the only possibilities for both are either 0 or ± 1. Once we put in possible values of a and b, we will realize that the stabilizer of i is exactly the subgroup generated by S.

On similar lines, we perform the calculation to find out the stabilizer of μ_3. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ so that $\gamma\mu_3 = (a\mu_3 + b)/(c\mu_3 + d) = \mu_3$. This further implies that $(a+c-d)\mu_3 + (b+c) = 0$. This helps us to conclude that the stabilizer of μ_3 consists of the following set:

$$\left\{ \begin{pmatrix} a & b \\ -b & -b+a \end{pmatrix} \;\middle|\; a^2 + b^2 - ab = 1, a \text{ and } b \in \mathbb{Z} \right\}$$

Notice that $a^2 + b^2 - ab = 1$ if and only if $(2a+b)^2 + 3b^2 = 4$. Since a and b are integers, possible solutions to the equation are: $a = 0, b = \pm 1$ or $a = \pm 1, b = 0$ or $a = \mp 1, b = \pm 1$. When we substitute the possible values we get exactly the matrices in the subgroup generated the matrix W. $\square$

It follows that the stabiliser $SL_2(\mathbb{Z})_\tau$ of any elliptic point τ are conjugates of subgroups generated by S or W, hence they are cyclic of order 2 or 3.

Suppose Γ is any congruence subgroup of $SL_2(\mathbb{Z})$. The elliptic points of Γ are a subset of the points $SL_2(\mathbb{Z})i$ and $SL_2(\mathbb{Z})\mu_3$. Since Γ has a finite index in $SL_2(\mathbb{Z})$, we can conclude that the number of elliptic points in $Y(\Gamma)$ is finite. Otherwise they would be in infinitely many Γ-orbits in $SL_2(\mathbb{Z})$ which is not possible. For any elliptic point τ, the stabiliser Γ_τ is a subgroup of $SL_2(\mathbb{Z})_\tau$. Therefore Γ_τ is finite cyclic. From the above discussion we arrive at a very important corollary:

Corollary 2.3.5. Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. For each elliptic point τ of Γ, its stabiliser Γ_τ is finite cyclic.

To each τ, we associate a positive integer e_τ such that e_τ counts the number of τ fixing transformations. Keeping in mind that $-I$ acts trivially, $e_\tau = |\Gamma_\tau|$ if $-I \notin \Gamma$. Otherwise, $e_\tau = |\Gamma_\tau|/2$, that is

$$e_\tau = |\{\pm\}\Gamma_\tau/\{\pm I\}|$$

Notice that $e_\tau > 1$ if and only if τ is elliptic. The integer e_τ is called the period of τ.

Remark 2.3.6. If τ is an elliptic point, the only possibilities for the value of e_τ are 2 and 3.

2.4 Coordinate charts for elliptic points

The goal of this section is to define coordinate charts around elliptic points. We generalize the argument given for the point i after the claim 2.2.1 to define a coordinate chart around elliptic points. The idea is as follows: Suppose τ is an elliptic point. We will take a neighborhood U of τ such that it contains no other elliptic points (such a neighborhood exists, as we will see later). We space out the Γ-equivalent points by a fixed angle $2\pi/e_\tau$. We then wrap the sector of angle $2\pi/e_\tau$ of U around a disc V. This will induce a homeomorphism from $\Pi(U) \to V$. Figure 2.2 shows the coordinate chart around the elliptic point i.

Claim 2.4.1. For any $\tau \in \mathbb{H}$, the period e_τ is well defined on the modular curve $Y(\Gamma)$.

Proof. We need to show that for any $\tau \in \mathbb{H}$, the period of $\delta\tau$ is same as the period of τ for all $\delta \in \Gamma$. Suppose $\gamma \in \Gamma_\tau$. We then have that $\gamma\tau = \tau$ if and only if $\delta\gamma\delta^{-1}(\delta\tau) = \delta\tau$. Thus $\delta\Gamma_\tau\delta^{-1} = \Gamma_{\delta\tau}$. Notice that $|\delta\Gamma_\tau\delta^{-1}| = |\Gamma_\tau|$. Therefore, $|\Gamma_\tau| = |\Gamma_{\delta\tau}|$ showing that the periods of τ and $\delta\tau$ are same. □

Suppose τ is any elliptic point. Take a neighborhood U of τ given in Corollary 2.0.3. Consider the transformation $\delta_\tau = \begin{pmatrix} 1 & -\tau \\ 1 & -\bar{\tau} \end{pmatrix}$. This takes $\tau \to 0$ and $\bar{\tau} \to \infty$ acting as the same way as the modular group. Notice that $\delta_\tau(U)$ is a neighborhood around the origin. The stabiliser of $\delta_\tau\tau$ is $\delta_\tau\Gamma_\tau\delta_\tau^{-1}$. We will use the following observation: Γ_τ acts on the point z in U in the same way as $\delta_\tau\Gamma_\tau\delta_\tau^{-1}$ acts on δz.

The transformation $\delta_\tau\Gamma_\tau\delta_\tau^{-1}$ fixes $\delta\tau = 0$ and $\delta\bar{\tau} = \infty$. Therefore, $\delta_\tau\Gamma_\tau\delta_\tau^{-1}$ is a transformation of the form $z \mapsto cz$ for some complex number c. By Corollary 2.3.5, we know

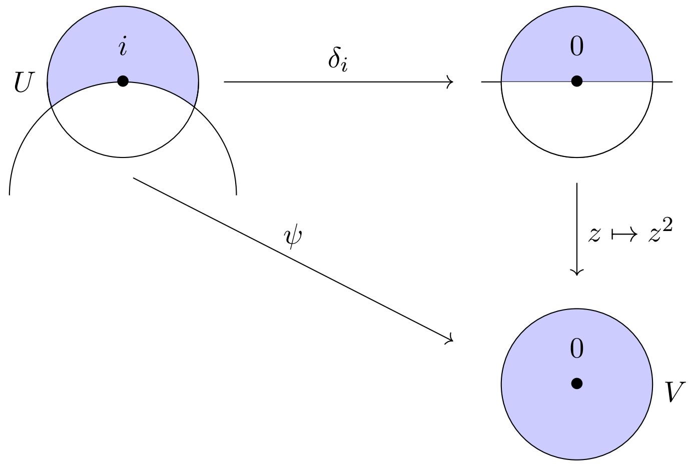

FIGURE 2.2: Coordinate chart around i

that $\delta_\tau \Gamma_\tau \delta_\tau^{-1}$ is a finite cyclic group of order e_τ. It is therefore the group of rotations by $2\pi/e_\tau$ about 0. Thus Γ_τ corresponds to a rotation of $2\pi/e_\tau$ about τ. When we transform U via the map δ_τ, the Γ-equivalent points are separated by a fixed angle. We next wrap the sector around a disc V via the map $z \mapsto z^{e_\tau}$. Call this map ρ. Define $\psi = \rho \circ \delta$. Since it is clear that δ_τ is defined for a particular point τ, for simplicity, we drop the subscript. Note that by the Open mapping theorem ψ is an open map. Our final claim for this section is the following:

Claim 2.4.2. The projection $\Pi : U \to \Pi(U)$ and the map $\psi : U \to V$ identify the same points.

Proof. We need to show that for any two points τ_1 and τ_2 in U, $\Pi(\tau_1) = \Pi(\tau_2)$ if and only if $\psi(\tau_1) = \psi(\tau_2)$. Notice that $\Pi(\tau_1) = \Pi(\tau_2)$ if and only if $\tau_1 \in \Gamma\tau_2$. Since U is a neighborhood as in Corollary 2.0.3, we see that $\tau_1 \in \Gamma_\tau \tau_2$. This happens if and only if $\delta\tau_1 \in \delta\Gamma_\tau\delta^{-1}(\delta\tau_2)$ if and only if $\delta\tau_1 = e^{2\pi i d/e_\tau}\delta\tau_2$ for some positive integer d. This implies that $\Pi(\tau_1) = \Pi(\tau_2)$ if and only if $(\delta\tau_1)^{e_\tau} = (\delta\tau_2)^{e_\tau}$ if and only if $\psi(\tau_1) = \psi(\tau_2)$. □

The above claim helps us to induce an injection $\phi : \Pi(U) \to V$ such that the diagram below commutes. In fact ϕ is an open map and a surjection as well since Π and ψ are so. Therefore, ϕ is a homeomorphism from $\Pi(U)$ to V.

$$\begin{tikzcd} & U \arrow[dl, "\Pi"'] \arrow[dr, "\psi"] & \\ \Pi(U) \arrow[rr, "\phi"'] & & V \end{tikzcd}$$

We need to check that the map defined above agrees on the overlaps. This is seen as follows: The reader is encouraged to draw a picture to have a better idea of what is going on. We first need a small result:

Claim 2.4.3. The neighborhood U of the elliptic point τ which satisfies the properties of Corollary 2.0.3 does not have any other elliptic point.

Proof of claim 2.4.3. Notice that $z = (az+b)/cz+d)$ is a degree two equation $cz^2 + (d-a)z-b = 0$. This has only two solutions. If any matrix γ fixes a point τ in the upper half plane, then it also fixes $\overline{\tau}$. Therefore, γ will have only one fixed point in the upper half plane. Now suppose U is a neighborhood around the elliptic point τ as in Corollary 2.0.3. Suppose τ' is another elliptic point in U such that $\delta\tau' = \tau'$. Then $\delta U \cap U \neq \varphi$. By Corollary 2.0.3, $\delta \in \Gamma_\tau$ that is δ fixes τ. However, δ cannot have two fixed points in the upper half plane, and so $\tau' = \tau$.

□

Next, let U_1 and U_2 be neighborhoods as above, of elliptic points $\widetilde{\tau}_1$ and $\widetilde{\tau}_2$ with periods h_1 and h_2 respectively. We need to check that the restriction of the map $\phi_2 \circ \phi_1^{-1}$ on the overlap $\phi_1(\Pi_1(U_1) \cap \Pi_2(U_2))$ is holomorphic. Let $V_{1,2} = \phi_1((\Pi_1(U_1) \cap \Pi_2(U_2))$ and $V_{2,1} = \phi_2((\Pi_1(U_1) \cap \Pi_2(U_2))$. We then have the following commutative diagram:

$$\begin{tikzcd} & \Pi(U_1) \cap \Pi(U_2) \arrow[dr, "\phi_2"] & \\ V_{1,2} \arrow[ur, "\phi_1^{-1}"] \arrow[rr, "\phi_2 \circ \phi_1^{-1}"'] & & V_{2,1} \end{tikzcd}$$

For each $x \in \Pi(U_1) \cap \Pi(U_2)$, it is enough to check holomorphy of the transition map in some neighborhood of $\phi_1(x)$. Write $x = \Pi(\tau_1) = \Pi(\tau_2)$ for some τ_1 and τ_2 in U_1 and U_2 respectively, such that and $\tau_1 = \gamma\tau_2$ for some $\gamma \in \Gamma$. Let $U_{1,2} = U_1 \cap \gamma^{-1}(U_2)$. Then $\Pi(U_{1,2})$ is a neighborhood of x in $\Pi(U_1) \cap \Pi(U_2)$ and so $\phi_1(\Pi(U_{1,2}))$ is a neighborhood of $\phi_1(x)$ in $V_{1,2}$. Keeping in mind the following local chart maps from U_1 and U_2, we divide the argument in 3 parts.

$$\begin{tikzcd} & U_i \arrow[dl, "\Pi"'] \arrow[dr, "\psi_i"] & \\ \Pi(U) \arrow[rr, "\phi_i"'] & & V \end{tikzcd} \qquad \begin{tikzcd} & U_i \arrow[dl, "\Pi"'] \arrow[dr, "\psi_i"] & \\ \delta_i(U) \arrow[rr, "z \mapsto z^{h_i}"'] & & V_i \end{tikzcd}$$

Part 1 of the argument: Suppose that $\phi_1(x) = 0$. That is $\phi_1(\Pi(\tau_1)) = 0$. This implies that $\psi(\tau_1) = 0$. From the diagram above, we have that $\delta_1(\tau_1) = 0$. Since δ_1 is a homoemorphism from U_1, this implies that $\tau_1 = \widetilde{\tau_1}$ is the elliptic point in U_1. Let $q = \phi_1(x')$ where x' is any arbitrary point in the neighborhood $\Pi(U_{1,2})$ of x. Notice that $q = \delta_1(\tau')^{h_1}$ for some $\tau' \in U_{1,2}$. Finally the task is to check the holomorphy at $\phi_2 \circ \phi_1^{-1}(\phi_1(x')) = \phi_2(x')$. Notice that $\tau' \in U_{1,2}$ and so $\gamma\tau' \in U_2$. Therefore it is well defined to write $\phi_2(x') = \phi_2(\Pi(\gamma\tau')) = \psi_2(\gamma\tau')$. Write

$$\begin{aligned}\phi_2(x') &= \psi_2(\gamma\tau')\\ &= \delta_2(\gamma(\tau'))^{h_2}\\ &= (\delta_2\gamma\delta_1^{-1})(\delta_1(\tau'))^{h_2}\\ &= (\delta_2\gamma\delta_1^{-1})(q^{1/h_1})^{h_2}\end{aligned}$$

If $h_1 = 1$, then we do not have a problem as $q \mapsto (\delta_2\gamma\delta_1^{-1})(q^{h_2})$ is a holomorphic map. The case when $h_1 > 1$ is taken care of as follows: If $h_1 > 1$, then τ_1 is elliptic. Since $\tau_2 = \gamma\tau_1$, we see that τ_2 is elliptic with the same period. By construction, U_2 is a neighborhood as in Corollary 2.0.3 around the elliptic point $\widetilde{\tau_2}$. By claim 2.4.3, it has only one elliptic point. This implies that $\widetilde{\tau_2} = \tau_2$ and $h_2 = h_1$. Thus $\delta_2\gamma\delta_1^{-1}$ is a map such that:

$$0 \xmapsto{\delta_1^{-1}} \tau_1 \overset{\gamma}{\mapsto} \tau_2 \xmapsto{\delta_2} 0.$$

$$\infty \xmapsto{\delta_1^{-1}} \tau_1 \overset{\gamma}{\mapsto} \tau_2 \xmapsto{\delta_2} \infty.$$

As previously seen this implies that $\delta_2\gamma\delta_1^{-1}(z) = cz$ for some complex number c. Finally we get that $\phi_2 \circ \phi_1^{-1}$ is the map: $q \mapsto (cq^{1/h_1})^{h_1} = c^{h_1}q$ which is clearly holomorphic.

Part 2 *of the argument*: So far the argument assumes that $\phi_1(x) = 0$, But it also covers the case when, $\phi_2(x) = 0$. In this case, the map to consider is $\phi_1^{-1} \circ \phi_2$. However this is just the inverse of the previous map $\phi_2^{-1} \circ \phi_1$. Since inverse of a holomorphic bijection is holomorphic, this case follows easily.

Part 3 *of the argument*: For the general case, take a local chart U_3 around x such that $\phi_3 : \Pi(U_3) \to V_3$ is the local map and $\phi_3(x) = 0$. Write $\phi_2 \circ \phi_1^{-1} = (\phi_2 \circ \phi_3^{-1}) \circ (\phi_3 \circ \phi_1^{-1})$. The map on the right hand side is a composition of holomorphic maps by part 1 and 2 and so the left hand side defines a holomorphic map as well.

2.5 Cusps

In order to make the modular curve compact, we add the points $\mathbb{Q} \cup \{\infty\}$ to the upper half plane. Call $\mathbb{H} \cup \{\infty\} \cup \mathbb{Q}$ the extended upper half plane and denote it by $\mathbb{H}^*$. Action of $SL_2(\mathbb{Z})$ on rationals is defined as:

$$\begin{pmatrix} a & b \\ c & b \end{pmatrix} \left(\frac{m}{n}\right) = \frac{am+bn}{cm+dn}$$

The point ∞ goes to a/c. Also notice that $-c/d$ goes to ∞. The point infinity is only identified with rationals up to the action of $SL_2(\mathbb{Z})$ or any of its subgroups. This justifies adding rationals to the upper half plane. Notice that the stabiliser of the point ∞ is the subgroup:

$$SL_2(\mathbb{Z})_\infty = \left\{ \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \,\middle|\, m \in \mathbb{Z} \right\}$$

These are exactly the translations in $SL_2(\mathbb{Z})$ as they take τ to $\tau + m$.

Define the compact modular curve $X(\Gamma)$ to be $\mathbb{H}^*/\Gamma$. The points $\{\infty\} \cup \mathbb{Q}$ up to Γ-equivalence are called cusps of $X(\Gamma)$.

Proposition 2.5.1. The modular curve $X(1)$ has one cusp.

Proof. Let $s = a/c$ be a cusp such that $\gcd(a, c) = 1$. Then there exist integers b, d such that $ad - bc = 1$. Consider the matrix:

$$\gamma_s = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then $\gamma_s(\infty) = s$ and $\gamma_s \in SL_2(\mathbb{Z})$. Thus every cusp is $SL_2(\mathbb{Z})$-equivalent to the point ∞ and we have just one cusp in $X(1)$. $\square$

Since any congruence subgroup Γ has a finite index in $SL_2(\mathbb{Z})$, a similar argument as done for elliptic points shows that $X(\Gamma)$ has finite number of cusps.

2.6 Defining a topology on the extended upper half plane

In order to define coordinate charts around the cusps, we first need to define open sets around them. For any arbitrary $M > 0$, define $\mathcal{N}_M = \{\tau \in \mathbb{H} \mid \text{Im}(\tau) > M\}$. Take $\mathcal{N}_M \cup \{\infty\}$ to be the open sets around the point infinity. For the rationals, we take all the sets of the form $\alpha(\mathcal{N}_M \cup \{\infty\})$ where $\alpha \in SL_2(\mathbb{Z})$. The transformation α takes the upper half plane at $\text{Im}(\tau) = M$ to a disc tangent at a rational number s where s is such that $\alpha(\infty) = s$. We next induce quotient topology to the modular curve $X(\Gamma)$ via the natural map $\Pi : \mathbb{H}^* \to X(\Gamma)$.

Proposition 2.6.1. The modular curve $X(\Gamma)$ is Hausdorff, connected and compact.

Proof. Hausdorffness: We need to show any two points on the modular curve can be separated by disjoint open neighborhoods. We will consider 2 cases here.

1. Let $x_1 = \Gamma s_1$ where $s_1 \in \mathbb{Q}$ be a cusp and $x_2 = \Gamma\tau_2$ with $\tau_2 \in \mathbb{H}$ be a point of $X(\Gamma)$. We know that $s_1 = \alpha(\infty)$ for some $\alpha \in SL_2(\mathbb{Z})$. Let U_2 be any neighborhood of τ_2 with a compact closure K. We first want to show that the set $\{\operatorname{Im}\gamma K \mid \gamma \in SL_2(\mathbb{Z})\}$ is bounded. In order to show this, we need a small result.

   Lemma 2.6.2. $\operatorname{Im}(\gamma\tau) \leq \max\{\operatorname{Im}(\tau), 1/\operatorname{Im}(\tau)\}$.

   Proof. To the contrary, suppose otherwise. By the proof of Lemma 1.0.2, we know that $\operatorname{Im}(\gamma\tau) > \operatorname{Im}(\tau)$ implies that $|c\tau + d| < 1$ while $\operatorname{Im}(\gamma\tau) > 1/\operatorname{Im}(\tau)$ implies that $\operatorname{Im}(\tau) > |c\tau + d|$. Overall, we get

$$
\begin{aligned}
|c\tau + d| &= \geq \operatorname{Im}(|c\tau + d|) \\
&= |c|\operatorname{Im}(\tau) \\
&> |c||c\tau + d|
\end{aligned}
$$

   giving us that that $c = 0$. The inequality in the beginning helps us conclude that $d = 0$. With $c = d = 0$, $\gamma \notin SL_2(\mathbb{Z})$, a contradiction. □

   Coming back to the proof of the proposition, since K is compact, the set $\{\operatorname{Im}(\tau) \mid \tau \in K\} \subseteq [c_1, c_2]$ with $c_1 > 0$. This gives us that $\operatorname{Im}(\tau) \leq c_2$ and $1/\operatorname{Im}(\tau) \leq c_1$ for all $\tau \in \mathbb{H}$. By Lemma 2.6.2 it follows that $\{\operatorname{Im}(\gamma K) \mid \gamma \in SL_2(\mathbb{Z})\}$ is a bounded set. Therefore, we can find a suitable M such that $SL_2(\mathbb{Z})K \cap \mathcal{N}_M = \phi$. Now take the neighborhood $U_1 = \alpha(\mathcal{N}_M \cup \{\infty\})$ of s_1.

   Claim 2.6.3. The neighborhoods $\Pi(U_1)$ and $\Pi(U_2)$ of x_1 and x_2 respectively are disjoint in $X(\Gamma)$.

   Suppose $\Pi(U_1) \cap \Pi(U_2) \neq \phi$ Then, there exists $\tau_2 \in U_2 \subset K$ such that $\Pi(\tau_2) \in \Pi(U_1)$. This implies that there exists $\gamma \in \Gamma$ and $\tau_1 \in \mathcal{N}_M \cup \{\infty\}$ with $\tau_2 = \gamma(\alpha\tau_1)$. But then $\tau_1 = \alpha^{-1}\gamma^{-1}\tau_2$, a contradiction to the observation that $SL_2(\mathbb{Z})K \cap \mathcal{N}_M = \phi$.

2. Let $x_1 = \Gamma s_1$ and $x_2 = \Gamma s_2$ both be distinct cusps of $X(\Gamma)$. There exists α_1 and $\alpha_2 \in SL_2(\mathbb{Z})$ such that $s_1 = \alpha_2(\infty)$ and $s_2 = \alpha_2(\infty)$. Let $U_1 = \alpha_1(\mathcal{N}_2 \cup \{\infty\})$ and $U_2 = \alpha_2(\mathcal{N}_2 \cup \{\infty\})$. We chose the neighborhood $\mathcal{N}_2$ because of the following two observations which will be used in the proof and in the later sections as well.

   Lemma 2.6.4. The region $\mathcal{N}_2$ in $\mathbb{H}$ has no elliptic points.

Proof. Recall that the elliptic points of $SL_2(\mathbb{Z})$ are given by the set $\{\gamma i, \gamma\mu_3 \mid \gamma \in SL_2(\mathbb{Z})\}$. Suppose $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. Then $\mathrm{Im}(\gamma i) = \mathrm{Im}(i)/|ci+d|^2 = 1/(c^2+d^2)^2 \leq 1$. Similarly we calculate

$$\begin{aligned} \mathrm{Im}(\gamma\mu_3) &= \mathrm{Im}(\mu_3)/|c\mu_3+d|^2 \\ &= (\sqrt{3}/2)/|c(-1/2+\sqrt{3}/2)+d|^2 \\ &< 1/((d-c/2)^2+3c^2/4)^2 \\ &\leq 1 \end{aligned}$$

Therefore all the elliptic points lie below $\mathcal{N}_2$. □

Lemma 2.6.5. Suppose τ_1 and $\tau_2 \in \mathcal{N}_2$ are distinct points such that $\tau_1 = \gamma\tau_2$ for some $\gamma \in SL_2(\mathbb{Z})$. Then γ is a translation.

Proof. Observe that with the given conditions of the lemma, τ_1 and τ_2 correspond to the same points in the fundamental domain $\mathcal{D}$. Since $\tau_1 \neq \tau_2$, this implies that they must differ by a translation, otherwise they would correspond to different points in $\mathcal{D}$. □

Coming back to the argument, we next prove that with U_1 and U_2 as defined above, the intersection $\Pi(U_1) \cap \Pi(U_2) = \phi$. Suppose not, then there exists $\gamma \in \Gamma$ and $\tau_1, \tau_2 \in \mathcal{N}_2$ such that $\gamma(\alpha_1(\tau_1)) = \alpha_2\tau_2$. It follows that $\alpha_2^{-1}\gamma\alpha_1(\tau_1) = \tau_2$. Notice that if $\tau_1 = \tau_2$, then $\alpha^{-1}\gamma\alpha_1$ fixes τ_1 implying that $\mathcal{N}_2$ has elliptic points. This is not possible by Lemma 2.6.4 and so $\tau_1 \neq \tau_2$ and they correspond to the same point in $\mathcal{D}$. By Lemma 2.6.5, $\alpha_2^{-1}\gamma\alpha_1$ is a translation. Finally this implies that $\alpha_2^{-1}\gamma\alpha_1(\infty) = \infty$, giving us that $\gamma s_1 = s_2$, a contradiction to the fact that Γs_1 and Γs_2 are distinct points.

3. The case when both the points are in $\mathbb{H}^*$ has been dealt with in Corollary 2.0.8.

Connectedness: Suppose that $\mathbb{H}^*$ is a disjoint union of open subsets $\mathcal{O}_1$ and $\mathcal{O}_2$. Consider $(\mathcal{O}_1 \cup \mathcal{O}_2) \cap \mathbb{H}$. Since $\mathbb{H}$ is connected, $\mathbb{H} \subset \mathcal{O}_1$ (say) and $\mathcal{O}_2 \in \mathbb{Q} \cup \{\infty\}$. But then $\mathcal{O}_2$ is open and so it has to be empty. It follows that $\mathbb{H}^*$ is connected.

Compactness: We first show that $\mathcal{D} \cup \{\infty\} = \mathcal{D}^*$ is compact. Consider an open cover of $\mathcal{D}^*$ in the subspace topology inherited by $\mathbb{H}^*$. One of the open sets, say U in the cover must contain the point ∞ and therefore, it must contain the neighborhood $(\mathcal{N}_M \cup \{\infty\}) \cap \mathcal{D}^*$ for some M. Now $\mathcal{D}^* \backslash (\mathcal{N}_M \cup \{\infty\} \cap \mathcal{D}^*)$ is closed and bounded, and so compact in

$\mathbb{H}$. Thus we have a finite sub cover of this space. This sub cover, together with U is a finite cover of $\mathcal{D}^*$. Next, notice that

$$\mathbb{H}^* = SL_2(\mathbb{Z})\mathcal{D}^* = \bigcup_j \Gamma\gamma_j(\mathcal{D}^*)$$

where γ_j are the coset representatives of Γ in $SL_2(\mathbb{Z})$. This gives us that

$$X(\Gamma) = \bigcup_j \Pi(\gamma_j(\mathcal{D}^*))$$

Since this is a finite union and Π is a continuous map, we conclude that $X(\Gamma)$ is compact. $\square$

We next define charts around cusps to make $X(\Gamma)$ into a compact Riemann surface.

2.7 Defining coordinate charts for cusps

At the cusps, infinitely many sectors come together. See figure 2.3. Let s be a cusp. In order to define a local coordinate at s, we first take the neighborhood $\mathcal{N}_2 \cup \{\infty\}$ of ∞. Let δ be a transformation such that $\delta s = \infty$. Let U_s be the neighborhood $\delta^{-1}(\mathcal{N}_2 \cup \{\infty\})$. In this case δ straightens the neighborhood by making the identified sectors differ by a "translation" by a positive integer h. We then act the transformation $\rho = e^{2\pi i\tau/h}$ which "wraps" the neighborhood of infinity into a disc with centre 0. To see this precisely, we first introduce the notion of the width of a cusp.

Definition 2.7.1 (Width of a cusp). Suppose s is a cusp for the modular curve $X(\Gamma)$. Let δ be a transformation taking s to ∞. We define:

$$h_s = |SL_2(\mathbb{Z})_\infty / (\delta\{\pm\}\Gamma\delta^{-1})_\infty|.$$

Proposition 2.7.2. The width h_s is characterized by the following properties:

1. It is finite.
2. It is independent of the transformation δ.
3. It is well defined on $X(\Gamma)$.

Proof of 1. Since $\Gamma(N) \subseteq \Gamma$, for some N and $\Gamma(N)$ is a normal subgroup of $SL_2(\mathbb{Z})$, we have that $\delta^{-1}\Gamma(N)\delta = \Gamma(N) \subseteq \delta^{-1}\Gamma(N)\delta$. This implies that

$$\begin{aligned} h_s &= |SL_2(\mathbb{Z})_\infty/(\{\pm\}\delta\Gamma\delta^{-1})_\infty| \\ &\leq |SL_2(\mathbb{Z})_\infty/(\{\pm\}\Gamma(N)_\infty| \end{aligned}$$

Since, $\pm\Gamma(N)_\infty = \left\{\pm\begin{pmatrix}1 & mN \\ 0 & 1\end{pmatrix} \;\middle|\; m \in \mathbb{Z}\right\}$, the quotient has order N, giving us that h_s is finite.

Proof of 2. We will show that $h_s = |SL_2(\mathbb{Z})_s/\{\pm\}\Gamma_s|$, where $SL_2(\mathbb{Z})_s$ denotes the stabiliser of s with respect to $SL_2(\mathbb{Z})$. Observe that if $\delta(s) = \infty$ and $\gamma \in \Gamma$, then $\delta\gamma\delta^{-1}(\infty) = \infty$ if and only if $\gamma(s) = s$. Therefore, $(\delta\Gamma\delta^{-1})_\infty = \delta\Gamma_s\delta^{-1}$. The width is now expressed as:

$$\begin{aligned} h_s &= |SL_2(\mathbb{Z})_\infty/(\{\pm\}\delta\Gamma\delta^{-1})_\infty| \\ &= |\delta SL_2(\mathbb{Z})_s\delta^{-1}/|(\delta\{\pm\}\Gamma_s\delta^{-1}| \\ &= |SL_2(\mathbb{Z})_s/\{\pm\}\Gamma_s| \end{aligned}$$

This shows that h_s is independent of δ.

Proof of 3. We will show that if $s \in \mathbb{Q} \cup \{\infty\}$ and $\gamma \in SL_2(\mathbb{Z})$, then $h_{s,\Gamma} = h_{\gamma(s),\gamma\Gamma\gamma^{-1}}$. In particular, if $\gamma \in \Gamma$, then we get that $h_{s,\Gamma} = h_{\gamma(s),\Gamma}$. By part 2 above it follows that

$$\begin{aligned} h_{\gamma(s),\gamma\Gamma\gamma^{-1}} &= |SL_2(\mathbb{Z})_{\gamma(s)}/(\{\pm\}\gamma\Gamma\gamma^{-1})_{\gamma(s)}| \\ &= |\gamma SL_2(\mathbb{Z})_s\gamma^{-1}/(\{\pm\}\gamma\Gamma_s\gamma^{-1})| \\ &= |SL_2(\mathbb{Z})_s|/\{\pm\}\Gamma_s| \\ &= h_s \end{aligned}$$

□

Next, define $\psi = \rho \circ \delta$ where ρ and δ are transformations described in the beginning of this section. We next claim that:

Claim 2.7.3. The maps ψ and Π carry out the same identification in the neighborhood U_s of s.

The argument given below will also help in realizing the geometric significance of h_s.

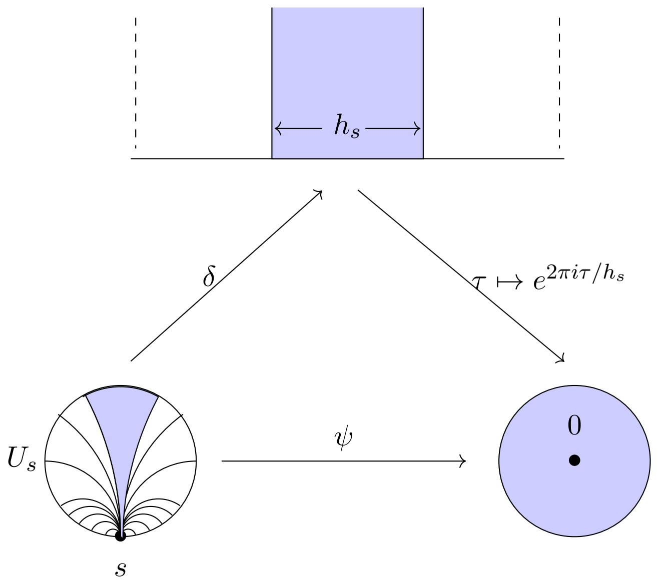

FIGURE 2.3: Coordinate chart around a cusp s.

Proof. Suppose τ_1 and τ_2 are two points in U_s, then $\Pi(\tau_1) = \Pi(\tau_2)$ if and only of $\tau_1 = \gamma\tau_2$ for some $\gamma \in \Gamma$. This happens if and only if $\delta\tau_1 = (\delta\gamma\delta^{-1})(\delta\tau_2)$. Since $\delta\tau_1$ and $\delta\tau_2$ both lie in $\mathcal{N}_2$, the transformation $\delta\gamma\delta^{-1}$ is a translation. We can now conclude that $\delta\gamma\delta^{-1} \in \delta\Gamma\delta^{-1} \cap SL_2(\mathbb{Z})_\infty = (\delta\Gamma\delta^{-1})_\infty$. This is exactly the subgroup generated by $\pm\begin{pmatrix} 1 & h_s \\ 0 & 1 \end{pmatrix}$ by the definition of the width. This implies that $\Pi(\tau_1) = \Pi(\tau_2)$ if and only if $\delta(\tau_1) = \delta(\tau_2) + mh_s$ for some $m \in \mathbb{Z}$, if and only if $\psi(\tau_1) = \psi(\tau_2)$. $\square$

As discussed at the beginning of the section, it is clear now that δ is spacing out the Γ-equivalent sectors by a translation by h_s. Claim 2.7.3 helps us to induce an injection ψ as we did in the case of elliptic points, from $\Pi(U_s)$ to V. Arguing similarly as before, ψ is a local homeomorphism from $\Pi(U_s)$ to V.

In the case of cusps as well, we need to check that the transition maps are holomorphic on the overlaps. The argument is divided into two cases.

Case 1*:* Suppose $\widetilde{\tau}_1$ is an elliptic point with period h_1 and s_2 is a cusp with width h_2. As done previously in case of elliptic points, let U_1 be the neighborhood around $\widetilde{\tau}_1$ with the corresponding map $\delta_{\widetilde{\tau}}$, and $U_2 = \delta_2^{-1}(\mathcal{N}_2 \cup \{\infty\})$. Here, $\delta_{\widetilde{\tau}} = \begin{pmatrix} 1 & -\overline{\widetilde{\tau}} \\ 1 & -\widetilde{\tau} \end{pmatrix}$ and δ_2 is the transformation, which maps s to ∞. As before, for each $x \in \Pi(U_1) \cap \Pi(U_2)$ write $x = \Pi(\tau_1) = \Pi(\tau_2)$ with $\tau_1 \in U_1$, $\tau_2 \in U_2$ and $\tau_2 = \gamma(\tau_1)$ for some $\gamma \in \Gamma$. We need to check the holomorphy of the map $\phi_2 \circ \phi_1^{-1}$ on the overlaps, but here ϕ_1 and ϕ_2 are maps corresponding to the coordinate chart of an elliptic point and a cusp respectively. Let

$U_{1,2} = U_1 \cap \gamma^{-1}(U_2)$, a neighborhood of $\tau_1 \in \mathbb{H}$. Then $\phi_1(\Pi(U_{1,2}))$ is a neighborhood of $\phi_1(x)$ in $\phi_1(\Pi(U_1) \cap \Pi(U_2)$. Suppose $x' \in \Pi(U_{1,2})$. Let $q = \phi_1(x') = \delta_{\widetilde{\tau}}(\tau')^{h_1}$ where $\tau' \in U_{1,2}$ such that $\Pi(\tau') = x'$. We compute

$$\begin{aligned} \phi_2 \circ \phi_1^{-1}(q) &= \phi_2(x') \\ &= \phi_2(\Pi(\gamma\tau')) \\ &= \psi_2(\gamma\tau') \\ &= \exp(2\pi i \delta_2 \gamma(\tau')/h_2) \\ &= \exp(2\pi i \delta_2 \gamma \delta^{-1}(\delta(\tau'))/h_2) \\ &= \exp(2\pi i \delta_2 \gamma \delta^{-1}(q^{1/h_1})/h_2) \end{aligned}$$

Since the map $z \mapsto z^{1/h}$ is analytic at all points except 0, the only case where the map $q \mapsto \phi_2 \circ \phi_1^{-1}(q)$ might not be holomorphic is when $h_1 > 1$ and $0 \in \phi_1(\Pi(U_{1,2}))$. However these two cannot happen at the same time. To see this we observe that if $h > 1$, then $\widetilde{\tau}_1 \notin U_{1,2}$. Otherwise, $\delta_2(\gamma\widetilde{\tau}_1)$ is an elliptic point in $\mathcal{N}_2$. We conclude that if $h > 1$, then $0 \notin (\delta_{\widetilde{\tau}}(U_{1,2}))^{h_1} = \psi(U_{1,2}) = \phi_1(\Pi(U_{1,2}))$.

Case 2: Suppose s_1 and s_2 are both cusps with width h_1 and h_2 respectively. Let $U_1 = \delta_1^{-1}(\mathcal{N}_2 \cup \{\infty\})$ and $U_2 = \delta_2^{-1}(\mathcal{N}_2 \cup \{\infty\})$. Here $\delta_1(s_1) = \infty$ and $\delta_1(s_2) = \infty$. Suppose $\Pi(U_1) \cap \Pi(U_2) \neq \phi$. This implies that for some $\gamma \in \Gamma$, $\gamma^{-1}\delta_1^{-1}(\mathcal{N}_2 \cup \{\infty\}) \cap \delta_2^{-1}(\mathcal{N}_2 \cup \{\infty\}) \neq \phi$. Thus, $\delta_2\gamma\delta_1^{-1}(\tau_1) = \tau_2$ for some τ_1 and $\tau_2 \in \mathcal{N}_2 \cup \{\infty\}$. By Lemma 2.6.5, $\delta_2\gamma\delta_1^{-1} = \pm \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$ for some $m \in \mathbb{Z}$. We finally get that

$$\begin{aligned} \gamma(s_1) &= \gamma(\delta_1^{-1}(\infty)) \\ &= \pm\delta^{-1} \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} (\infty) \\ &= s_2 \end{aligned}$$

since $\gamma(s_1) = s_2$ and $h_1 = h_2 = h$. Using the same notation as in case 1, the transition map $\phi_2 \circ \phi^{-1}$ takes a point $q = \Pi(\phi_1(\tau')) = \psi(\tau') = \exp(2\pi i \delta_1(\tau')/h)$, where $\tau' \in U_{1,2}$

to $\phi_2 \circ \phi_1 - 1(q)$ which is computed as follows:

$$\begin{aligned}
\phi_2 \circ \phi_1^{-1}(q) &= \phi_2(q) \\
&= \phi_2(\Pi(\tau')) \\
&= \phi_2(\Pi(\gamma\tau') \\
&= \psi_2(\gamma\tau') \\
&= \exp(2\pi i \delta_2 \gamma(\tau')/h_{)} \\
&= \exp(2\pi i \delta_2 \gamma \delta^{-1}(\delta(\tau'))/h) \\
&= \exp(2\pi i(\delta_1(\tau) \pm m)/h) \\
&= q\exp(2\pi i m/h)
\end{aligned}$$

The map $q \mapsto q\exp(2\pi i m/h)$ is clearly holomorphic.

2.8 Genus

We have finally established a modular curve as a compact Riemann surface. We noted in section 2.2 that a g holed surface is an example of a Riemann surface. We in fact have that:

Fact 2.8.1. Every compact Riemann surface looks like a g holed surface for some positive integer g. The number g is called the genus of the surface.

After doing the identifications on the fundamental domain, we see that the modular curve $Y(1)$ which is topologically equivalent to a punctured sphere. When we add the cusp to make it compact, we get a sphere. Therefore $X(1)$ is a surface of genus 0. In fact the j function defined in section 1.5 is a well defined homeomorphism from $X(\Gamma)$ to the complex sphere. In order to prove this we first introduce some notions related to maps between compact Riemann surfaces. We might need some results from complex analysis which we will only state and not prove.

Lemma 2.8.2. Any holomorphic map between compact Riemann surfaces is either a constant or a surjection.

Proof. Let X and Y be compact Riemann surfaces. Let $f : X \mapsto Y$ is holomorphic. Notice that X is compact and connected. Since f is continuous, $f(X)$ is compact, and hence, closed and connected. By the Open mapping theorem for compact Riemann surfaces, $f(X)$ is open. Therefore, the image being connected is either a single point or all of Y. $\square$

We next characterize the non-constant maps. Since a Riemann surface locally "looks like" the complex plane, we want to visualize how do these maps look locally. This is given by a result known as the local mapping theorem which we explain below. Let f be as before and U and V be neighborhoods of points x and $f(x)$ in X and Y respectively. We have the corresponding local maps between open neighborhoods around 0 in $\mathbb{C}$ as shown in the diagram. The local mapping tells us that diagram below commutes. In other words, for each x, locally f is the map $z \mapsto z^n$ for some n, in a suitable neighborhood of x.

$$\begin{tikzcd} U \arrow[r, "f"] \arrow[d, "\phi"'] & V \arrow[d, "\psi"] \\ U' \arrow[r, "z\mapsto z^n"'] & V' \end{tikzcd}$$

Definition 2.8.3 (Ramification degree). For each $x \in X$, the ramification degree e_x of x is the multiplicity with which f takes 0 to 0 in the local coordinate charts. That is, f is an e_x to 1 map around x.

We first claim that there are only finitely many points with ramification degree greater than 1. This can be seen by realizing that the above definition is equivalent to saying that if e_x is the ramification degree then the function, $f(z) - f(x)$ vanishes at x to the order of e_x. Mathematically, this means that $f(z) - f(x) = (z - x)^{e_x} g(z)$ for some function $g(z)$. Suppose there are infinitely many points with ramification greater than 1. Then $f'(z) = 0$ at these points. Since X is compact, we can find a converging sequence of points where the derivative f' vanishes, and hence, $f'(z) = 0$ for all $z \in X$ by the Identity theorem for Riemann surfaces. So f is constant, which is not the case. Now, for every $y \in Y$, $f^{-1}(y)$ is discrete. Since X is compact, $f^{-1}(y)$ is finite.

Theorem 2.8.4. There exists a positive integer d such that

$$d = \sum_{x \in f^{-1}(y)} e_x$$

This theorem asserts that $|f^{-1}(y)|$ is constant in the case when all the pre-images $x \in f^{-1}(y)$ are unramified. For the points where ramification happens, we have to count the multiplicity of the ramified points as well. When we do that, we again end up with the same constant. We will call d, the *degree* of the map f.

Proof. Let $S = \{x \in X \mid e_x > 1\}$ denote the finite set of ramified points. Let $X' = X \backslash S$ and $Y' = Y \backslash f(S)$ be the Riemann surfaces obtained by deleting these finitely many points. Now, for each $x \in f^{-1}(y), e_x = 1$, so there exists a neighborhood U_x where f is locally bijective. We can shrink these neighborhoods to make them disjoint. This implies that for every y, there exists a neighborhood U_y such that $f^{-1}(U_y) = \cup_{x\in f^{-1}(y)} U_x$, where each U_x is disjoint from the other and $f|_{U_x} : U_x \to U_y$ is a bijection. Therefore, $y \mapsto |f^{-1}(y)|$ defines a continuous function from Y' to $\mathbb{Z}$. Since Y' is connected, the function has to be constant. To account for points which are ramified, we yet again look at the definition of the ramification degree. If $y = f(x)$ is the image of a ramified point $x \in X$, then definition 2.8.3 implies that every point y' in a small enough neighborhood of y is the image of e_x many points x' with $e_{x'} = 1$ near x. Therefore in this case as well,

$$\sum_{x\in f^{-1}(y)} e_x = |f^{-1}(y')| = d$$

□

Having introduced the degree of a map, let us get back to the j function. From the definition given in (1.5.3), we know that the j function has poles wherever the Δ function vanishes. Using the following theorem, we deduce that the j function is holomorphic at all of $\mathbb{H}$ and meromorphic only at the cusp ∞ of $X(\Gamma)$.

Theorem 2.8.5. The Δ function does not vanish for any $\tau \in \mathbb{H}$.

This is a non trivial result and comes from the theory of elliptic curves. One can refer to the Appendix to see a brief idea of the proof. Since the Δ function is a cusp form, it vanishes at the unique cusp ∞ of $X(1)$. So the only pole of the j function is at ∞. It follows that the j function is a well defined meromorphic map from $X(1)$ to $\mathbb{C}$ and so defines a holomorphic map to the Riemann sphere $\widehat{\mathbb{C}}$. As a map between Riemann surfaces, by Lemma 2.8.2, it is a surjection. Its q expansion given in 1.5.4 tells us that the only pole it has, is simple. Therefore, it is a degree 1 map and hence an injection. This completes the argument that the simplest modular curve we defined is a complex sphere.

2.9 Order of modular forms on $X(\Gamma)$

Definition 2.9.1 (Order of f). Suppose $f \in \mathcal{A}_k(\Gamma)$. About any point $\tau \in V \subset \mathbb{H}$, the function f has a Laurant series expansion

$$f(z) = \sum_{n=m}^{\infty} a_n (z-\tau)^n \tag{2.9.2}$$

Define the order of vanishing at the point τ to be m. We denote it by $ord_\tau(f)$.

Clearly if f is holomorphic at τ then $ord_\tau(f) \geq 0$, otherwise if τ is a pole then $ord_\tau(f) < 0$.

Definition 2.9.3 (Order at the cusps). For a point $s \in \mathbb{Q} \cup \{\infty\}$, define the order of f to be $ord_s(f) = ord_\infty(f[\alpha]_k)$ where α is such that $\alpha(\infty) = s$ and $\alpha \in SL_2(\mathbb{Z})$.

We need to make sure the order at the cusps is well defined. To see that it is independent of the transformation α, let δ be another such transformation such that $\delta(\infty) = s$. Then $\alpha^{-1}\delta$ fixes ∞. This implies that $\delta = \alpha \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$. So, $(f[\delta]_k)(\tau) = f[\alpha]_k(\tau + m)$ and we get that $ord_\infty(f[\alpha]_k) = ord_\infty(f[\delta]_k)$. To check that the definition of the order is well defined on the quotient $X(\Gamma)$, notice that if $\gamma \in \Gamma$, then $ord_s(f) = ord_{\gamma s}(f)$ because $f[\alpha]_k = f[\gamma\alpha]_k$.

Remark 2.9.4. By definition, the order of f at a point τ in $\mathbb{H}^*$ is always integral. However, when we define it on the modular curve, one sees that this may not hold true.

We know that any meromorphic modular form is not a well defined function on the modular curve $X(\Gamma)$. In order to make sense of the order of vanishing of the modular form at a point on the modular curve, we need to take the local structure into account.

1. Defining the order at a non cusp: Let $\Pi(\tau) \in X(\Gamma)$ be a non cusp. Suppose the Laurant series expansion of f about τ is as given in (2.9.2) such that $ord_\tau(f) = m$. The local coordinate at $\Pi(\tau)$ is $q = (z - \tau)^e$ where e is the period of τ. In local coordinates,
$$f(z) = a_m q^{m/e} + \dots$$
We naturally define the order of f at $\Pi(\tau)$ to be
$$ord_{\Pi(\tau)}(f) = ord_\tau(f)/e \tag{2.9.5}$$
where e is the period of τ. Notice that for a non elliptic point, $e = 1$ and so the order remains the same, in agreement with the fact that Π is a local homeomorphism in this case.

2. Defining the order at a cusp: We first work out the definition for the cusp $\Pi(\infty)$. Suppose the q expansion of f about 0 is
$$\sum_{n=m}^{\infty} a_n q_{h_1}^n \text{ ; where } h_1 \text{ is the period of } f$$
The local coordinates about the cusp are given by $q_h = e^{2\pi i \tau/h}$ where h is the width of the cusp. In order to write f in terms of the local coordinate, we need

to relate the period h_1 of f with the width of $\Pi(\infty)$. Recall that the width is characterized by the index of the subgroup $\{\pm I\}\Gamma_\infty = \{\pm\}\left\langle\begin{pmatrix}1 & h\\ 0 & 1\end{pmatrix}\right\rangle; h \in \mathbb{Z}_+$. Depending on whether $-I$ belongs to Γ or not, three cases arise.

(a) The subgroup $\Gamma_\infty = \{\pm I\}\left\langle\begin{pmatrix}1 & h\\ 0 & 1\end{pmatrix}\right\rangle$; Recall that the period h_1 of f is the smallest positive integer such that the tranformation $\begin{pmatrix}1 & h_1\\ 0 & 1\end{pmatrix} \in \Gamma$. This, along with the description of the subgroup Γ_∞ above, gives us that $h_1 = h$. So in terms of the local coordinates as well, we get the same q expansion of f about 0 and thus the order remains the same on $X(\Gamma)$.

(b) When $\Gamma_\infty = \left\langle\begin{pmatrix}1 & h\\ 0 & 1\end{pmatrix}\right\rangle$; This case is similarly handled as in 1 and we get the period f to be the same as h.

(c) If $\Gamma_\infty = \left\langle -\begin{pmatrix}1 & h\\ 0 & 1\end{pmatrix}\right\rangle$; This case arises only when $-I \notin \Gamma$. Denote the generator as γ. Then notice that,

$$f(\tau + h) = f(\gamma\tau) = (-1)^k f(\tau).$$

When k is even, we again get that the period is h. However, if k is odd, then the period is $2h$. In this case:

$$f(q) = \sum_{n=m}^{\infty} a_n q_{2h}^n$$

and the order of f on the modular curve is $m/2$.

In general, for any cusp $\Pi(s)$, we take a transformation α which takes ∞ to s and run the same argument with $f[\alpha]_k$ instead of f and $(\alpha^{-1}\Gamma\alpha)_\infty$ in place of Γ_∞.

To summarize the above discussion:

$$ord_{\Pi(s)}(f) = \begin{cases} ord_s(f)/2 & \text{if } (\alpha^{-1}\Gamma\alpha)_\infty = \left\langle -\begin{pmatrix}1 & h\\ 0 & 1\end{pmatrix}\right\rangle \text{ and } k \text{ is odd} \\ ord_s(f) & \text{Otherwise} \end{cases} \tag{2.9.6}$$

When $ord_{\Pi(s)}(f)$ is half integral, we call these cusps as irregular, otherwise they are called regular cusps.

Having suitably defined the order on $X(\Gamma)$, we move on to studying a very important relation between modular forms and differential forms on $X(\Gamma)$.

2.10 A formal introduction to differential forms.

Definition 2.10.1 (Meromorphic differential form). A Meromorphic differential form on an open subset U of $\mathbb{C}$ is an expression of the form $f(q)dq$ where f is a meromorphic function on U.

Suppose $\phi : U_1 \to U_2$ is a mapping between two open sets in $\mathbb{C}$. Let $q_2 = \phi(q_1)$, where q_1 and q_2 denote the local coordinates on U_1 and U_2 respectively. Let $\omega = f(q_2)dq_2$ be a differential form on U_2. We define $\phi^*(\omega)$ to be a differential form defined as $f(\phi(q_1))\phi'(q_1)dq_1$ on U_1. The differential from $\phi^*(\omega)$ is called the pullback of ω.

Consider a differential form $\omega = f(z)dz$ on $\mathbb{H}$ where $f(z)$ is a meromorphic function. Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$. When we pull back the differential ω via the action of γ, we get that:

$$\gamma^*(\omega) = f(\gamma z)d(\gamma z) = f(\gamma z)(cz + d)^{-2}.$$

It follows that $\gamma^*(\omega) = \omega$ if and only if $f(\gamma z) = (cz + d)^2 f(z)$ i.e., f is a meromorphic modular form of weight 2. We in fact have a one-one correspondence between meromorphic modular forms of weight 2 and meromorphic differential forms on $\mathbb{H}$ that are invariant under the action of $SL_2(\mathbb{Z})$. These are precisely the meromorphic differentials on $X(1)$. We generalize this to weight k modular forms with respect to the subgroup Γ by formally defining differential forms of degree k.

Definition 2.10.2 (degree k differential forms). Define a differential form of degree k by the symbol $f(q)(dq)^k$ on an open set $U \subset \mathbb{C}$ where f is a meromorphic function on U.

Denote the space of k-fold meromorphic differential forms on U by $\Omega^k(U)$. This forms a vector space over $\mathbb{C}$ under the natural definition of addition and scaler multiplication:

$$f(q)dq^k + g(q)dq^k = (f + g)(q)(dq)^k$$

$$c(f(q))(dq)^k = cf(q)(dq)^k$$

Since a Riemann surface also involves transition maps, as introduced before, we need to define the pullback of local differentials from one open set to another. Any holomorphic map $\phi_1 : V_1 \to V_2$, between open sets in $\mathbb{C}$ induces the pullback as follows:

$$\phi^* : \Omega^k(V_1) \to \Omega^k(V_1)$$

$$f(q_2)(dq_2)^k \mapsto f(\phi(q_1))(\phi'(q_1))^k(dq_1)^k$$

Properties of the pullback are summarized with proofs in the following lemmas.

Lemma 2.10.3. The pullback is contravariant. This means that $(\phi_1 \circ \phi_2)^* = \phi_2^* \circ \phi_1^*$.

Proof. Let $\phi_1 : V_1 \to V_2$ and $\phi_2 : V_2 \to V_3$ be maps between open sets such that $f(q_i)dq_i$ are differential forms defined on each V_i respectively for $i = 1, 2$.

The map $(\phi_2 \circ \phi_1)^* : \Omega^k(V_3) \to \Omega^k(V_1)$ acts as follows:

$$\begin{aligned}
(\phi_2 \circ \phi_1)^*(f(q_3)(dq_3)^k) &= f(\phi_2 \circ \phi_1(q_1))((\phi_2 \circ \phi_1)'(q_1))^k(dq_1)^k \\
&= f(\phi_2(\phi_1(q_1)))(\phi_2'(\phi_1(q_1))^k(\phi_1'(q_1))^k(dq_1)^k \\
&= \phi_1^*(f(\phi_2(q_2)))(\phi_2'(q_2))^k(dq_2)^k \\
&= (\phi_1^* \circ \phi_2^*)(f(q_3)dq_3^k)
\end{aligned}$$

Therefore, we conclude that $(\phi_2 \circ \phi_1)^* = \phi_1^* \circ \phi_2^*$. □

Lemma 2.10.4. If $V_1 \subset V_2$ and $i : V_1 \to V_2$ is the inclusion map between open sets, then its pull back $i^* : \Omega^k(V_2) \to \Omega^k(V_1)$ is the restriction map, that is, $i^*(\omega) = \omega|_{V_1}$.

Proof. Let $\omega = f(q)(dq)^k$ be a differential form of degree k on V_2. When we pullback ω on V_2 we get that:

$$\begin{aligned}
i^*(f(q)(dq)^k &= f(i(q))(i'(q))^k(dq)^k \\
&= (f \circ i)(q)(dq)^k \\
&= (f|_{V_1})(q)(dq)^k \\
&= \omega|_{V_1}
\end{aligned}$$

□

Lemma 2.10.5. If ϕ is a holomorphic bijection, then $(\phi^*)^{-1} = (\phi^{-1})^*$.

Proof. Recall that if ϕ is a holomorphic bijective map, then by the theory of complex analysis ϕ^{-1} is also holomorphic. Let $\phi_1 : V_1 \to V_2$ be a holomorphic bijection between open sets. Then $(\phi^{-1})^* : \Omega^k(V_1) \to \Omega^k(V_2)$ is such that $(\phi^{-1})^*(f(q_1)(dq_1)^k) = f(\phi^{-1}(q_2))(\phi^{-1})'(q_2)^k(dq_2)^k$. We will show that the inverse of ϕ^* is this map. Indeed, if we compute:

$$\begin{aligned}
\phi^*((\phi^{-1})^*(f(q_1)(dq_1)^k)) &= \phi^*(f(\phi^{-1}(q_2))(\phi^{-1})'(q_2)^k(dq_2)^k) \\
&= f(\phi^{-1}(\phi(q_1)))((\phi^{-1})'(\phi(q_1)))^k(\phi'(q_1))^k(dq_1)^k \\
&= f(q_1)((\phi \circ \phi{-}1)'(q_1))^k(dq_1)^k \\
&= f(q_1)(dq_1)^k
\end{aligned}$$

Similarly one can easily see that $(\phi^{-1})^* \circ \phi^* = I|_{\Omega^k(V_2)}$. □

Lemma 2.10.6. If $\pi : V_1 \to V_2$ is a surjection of open sets in $\mathbb{C}$, then π^* is an injection.

Proof. The map $\pi^* : \Omega^k(V_2) \to \Omega^k(V_1)$ is given by:

$$f(q_2)(dq_2)^k \mapsto f(\pi(q_1))\pi'(q_1)^k(dq_1)^k$$

Suppose that $f(\pi(q_1))\pi'(q_1)^k(dq_1)^k = g(\pi(q_1))\pi'(q_1)^k(dq_1)^k$. This implies that $f(\pi(q_1)) = g(\pi(q_1))$. Since π is a surjection, $f(q_1) = f(q_2)$. It follows that π^* is an injection.

□

Suppose X is a Riemann surface. Let $(U_i, V_i, \phi_i)_i$ be its coordinate charts with U_i open in X, the set V_i open in $\mathbb{H}$ for each i and $\phi_i : U_i \to V_i$ the local homeomorphism. We define meromorphic differentials on X as follows:

Definition 2.10.7. A meromorphic differential of degree k, denoted by ω on X is a collection of local meromorphic differentials $(\omega_j)_{j\in J} \in \prod_{j\in J} \Omega^k(V_j)$ which satisfy the following compatibility criteria: Let $V_{j,k} = \phi_j(U_j \cap U_k)$ and $V_{k,j} = \phi_k(U_j \cap U_k)$. Then the compatibility criteria says that when we consider the pullback of the transition map:

$$\phi_{k,j} = \phi_k \circ \phi_j^{-1} : V_{j,k} \to V_{k,j}$$

to pull back a differential form ω_k restricted to $V_{k,j}$, we get exactly the differential form on V_j restricted to $V_{j,k}$. In mathematical terms,

$$\phi_{k,j}^*(\omega_k|_{V_{k,j}}) = \omega_j|_{V_{j,k}}$$

We denote the set of differential forms of degree k on the modular curve $X(\Gamma)$ by $\Omega^k(X(\Gamma))$. This clearly forms a complex vector space. We next state a very important theorem which helps us to associate weight $2k$ modular forms with k-fold differential forms.

2.11 Viewing modular forms as differential forms

As seen previously, if f is a modular form of weight $2k$ with respect to Γ, then formally, the weight k differential form $f(q)(dq)^k$ is Γ invariant. Therefore, it is natural to guess that modular forms can be viewed as differential forms on the modular curve $X(\Gamma)$. We will make this explicit in this section.

Theorem 2.11.1. Let $k \in \mathbb{Z}$. Let Γ be a congruence subgroup of $SL_2(\mathbb{Z})$. Let $\Pi : \mathbb{H}^* \to X(\Gamma)$ be the projection map. Then the map:

$$\Omega^k(X(\Gamma)) \to \mathcal{A}_{2k}(\Gamma)$$

is an isomorphism of complex vector spaces. Under this isomorphism any meromorphic modular form $f \mapsto (\omega_j)_{j\in J}$ where $(\omega_j)_{j\in J}$ pulls back to the differential form $f(\tau)d\tau^k \in \Omega^k(\mathbb{H}^*)$ via Π^*.

We note that $\Pi^* : \Omega^k(X(\Gamma)) \to \Omega^k(\mathbb{H}^*)$ pulls back a differential form on $X(\Gamma)$ to a differential form on $\mathbb{H}^*$.

Proof. We start by mapping each meromorphic differential form ω on $X(\Gamma)$ to a meromorphic differential from $f(\tau)(d\tau)^k$ on $\mathbb{H}$ and we will see that the function f is indeed a meromorphic modular form. We define such an association as follows: Let $\{\Pi(U_j) \mid j \in J\}$ be a collection of coordinate neighborhoods on $X(\Gamma)$ where each $U_j \subset \mathbb{H}^*$ is a neighborhood of a point $\tau_j \in \mathbb{H}$ or of a cusp $\in \mathbb{Q} \cup \{\infty\}$. The corresponding maps are as follows:

$$\begin{tikzcd} & U_j \arrow[dl, "\Pi"'] \arrow[dr, "\psi_j"] & \\ \Pi(U_j) \arrow[rr, "\phi_j"'] & & V_j \end{tikzcd}$$

Let $\omega = (\omega_j)_{j\in J}$ be a meromorphic differential on $X(\Gamma)$. For each $j \in J$, let $U'_j = U_i \cap \mathbb{H}$, $V'_j = \psi(U'_j)$ and $\omega'_j = \omega_j|_{V'_j}$. Observe that for all points except the cusps $U_j = U'_j$. Since the above diagram commutes and ϕ_j is a homeomorphism from $\Pi(U_j)$ to V_j, the most natural way to define $\Pi^*(\omega)$ locally on $\mathbb{H}$ is via the map ψ^*. Therefore, for all $j \in J$ define:

$$\Pi^*(\omega)|_{U'_j} = \psi^*(\omega'_j)$$

Our first task is to check that $\Pi^*(\omega)$ gives a well defined global meromorphic differential from on all of $\mathbb{H}$. Consider the following commutative diagram, depicting the maps on the overlaps.

$$\begin{tikzcd} & U_j \cap U_k \arrow[dl, "\psi_j"'] \arrow[d, "\Pi"] \arrow[dr, "\psi_k"] & \\ V_{j,k} \arrow[r, "\phi_j^{-1}"'] & \Pi(U_j \cap U_k) \arrow[r, "\phi_k"'] & V_{k,j} \end{tikzcd}$$

Claim 2.11.2. Pulling back the differential form $\omega_k|_{V_{j,k}}$ via ψ_j^* and pulling back $\omega_j|_{V_{k,j}}$ via ψ_k^* gives us the same form on $U_j \cap U_k$.

By the commutativity of the above diagram, the map $\phi_{k,j} = \phi_k \circ \phi_j^{-1} : V_{j,k} \to V_{k,j}$ satisfies $\phi_{k,j} \circ \psi_j = \psi_k$ on $U_j \cap U_k$. Thus by Lemma 2.10.3, $\psi_k^* = \psi_j^* \circ \phi_{k,j}^*$. Writing $V'_{j,k} = \psi_j(U'_j \cap U'_k)$ and $V'_{k,j} = \psi_k(U'_j \cap U'_k)$, we have that:

$$\begin{aligned} \psi_k^*(\omega_k|_{V'_{k,j}}) &= \psi_j^*(\phi_{k,j}^*(\omega_k|_{V'_{k,j}})) \\ &= \psi_j^*(\omega_j|_{V'_{j,k}}) \end{aligned}$$

where the last equality is due to the compatibility criteria satisfied by the differential forms on $X(\Gamma)$.

Since the pullback agree on the overlaps, $\Pi^*(\omega)$ defines a global weight k differential form on $\mathbb{H}$ which we denote by $f(\tau)(d\tau)^k$. Next, we need to check that the function f is a meromorphic modular form. The function f is clearly meromorphic. We check that it satisfies the remaining two conditions in the definition of a modular form as well.

1. Modularity condition: For any $\gamma \in \Gamma$ such that $\gamma : \mathbb{H} \to \mathbb{H}$ and γ^* pulls back a differential form to $\mathbb{H}$. The condition $\Pi^*(\omega) = (\Pi \circ \gamma)^*(\omega) = \gamma^*(\Pi^*(\omega))$ gives that

$$\begin{aligned} f(\tau)(d\tau)^k &= \gamma^*(f(\tau)(d\tau)^k) \\ &= f(\gamma(\tau))(\gamma'(\tau))^k(d\tau)^k \\ &= \jmath(\gamma,\tau)^{-2k} f(\gamma(\tau))(d\tau)^k \\ &= (f[\gamma]_{2k})(\tau)(d\tau)^k \end{aligned}$$

It follows that $f(\tau) = (f[\gamma]_{2k})(\tau)$ giving us that f satisfies the modularity condition with respect to Γ.

2. Holomorphy at cusps: This requires us to show that $f[\alpha]_{2k}$ is meromorphic at ∞ for any $\alpha \in SL_2(\mathbb{Z})$. Recall that the local structure at a cusp s is given by the composition of the maps $\psi = \rho \circ \delta$ explained as follows.

$$\tau \mapsto z \mapsto q = e^{2\pi i z/h}$$

$$s \mapsto \infty \mapsto 0$$

Let δ be the transformation which takes s to ∞. Suppose $\alpha = \delta^{-1}$ such that $s = \alpha(\infty)$. Since ω is a meromorphic at the cusps of $X(\Gamma)$, locally it is of the form $g(q)(dq)^k$ where g is meromorphic at 0. Let $\Pi(U)$ denote the local chart around s and $V = \psi(U)$. The differential form $f(\tau)(d\tau)^k|_{U\setminus\{s\}}$ we constructed above is the pullback of $\omega|_{V\setminus\{0\}}$ via ψ^*.

This implies that,

$$\begin{aligned} f(\tau)(d\tau)^k &= \psi^*(g(q)(dq)^k) = g(\psi(\tau))(\psi'(\tau))^k(d\tau)^k \\ &= g(q)(\rho'(\delta(\tau))\delta'(\tau))^k(d\tau)^k \\ &= g(q)q^k(2\pi i/h)^k \jmath(\delta,\tau)^{-2k}(d\tau)^k \\ &= \widetilde{f}[\delta]_{2k}(\tau)(d\tau)^k \end{aligned}$$

where $\widetilde{f}(z) = g(q)q^k(2\pi i/h)^k; q = e^{2\pi iz/h}$. This implies that $f = \widetilde{f}[\delta]_{2k}$. Therefore $f[\alpha]_{2k} = \widetilde{f}[\alpha\delta]_{2k} = \widetilde{f}$. Since g is meromorphic at 0, thus $\widetilde{f}$, and so $f[\alpha]_{2k}$ is meromorphic at ∞.

We have established that every meromorphic differential form of degree k on $X(\Gamma)$ pulls back to a meromorphic differential form $f(\tau)(d\tau)^k$ where f is a meromorphic modular form. We now show that the above association is in fact surjective. That is, given an automorphic form $f \in \mathcal{A}_{2k}(\Gamma)$, we construct a meromorphic differential form ω such that pulling back ω under Π^* gives the differential form $f(\tau)(d\tau)^k$. By the following lemma it is enough to construct local differential forms $(\omega_j)_{j\in J}$ that pull back to restrictions of some meromorphic differential from $f(\tau)(d\tau)^k$ on $\mathbb{H}$.

Lemma 2.11.3. Given a collection $(\omega_j)_{j\in J} \in \prod_{j\in J} \Omega^k(V_j)$ of local meromorphic differentials on $X(\Gamma)$. As before, let $U'_j = U_j \cap \mathbb{H}$, $V'_j = \psi(U'_j)$ and $\omega'_j = \omega_j|_{V'_j}$. If $\omega_j|_{V'_j}$ pull back under ψ^*_j to the restriction of some meromorphic differential, say, $f(\tau)(d\tau)^k$ on $\mathbb{H}$, then $(\omega_j)_{j\in J}$ satisfy the compatibility criteria. In other words, $\omega = (\omega_j)_{j\in J}$ is a well defined differential form on $X(\Gamma)$.

Proof. We will essentially repeat the argument done in the first half of the proof backwards. Since ω'_j pulls back to a global differential form on $\mathbb{H}$, the pullback under ψ^* gives the same value on the intersection. This implies that $\psi^*_k(\omega_k|_{V'_{k,j}}) = \psi^*_j(\omega_j|_{V'_{j,k}})$. Again, using the commutative diagram above claim 2.11.2, we see that

$$\psi^*_j(\omega_j|_{V'_{j,k}}) = \psi^*_k(\omega_k|_{V'_{k,j}}) = \psi^*_j(\phi^*_{k,j}(\omega_k|_{V'_{k,j}}))$$

Since $\psi^*_j|_{U'_j\cap U'_k}$ is a surjection onto $V'_{j,k}$, Lemma 2.10.6 implies that ψ^*_j is an injection. Therefore, $\omega_j|_{V'_{j,k}} = \phi^*_{k,j}(\omega_k|_{V'_{k,j}})$. This is exactly the compatibility criteria. □

First we extend the weight k operator to matrices in $GL_2^+(\mathbb{C})$ by the following definition.

$$(f[\gamma]_k)(\tau) = (\det\gamma)^{k/2}\jmath(\gamma,\tau)^{-k}f(\gamma\tau)$$

for any $\gamma \in GL_2^+(\mathbb{C})$. We will later see that the extra factor of $(\det\gamma)^{k/2}$ is indeed useful in this context. However when we introduce Hecke operators, we will redefine the operator and work with a different factor. Recall the local structure for elliptic points and cusps shown in figure 2.2 and 2.3 respectively.

$$\text{For an elliptic point: } \tau \xmapsto{\delta_j = \begin{pmatrix} 1 & -\bar{\tau} \\ 1 & -\tau \end{pmatrix}} 0 \xmapsto{\rho_j(z)=z^h} 0$$

$$\text{For a cusp: } s \xmapsto{\delta_j} \infty \xmapsto{\rho_j(z)=e^{2\pi i z/h}} 0$$

In both cases δ_j is invertible. So, let $U_j' = U_j \cap \mathbb{H}$. Then $f(\tau)(d\tau)^k|_{U_j'}$ is the pullback $\delta_j^*(\lambda_j)$ where λ_j is the differential form obtained by the pull back of $f(\tau)(d\tau)^k$ by $(\delta_j^{-1})^*$. Let $\alpha = \delta_j^{-1}$, then

$$\begin{aligned} \lambda_j &= \alpha^*(f(\tau)(d\tau)^k) \\ &= f(\alpha(z)(d(\alpha z))^k \\ &= f(\alpha z)((\det \alpha)/j(\alpha, z)^2)^k (dz)^k \\ &= f[\alpha]_{2k}(z)(dz)^k \end{aligned}$$

Thus, $f[\alpha]_{2k}(z)(dz)^k$ pulls back to $f(\tau)(d\tau)^k$ under δ_j. Next, we need to find a differential form on V_j which in turn pulls back to $f[\alpha]_{2k}(z)(dz)^k$ under $\rho*$. Notice that the map ρ is not invertible so we cannot do the same trick as we did in the case of δ_j. We handle this part separately for an elliptic point and a cusp. Observe that the function $f[\alpha]_{2k}$ satisfies the modularity condition for the subgroup $\alpha^{-1}\Gamma\alpha = \delta_j\Gamma\delta_j^{-1}$. Thus, the differential form $\lambda_j = (f[\alpha]_{2k})(z)(dz)^k$ is $\delta_j\Gamma\delta_j^{-1}$ invariant. In the case where $\tau_j \in U_j$ is not a cusp, $\delta_j(\tau_j) = 0$ and $\{\pm I\}(\delta_j\Gamma\delta_j^{-1})_0/\{\pm I\}$ is cyclic of order h where h is the period of τ_j. The group is generated by the map $r_h : z \mapsto \mu_h z$ where $\mu_h = e^{2\pi i/h}$. Because λ_j is invariant under $\delta_j\Gamma\delta_j^{-1}$, we have that $r_h^*(\lambda_j) = \lambda_j$ which gives us that $(f[\alpha]_{2k})(z)(dz)^k = (f[\alpha]_{2k})(\mu_h z)\mu_h^k(dz)^k$ or equivalently $z^k(f[\alpha]_{2k})(z) = (\mu_h z)^k(f[\alpha]_{2k})(\mu_h z)$. This implies that $z^k(f[\alpha]_{2k})(z) = g_j(z^h)$ for some meromorphic function g_j, using the fact that if a meromorphic function t satisfies the property that $t(z) = t(\mu_h z)$, then $t = g(z^h)$ for some meromorphic function g. This can be seen by writing the Laurent series expansion for t and then plugging in the condition. Now define a local meromorphic differential ω_j on V_j in the q coordinate where $q = z^h$ in this case.

$$\omega_j = \frac{g_j(q)}{(hq)^k}(dq)^k \tag{2.11.4}$$

Claim 2.11.5. The pull back $\rho^*(\omega_j) = \lambda_j$ where ρ is the map $z \mapsto z^h$.

$$\begin{aligned}
\rho*(\omega_j) &= \frac{g_j(\rho(z))}{(h\rho(z))^k}\rho'(z)^k(dz)^k \\
&= \frac{g_j(q)}{h^k(z^{hk})}(hz^{h-1})^k(dq)^k \\
&= \frac{g_j(q)}{z^k}(dz)^k \\
&= \lambda_j
\end{aligned}$$

Therefore, $\delta^*(\rho^*(\omega_j)) = \delta^*(\lambda_j) = f(\tau)(d\tau)^k|_{U_j{}'}$.

In the case when U_j contains a cusp s_j, we know that $\delta_j(s_j) = \infty$. Recall that $(\delta_j\Gamma\delta_j^{-1})_\infty$ is the group of translations by mh where $m \in \mathbb{Z}$ and h is the width of the cusp h. This group is generated by $t_h = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix}$ which maps z to $z + h$. Arguing on similar lines as above, we see that λ_j is t_h invariant. The equaility $t_h^*(\lambda) = \lambda_j$ implies that $f[\alpha]_{2k}(z)(dz)^k = f[\alpha]_{2k}(z + h)(dz)^k$. Therefore, $f[\alpha]_{2k}(z)$ is of the form $g_j(e^{2\pi iz/h})$ for some meromorphic function g_j. Set

$$\omega_j = \frac{g_j(q)}{2\pi iq/h)^k}(dq)^k \tag{2.11.6}$$

where $q = e^{2\pi iz/h}$. Our final task is to check that ω_j is pulled back via ρ^* to $f[\alpha]_{2k}(z)$, where the map ρ_j in this case is $z \mapsto e^{2\pi iz/h}$. This is easily computed in the following calculation.

$$\begin{aligned}
\rho*(\omega_j) &= \frac{g_j(\rho(z))}{(2\pi i\rho(z)/h)^k}\rho'(z)^k(dz)^k \\
&= \frac{g_j(q)}{(2\pi iq/h)^k}q^k(2\pi i/h)^k(dq)^k \\
&= g_j(q) \\
&= f[\alpha]_{2k}(z)
\end{aligned}$$

It is clear that the above map between the space of meromorphic modular forms and meromorphic differential forms is injective and $\mathbb{C}$ linear. It follows that they are isomorphic as vector spaces. □

In summary, we have established a remarkable correspondence between the set of meromorphic modular forms of weight $2k$ and the set of weight k differential forms on the modular curve $X(\Gamma)$. We will use this isomorphism of vector spaces to find the dimension of the space $\mathcal{M}_{2k}(\Gamma)$.

Let $C(X(\Gamma))$ denote the field of meromorphic functions on $X(\Gamma)$. For any fixed non zero element f in $\mathcal{A}_k(\Gamma)$, it is easy to deduce that $\mathcal{A}_k(\Gamma) = \{f_0 f \mid f_0 \in C(X(\Gamma))\}$. This is because if $g \in \mathcal{A}_k(\Gamma)$ is arbitrary, then g/f is a weight 0 meromorphic modular form and so a well defined function on $X(\Gamma)$. From the isomorphism in Theorem 2.11.1, we can conclude that $\Omega^{k/2}(X(\Gamma)) = C(X(\Gamma))\omega_0$ for some non zero differential form ω_0.

2.11.1 Order of a modular form and its corresponding differential form

We know that the local homeomorphism around the elliptic points "looks like" $z \mapsto z^e$, similarly for the cusps the local map looks like $z \mapsto e^{2\pi i z/h}$. Therefore when we look at the corresponding differential form, the order of the differential form at the cusps and the elliptic points will not be the same as the order of the modular form at those points. This can be seen directly from the construction of the corresponding differential form for a modular form we did in the proof of Theorem 2.11.1.

Let τ_j be a non cusp and ω_j as given in (2.11.4) be the corresponding differential from on V, then the order of vanishing at 0 of ω_j is given by

$$ord_0(\omega_j) = ord_0(g_j) - k = ord_{\tau_j}(f)/h + k(1 - 1/h) \tag{2.11.7}$$

using the fact that $g_j(z^h) = z^n f[\alpha]_{2k}$ to compare orders at the last step. Performing a sanity check for $h = 1$ case, we see that for non elliptic points the order of f and its corresponding differential form match.

On similarly lines, take s_j to be a cusp and its corresponding differential form ω_j in (2.11.6). Using that $g_j(e^{2\pi i z/h}) = f[\alpha]_{2k}(z)$, we get

$$ord_0(\omega_j) = ord_0(g_j) - k = ord_{s_j}(f) - k \tag{2.11.8}$$

Remark 2.11.9. Since $\Pi(U)$ is homeomorphic to V, we can consider the differential form to be on $\Pi(U)$ so that $ord_0(\omega_j) = ord_{\Pi(\tau_j)}(\omega_j)$.

For even weights, putting together equations (2.11.7) and (2.11.8), along with the formula in 2.9.6, we arrive at the following theorem:

Theorem 2.11.10. Let $f \in \mathcal{A}_{2k}(\Gamma)$ and let ω be the corresponding k-fold differential form on $X(\Gamma)$. Let Π be the projection map as defined previously. Suppose $\tau \in \mathbb{H}^*$.

1. If τ is an elliptic point with period e, then $ord_{\Pi(\tau)}(f) = ord_{\Pi(\tau)}(\omega) + k(1 - 1/e)$
2. If τ is a cusp point with width h, then $ord_{\Pi(\tau)}(f) = ord_{\Pi(\tau)}(\omega) + k$
3. If τ is any other point then $ord_{\Pi(\tau)}(f) = ord_{\Pi(\tau)}(\omega)$

2.12 Divisors and the Riemann Roch Theorem

We next introduce the Riemann Roch Theorem. The aim of this section is to formally develop the necessary definitions in order to make sense of the Riemann Roch Theorem. We first state some facts regarding the zeroes and the poles of functions on a compact Riemann surface X.

Fact 2.12.1. Any holomorphic function on X is constant.

Fact 2.12.2. A meromorphic function f on X has finitely many poles and zeroes.

Fact 2.12.3. Any meromorphic function f on X has the same number of poles as the zeroes counting with multiplicities.

Definition 2.12.4 (Divisor). Let X be a compact Riemann surface. A divisor D on X is a finite sum over the points of X formally written as $\sum_i n_i \cdot P_i$ where $n_i \in \mathbb{Z}$ and P_i are points of X.

The divisor $D = \sum n_i \cdot P_i$ is called positive if $n_i \geq 0$ for all i. In such a case we write $D \geq 0$.

Definition 2.12.5 (Order of f at a point). Let f be a meromorphic function on X. For any point $P \in X$, order of f at P is the order of the pole or zero at P. It is denoted by $\mathrm{ord}_P(f)$. It is positive if P is a zero of f, negative if P is a pole or zero otherwise.

Definition 2.12.6 (Divisor of f). Let f be a non zero meromorphic function on X. The divisor of f is the formal sum $\displaystyle\sum_P \mathrm{ord}_P(f) \cdot P$. We denote it by $\mathrm{div}(f)$.

We note that the sum makes sense because of fact 2.12.2.

Remark 2.12.7. We will see in section 2.13.5 that the divisor associated to a holomorphic modular form is not integral i.e., its coefficients are not integral. In such a case we take the greatest integer of the coefficients. This makes sense because $\mathrm{div}(f) \geq 0$ if and only if $\lfloor \mathrm{div}(f) \rfloor \geq 0$.

Definition 2.12.8 (Degree of a divisor). For any divisor $D = \sum_i n_i \cdot P_i$, the degree of D is the sum $\sum_i n_i$.

The fact in 2.12.3 implies that the degree of a meromorphic function is 0.

We next associate a divisor to a differential form. Let $P \in X$ and U be a coordinate neighborhood around P. Let z denote the local coordinate on U. The differential form on U is of the form $f(z)dz$. Define the order of ω at P by $\mathrm{ord}_P(\omega) = \mathrm{ord}_P(f)$. It now makes sense to write $\mathrm{div}(\omega) = \sum_P \mathrm{ord}_P(\omega) \cdot P$. It is easy to see that $\mathrm{div}(f\omega) = \mathrm{div}(f) + \mathrm{div}(\omega)$.

Let $\mathcal{M}(X)$ denote the set of meromorphic functions on X. For any divisor D on X, we define the set

$$L(D) = \{f \in \mathcal{M}(X) \mid \mathrm{div}(f) + D \geq 0\} \cup \{0\}.$$

It is easy to see that this is a vector space. In fact $L(D)$ is finite dimensional. Its dimension is denoted by $l(D)$.

Definition 2.12.9 (Canonical divisor). For any compact Riemann surface X, a canonical divisor is a divisor $\mathrm{div}(\omega)$ where ω is a one form on X.

The Riemann Roch Theorem helps us to calculate the dimension of the space $L(D)$ for any divisor D. In this way, given a specific number of zeroes and poles of a function say $\widetilde{f}$, we can define a divisor D on X with appropriate coefficients so that the space $L(D)$ contains functions f vanishing with high enough order so as to make the product $f\widetilde{f}$ holomorphic. The Riemann Roch Theorem helps us calculate the dimension of the space of such functions. The following is the statement of the theorem:

Theorem 2.12.10 (Riemann-Roch). Let X be a compact Riemann surface. Let g denote the genus of the surface. Then for any divisor D on X and K a canonical divisor on X, we have that

$$l(D) = deg(D) + 1 - g + l(K - D)$$

2.13 Consequences of the Riemann Roch Theorem

In this section we prove three immediate corollaries of the Riemann Roch Theorem which will help us prove the dimension formula.

Corollary 2.13.1. A canonical divisor K has degree $2g - 2$ and $l(K) = g$.

Proof. If we take $D = 0$, the space $L(D)$ is the space of holomorphic functions. By fact 2.12.1, we have that $L(D)$ consists of constant functions. Thus $l(0) = 1$. Substituting everything in Theorem 2.12.10 we get that $l(K) = g$.

When we put $D = K$, then $l(K) = \deg(K) + 1 - g + l(0)$. Putting $l(K) = 1$, we have that $\deg(K) = 2g - 2$. □

Corollary 2.13.2. The degree of a k-fold differential form is $k(2g - 2)$.

Proof. Suppose that $\lambda \in \Omega^1(X(\Gamma))$ is a non zero 1-form. We know that its degree is $2g - 2$. Notice that $\lambda^k \in \Omega^k(X(\Gamma))$ and its divisor has degree $k(2g - 2)$. This implies that $\Omega^k(X(\Gamma)) = C(X(\Gamma))\lambda^k$. Since for all $f \in C(X(\Gamma))$, $\deg(\mathrm{div}(f)) = 0$, it follows that every non zero differential form $\omega \in \Omega^k(X(\Gamma))$ has degree $k(2g - 2)$. □

Corollary 2.13.3. If $deg(D) > 2g - 2$, then $l(D) = deg(D) + 1 - g$

Proof. If $\deg(D) > 2g - 2$, then $\deg(K - D) < 0$. This implies that for any non zero $f \in \mathcal{M}(X)$, $\deg(f + K - D) < 0$. Therefore, $\operatorname{div}(f + K - D)$ will never be a positive divisor. Thus, $L(K - D) = 0$ so that $l(K - D) = 0$ and the result follows. □

Corollary 2.13.4. If $deg(D) < 0$, then $l(D) = 0$.

Proof. The proof of this corollary is along similar lines to the proof of Corollary 2.13.3. □

Keeping the above corollaries in mind, we finally move on to find out the dimension of the space $\mathcal{M}_{2k}(\Gamma)$.

2.13.1 The dimension formula for even weights

All throughout this section, f is a modular form of weight $2k$. The aim of this section is to prove the following theorem:

Theorem 2.13.5. Suppose the modular curve $X(\Gamma)$ has a genus g for some congruence subgroup Γ. Let ϵ_∞ denote the number of cusps of $X(\Gamma)$. Let ϵ_2 and ϵ_3 be the number of elliptic points of period 2 and 3 respectively on $X(\Gamma)$. Then the dimension of $\mathcal{M}_{2k}(\Gamma)$ is given by:

$$\dim(\mathcal{M}_{2k}(\Gamma)) = \begin{cases} 0 & \text{if } k < 0 \\ 1 & \text{if } k = 0 \\ (2k-1)(g-1) + \epsilon_\infty k + \lfloor k/2 \rfloor \epsilon_2 + \lfloor 2k/3 \rfloor \epsilon_3 & \text{if } k > 0 \end{cases}$$

Proof. We will first characterize the image of the subspace $\mathcal{M}_{2k}(\Gamma) \subset \mathcal{A}_{2k}(\Gamma)$ under the isomorphism in Theorem 2.11.1. Suppose $f \in \mathcal{M}_{2k}(\Gamma)$. Let ω be its corresponding k-fold differential form on $X(\Gamma)$. Since f is holomorphic, $\operatorname{ord}_\tau(f) \geq 0$ for all $\tau \in \mathbb{H}^*$. Let ω_0 be a nonzero k-fold differential form so that $\omega = h\omega_0$ for some function $h \in C(X(\Gamma))$. Thus $\operatorname{ord}_{\Pi(\tau)}(\omega) = \operatorname{ord}_{\Pi(\tau)}(h\omega_0)$. From the previous section relating the order of zeroes and poles of ω with f we get the following inequalities.

When τ is an elliptic point: $\operatorname{ord}_{\Pi(\tau)}(h) + \operatorname{ord}_{\Pi(\tau)}(\omega_0) + k(1 - 1/e) \geq 0$

When τ is a cusp: $\operatorname{ord}_{\Pi(\tau)}(h) + \operatorname{ord}_{\Pi(\tau)}(\omega_0) + k \geq 0$

Otherwise: $\operatorname{ord}_{\Pi(\tau)}(h) + \operatorname{ord}_{\Pi(\tau)}(\omega_0) \geq 0$

The above three equations characterize the subspace $\mathcal{M}_{2k}(\Gamma)$. Combining the three equations and rewriting them in the language of divisors we arrive at the following expression.

$$\operatorname{div}(h) + \operatorname{div}(\omega_0) + \sum_{\tau \text{ cusp}} k \cdot \tau + \sum_{\tau_i \text{ elliptic}} \lfloor k(1 - 1/e_i) \rfloor \cdot \tau_i \geq 0 \tag{2.13.6}$$

The symbol e_i refers to the period of the elliptic point τ_i. Let D be the divisor given by

$$D = \operatorname{div}(\omega_0) + \sum_{\tau \text{ cusp}} k \cdot \tau + \sum_{\tau_i \text{ elliptic}} \lfloor k(1 - 1/e_i) \rfloor \cdot \tau_i \tag{2.13.7}$$

Then the above condition characterizing the space of holomorphic modular forms reduces to the following:

$$\operatorname{div}(h) + D \geq 0$$

In other words, because of the isomorphism in Theorem 2.11.1 we have a one to one correspondence between the sets given below.

$$\{f \mid f \in \mathcal{M}_{2k}(\Gamma)\} \leftrightarrow \{h \in C(X(\Gamma)) \mid \operatorname{div}(h + D) \geq 0\}.$$

We therefore deduce that $L(D) \cong \mathcal{M}_{2k}(\Gamma)$. We are now ready to apply the Riemann Roch Theorem to find the dimension $l(D)$ which will give us the dimension for the space $\mathcal{M}_{2k}(\Gamma)$. Since ω_0 is a k-fold differential form, its degree is $k(2g-2)$ by Corollary 2.13.2. This implies that $\deg(D) = k(2g-2) + \epsilon_\infty k + \lfloor k/2 \rfloor \epsilon_2 + \lfloor 2k/3 \rfloor \epsilon_3$. Notice that when $k > 0, \deg(D) > (2g-2)$. This implies that $l(D) = \deg(D) - g + 1$ by Corollary 2.13.3. It follows that for $k > 0$,

$$\dim(\mathcal{M}_{2k}(\Gamma)) = l(D) = (2k-1)(g-1) + \epsilon_\infty k + \lfloor k/2 \rfloor \epsilon_2 + \lfloor 2k/3 \rfloor \epsilon_3.$$

When $k < 0, \deg(D) < 0$. By Corollary 2.13.4 $l(D) = 0$. In the case when $k = 0$, modular forms are well defined holomorphic functions on the modular curve. By fact 2.12.1, they are constant functions. Therefore $\dim(\mathcal{M}_0(\Gamma)) = 1$. □

Following a similar argument we compute the dimension of the subspace of cusp forms $\mathcal{S}_k(\Gamma)$. For any cusp form $f \in \mathcal{S}_{2k}(\Gamma)$, $\operatorname{ord}_s(f) \geq 1$ where s represents a cusp in $X(\Gamma)$. As argued before, suppose $\omega = h\omega_0$ is the corresponding differential form. In this case $\operatorname{div}(f) - \sum_{x_i \text{ cusp}} x_i \geq 0$ so that there exists a one to one correspondence between the sets:

$$\{f \mid f \in \mathcal{S}_{2k}(\Gamma)\} \leftrightarrow \left\{h \in C(X(\Gamma)) \,\middle|\, \operatorname{div}\left(h + D - \sum_{x_i \text{ cusp}} x_i\right) \geq 0\right\}.$$

where D is the divisor as given (2.13.7). It follows that the dimension of $\mathcal{S}_k(\Gamma) = l\left(D - \sum_{x_i \text{ cusp}} x_i\right)$. Observe that $\deg\left(D - \sum_{x_i \text{ cusp}} x_i\right) = \deg(D) - \epsilon_\infty > 2g - 2$ when $k > 1$. By the Riemann Roch Theorem, $\dim(\mathcal{S}_{2k}(\Gamma)) = \dim(\mathcal{M}_{2k}(\Gamma)) - \epsilon_\infty$. $S_0(\Gamma)$ consists of constant functions. We handle the case of $\mathcal{S}_2$ below separately.

Lemma 2.13.8. The space of cusp forms of weight 2, $\mathcal{S}_2(\Gamma)$ is isomorphic to the holomorphic differentials of degree 1 on $X(\Gamma)$, denoted by $\Omega^1_{hol}(X(\Gamma))$.

Proof. Let $f \in \mathcal{S}_2(\Gamma)$ and ω be the associated 1 form on $X(\Gamma)$. Using Theorem 2.11.10 for the $k = 1$ case, we characterize the image of $\mathcal{S}_2(\Gamma)$ under the isomorphism map between $\mathcal{M}_2(\Gamma)$ and $\Omega^1(X(\Gamma))$. Observing the fact that $\mathrm{ord}_\tau(f) \geq 1$ at cusps, while it is greater than or equal to 0 at other points, we see that, $\mathrm{ord}_{\Pi(\tau)}(\omega)$ being integral is greater than or equal to 0 at all points of $X(\Gamma)$. Thus, $\omega \in \Omega^1_{hol}(X(\Gamma))$. □

Next, recall that $\Omega^1(X(\Gamma)) = C(X(\Gamma))\lambda$ where λ is a non zero one form on $X(\Gamma)$. Define a map as follows:

$$\Omega^1(X(\Gamma)) \to C(X(\Gamma))$$

$$f_0\lambda \mapsto f_0$$

This is clearly a well defined, vector space isomorphism. Any divisor $f_0\lambda \in \Omega^1_{hol}(X(\Gamma))$ if and only if $\mathrm{div}(f_0\lambda) \geq 0$ if and only if $\mathrm{div}(f_0)+\mathrm{div}(\lambda) \geq 0$ if and only if $f_0 \in L(\lambda)$. Thus, under the above isomorphism, $\Omega^1_{hol}(X(\Gamma))$ is mapped to $L(\lambda)$. So dimension of $\mathcal{S}_2(\Gamma) = l(\lambda) = g$ by Corollary 2.13.1.

To summarize the above discussion:

$$\dim(\mathcal{S}_{2k}(\Gamma)) = \begin{cases} 1 & \text{if } k = 0 \\ g & \text{if } k = 1 \\ (2k-1)(g-1) + \epsilon_\infty(k-1) + \lfloor k/2 \rfloor \epsilon_2 + \lfloor 2k/3 \rfloor \epsilon_3 & \text{if } k > 1 \end{cases}$$

2.14 Dimension formula for odd weights

When $-I \notin \Gamma$, then odd weight modular forms may exist. The isomorphism in Theorem 2.11.1 made sense for even weights and not for odd weight forms. To take this into account we will slightly modify the argument as done for even weights. Throughout this section, we assume that f is a modular form of weight k where k is an odd integer. Let $\omega \in \Omega^k(X(\Gamma))$ be the differential form which pulls back to the form $f(\tau)^2(d\tau)^k$. Recall that $\mathcal{A}_k(\Gamma) = \{f_0 f \mid f_0 \in C(X(\Gamma))\}$. The space $\mathcal{M}_k(\Gamma)$ is identified to $L(\mathrm{div}(f))$ since

$$\begin{aligned}\mathcal{M}_k(\Gamma) &= \{f_0 \mid \mathrm{div}(f_0 f) \geq 0\} \\ &= \{f_0 \mid \mathrm{div}(f_0) + \mathrm{div}(f)) \geq 0\} \\ &= L(\mathrm{div}(f))\end{aligned}$$

We need to find $l(\mathrm{div}(f))$. Essentially we will follow a similar argument as done for even weights but with slight modifications. Observe that $2\mathrm{ord}_\tau(f) = \mathrm{ord}_\tau(\omega)$. As before, Before writing the divisor of f in terms of the divisor of ω, we make the following observations for the case when $-I \notin \Gamma$.

1. Along with the regular cusps, $X(\Gamma)$ might have irregular cusps as well.
2. The modular curve does not contain any elliptic points of period 2. This is seen as follows: Suppose otherwise. Then any elliptic point of period two of the form γi is fixed by an element $\gamma S^j \gamma^{-1}$ with $j \in \{1, 3\}$, $\gamma \in SL_2(\mathbb{Z})$ and S as in 1.0.3. But then either $(\gamma S \gamma^{-1})^2 = -I \in \Gamma$ or $(\gamma S^3 \gamma^{-1})^2 = -I \in \Gamma$. This is not possible when $-I \notin \Gamma$.

Keeping in mind Theorem 2.11.10 and writing everything in terms of divisors, we see that

$$\mathrm{div}(f) = 1/2\mathrm{div}(\omega) + \sum_{\text{reg cusp}} k/2 \cdot \tau + \sum_{\text{irreg cusp}} k/2 \cdot \tau + \sum_{\tau \text{ elliptic}} k/3 \cdot \tau \qquad (2.14.1)$$

Since the coefficients of the divisor are not integers, we need to study $\lfloor \mathrm{div}(f) \rfloor$. This is done case by case depending on whether τ is a cusp, elliptic point or neither of the two.

1. Suppose that $\Pi(\tau)$ is neither a cusp nor an elliptic point. In this case $\mathrm{ord}_{\Pi(\tau)}(f)$ is integral and so $1/2(\mathrm{ord}_{\Pi(\tau)}(\omega))$ is integral.
2. When $\Pi(\tau)$ is an elliptic point of order 3 we write $\mathrm{ord}_{\Pi(\tau)}(f) = m + j/3$ where $m \in \mathbb{Z}$ and $0 \leq j < 3$.

$$\begin{aligned}1/2(\mathrm{ord}_{\Pi(\tau)}(\omega)) &= 1/2\Big(2(\mathrm{ord}_{\Pi(\tau)}(f)) - 2/3k\Big) \\ &= m + \frac{j-k}{3}\end{aligned}$$

Since $\mathrm{ord}_{\Pi(\tau)}(\omega)$ is integral, $j \equiv k \bmod 3$. Therefore $1/2(\mathrm{ord}_{\Pi(\tau)}(\omega))$ is integral giving us that

$$\lfloor \mathrm{ord}_{\Pi(\tau)}(f) \rfloor = 1/2(\mathrm{ord}_{\Pi(\tau)}(\omega)) + \left\lfloor \frac{k}{3} \right\rfloor.$$

3. When $\Pi(\tau)$ is an regular cusp, we see that

$$\begin{aligned} 1/2(\mathrm{ord}_{\Pi(\tau)}(\omega)) &= 1/2\Big(2(\mathrm{ord}_{\Pi(\tau)}(f)) - k\Big) \\ &= \mathrm{ord}_{\Pi(\tau)}(f) - k/2 \end{aligned}$$

   Since $\mathrm{ord}_{\Pi(\tau)}(f)$ is integral and k is odd, $1/2(\mathrm{ord}_{\Pi(\tau)}(\omega))$ is half integral and so $\lfloor \mathrm{ord}_{\Pi(\tau)}(f) \rfloor = 1/2(\mathrm{ord}_{\Pi(\tau)}(\omega)) + k/2)$, a sum of half integers.

4. If $\Pi(\tau)$ is an irregular cusp then we assume $\mathrm{ord}_{\Pi(\tau)}(f) = m/2$ for some odd integer m so that

$$\begin{aligned} 1/2(\mathrm{ord}_{\Pi(\tau)}(\omega)) &= \mathrm{ord}_{\Pi(\tau)}(f) - k/2 \\ &= \frac{m-k}{2} \end{aligned}$$

   Since k and m both are odd, $1/2(\mathrm{ord}_{\Pi(\tau)}(\omega))$ is integral giving us that $\lfloor \mathrm{ord}_{\Pi(\tau)}(f) \rfloor = 1/2\mathrm{ord}_{\Pi(\tau)}(\omega) + (k-1)/2$, a sum of half integers.

Let $\epsilon_\infty^{\mathrm{reg}}$ and $\epsilon_\infty^{\mathrm{irreg}}$ denote the number of regular and irregular cusps respectively. Putting together the above discussion, we conclude that

$$\begin{aligned} \deg(\lfloor \mathrm{div}(f) \rfloor) &= k(g-1) + \lfloor k/3 \rfloor \epsilon_3 + k/2\epsilon_\infty^{\mathrm{reg}} + (k-1)/2\epsilon_\infty^{\mathrm{irreg}} \\ &\geq (k-2)(g-1+\epsilon_3/3+\epsilon_\infty/2) + 2g - 2 \\ &> 2g-2 \text{ when } k \geq 3 \end{aligned}$$

We use a small observation that $\lfloor k/3 \rfloor \leq (k-2)/3$ in the second inequality. Applying Corollary 2.13.3 we arrive at the following result.

$$l(\lfloor \mathrm{div}(f) \rfloor) = (k-1)(g-1) + \lfloor k/3 \rfloor \epsilon_3 + k/2\epsilon_\infty^{\mathrm{reg}} + (k-1)/2\epsilon_\infty^{\mathrm{irreg}}$$

To study cusp forms of odd weight, we need to consider the fact that at regular cusps $\mathrm{ord}_\tau(f_0 f) \geq 1$ and at irregular cusps $\mathrm{ord}_\tau(f_0 f) \geq 1/2$. This amounts to studying the divisor div $\left\lfloor \left(f - \sum_{\text{reg cusp}} \tau + \sum_{\text{irreg cusp}} \tau \right) \right\rfloor$. Let us call this divisor D. Calculations similar to those done in case of even weight help us conclude the following result.

$$l(D) = (k-1)(g-1) + \lfloor k/3 \rfloor \epsilon_3 + (k-2)/2\epsilon_\infty^{\mathrm{reg}} + (k-1)/2\epsilon_\infty^{\mathrm{irreg}} \tag{2.14.2}$$

We now state the final theorem of this section, summarizing our discussion above.

Theorem 2.14.3. Let k be an odd integer. Suppose Γ is a congruence subgroup of $SL_2(\mathbb{Z})$. Let g be the genus of $X(\Gamma)$, ϵ_3 be the number of elliptic points of period 3, $\epsilon_\infty^{\mathrm{reg}}$ and $\epsilon_\infty^{\mathrm{irreg}}$ be regular and irregular cusps respectively. When $k < 0$, or if $-I \in \Gamma$, then

$\mathcal{M}_k(\Gamma) = \mathcal{S}_k(\Gamma) = \{0\}$. If $-I \notin \Gamma$ and $k > 0$, then the dimension of $\mathcal{M}_k(\Gamma)$ and $\mathcal{S}_k(\Gamma)$ is given by the following expression:

$$\dim(\mathcal{M}_k(\Gamma)) = (k-1)(g-1) + (k/2)\epsilon_\infty^{\text{reg}} + ((k-1)/2)\epsilon_\infty^{\text{irreg}} + \lfloor k/3 \rfloor \epsilon_3 \tag{2.14.4}$$

$$\dim(\mathcal{S}_k(\Gamma)) = (k-1)(g-1) + ((k-2)/2)\epsilon_\infty^{\text{reg}} + ((k-1)/2)\epsilon_\infty^{\text{irreg}} + \lfloor k/3 \rfloor \epsilon_3 \tag{2.14.5}$$

With this result, we end this chapter and conclude the dimension theory. We move on to studying modular forms via Hecke theory.

Chapter 3

The theory of Hecke operators

3.1 Modular curves and the corresponding Moduli space

We begin the chapter by establishing a very interesting correspondence between the points of modular curves and the isomorphism classes of complex elliptic curves.

The theory of complex tori deals with viewing an elliptic curve as a complex torus. In the process of establishing a one-one correspondence between the two mathematical objects, we work with doubly periodic functions, more specifically, the Weierstrass $\wp$ function. For a brief introduction, one can refer to the Appendix. In the view of this bijection, we will essentially consider the isomorphism class of elliptic curves as the class of complex tori up to isomorphism.

Let Λ and Λ' be lattices in $\mathbb{C}$. We know that two complex tori $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ are holomorphically isomorphic if and only if there exists a complex number m such that $\Lambda = m\Lambda'$. Let Λ_τ denote the lattice generated by 1 and τ. We aim to show that if τ and τ' are points in $\mathbb{H}$ such that $\tau = \gamma\tau'$ with $\gamma \in SL_2(\mathbb{Z})$, then Λ_τ is isomorphic to $\Lambda_{\tau'}$. More generally, we will show this correspondence for the congruence subgroups $\Gamma_0(N)$ and $\Gamma_1(N)$. However, to work in a general setting we will need to attach some additional information with the isomorphism classes of elliptic curves. We begin by defining relevant *torsion data* for the congruence subgroups.

Definition 3.1.1 (Enhanced elliptic curves for congruence subgroups). Let E denote an elliptic curve, C a cyclic subgroup of order N of E and Q a point of order N in E.

1. The pair (E, C) is called an enhanced elliptic curve for the subgroup $\Gamma_0(N)$. Two pairs (E, C) and (E', C') are equivalent if there exists some isomorphism from E to E' that takes C to C'. Denote the set of equivalence classes by $S_0(N)$. An element of $S_0(N)$ is written $[E, C]$.

2. An enhanced elliptic curve for the subgroup $\Gamma_1(N)$ is given by the pair (E, Q). Two pairs (E, Q) and (E', Q') are equivalent if there exists some isomorphism from E to E' that takes Q to Q'. The set of equivalence classes is written as $S_1(N)$. An element of $S_1(N)$ is denoted by $[E, Q]$.

Each $S_0(N)$ and $S_1(N)$ is known as the *moduli space* of isomorphism classes of elliptic curves with N torsion data.

The theorem below describes the one-to-one correspondence between points of moduli space and the modular curve mentioned in the beginning. We denote the modular curve $\mathbb{H}/\Gamma_1(N)$ as $Y_1(N)$ and $\mathbb{H}/\Gamma_0(N)$ as $Y_0(N)$.

Theorem 3.1.2. Let N be a positive integer. Suppose E_τ represents an elliptic curve biholomorphic to the complex torus $\mathbb{C}/\Lambda_\tau$.

1. The moduli space for $\Gamma_1(N)$ is given by the set

$$S_1(N) = \{[E_\tau, 1/N + \Lambda_\tau] \mid \tau \in \mathbb{H}\}$$

Two points $[E_\tau, 1/N + \Lambda_\tau] = [E_{\tau'}, 1/N + \Lambda_{\tau'}]$ in $S_1(N)$ if and only if $\Gamma_1(N)\tau = \Gamma_1(N)\tau'$ in $Y_1(N)$. Thus there is a bijection

$$\psi_1 : S_1(N) \to Y_1(N) \ ; \ [\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau] \mapsto \Gamma_1(N)\tau$$

2. The moduli space for $\Gamma_0(N)$ is given by the set

$$S_0(N) = \{[E_\tau, \langle 1/N + \Lambda_\tau \rangle] \mid \tau \in \mathbb{H}\}$$

Two points $[E_\tau, \langle 1/N + \Lambda_\tau \rangle] = [E_{\tau'}, \langle 1/N + \Lambda_{\tau'} \rangle]$ in $S_0(N)$ if and only if $\Gamma_0(N)\tau = \Gamma_0(N)\tau'$ in $Y_0(N)$. Thus there is a bijection

$$\psi_1 : S_0(N) \to Y_0(N) \ ; \ [\mathbb{C}/\Lambda_\tau, \langle 1/N + \Lambda_\tau \rangle] \mapsto \Gamma_0(N)\tau$$

Proof. We will prove the theorem for the subgroup $\Gamma_1(N)$ and second part follows by a similar idea. We first wish to show that for any class $[E, Q]$, we can in fact specifically choose the representative and write it as $[\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau]$. Consider $E = \mathbb{C}/\Lambda_{\tau'}$ for some $\tau' \in \mathbb{H}$ and $Q = \frac{(c\tau'+d)}{N} + \Lambda_{\tau'}$. We wish to show that there exists some element τ in $\mathbb{H}$ and an isomorphism between the tori $\mathbb{C}/\Lambda_\tau$ and $\mathbb{C}/\Lambda_{\tau'}$ which maps Q to $1/N + \Lambda_\tau$. Since Q has order N, $(c, d, N) = 1$ so that $ad - bc - kN = 1$ for some $a, b, k \in \mathbb{Z}$. This implies that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}/N\mathbb{Z})$. Take a lift γ of this matrix in $SL_2(\mathbb{Z})$. This will not change the order N point Q because the entries are changed up to a multiple of N. Let $\tau = \gamma\tau' = \frac{a\tau'+b}{c\tau'+d}$. Let $m = c\tau' + d$ and so $m\tau = a\tau' + b$. Observe that:

$$m\Lambda_\tau = m(\tau\mathbb{Z} \oplus \mathbb{Z}) = (a\tau' + b)\mathbb{Z} \oplus (c\tau' + d)\mathbb{Z} = \tau'\mathbb{Z} \oplus \mathbb{Z}.$$

It follows that $\mathbb{C}/\Lambda_{\tau'} \cong \mathbb{C}/\Lambda_\tau$. Under this isomorphism of complex tori, the point $\frac{1}{N} + \Lambda_\tau$ maps to $\frac{c\tau'+d}{N} + \Lambda'_\tau = Q$. Finally this allows to write $[E, Q] = [E_\tau, 1/N + \Lambda_\tau]$. Our next task is to show that if two points are equivalent in $S_1(N)$, then the corresponding points in the modular curve $Y_1(N)$ are equivalent.

First suppose that $[E_\tau, 1/N + \Lambda_\tau] = [E_{\tau'}, 1/N + \Lambda_{\tau'}]$ with τ and $\tau' \in \mathbb{H}$. Then $\mathbb{C}/\Lambda_{\tau'} \cong \mathbb{C}/\Lambda_\tau$. This means that for some $m \in \mathbb{C}$, $m\Lambda_\tau = \Lambda_{\tau'}$ such that, under this map

$$(z + \Lambda_\tau) \mapsto mz + \Lambda_{\tau'} \;;\; \frac{1}{N} + \Lambda_\tau \mapsto \frac{1}{N} + \Lambda_{\tau'}$$

From Lemma A.1.3 we conclude that:

$$\begin{pmatrix} m\tau \\ m \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \tau' \\ 1 \end{pmatrix} \text{ for some } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \tag{3.1.3}$$

This implies that $m = c\tau' + d$ and so $\frac{c\tau'+d}{N} + \Lambda_{\tau'} = \frac{1}{N} + \Lambda_{\tau'}$. Therefore, $(c, d) \equiv (0, 1) \mod N$. It follows that $\gamma \in \Gamma_1(N)$. The expression in 3.1.3 shows that $\tau = \gamma\tau'$.

Conversely, suppose that τ and $\tau' \in \mathbb{H}$ such that $\tau = \gamma\tau'$ for some $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1(N)$. Assume $m = c\tau' + d$. From the calculations done in the first part of the proof, we know that $m\Lambda_\tau = \Lambda_{\tau'}$ and under the "multiplication by m map", the point $1/N + \Lambda_\tau$ maps to $\frac{c\tau'+d}{N} + \Lambda_{\tau'}$. Moreover, $(c, d) \equiv (0, 1) \mod N$ implies that $\frac{c\tau'+d}{N} + \Lambda_{\tau'} = 1/N + \Lambda_{\tau'}$. This helps us conclude that $[E_\tau, 1/N + \Lambda_\tau] = [E_{\tau'}, 1/N + \Lambda_{\tau'}]$.

□

Theorem 3.1.2 helps us to translate maps between modular curves to maps between the corresponding moduli spaces. For example, since $\Gamma_1(N) \subset \Gamma_0(N)$, we have a natural map from $Y_1(N)$ to $Y_0(N)$ taking the orbit space $\Gamma_1(N)\tau$ to $\Gamma_0(N)\tau$. The corresponding map on the moduli space takes the point $[E, Q]$ in $S_1(N)$ to the point $[E, \langle Q \rangle]$ in $S_0(N)$.

The fact that $\Gamma_1(N)$ is normal in $\Gamma_0(N)$ is also used to define a map from $Y_1(N)$ to $Y_1(N)$ which maps the point $\Gamma_1(N)\tau$ to $\Gamma_1(N)\gamma\tau$ where $\gamma \in \Gamma_0(N)$. The corresponding map on the moduli space $S_1(N)$ turns out to be: $[E, Q] \mapsto [E, dQ]$ where d corresponds to the lower right entry of γ modulo N. This map will reappear as a Hecke operator which we introduce in section 3.4.

3.2 Dirichlet characters and the Diamond operator

In this section, we aim to introduce certain functions on the multiplicative group $(\mathbb{Z}/N\mathbb{Z})^*$ called the Dirichlet characters. As we go on, we will see that they play a fundamental role in our understanding of Hecke operators and L functions. For notational convenience, we denote $(\mathbb{Z}/N\mathbb{Z})^*$ as G_N.

Definition 3.2.1 (Dirichlet Character). A Dirichlet character modulo N is a group homomorphism $\chi : G_N \to \mathbb{C}^*$.

Since G_N is a finite group, the values taken by any Dirichlet character are finite order elements in $\mathbb{C}^*$. These are precisely the roots of unity. The set of Dirichlet characters mod N is a group under multiplication. To see this, let χ and ψ be two Dirichlet characters. Then, define the product $\chi\psi$ as the function, $\chi\psi(n) = \chi(n)\psi(n)$. Under this operation, the identity map is the trivial map $\mathbb{I}$ that maps every element of G_N to 1. The inverse of any character χ is the map $\overline{\chi}$ which acts by conjugation, that is, $\overline{\chi}(n) = \overline{\chi(n)}$. We denote this group by $\widehat{G_N}$.

If G_N is cyclic with a generator s, then any Dirichlet character $\chi \in \widehat{G_N}$ is completely determined by the image of $\chi(s)$. Since s has order N, $\chi(s)$ represents a N^{th} root of unity. Therefore, it is easy to verify in this case that $\chi \mapsto \omega$ is an isomorphism of $\widehat{G_N}$ and $(Z/N\mathbb{Z})^*$. In fact, in general we have the following proposition:

Proposition 3.2.2. The group $\widehat{G_N}$ is isomorphic to G_N.

The above proposition helps us conclude that the number of Dirichlet characters mod N is $\phi(N)$. We next prove two very important relations satisfied by the Dirichlet characters. These are called the orthogonality relations, for reasons coming from representation theory.

Proposition 3.2.3. The groups G_N and $\widehat{G_N}$ satisfy the orthogonality relations given by the following equations:

$$\sum_{n \in G_N} \chi(n) = \begin{cases} \phi(N) & \text{if } \chi = \mathbb{I} \\ 0 & \text{if } \chi \neq \mathbb{I} \end{cases} \tag{3.2.4}$$

$$\sum_{\chi \in \widehat{G_N}} \chi(n) = \begin{cases} \phi(N) & \text{if } n = 1 \\ 0 & \text{if } n \neq 1 \end{cases} \tag{3.2.5}$$

Proof. Suppose that $\chi \neq \mathbb{I}$. To prove (3.2.4), we choose any $n_0 \in G_N$ such that $\chi(n_0) \neq 1$. Now notice that the sum

$$\begin{aligned}\sum_{n \in G_N} \chi(n_0)\chi(n) &= \sum_{n \in G_N} \chi(nn_0) \\ &= \sum_{n \in G_N} \chi(n)\end{aligned}$$

The last equality comes from the observation that for a fixed $n_0 \in G_N$, as n runs over G_N, nn_0 runs over G_N as well. This gives us that,

$$\Big(\chi(n_0) - 1\Big) \sum_{n \in G_N} \chi(n) = 0$$

hence proving the relation (3.2.4) for $\chi \neq \mathbb{I}$. When $\chi = \mathbb{I}$, then the sum is clearly $\phi(N)$.

We proceed on similar lines to prove the next relation. It is clear that when $n = 1$, the sum is $\phi(N)$. When $n \neq 1$, let χ_0 be a Dirichlet character, not equal to $\mathbb{I}$. Then we have the following equality of sums:

$$\sum_{\chi \in \widehat{G_N}} (\chi_0\chi)(n) = \sum_{\chi \in \widehat{G_N}} \chi(n)$$

This implies that,

$$\Big(\chi_0 - \mathbb{I}\Big) \sum_{\chi \in \widehat{G_N}} \chi(n) = 0$$

giving us (3.2.5) for $n \neq 1$. □

Every character χ mod N extends to a function $\chi : \mathbb{Z}/N\mathbb{Z} \to \mathbb{C}$ by defining $\chi(n) = 0$ for non invertible elements $n \in \mathbb{Z}/N\mathbb{Z}$. This further extends to a function from $\mathbb{Z}$ to $\mathbb{C}$ if we define $\chi(n) = \chi(n \bmod N)$ for all $n \in \mathbb{Z}$. This way, $\chi(n) = 0$ for every n such that $(n, N) > 1$. Observe that χ is no longer a group homomorphism, but it still satisfies that $\chi(m)\chi(n) = \chi(mn)$ for $m, n \in \mathbb{Z}/N\mathbb{Z}$.

3.2.0.1 Lift of a Dirichlet character

Let N and d be positive integers such that $d \mid N$. Observe that any Dirichlet character χ modulo d can be lifted to a Dirichlet character χ_N modulo N by simply defining $\chi_N(n \bmod N) = \chi(n \bmod d)$ for all $n \in \mathbb{Z}$ such that $(n, N) = 1$. In other words, $\chi_N = \chi \circ \Pi_{N,d}$ where $\Pi_{N,d}$ is the is the natural projection $G_N \to G_d$.

Other way round, we define the *conductor* of χ modulo N to be the smallest possible integer d such that χ is the lift of some character χ_d mod d. For example, consider the character χ modulo 12 taking $1, 5$ to 1 and $7, 11$ to -1. Then χ has conductor 4 as $\chi = \chi_4 \circ \Pi_{12,4}$ where χ_4 maps 1 to 1 and 3 to -1. However, it is not always necessary that a character mod N comes from characters of lower levels. This is because χ mod N will not necessarily give a character mod d by the usual projection where $d \mid N$.

Definition 3.2.6 (Primitive characters). A character χ mod N is called primitive if its conductor is 1. In other words, χ is not the lift of any character.

Primitive characters will be very useful in the coming sections. One of the many interesting properties of primitive characters is that we can extend it from a function on $\mathbb{Z}$ to a smooth function on $\mathbb{R}$. To see how this works, we introduce the *Gauss sum*.

Definition 3.2.7. Let χ be a primitive Dirichlet character modulo N. The Gauss sum $\tau(\chi)$ is defined by the formula

$$\tau(\chi) = \sum_{n \bmod N} \chi(n) e^{2\pi i n/N}$$

We will need the following expression for Gauss sum.

$$\sum_{n \bmod N} \chi(n) e^{2\pi i n m/N} = \overline{\chi}(m)\tau(\chi). \tag{3.2.8}$$

Proof of the expression in equation (3.2.8). First consider the case when $(m, N) = 1$. Using the fact that $\chi(m)\overline{\chi(m)} = 1$, we write

$$\begin{aligned} \sum_{n \bmod N} \chi(n) e^{2\pi i n m/N} &= \sum_{n \bmod N} \chi(n)\chi(m)\overline{\chi(m)} e^{2\pi i n m/N} \\ &= \overline{\chi(m)} \sum_{n \bmod N} \chi(nm) e^{2\pi i n m/N} \end{aligned}$$

As n runs in $\mathbb{Z}/N\mathbb{Z}$, nm mod N runs over the same group as well and so the last expression in the equality equals $\overline{\chi(m)}\tau(\chi)$ in this case.

Suppose that $(m, N) = d$ with $d > 1$. Then $\overline{\chi(m)} = 0$. It remains to show that the left hand side of the expression in (3.2.8) vanishes. Let $m = dM$ and $N = dN_1$. We first show that there exists some integer c such that $c \equiv 1 \bmod (N_1)$ and $(c, N) = 1$ with the property that $\chi(c) \neq 1$. Suppose otherwise, then $\chi(c) = 1$ for all c such that $c \equiv 1 \bmod (N_1)$ and $(c, N) = 1$. This implies that if we have integers n and n' relatively prime to N such that $n \equiv n' \bmod N_1$, then $\chi(n) = \chi(n') \bmod N_1$. It follows that χ is well defined modulo N_1. That is, χ is a lift of a Dirichlet character modulo N_1 contradicting the primitivity of χ. Next, write:

$$\begin{aligned}\sum_{n \bmod N} \chi(n)e^{2\pi inm/N} &= \sum_{n \bmod N} \chi(n)e^{2\pi inM/N_1} \\ &= \sum_{r=0}^{N_1-1}\Bigg(\sum_{\substack{n=0 \\ n\equiv r \bmod N_1}}^{N-1} \chi(n)\Bigg)e^{2\pi irM/N_1}\end{aligned}$$

The expression in the last equality is justified because for any integers n_1 and n_2 between 0 to $N-1$, the expression $e^{2\pi in_1M/N_1} = e^{2\pi in_2M/N_1}$ if and only if $n_1 \equiv n_2 \bmod N_1$. So for each $r \in \mathbb{Z}/N_1\mathbb{Z}$, we sum n from 0 to $N-1$ with the condition that $n \equiv r \bmod N_1$ so as to club the coefficients with the same value of $e^{2\pi inM/N_1}$. For each r in 0 to N_1-1, consider the sum

$$\sum_{\substack{n=0 \\ n\equiv r \bmod N_1}}^{N-1} \chi(n) \tag{3.2.9}$$

The existence of an integer c with the desired property allows us to do a change of variable from n to cn. This just permutes the elements of $\mathbb{Z}/N\mathbb{Z}$ as $(c, N) = 1$. Moreover, since $c \equiv 1 \bmod N_1$, it preserves the condition that $n \equiv r \bmod N_1$. This allows us to write:

$$\sum_{\substack{n=0 \\ n\equiv r \bmod N_1}}^{N-1} \chi(n) = \sum_{\substack{n=0 \\ n\equiv r \bmod N_1}}^{N-1} \chi(cn) \tag{3.2.10}$$

This implies that,

$$\sum_{\substack{n=0 \\ n\equiv r \bmod N_1}}^{N-1} \chi(n)(\chi(c)-1) = 0$$

Since $\chi(c) \neq 1$, the sum in (3.2.10) must vanish, making the left hand side in (3.2.8) equal to 0. □

We next show that the Gauss sum does not vanish. More explicitly, the following proposition is true.

Proposition 3.2.11. Suppose χ is a primitive character modulo N and $\tau(\chi)$ denotes the Gauss sum defined above. Then $|\tau(\chi)| = \sqrt{N}$.

Proof. Consider the product:

$$\Bigg(\sum_{n \bmod N} \chi(n)e^{2\pi inm/N}\Bigg)\Bigg(\overline{\sum_{n \bmod N} \chi(n)e^{2\pi inm/N}}\Bigg)$$

We expand it as follows:

$$\begin{aligned}\Big|\sum_{n \bmod N} \chi(n)e^{2\pi inm/N}\Big|^2 &= \Big(\sum_{n \bmod N} \chi(n)e^{2\pi inm/N}\Big)\Big(\sum_{n \bmod N} \overline{\chi(n)}e^{-2\pi inm/N}\Big) \\ &= \sum_{\substack{n_1,n_2 \bmod N \\ (n_1,n_2,N_1)=1}} \chi(n_1)\overline{\chi(n_2)}e^{2\pi i(n_1-n_2)m/N} \\ &= (\overline{\chi(m)}\tau(\chi))(\chi(m)\overline{\tau(\chi)})\end{aligned}$$

The last expression equals $|\tau(\chi)^2|$ if $(m, N) = 1$, otherwise it is 0. Summing over $m \in \mathbb{Z}/N\mathbb{Z}$ we get,

$$\phi(N)|\tau(\chi)^2| = \sum_{m \bmod N} \Big(\sum_{\substack{n_1,n_2 \bmod N \\ (n_1,n_2,N_1)=1}} \chi(n_1)\overline{\chi(n_2)}e^{2\pi i(n_1-n_2)m/N}\Big) \tag{3.2.12}$$

Observe that

$$\sum_{m=0}^{N-1} e^{2\pi iam/N} = \begin{cases} 0 & \text{if } N \nmid a \\ N & \text{if } N \mid a \end{cases}$$

This helps us to conclude that the terms with $N \mid (n_1 - n_2)$ will contribute a factor of $\chi(n_1)\overline{\chi(n_2)}N$ in the right hand side of equation (3.2.12). Since $n_1 \equiv n_2 \bmod N$, and so $\chi(n_1) = \chi(n_2)$, the left hand side of the expression in (3.2.12) simplifies to

$$\sum_{\substack{n_1 \equiv n_2 \bmod N \\ (n_1,n_2,N_1)=1}} N = \phi(N)N$$

On comparing the two sides we get that $|\tau(\chi)^2| = N$. □

We need one more result about the Gauss sum which is easy to see:

$$\overline{\tau(\chi)} = \chi(-1)\tau(\overline{\chi}) \tag{3.2.13}$$

Putting every thing together from the results in (3.2.8), (3.2.11) and (3.2.13) and doing a little bit of algebra we arrive at the following expression:

$$\chi(n) = \frac{\chi(-1)\tau(\chi)}{N} \sum_{m \bmod N} \overline{\chi}(m)e^{2\pi inm/N} \tag{3.2.14}$$

Notice that the right hand side of the equation in (3.2.14) is defined for an arbitrary real number $n \in \mathbb{R}$. Therefore, the expression in (3.2.14) allows us to interpolate the character to view it as a function on $\mathbb{R}$ instead of just $\mathbb{Z}$. We will need these results in the later sections to come and for now leave the discussion on Dirichlet characters for a while.

Returning to modular forms, the results from now on will focus on the space $\mathcal{M}_k(\Gamma_1(N))$. The theory we develop in the next few sections will be dedicated to studying linear operators on this space so as to decompose this space into eigen subspaces and find a suitable basis of eigen vectors, which we call *eigenforms*. Once this is achieved, we will start seeing very interesting properties of L functions associated to these eigenforms. To that end, we start with defining the χ − eigen space of $\mathcal{M}_k(\Gamma_1(N))$.

Definition 3.2.15. For each Dirichlet character χ mod N, the χ-eigen space of $\mathcal{M}_k(\Gamma_1(N))$, denoted by $\mathcal{M}_k(N, \chi)$ is the subspace:

$$\{f \in \mathcal{M}_k(\Gamma_1(N)) \mid f[\gamma]_k = \chi(d_\gamma)f \text{ for all } \gamma \in \Gamma_0(N)\} \tag{3.2.16}$$

In order to motivate this definition, we introduce the diamond operator. The subgroup $\Gamma_0(N)$ acts on the space $\mathcal{M}_k(\Gamma_1(N))$ via the weight k operator. Clearly $\Gamma_1(N)$ acts trivially. Proposition 1.4.7 helps us conclude that the action can be considered to be of $(\mathbb{Z}/N\mathbb{Z})^*$. For any $d \in (\mathbb{Z}/N\mathbb{Z})^*$ denote the diamond operator as $\langle d\rangle$. Then for any $f \in \mathcal{M}_k(\Gamma_1(N)), \langle d\rangle f = f[\gamma]_k$ where $\gamma \in \Gamma_0(N)$ is such that $\gamma = \begin{pmatrix} a & b \\ c & d' \end{pmatrix}$ and $d' \equiv d \bmod N$. This action is independent of the lift $\gamma \in \Gamma_0(N)$ because of Proposition 1.4.7 as remarked earlier.

Now it is clear that the subspace in (3.2.16) is precisely the χ-eigen space of the diamond operator. We would like to decompose the subspace $\mathcal{M}_k(\Gamma_1(N))$ into these subspaces. In order to do this, for each Dirichlet character χ mod N, we define an operator π_χ on the space $\mathcal{M}_k(\Gamma_1(N))$ given by the following expression:

$$\pi_\chi = \frac{1}{\phi(N)} \sum_{d \in (\mathbb{Z}/N\mathbb{Z})^*} \chi(d)^{-1}\langle d\rangle$$

We wish to use the following theorem from linear algebra:

Theorem 3.2.17. Let I denote the identity operator and let $E_1 \dots E_k$ be k linear operators on a vector space V which satisfy the following three conditions.

1. Each E_i is a projection.
2. $E_iE_j = 0$ for $1 \le i, j \le k$ and $i \neq j$.
3. $E_1 + \cdots + E_k = I$

If W_i is the range of E_i for $1 \le i \le k$ then, $V = \bigoplus_{i=1}^{k} W_i$.

We proceed by checking the three conditions in Theorem 3.2.17 for the operator π_χ.

1. In order to check that $\pi_\chi^2 = \pi_\chi$, expand

$$\begin{aligned}
\pi_\chi^2(f) &= \pi_\chi\bigg(\frac{1}{\phi(N)}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d)^{-1}\langle d\rangle f\bigg)\\
&= \frac{1}{\phi(N)^2}\bigg(\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d')^{-1}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d)^{-1}\langle d'\rangle(\langle d\rangle f)\bigg)\\
&= \frac{1}{\phi(N)^2}\bigg(\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(dd')^{-1}\langle dd'\rangle f\bigg)\\
&= \frac{1}{\phi(N)^2}\bigg(\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d)^{-1}\langle d\rangle f\bigg)\\
&= \pi_\chi
\end{aligned}$$

For the second last equality we use the observation that for a fixed $d' \in (\mathbb{Z}/N\mathbb{Z})^*$, as d runs over $(\mathbb{Z}/N\mathbb{Z})^*$, dd' runs over $(\mathbb{Z}/N\mathbb{Z})^*$ as well. This proves that the operator π_χ is a projection.

2. We next show that $\pi_{\chi'} \circ \pi_\chi = 0$ whenever $\chi \neq \chi'$.

$$\begin{aligned}
(\pi_{\chi'}\circ\pi_\chi)(f) &= \pi_{\chi'}\bigg(\frac{1}{\phi(N)}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d)^{-1}\langle d\rangle f\bigg)\\
&= \frac{1}{\phi(N)^2}\bigg(\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi'(d')^{-1}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d)^{-1}\langle d'\rangle(\langle d\rangle f)\bigg)\\
&= \frac{1}{\phi(N)^2}\bigg(\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi'(d')^{-1}\chi(d')\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(dd')^{-1}\langle dd'\rangle f\bigg)\\
&= \frac{1}{\phi(N)^2}\bigg(\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d)^{-1}\langle d\rangle f\bigg)\bigg(\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi'(d')^{-1}\chi(d')\bigg)
\end{aligned}$$

Since $\chi' \neq \chi$, we can find some $d_0 \in (\mathbb{Z}/N\mathbb{Z})^*$ so that $\chi'(d_0)^{-1}\chi(d_0) \neq 1$. Observe that the sum in the last equality can written as

$$\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi'(d')^{-1}\chi(d') = \sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi'(d_0d')^{-1}\chi(d_0d')$$

This implies that

$$\bigg(\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi'(d')^{-1}\chi(d')\bigg)\bigg(\chi'(d_0)^{-1}\chi(d_0)-1\bigg) = 0$$

It follows that $\sum_{d'\in(\mathbb{Z}/N\mathbb{Z})^*}\chi'(d')^{-1}\chi(d') = 0$ and so $(\pi_{\chi'}\circ\pi_\chi)(f) = 0$.

3. To check that the projections add up to the identity map we take the sum over the set of Dirichlet characters modulo N.

$$\begin{aligned}\left(\sum_{\chi\in\widehat{G_N}}\pi_\chi\right)(f) &= \sum_{\chi\in\widehat{G_N}}\left(\frac{1}{\phi(N)}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\chi(d)^{-1}\langle d\rangle f\right)\\ &= \frac{1}{\phi(N)}\sum_{d\in(\mathbb{Z}/N\mathbb{Z})^*}\left(\sum_{\chi\in\widehat{G_N}}\chi(d)^{-1}\langle d\rangle f\right)\end{aligned}$$

Using the orthogonality relation in (3.2.5), we get that the above sum is equal to $f[\gamma_1]$ where γ_1 is the lift in $\Gamma_0(N)$ corresponding to $d=1$. We can as well take this lift to be identity to get that $\left(\sum_\chi \pi_\chi\right)(f) = f$.

Having proved the three conditions for our operators, our final task is to find the image of these operators. Observe that if $f\in\mathcal{M}_k(N,\chi)$, then $\pi_\chi(f)=f$. We finally make the following claim:

Claim 3.2.18. The image $\pi_\chi(\mathcal{M}_k(\Gamma_1))\subseteq\mathcal{M}_k(N,\chi)$.

Proof of claim. Let $f\in\mathcal{M}_k(\Gamma_1(N))$ and $\gamma_{d'}\in\Gamma_0(N)$ with lower right entry congruent to d' modulo N. We wish to show that $(\pi_\chi(f))[\gamma_{d'}]_k=\chi(d')(\pi_\chi)f$.

$$\begin{aligned}(\pi_\chi(f))[\gamma_{d'}]_k &= \frac{1}{\phi(N)}\sum_{d\in\mathbb{Z}/N\mathbb{Z}^*}\chi(d)^{-1}\langle dd'\rangle(f)\\ &= \frac{\chi(d')}{\phi(N)}\sum_{d\in\mathbb{Z}/N\mathbb{Z}^*}\chi(d'd)^{-1}\langle dd'\rangle(f)\\ &= \chi(d')\pi_\chi(f)\end{aligned}$$

This implies that $\pi_\chi(f)\in\mathcal{M}_k(N,\chi)$. ☐

It follows that the projection is onto the space $\mathcal{M}_k(N,\chi)$. Theorem 3.2.17 now helps us conclude the following theorem.

Theorem 3.2.19. Let N be a positive integer. Then $\mathcal{M}_k(\Gamma_1(N))$ decomposes into $\chi-$ eigen spaces as follows:

$$\mathcal{M}_k(\Gamma_1(N))=\bigoplus_\chi\mathcal{M}_k(N,\chi)$$

Similarly for the subspace $\mathcal{S}_k(\Gamma_1(N))$, since the projection will be onto cusp forms we see that $\mathcal{S}_k(\Gamma_1(N))=\bigoplus_\chi\mathcal{S}_k(N,\chi)$.

3.3 Double coset operators

Throughout this section Γ_1 and Γ_2 denote congruence subgroups of $SL_2(\mathbb{Z})$. The group $GL_2^+(\mathbb{Q})$ denotes invertible 2×2 matrices in $\mathbb{Q}$ with a positive determinant.

In addition to the diamond operator, we wish to introduce a special kind of operator which we call as the T_n operator. It turns out that these two operators come under a much generalized class of operators known as the Hecke operators. We will study these via the theory of double coset operators.

Definition 3.3.1. (Double Coset) Let $\alpha \in GL_2^+(\mathbb{Q})$. The set $\Gamma_1\alpha\Gamma_2 = \{\gamma_1\alpha\gamma_2 \mid \gamma_1 \in \Gamma_1, \gamma_2 \in \Gamma_2\}$ is called a double coset.

The subgroup Γ_1 acts naturally on the double coset $\Gamma_1\alpha\Gamma_2$ by left multiplication. Let $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$ denote the orbit space under this action such that $\Gamma_1\alpha\Gamma_2 = \cup_j\Gamma_1\beta_j$. Here, β_j denote the orbit representatives of the double coset.
Our first task is to show that this union is finite. This requires a series of small results, which we prove below.

Lemma 3.3.2. Let Γ be a congruence subgroup and let $\alpha \in GL_2^+(\mathbb{Q})$. Then $\alpha^{-1}\Gamma\alpha \cap SL_2(\mathbb{Z})$ is a congruence subgroup of $SL_2(\mathbb{Z})$.

Proof. We need to show that for some positive integer $N > 1$, the subgroup $\Gamma(N) \subseteq \alpha\Gamma\alpha^{-1} \cap SL_2(\mathbb{Z})$. Since Γ is a congruence subgroup, $\Gamma(M) \subseteq \Gamma$ for some $M > 1$. Let l denote the least common multiple of the denominators of the entries in the matrix α and α^{-1}. Take $\widetilde{N} = Ml$. Then $\widetilde{N}\alpha$ and $\widetilde{N}\alpha^{-1}$ both belong to $M_2(\mathbb{Z})$. Observe that $\Gamma(\widetilde{N}) \subseteq \Gamma$.

Claim 3.3.3. Let $\widetilde{N}^3 = N$. Then $\Gamma(N) \subseteq \alpha\Gamma\alpha^{-1} \cap SL_2(\mathbb{Z})$.

Consider the subgroup $\alpha\Gamma(N)\alpha^{-1} \subseteq (SL_2(\mathbb{Z})$. We also have that:
$\alpha\Gamma(N)\alpha^{-1} = \alpha\Gamma(\widetilde{N}^3)\alpha^{-1} \subseteq \alpha(I + \widetilde{N}^3M_2(\mathbb{Z}))\alpha^{-1}$. The last term further simplifies to: $I + \alpha\widetilde{N} \cdot \widetilde{N} \cdot M_2(\mathbb{Z}) \cdot \widetilde{N}\alpha^{-1} = I + \widetilde{N} \cdot M_2(\mathbb{Z}) \subseteq \Gamma(\widetilde{N})$. We finally have that $\alpha\Gamma(N)\alpha^{-1} \subseteq \Gamma(\widetilde{N})$. Therefore $\Gamma(N) \subseteq \alpha\Gamma(\widetilde{N})\alpha^{-1} \subseteq \alpha\Gamma\alpha^{-1}$. This proves the claim. □

Lemma 3.3.4. Let $\alpha \in GL_2^+(\mathbb{Q})$ as above. Set $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2$. Let φ be the map defined as follows:

$$\varphi : \Gamma_2 \to \Gamma_1\backslash\Gamma_1\alpha\Gamma_2$$

$$\gamma_2 \mapsto \Gamma_1\alpha\gamma_2$$

The above map induces a bijection from $\Gamma_3\backslash\Gamma_2 \to \Gamma_1\backslash\Gamma_1\alpha\Gamma_2$. In other words, the set $\{\gamma_{2,j}\}$ are the orbit representatives of $\Gamma_3\backslash\Gamma_2$ if and only if $\{\alpha\gamma_{2,j}\}$ represent the orbit space $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$.

Proof. The map is clearly surjective. To see that the map induces an injection from the orbit space $\Gamma_3\backslash\Gamma$ observe the following. $\phi(\gamma_2') = \phi(\gamma_2)$ if and only if $\gamma_2'\gamma_2^{-1} \in \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2 = \Gamma_3$ if and only if $\Gamma_3\gamma_2' = \Gamma_3\gamma_2$. □

Lemma 3.3.5. Let Γ_1 and Γ_2 be any two congruence subgroups. Then the index $[\Gamma_1 : \Gamma_1 \cap \Gamma_2]$ and $[\Gamma_2 : \Gamma_1 \cap \Gamma_2]$ is finite.

Proof. First, notice that if $\Gamma(N_1) \subseteq \Gamma_1$ and $\Gamma(N_2) \subseteq \Gamma_2$ for some positive integers N_1 and $N_2 > 1$, then $\Gamma(N_1N_2) \subseteq \Gamma_1 \cap \Gamma_2$. We also have that,

$$
\begin{aligned}
[\Gamma_1 : \Gamma_1 \cap \Gamma_2] &\leq [\Gamma_1 : \Gamma(N_1N_2)] \\
&\leq [SL_2(\mathbb{Z}) : \Gamma(N_1N_2)]
\end{aligned}
$$

which is finite. Arguing the same for Γ_2 instead helps us conclude that $[\Gamma_2 : \Gamma_1 \cap \Gamma_2]$ is finite as well. □

Let Γ_3 be as defined above. By Lemma 3.3.2, $\alpha^{-1}\Gamma\alpha \cap SL_2(\mathbb{Z})$ is a congruence subgroup. Applying Lemma 3.3.5 to the subgroups $\alpha^{-1}\Gamma\alpha \cap SL_2(\mathbb{Z})$ and Γ_2, we get that $\Gamma_3\backslash\Gamma_2$ is finite. Because of the bijection in 3.3.4, it follows that $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$ is finite.

Having established the finiteness of the orbit space of the double coset under the actin of Γ_1, we can finally move on to introducing the double coset operator acting on the space $\mathcal{M}_k(\Gamma_1)$.

So far we have defined the weight k operator for matrices in $SL_2(\mathbb{Z})$. However in order to work with double coset operator we need to define the weight k operator for any arbitrary matrix in $GL_2^+(\mathbb{Q})$.

Definition 3.3.6. Let $\beta \in GL_2^+(\mathbb{Q})$ and $k \in \mathbb{Z}$. The weight k operator on a function $f : \mathbb{H} \to \mathbb{C}$ is given by:

$$
(f[\beta]_k)(\tau) = (\det \beta)^{k-1} \jmath(\beta, \tau)^{-k} f(\beta\tau)
$$

Definition 3.3.7. (Double coset operator) Let Γ_1, Γ_2 and α be as defined as above. Let β_j be the coset representatives of $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$ such that $\Gamma_1\alpha\Gamma_2 = \cup_j \Gamma_1\beta_j$. Define the double coset operator $[\Gamma_1\alpha\Gamma_2]_k$ on $\mathcal{M}_k(\Gamma_1)$ as follows:

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

In order to justify the definition, we must prove that it is independent of the choice of representatives β_j. To see this, let γ_j be another set of representatives. Then $\gamma_j = \alpha_j\beta_j$

for some $\alpha_j \in \Gamma_1$. This implies that,

$$\sum_j f[\gamma_j]_k = \sum_j (f[\alpha_j])[\beta_j]_k = \sum_j f[\beta_j]_k$$

Lemma 3.3.8. The map $[\Gamma_1\alpha\Gamma_2]_k$ maps $\mathcal{M}_k(\Gamma_1)$ to $\mathcal{M}_k(\Gamma_2)$.

Proof. We first show that $(f[\Gamma_1\alpha\Gamma_2]_k)[\gamma_2]_k = f[\Gamma_1\alpha\Gamma_2]_k$ for any matrix $\gamma_2 \in \Gamma_2$. Notice that for each $\gamma_2 \in \Gamma_2$, the "multiplication by γ_2" map

$$\gamma_2 : \Gamma_1\backslash\Gamma_1\alpha\Gamma_2 \to \Gamma_1\backslash\Gamma_1\alpha\Gamma_2$$

$$\Gamma_1\beta \mapsto \Gamma_1\beta\gamma_2$$

is well defined and bijective. Therefore, if the set $\{\beta_j\}$ represents the orbit space $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$, then the set $\{\beta_j\gamma_2\}$ represents it as well. Now, for any $\gamma_2 \in \Gamma_2$,

$$(f[\Gamma_1\alpha\Gamma_2]_k)[\gamma_2]_k = \sum_j f[\beta_j\gamma_2]_k = f[\Gamma_1\alpha\Gamma_2]_k$$

We also need to make sure that the function $f[\Gamma_1\alpha\Gamma_2]_k$ is holomorphic at cusps. This requires a small result which will help in proving holomorphy at cusps:

Claim 3.3.9. If $\gamma \in GL_2^+(\mathbb{Q})$, then $\gamma = \alpha\gamma'$ where $\alpha \in SL_2(\mathbb{Z})$ and $\gamma' = r\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ for some $r \in \mathbb{Q}$.

Suppose $\gamma = q\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ where $a', b', c', d' \in \mathbb{Z}$ and $q \in \mathbb{Q}$. Now let g and h be such that $a'/c' = h/g$ such that $(h, g) = 1$. This implies that $a'g - c'h = 0$. Also, there exists integers h' and g' such that $hh' - gg' = 1$. This helps us to find a matrix in $SL_2(\mathbb{Z})$ with the desired property.

$$q\begin{pmatrix} h & h' \\ g & g' \end{pmatrix}\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = \begin{pmatrix} * & * \\ 0 & * \end{pmatrix}.$$

Coming back to the proof, consider the function $f[\beta_j]_k$. Using the above result we write it as $(f[\alpha_j]_k)[\delta_j]_k$ for some $\alpha_j \in SL_2(\mathbb{Z})$ by the previous calculation. The modular form f is holomorphic at cusps. By definition this means that $f[\gamma]_k$ is holomorphic at ∞ for all $\gamma \in SL_2(\mathbb{Z})$. Therefore, we get a Fourier series expansion:

$$f[\alpha_j]_k = \sum_{n=1}^{\infty} a_n e^{2\pi i n\tau/h}$$

If $\delta_j = r\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ for some $a, b, d \in \mathbb{Z}$ and $r \in \mathbb{Q}$, we have:

$$(f[\alpha_j]_k)[\delta_j]_k = \left(\sum_{n=1}^{\infty} a_n e^{2\pi i n\tau/h}\right)[\delta_j]_k = \left(d^{-k}(\det \delta_j)^{k-1} \sum_{n=1}^{\infty} a_n e^{2\pi i n(a\tau+b)/dh}\right)$$

This is further equal to $\sum_{n=1}^{\infty} a'_n e^{2\pi i n a\tau/dh}$ where $a'_n = d^{-k}(\det \delta_j)^{k-1} a_n e^{2\pi i b n/hd}$. This shows that $f[\beta_j]_k$ is holomorphic at ∞ with period hd. Holomorphy of the function $f[\Gamma_1\alpha\Gamma_2] = \sum_j f[\beta_j]_k$ now follows from the following observation: If $g_1, \dots g_d$ are holomorphic functions on $\mathbb{H}$, then their sum is holomorphic on ∞ as well. To see this, assume that for every j, g_j has a period h_j so that $g_j = \sum_n b_n e^{2\pi i\tau/h_j}$. Let h be the least common multiple of $\{h_j\}_j$ and for every j, write $l_j = h/h_j$. This way, we can rewrite each $g_j = \sum_n b_n e^{2\pi i l_j n\tau/h}$. Now taking the usual term wise sum of the powers series of g_j, we get a power series expansion of the sum $\sum_j g_j$, with period h.

□

Lemma 3.3.10. The map $[\Gamma_1\alpha\Gamma_2]_k$ maps $\mathcal{S}_k(\Gamma_1)$ to $\mathcal{S}_k(\Gamma_2)$.

Proof. If $f \in \mathcal{S}_k(\Gamma_1)$, then the a_0 term in the Fourier series expansion of $f[\alpha]_k$ is 0 for all $\alpha \in SL_2(\mathbb{Z})$. From the calculation done at the end of Lemma 3.3.8, we get that the a_0 term of the Fourier expansion of $f[\beta_j]$, and hence that of $f[\Gamma_1\alpha\Gamma_2]$ is 0. □

3.3.1 Special cases

In this section, we will see three special cases of double coset operators and prove that any operator is a composition of these. The Hecke operators, which we introduce in the coming section, are examples of these special cases. Throughout this section $f \in \mathcal{M}_k(\Gamma_1)$.

1. $\Gamma_2 \subseteq \Gamma_1$ and $\alpha = I$: In this case $\Gamma_1\alpha\Gamma_2 = \Gamma_2$. Therefore $f[\Gamma_1\alpha\Gamma_2]_k = f[I]_k = f$ so that $[\Gamma_1\alpha\Gamma_2]_k : \mathcal{M}_k(\Gamma_1) \hookrightarrow \mathcal{M}_k(\Gamma_2)$ is the usual inclusion map.

2. $\alpha^{-1}\Gamma_1\alpha = \Gamma_2$; $\alpha \in GL_2(\mathbb{Q})$: Here, $\Gamma_1\alpha\Gamma_2 = \Gamma_1\alpha$ so that $f[\Gamma_1\alpha\Gamma_2]_k = f[\alpha]_k$ gives us a map: $\mathcal{M}_k(\Gamma_1) \to \mathcal{M}_k(\Gamma_2)$. Notice that since $\Gamma_1 = \alpha\Gamma_2\alpha^{-1}$, the operator $[\Gamma_2\alpha\Gamma_1]_k = f[\alpha^{-1}]_k$ is a map from $\mathcal{M}_k(\Gamma_2)$ to $\mathcal{M}_k(\Gamma_1)$. These two are inverses of each other and so $\mathcal{M}_k(\Gamma_1)$ is isomorphic to $\mathcal{M}_k(\Gamma_2)$.

3. $\Gamma_1 \subseteq \Gamma_2$ and $\alpha = I$: In this case we have $\Gamma_1\alpha\Gamma_2 = \cup_j \Gamma_1\beta_j$ where β_j are the coset representatives of the orbit space $\Gamma_1\backslash\Gamma_2$. Thus, $[\Gamma_1\alpha\Gamma_2]$ is the map $\sum_j f[\beta_j]_k$.

This is called the *Trace operator*. This map is a surjection onto $\mathcal{M}_k(\Gamma_2)$. To see this, notice that $\mathcal{M}_k(\Gamma_2) \subseteq \mathcal{M}_k(\Gamma_1)$. If $f \in \mathcal{M}_k(\Gamma_2)$ then $f[\Gamma_1\alpha\Gamma_2] = [\Gamma_2 : \Gamma_1]f$. Therefore, $(1/[\Gamma_2 : \Gamma_1])f \mapsto f$ via the trace operator.

Proposition 3.3.11. Any double coset operator is a combination of the operators described in 1, 2, and 3.

Proof. Let $\Gamma_1, \Gamma_2, \Gamma_3$ and α be as described in Lemma 3.3.4. Suppose $[\Gamma_1\alpha\Gamma_2]_k$ is any arbitrary double coset operator. Let $\Gamma_3' = \alpha\Gamma_3\alpha^{-1} = \Gamma_1 \cap \alpha\Gamma_2\alpha^{-1}$. Now, $\Gamma_3 \subseteq \Gamma_2$, $\Gamma_3' \subseteq \Gamma_1$ and $\alpha^{-1}\Gamma_3'\alpha = \Gamma_3$. If $f \in \mathcal{M}_k(\Gamma_1)$, then via the operators described in 1, 2, and 3, f is mapped as follows:

$$\mathcal{M}_k(\Gamma_1) \hookrightarrow \mathcal{M}_k(\Gamma_3') \xrightarrow{\sim} \mathcal{M}_k(\Gamma_3) \twoheadrightarrow \mathcal{M}_k(\Gamma_2)$$

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

where $\gamma_{2,j}$ represent orbits of $\Gamma_3\backslash\Gamma_2$. By Lemma 3.3.4, $\{\alpha\gamma_{2,j}\}$ represent orbits of $\Gamma_1\backslash\Gamma_1\alpha\Gamma_2$. Therefore, f finally gets mapped to $\sum_j f[\beta_j]_k = f[\Gamma_1\alpha\Gamma_2]_k$. □

3.3.2 Maps at the level of modular curves

As seen in section 2, to every space of modular forms, we can associate a modular curve which is in fact a Riemann surface. We would like to geometrically translate the maps of double coset operators from the space of modular forms to the corresponding modular curves. This will help us to geometrically interpret the operators in terms of points on the Riemann surface.

Let $\Gamma_1, \Gamma_2, \Gamma_3$ and Γ_3' be as described in the proof of Proposition 3.3.11. The following diagram shows the composition of maps between groups as in Proposition 3.3.11.

$$\begin{array}{ccc} \Gamma_3 & \xrightarrow{\sim} & \Gamma_3' \\ \downarrow & & \downarrow \\ \Gamma_2 & & \Gamma_1 \end{array}$$

The isomorphism between Γ_3 and Γ_3' is the map given by $\gamma \mapsto \alpha\gamma\alpha^{-1}$. The corresponding map on the modular curves is:

$$\begin{array}{ccc} X_3 & \xrightarrow{\sim} & X_3' \\ \downarrow{\scriptstyle \pi_1} & & \downarrow{\scriptstyle \pi_2} \\ X_2 & & X_1 \end{array}$$

Inclusions of congruence subgroups induce surjections π_1 and π_2 between the corresponding modular curves. The isomorphism between the modular curve Γ_3 and Γ'_3 is the map:

$$\alpha : X_3 \to X'_3$$

$$\Gamma_3\tau \mapsto \Gamma'_3\alpha(\tau)$$

We check that this map is well defined: If $\Gamma_3\tau = \Gamma_3\tau'$, then there exists some $\gamma \in \Gamma_3$ such that $\gamma\tau = \tau'$. This implies that $\Gamma'_3\alpha\tau' = \Gamma'_3\alpha\gamma(\tau)$. Recall that $\Gamma_3 = \alpha^{-1}\Gamma_1\alpha \cap \Gamma_2$ so that $\gamma \in \Gamma_3$ is of the form $\alpha^{-1}\delta\alpha$ where $\delta \in \Gamma_1$. This gives us that $\Gamma'_3\alpha\gamma(\tau) = \Gamma'_3\delta\alpha(\tau)$. Notice that $\delta \in \Gamma_1 \cap \alpha\Gamma_2\alpha^{-1} = \Gamma'_3$ and so $\Gamma'_3\alpha\tau' = \Gamma'_3\alpha\tau$. More explicitly, the map on the points of the modular curve is described as follows: Let $\Gamma_3\backslash\Gamma_2 = \cup_j \Gamma_3\gamma_{2,j}$ and $\beta_j = \alpha\gamma_{2,j}$ such that $\Gamma_1\alpha\Gamma_2 = \cup_j \Gamma_3\beta_j$. Each point of X_2 is taken to a set of points of X_1 via $\pi_1\alpha\pi_2^{-1}$ to a set of points described in the following diagram:

$$\begin{tikzcd}
\{\Gamma_3\gamma_{2,j}(\tau)\} \arrow[r, "\alpha"] & \{\Gamma'_3\beta_j(\tau)\} \arrow[d, "\pi_2"] \\
\Gamma_2(\tau) \arrow[u, "\pi_1^{-1}"] & \{\Gamma_1\beta_j(\tau)\}
\end{tikzcd}$$

We can see that the operator $[\Gamma_1\alpha\Gamma_2]_k$ does not induce a well defined map from X_2 to X_1. However, when we view the operator at the level of divisor groups of the modular curves, it is a well defined divisor group homomorphism from $Div(X_2)$ to $Div(X_1)$. In the special cases of section 3.3.1, the corresponding map on the divisor groups are described explicitly as follows.

1. When $\Gamma_2 \subseteq \Gamma_1$ and $\alpha = I$, then $\Gamma_3 = \Gamma_2 = \Gamma'_3$. In this case $\gamma_{2,j} = \beta_j = I$ and $j = 1$. The diagram above specializes to:

$$\begin{tikzcd}
\{\Gamma_2\tau\} \arrow[r, "I"] & \{\Gamma_2\tau\} \arrow[d, "\pi_2"] \\
\Gamma_2\tau \arrow[u, "\pi_1^{-1}"] & \{\Gamma_1\tau\}
\end{tikzcd}$$

   This is the natural map $\Gamma_2\tau \mapsto \Gamma_1\tau$ which induces a surjection between the modular curves and thus between the corresponding divisor groups.

2. If $\alpha^{-1}\Gamma_1\alpha = \Gamma_2$, then $\Gamma_3 = \Gamma_2$ and $\Gamma'_3 = \Gamma_1$. In this case we have the following diagram:

$$\begin{tikzcd}
\{\Gamma_2\tau\} \arrow[r, "\sim"] & \{\Gamma_1\alpha\tau\} \arrow[d, "\pi_2"] \\
\Gamma_2\tau \arrow[u, "\pi_1^{-1}"] & \{\Gamma_1\alpha\tau\}
\end{tikzcd}$$

On the Divisor groups this is the map:

$$\varphi : Div(X_2) \to Div(X_1)$$

$$\Gamma_2\tau \mapsto \Gamma_1\alpha\tau$$

There exists a well defined inverse from $Div(X_1)$ to $Div(X_2)$ given by,

$$\Gamma_1\tau \mapsto \Gamma_2\alpha^{-1}\tau$$

The composition of the two maps give us the identity map and so at the level of divisor groups we get an isomorphism.

3. When $\Gamma_1 \subseteq \Gamma_2$ and $\alpha = I$, then $\Gamma_3 = \Gamma_1 = \Gamma_3'$. In this case $\gamma_{2,j}$ represent the cosets of Γ_1/Γ_2 and $j = [\Gamma_2 : \Gamma_1]$. On the divisor group we have the map:

$$\varphi : Div(X_2) \to Div(X_1)$$

$$\Gamma_2\tau \mapsto \sum_j \Gamma_1\gamma_{2,j}\tau$$

We claim that this is an injection of divisor groups. Suppose τ and τ' map to the same set of points in X_1, then for some j and k, $\Gamma_1\gamma_{2,j}\tau = \Gamma_1\gamma_{2,k}\tau'$. So there exists some $\delta \in \Gamma_1$ such that $\gamma_{2,j}\tau = \delta\gamma_{2,k}\tau'$. Since $\Gamma_1 \subseteq \Gamma_2$ and $\gamma_{2,j} \in \Gamma_2$, we see that $\Gamma_2\tau = \Gamma_2\tau'$.

3.4 The T_p and $\langle d\rangle$ Operators.

This section will be dedicated to defining the T_p operator and then studying its various properties. Parallely, we will also study the diamond operator which we defined in section 3.2.

Redefining the diamond operator in terms of double coset operators, we will see that we get back same definition. Taking $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$ and $\alpha \in \Gamma_0(N)$ in definition 3.3.7, we see that if $f \in \mathcal{M}_k(\Gamma_1(N))$, then, $f[\Gamma_1(N)\alpha\Gamma_1(N)]_k = f[\alpha]_k$ using the fact that $\Gamma_0(N)$ is normal in $\Gamma_1(N)$. Since $\Gamma_1(N)$ acts trivially, as discussed before the quotient acts on $\mathcal{M}_k(\Gamma_1(N))$. The action of α is therefore determined by its right entry d mod N.

Next, to define the T_p operator, let $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$ and $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ where p is a prime. Define the T_p operator by the following double coset operator:

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

Our task is to study this operator. We will make use of the following fact:

Fact 3.4.1. $\Gamma_1(N)\alpha\Gamma_1(N) = \left\{\gamma \in M_2(\mathbb{Z}) \,\middle|\, \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & p \end{pmatrix} \bmod N;\ \det\gamma = p\right\}$

In order to give an explicit representation of T_p we need to first find out $\Gamma_3 = \alpha^{-1}\Gamma_1(N)\alpha \cap \Gamma_1(N)$. Once we have Γ_3, we will find its coset representatives $\{\gamma_{2,j}\}_j$ in $\Gamma_1(N)$. By Lemma 3.3.4, the set $\{\alpha\gamma_{2,j}\}_j$ will represent the cosets of $\Gamma_1(N)$ in $\Gamma_1(N)\alpha\Gamma_1(N)$. To that end, let $\Gamma^0(p)$ be the following set.

$$\Gamma^0(p) = \left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix} = SL_2(\mathbb{Z}) \,\middle|\, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \bmod p\right\}$$

Claim 3.4.2. Let $\Gamma_1^0(p)$ denote the subgroup $\Gamma^0(p) \cap \Gamma_1(N)$. Then $\Gamma_3 = \Gamma_1^0(p)$.

Proof of claim. Let $\gamma = \begin{pmatrix} aN+1 & bp \\ cN & dN+1 \end{pmatrix} \in \Gamma_1^0(N)$. We wish to show that $\gamma = \alpha^{-1}\gamma'\alpha$ for some matrix $\gamma' \in \Gamma_1(N)$. Take γ' to be as follows.

$$\gamma' = \begin{pmatrix} aN+1 & bp \\ cNp & d_N+1 \end{pmatrix}$$

Clearly γ' belongs to $\in \Gamma_1(N)$ and $\gamma = \alpha^{-1}\gamma'\alpha$. $\Gamma_1^0(p)$. Conversely, it is easy to check that if δ is a matrix in $\Gamma_1(N)$, then $\alpha^{-1}\delta\alpha \in \Gamma_1^0(N)$. This establishes the claim. $\square$

Next, we find the coset representatives of $\Gamma_1^0(p)$ in $\Gamma_1(N)$. Since the upper right entry b is congruent to 0 mod p in $\Gamma_1^0(p)$, our first guess for the representatives should be $\gamma_{2,j} = \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix}$ for $0 \leq j < p$. With these, we check that for any arbitrary element $\gamma = \begin{pmatrix} a & b \\ c & b \end{pmatrix} \in \Gamma_1(N)$, the matrix $\gamma\gamma_{2,j}^{-1} \in \Gamma_1^0(p)$ for some $0 \leq j < p$. If this is true, then $\{\gamma_{2,j}\}_j$ is a complete set of representatives. If not, we would have to find more. Observe that,

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

This clearly belongs to $\Gamma_1(N)$. For this matrix to belong in $\Gamma_1^0(p)$, we need to find some $0 \leq j < p$ such that $b - aj \equiv 0 \bmod p$. When $p \mid N$ then, $p \nmid a$. In that case $j = ba^{-1}$ works. However when $p \mid a$, then $b - aj \not\equiv 0 \bmod p$ for any j. Otherwise, $p \mid (b - aj)$ implies that $p \mid b$, and so $p \mid (ad - bc) = 1$, which is not possible. Cases of p dividing a might occur when p divides N. Therefore, in this case, $\{\gamma_{2,j}\}_j$ fail to be the complete set of coset representatives. However to this set, add $\gamma_{2,\infty} = \begin{pmatrix} mp & n \\ N & 1 \end{pmatrix}$ where m and n are chosen such that $mp - nN = 1$. It is easily seen that if $p \mid a$, then $\gamma\gamma_{2,\infty}^{-1} \in \Gamma_1^0(p)$. Thus in this case $\{\gamma_{2,j}\}_j \cup \{\gamma_{2,\infty}\}$ are the set of coset representatives of $\Gamma_1^0(p)$ in $\Gamma_1(N)$. By Lemma 3.3.4, we multiply in each case, every element in the set by α to get the representatives of $\Gamma_1(N)$ in $\Gamma_1(N)\alpha\Gamma_1(N)$. By the definition of a double coset operator, we arrive at the following representation of T_p :

Theorem 3.4.3. Let N be a positive integer. Let $\Gamma_1 = \Gamma_2 = \Gamma_1(N)$ and let $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ where p is a prime. Then the operator $T_p = [\Gamma_1\alpha\Gamma_2]_k$ on $\mathcal{M}_k(\Gamma_1(N))$ is given by:

$$T_p(f) = \begin{cases} \displaystyle\sum_{j=0}^{p-1} f\left[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\right]_k & \text{if } p \mid N \\ \displaystyle\sum_{j=0}^{p-1} f\left[\begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix}\right]_k + f\left[\begin{pmatrix} m & n \\ N & p \end{pmatrix}\right]_k ; mp - nN = 1 & \text{if } p \nmid N \end{cases}$$

Proposition 3.4.4. The two Hecke operators $\langle d \rangle$ and T_p commute.

Proof. Suppose $f \in \mathcal{M}_k(\Gamma_1(N))$ and $\gamma \in \Gamma_0(N)$ with the lower right entry congruent to $d \bmod N$. Then we want to show that,

$$(T_p(f))[\gamma]_k = T_p(f[\gamma]_k). \tag{3.4.5}$$

Let β_j denote the coset representatives of Γ_1 in $\Gamma_1(N)\alpha\Gamma_1(N)$ such that T_p acts as $\sum_j [\beta_j]_k$. Then the equation in (3.4.5) amounts to showing that,

$$\sum_j f[\beta_j\gamma]_k = \sum_j f[\gamma\beta_j]_k \tag{3.4.6}$$

This means we have to prove the following lemma:

Lemma 3.4.7. Suppose $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$ such that the double coset $\Gamma_1(N)\alpha\Gamma_1(N) = \bigcup_j \Gamma_1(N)\beta_j$. Then for any matrix $\gamma \in \Gamma_0(N)$, $\bigcup_j \Gamma_1(N)\beta_j\gamma = \bigcup_j \Gamma_1(N)\gamma\beta_j$.

Proof of Lemma 3.4.7. By the representation of the double coset $\Gamma_1(N)\alpha\Gamma_1(N)$ in fact 3.4.1, it is easy to see that replacing α with any matrix in the given set does not change the double coset. For any matrix $\gamma \in \Gamma_0(N)$, one can check that,

$$\gamma\alpha\gamma^{-1} \equiv \begin{pmatrix} 1 & * \\ 0 & p \end{pmatrix} \mod p$$

so that $\gamma^{-1}\alpha\gamma \in \Gamma_1(N)\alpha\Gamma_1(N)$. Therefore replacing α with $\gamma\alpha\gamma^{-1}$ in the double coset and using the fact that $\Gamma_0(N)$ is normal in $\Gamma_1(N)$, we get that,

$$\begin{aligned} \Gamma_1(N)\alpha\Gamma_1(N) &= \Gamma_1(N)\gamma^{-1}\alpha\gamma\Gamma_1(N) \\ &= \gamma\Gamma_1(N)\alpha\Gamma_1(N)\gamma^{-1} \\ &= \gamma\bigcup_j \Gamma_1(N)\beta_j\gamma^{-1} \\ &= \bigcup_j \Gamma_1(N)\gamma\beta_j\gamma^{-1} \end{aligned}$$

On comparing the decomposition of the double coset, it follows that, $\bigcup_j \Gamma_1(N)\beta_j\gamma = \bigcup_j \Gamma_1(N)\gamma\beta_j$. □

This completes the proof of Proposition 3.4.5. □

We next determine the action of the T_p operator on the Fourier coefficients of a modular form. This is a direct calculation using Theorem 3.4.3. Observe that $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_1(N)$ and so any $f \in \mathcal{M}_k(\Gamma_1(N))$ has period 1.

Proposition 3.4.8. Let $f \in \mathcal{M}_k(\Gamma_1(N))$ such that the Fourier expansion of f is given by the following expression:

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

Let $\mathbb{I}_N$ be the trivial character mod N. Then the Fourier expansion of $T_p(f)$ is given by the expression below:

$$T_pf(\tau) = \sum_{n=0}^{\infty} a_{np}(f)q^n + \mathbb{I}_N(p)p^{k-1}\sum_{n=0}^{\infty} a_n(\langle p\rangle f)q^{np}$$

Proof. We set some notation first. Let $\beta_{2,j}$ with $0 \le j < p$ and $\beta_{2,\infty}$ be matrices given below.

$$\beta_{2,j} = \begin{pmatrix} 1 & j \\ 0 & p \end{pmatrix} \ ; \ \beta_{2,\infty} = \begin{pmatrix} m & n \\ N & p \end{pmatrix} \tag{3.4.9}$$

The calculation splits into two parts, depending on whether p divides N or not. In the former case,

$$T_p(f) = \sum_{j=0}^{p-1} f[\beta_{2,j}]_k \tag{3.4.10}$$

In terms of the Fourier coefficients of f, the Fourier series of $f[\beta_{2,j}]_k$ for every j is given by the expression:

$$f[\beta_{2,j}]_k = \frac{1}{p}\sum_{n=0}^{\infty} a_n(f)e^{2\pi in(\tau/p)}e^{2\pi in(j/p)} \tag{3.4.11}$$

Write $q_p = e^{2\pi in(\tau/p)}$ and $\mu_p = e^{2\pi i/p}$. Summing over j in the expression (4.4.10) we get,

$$\sum_{j=0}^{p-1} f[\beta_{2,j}]_k = \left(\frac{1}{p}\sum_{n=0}^{\infty} a_n(f)q_p^n\right)\sum_{j=0}^{p-1}\mu_p^{nj} \tag{3.4.12}$$

Notice that, the geometric series $\sum_j \mu_p^{nj}$ evaluates to 0 when $p \nmid n$ while when $p \mid n$, then it is equal to p. This gives us that the expression in (4.3.18) equals $\sum_{n=0}^{\infty} a_{np}(f)q^n$.

In the case $p \nmid N$, we need to also consider the term $f[\beta_{2,\infty}]_k$.

$$\begin{aligned} f[\beta_{2,\infty}]_k &= f\left[\begin{pmatrix} m & n \\ N & p \end{pmatrix}\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\right]_k \\ &= p^{k-1}\sum_{n=0}^{\infty} a_n(\langle p\rangle f)q^{np} \end{aligned}$$

Clubbing the two, we get the required expression.

□

Next, suppose that $f \in \mathcal{M}_k(N,\chi)$. Using Proposition 3.4.5, we see that,

$$\begin{aligned} \langle d\rangle(T_pf) &= T_p(\langle d\rangle f) \\ &= \chi(d)(T_pf) \end{aligned}$$

This implies that $T_pf \in \mathcal{M}_k(N,\chi)$. Proposition 3.4.8 can therefore be specialized for f in $\mathcal{M}_k(N,\chi)$ to give the following corollary:

Corollary 3.4.13. Let $\chi : (\mathbb{Z}/N\mathbb{Z})^*$ be Dirichlet character modulo N. Let $f \in \mathcal{M}_k(N,\chi)$ such that the Fourier expansion of f is given by the following expression:

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

Then $T_p(f) \in \mathcal{M}_k(N,\chi)$ and the Fourier expansion of $T_p(f)$ is given by the expression below:

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

A very useful property of the the Hecke operators is that they commute. We already showed that $\langle d\rangle$ and T_p commute with each other. We further prove the following proposition:

Proposition 3.4.14. Let d and e belong to $(\mathbb{Z}/N\mathbb{Z})^*$ and p and q be primes. Then the operators $\langle d\rangle$ and $\langle e\rangle$ commute. That is, $\langle d\rangle\langle e\rangle = \langle e\rangle\langle d\rangle$ Similarly, $T_pT_q = T_qT_p$.

Proof. Because of Theorem 3.2.19, it suffices to check commutativity of the operators for any $f \in \mathcal{M}_k(N,\chi)$ where χ is a Dirichlet character modulo N. This way the first part is clear.

For the T_p operator, using the calculations done in Theorem 3.4.8, we can write $a_n(T_pf) = a_np(f) + \chi(p)p^{k-1}a_{n/p}(f)$ where $a_{n/p}$ is 0 whenever n/p is not an integer. Therefore,

$$a_n(T_p(T_qf)) = a_np(T_pf) + \chi(p)p^{k-1}a_{n/p}(T_qf) =$$
$$a_{npq}(f) + \chi(q)q^{k-1}a_{np/q}(f) + \chi(p)p^{k-1}a_{nq/p}(f) + \chi(p)\chi(q)(pq)^{k-1}a_{n/qp}(f)$$

Since the last expression is symmetric in p and q, we would have landed up with the same coefficient if we had computed instead $T_q(T_pf)$. ☐

3.4.1 Interpretation of T_p and $\langle d\rangle$ in terms of in terms of Divisor groups of Moduli spaces.

Via the one one correspondence between the points on the modular curve and its corresponding moduli space observed in section 3.1, we can interpret T_p as an operator between the moduli spaces as well.

Let $Div(S_1(N))$ denote the divisor group of $S_1(N)$. Let $\beta_{2,j}$ and $\beta_{2,\infty}$ be matrices which appear in the definition T_p given in (3.4.9). The modular curve interpretation of T_p is:

$$T_p : Div(X_1(N)) \to Div(X_1(N))$$

$$\Gamma_1(N)\tau \mapsto \sum_j \Gamma_1(N)\beta_{2,j}(\tau)$$

We include $\beta_{2,\infty}$ when $p \nmid N$. The corresponding map on the moduli space is:

$$[\mathbb{C}/\Lambda_\tau, 1/N + \Lambda_\tau] \mapsto \sum_j [\mathbb{C}/\Lambda_{\beta_{2,j}(\tau)}, 1/N + \Lambda_{\beta_{2,j}(\tau)}]$$

Writing $\mathbb{C}/\Lambda_\tau$ as E and the order N point Q, we show that this map is equivalent to the following map:

$$\psi : Div(S_1(N)) \to Div(S_1(N))$$

$$[E, Q] \mapsto \sum_C [E/C, Q + C]$$

where the sum is taken over all the order p subgroups $C \subset E$ such that $C \cap \langle Q \rangle = \{0_E\}$. In order to show this, to each $\beta_{2,j}$, we associate a subgroup to C_j such that $\mathbb{C}/\Lambda_{\beta_{2,j}(\tau)} \cong E/C_j$. We then show that these are all the subgroups of order p in E satisfying $C \cap \langle Q \rangle = \{0_E\}$. As a start, consider the lattice $\Lambda_{\beta_{2,j}} = \frac{\tau+j}{p}\mathbb{Z} \oplus \mathbb{Z}$. We next establish the following claim:

Claim 3.4.15. Let $0 \leq j < (p-1)$. Then $\frac{\tau+j}{p}\mathbb{Z} \oplus \mathbb{Z} = \frac{\tau+j}{p}\mathbb{Z} + \tau\mathbb{Z} \oplus \mathbb{Z}$.

Proof of the claim. One way containment is clear. To show that $\frac{\tau+j}{p}\mathbb{Z} + \tau\mathbb{Z} \oplus \mathbb{Z} \subset \frac{\tau+j}{p}\mathbb{Z} \oplus \mathbb{Z}$, write $\tau = \frac{p(\tau+j)}{p} - j$. Therefore,

$$\begin{aligned} \frac{\tau+j}{p}\mathbb{Z} + \tau\mathbb{Z} \oplus \mathbb{Z} &= \frac{\tau+j}{p}\Big(\mathbb{Z} + p\mathbb{Z}\Big) - j\mathbb{Z} \oplus \mathbb{Z}. \\ &= \frac{\tau+j}{p}\Big(\mathbb{Z} + p\mathbb{Z}\Big) \oplus \mathbb{Z} - j\mathbb{Z} \\ &\subseteq \frac{\tau+j}{p}\mathbb{Z} \oplus \mathbb{Z} \end{aligned}$$

In the second last equality the $j\mathbb{Z}$ term gets absorbed in the right hand side, because, the direct sum of two disjoint sets here is essentially the same as their complex addition. $\square$

It follows that $\Lambda_{\beta_{2,j}}$ corresponds to the set $\left\langle \frac{\tau+j}{p} \right\rangle + \Lambda_\tau$ in $\mathbb{C}$. This set contains Λ_τ and hence, is a superlattice of Λ_τ. To see that it corresponds to the unique subgroup $C_j = \left\langle \frac{\tau+j}{p} + \lambda_\tau \right\rangle$ in $\mathbb{C}/\Lambda_\tau$, we may define a map,

$$\psi : \mathbb{C}/\Lambda_\tau \to \mathbb{C}/\Lambda_{\beta_{2,j}(\tau)}$$

$$z + \Lambda_\tau \mapsto z + \Lambda_{\beta_{2,j}(\tau)}$$

The map is clearly surjective and the kernel is the cyclic group C_j. When $p \nmid N$, we need to consider the lattice $\Lambda_{\beta_{2,\infty}}$ as well in the sum.

Let $y = Np\tau + p$. Then, $y\Lambda_{\beta_{2,\infty}} = (Np\tau + p)\mathbb{Z} \oplus (mp\tau + n)\mathbb{Z} = \mathbb{Z} \oplus (p\tau)\mathbb{Z}$. This gives us that $y/p(\Lambda_{\beta_{2,\infty}}) = (1/p)\mathbb{Z} \oplus \tau\mathbb{Z}$. Clearly, this is equal to the set $\langle 1/p \rangle + \Lambda_\tau$ in $\mathbb{C}$. Let

$C_\infty = \langle 1/p + \Lambda_\tau \rangle$ in $\mathbb{C}/\Lambda_\tau$. Using the same argument done in the case of $\Lambda_{\beta_{2,j}}$, we see that E/C_∞ isomorphic to $\mathbb{C}/\Lambda_{\beta_{2,\infty}}$.

By noticing the non-trivial term of τ/p in every element of C_j, it is easy to see that $C_j \cap \langle 1/N + \Lambda_\tau \rangle = \{0\}$ for $0 \leq j < p$. When $p \nmid N$, we further show that $C_\infty \cap \langle 1/N + \Lambda_\tau \rangle = 0$. Suppose that $n/p + \Lambda_\tau = m/N + \Lambda_\tau$ for some integers n and m. Then $(nN - mp)/pN \in \Lambda_\tau$. Thus, $pN \mid (nN - mp)$ giving us that $pNx = nN - mp$ for some integer x. Since $p \nmid N$, p must divide n and so $n/p + \Lambda_\tau = m/N + \Lambda_\tau = 0$ in $\mathbb{C}/\Lambda_\tau$.

One can easily check that the p-cyclic subgroup C_j for each $0 \leq j < p$ is disjoint from C_∞ because of the non zero term of some multiple of τ/p in each element of C_j. To check that $C_j \cap C_i = 0$ for $0 \leq i, j < p$, and $i \neq j$, write $m(\tau + j)/p + \Lambda_\tau = n(\tau + i)/p + \Lambda_\tau$ for some integers m and n. This gives us that $p \mid (m - n)$ and $p \mid (jm - in)$. Write $m - n = pt$ and $px = jm - ni$ for some integers t and x. Substituting the first expression in the second, it follows that either $p \mid n$ or $p \mid (j - i)$. Since $0 < i - j < p$, the latter case is not possible. Therefore, $p \mid n$ and $p \mid m$ giving us that $C_j \cap C_i = 0$ in $\mathbb{C}/\Lambda_\tau$.

Finally we wish to show that if C is any subgroup of E such that $C \cong \mathbb{Z}/p\mathbb{Z}$ satisfying that $C \cap Q = 0_E$, then C has to be one $C_j; 0 \leq j < p$ or C_∞. This follows from the following claim:

Claim 3.4.16. Let $E_\tau[p]$ denote p torsion points of of E, a subgroup isomorphic to $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. Then $E_\tau[p] = C_1 \oplus \cdots \oplus C_\infty$.

Proof of the claim. The number of elements in $E_\tau[p]$ is p^2. Since the p cyclic subgroups we discovered are pairwise disjoint with the only common element to be 0, counting the number of elements in all of them comes out to be $1 + (p-1)(p+1) = p^2$. This establishes that $E_\tau[p] = C_1 + C_2 \cdots + C_\infty$. Now, suppose that $x \in C_i \cap (C_1 + \cdots + C_{i-1} + C_{i+1} + \cdots + C_\infty)$. This implies that $\langle x \rangle = C_i \subseteq (C_1 + \cdots + C_{i-1} + C_{i+1} + \cdots + C_\infty)$, which gives us that $|E_\tau[p]| = |C_1 + \cdots + C_{i-1} + C_{i+1} + \cdots + C_\infty| = p^2 - (p-1)$. This is a contradiction to the counting argument done at the beginning of the proof. $\square$

From the discussion above, it now follows that T_p as a map on the divisor group of moduli spaces maps $[E, C]$ to $\sum_C [E/C, Q + C]$ where C are all the cyclic subgroups of order p not intersecting $\langle Q \rangle$.

It is easier to view the map $\langle d \rangle$ on the divisor group of the corresponding moduli space. Suppose $\gamma \in \Gamma_0(N)$ with lower right entry as d mod N. As seen before, the diamond operator is the isomorphism:

$$X(\Gamma_1(N)) \to X(\Gamma_1(N))$$

$$\Gamma_1(N)\tau \mapsto \Gamma_1(N)\gamma\tau$$

The corresponding map on the moduli space $S_1(N)$ is:

$$[C/\Lambda_\tau, 1/N + \Lambda_\tau] \mapsto [C/\Lambda_{\gamma\tau}, 1/N + \Lambda_{\gamma\tau}]$$

We show that this map is equivalent to the mapping: $[E, Q] \mapsto [E, dQ]$. Using the methods outlined in section 3.1, we proceed in the same manner. Suppose that $\gamma = \begin{pmatrix} a & b \\ cN & d+Nk \end{pmatrix} \in \Gamma_0(N)$ be the representative of $\langle d \rangle$. Let $m = Nc\tau + (d + Nk)$. Then $m\Lambda_{\gamma\tau} = Nc\tau\mathbb{Z} \oplus (a\tau + (d + Nk))\mathbb{Z} = \mathbb{Z} \oplus \tau\mathbb{Z}$. This implies that $\Lambda_{\gamma\tau} \cong \Lambda_\tau$ under the "multiplication by m" map. Under this map $1/N + \Lambda_\tau \mapsto (cN\tau + d + Nk)/N + \Lambda_\tau$. This is exactly the point $d/N + \Lambda_\tau$. It follows that the diamond operator maps the point $[E, Q]$ to the point $[E, dQ]$ in $S_1(N)$.

Therefore., there is more than one way to interpret the T_p and the diamond operator. In addition to viewing them as operators on the space of modular forms, we can also view them as maps between the corresponding moduli spaces. For the topics covered ahead, we will however adopt the former interpretation of the operators.

3.5 The T_n and $\langle n \rangle$ operator for $n \in \mathbb{Z}^+$.

We aim to extend the definition of T_p and $\langle d \rangle$ to all the positive integers. In case of the diamond operator, this is easily extended to any $n \in \mathbb{Z}^+$ with $(n, N) = 1$ by defining $\langle n \rangle = \langle n \bmod N \rangle$. For n such that $(n, N) > 1$, we define $\langle n \rangle$ to be just the 0 operator on $\mathcal{M}_k(\Gamma_1(N))$.

We extend the T_p operator inductively for prime powers, by defining the relation:

$$T_1 = \mathbb{I} \; ; \; T_{p^r} = T_p T_{p^{r-1}} - p^{k-1}\langle p \rangle T_{p^{r-1}} \quad \text{for } r \geq 2 \tag{3.5.1}$$

By induction on r and s, Proposition 3.4.14 helps us conclude that $T_{p^r}T_{q^s} = T_{q^s}T_{p^r}$. For all $n \in \mathbb{Z}^+$, we extend the definition multiplicatively to all positive integers as follows: Suppose $n = \prod_i p_i^{e_i}$ where p_i denote the distinct primes dividing n. Then define

$$T_n = \prod_i T_{p_i{}^{e_i}} \tag{3.5.2}$$

It follows that if n and m are positive integers with $(n, m) = 1$, then,

$$T_{nm} = T_m T_n. \tag{3.5.3}$$

We will need the relation between Fourier coefficients of $T_n(f)$ and that of f which we compute below.

Proposition 3.5.4. Let $f \in \mathcal{M}_k(\Gamma_1(N))$ with the Fourier expansion given by the following expression:

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

Then for all $n \in \mathbb{Z}^+$, the coefficients of the Fourier expansion of $T_n(f)$ is given by the expression below:

$$a_m(T_n f) = \sum_{d|(m,n)} d^{k-1} a_{mn/d^2}(\langle d\rangle f) \tag{3.5.5}$$

In particular, if χ is a Dirichlet character modulo N and $f \in \mathcal{M}_k(N, \chi)$, then,

$$a_m(T_n f) = \sum_{d|(m,n)} \chi(d) d^{k-1} a_{mn/d^2}(f) \tag{3.5.6}$$

Proof. Because of the decomposition in Theorem 3.2.19, we just show the particular case in (3.5.6). We first prove it for prime powers and then generalize the result for any positive integer n. The result is trivially true when $n = 1$. When $n = p$, the right hand side is given by the expression: $\sum_{d|(m,p)} \chi(d)d^{k-1}a_{mp/d^2}(f)$. This equals $a_{mp}(f)$ when $p \nmid m$. While, when $p \mid m$, we get that $a_m(T_p f) = a_{mp}(f) + \chi(p)p^{k-1}a_{n/p}(f)$. Putting together the two cases we exactly the expression of $a_m(T_p f)$ as calculated in Corollary 3.4.13.

To see the result for prime powers, we use induction. Assume that the result holds for $n = 1, p, \ldots, p^{r-1}$. For $r \geq 2$, we use the recursion formula in (3.5.1) to expand $a_m(T_{p^r} f) = a_m(T_p T_{p^{r-1}} f) - p^{k-1} a_m(\langle p\rangle T_{p^{r-1}} f)$. Using the formula for T_p we further write this as:

$$a_{mp}(T_{p^{r-1}} f) + \chi(p)p^{k-1}a_{m/p}(T_{p^{r-1}} f) - \chi(p)p^{k-1}a_m(T_{p^{r-2}} f). \tag{3.5.7}$$

At this stage we use the induction hypothesis to expand (3.5.7) to get three summands:

$$\sum_{d|(mp,p^{r-1})} \chi(d)d^{k-1}a_{mp^r/d^2}(f) + \chi(p)p^{k-1} \sum_{d|(m/p,p^{r-1})} \chi(d)d^{k-1}a_{mp^{r-2}/d^2}(f)$$
$$- \chi(p)p^{k-1} \sum_{d|(m,p^{r-2})} \chi(d)d^{k-1}a_{mp^{r-2}/d^2}(f)$$

We look at these three summands separately. Write the first summand as:

$$\sum_{d|(mp,p^{r-1})} \chi(d)d^{k-1}a_{mp^r/d^2}(f) = a_{mp^r} + \sum_{\substack{d|(mp,p^{r-1}) \\ d>1}} \chi(d)d^{k-1}a_{mp^{r-2}/d^2}(f) \tag{3.5.8}$$

Compute the last sum as:

$$\chi(p)p^{k-1}\sum_{d|(m,p^{r-2})}\chi(d)d^{k-1}a_{\frac{mp^{r-2}}{d^2}}(f)=\sum_{d|(m,p^{r-2})}\chi(pd)(pd)^{k-1}a_{\frac{mp^{r-2}}{d^2}}(f)$$

Changing the variable shows that this is equal to the sum as in (3.5.8):

$$\sum_{\substack{h|(m,p^{r-1})\\ h>1}}\chi(h)(h)^{k-1}a_{\frac{mp^{r-1}}{d^2}}(f)$$

and so it gets cancelled in the overall sum. Combining the calculations done so far we get,

$$\begin{aligned}
a_m(T_{p^r}) &= a_{mp^r}+\chi(p)p^{k-1}\sum_{d|(m/p,p^{r-1})}\chi(d)d^{k-1}a_{mp^{r-2}/d^2}(f)\\
&= a_{mp^r}+\sum_{d|(m/p,p^{r-1})}\chi(pd)(pd)^{k-1}a_{mp^{r-2}/d^2}(f)\\
&= a_{mp^r}+\sum_{\substack{h|(m/p,p^{r-1})\\ h>1}}\chi(h)h^{k-1}a_{mp^{r-2}/d^2}(f)
\end{aligned}$$

The last expression is exactly the result we want. To generalize to arbitrary n, take $(n_1,n_2)\in\mathbb{Z}^+$ such that $(n_1,n_2)>1$. Compute,

$$\begin{aligned}
a_m(T_{n_1n_2}f) &= \sum_{d|(m,n_1)}\chi(d)d^{k-1}a_{mn_1/d^2}(T_{n_2}f)\\
&= \sum_{de|(m,n_1)}\sum_{e|(m,n_1/d^2)}\chi(de)(de)^{k-1}a_{mn_1n_2/(de)^2}(f)\\
&= \sum_{de|(m,n_1n_2)}\chi(de)(de)^{k-1}a_{mn_1n_2/(de)^2}(f)\\
&= \sum_{h|(m,n_1n_2)}\chi(h)h^{k-1}a_{mn_1n_2/h^2}(f)
\end{aligned}$$

The formula for $a_m(T_n)$ for any arbitrary $n\in\mathbb{Z}^+$ now follows clearly. □

3.6 A word on Eisenstein series of level N

The theory developed in the next few sections will focus more on the space of cusp forms. This section provides some justification for this so that the reader does not lose any continuity of thought. To that end, we digress a bit to make some remarks on Eisenstein series discussed in section 1.3. These are modular forms with respect to the whole group $SL_2(\mathbb{Z})$ and therefore we call them *Eisenstein series of level* 1. We explicitly calculate the Fourier series of $G_k(\tau)$. In addition to giving some intuition about Eisenstein series of level N, it will also be useful in chapter 4. Write,

$$\begin{aligned} G_k(\tau) &= \sum_{\substack{m,n\in\mathbb{Z}\\(m,n)\neq(0,0)}} \frac{1}{(m\tau+n)^k} \\ &= 2\left(\sum_{n\geq 1}\frac{1}{n^k} + \sum_{m\geq 1}\left(\sum_{n\in\mathbb{Z}}\frac{1}{(m\tau+n)^k}\right)\right) \end{aligned}$$

From the above expression, we see that as $\tau \to i\infty$, G_k takes the value $2\zeta(k)$, where the symbol ζ denotes the zeta function $\sum_{n\geq 0}\frac{1}{n^k}$. This is the first term a_0 of the Fourier expansion of G_k for $k \geq 3$. To find out the other terms, we need the following formula from complex analysis for $\tau \in \mathbb{H}$ and $k \geq 3$:

$$\sum_{n\in\mathbb{Z}}\frac{1}{(\tau+n)} = \frac{(-2\pi i)^k}{(k-1)!}\sum_{m=1}^{\infty} m^{k-1}q^m \ ; \ q = e^{2\pi i\tau}. \tag{3.6.1}$$

Using the formula in 3.6.1, we see that $G_k(\tau)$ can be written as:

$$\begin{aligned} G_k(\tau) &= 2\zeta_k + \left(\frac{2(-2\pi i)^k}{(k-1)!}\sum_{m\geq 1}\sum_{n\geq 1} n^{k-1}e^{2\pi in(m\tau)}\right) \\ &= 2\zeta(k) + \frac{2(-2\pi i)^k}{(k-1)!}\sum_{r\geq 1}\left(\sum_{d|n} d^{k-1}\right)e^{2\pi ir\tau} \\ &= 2\zeta(k) + C_k\sum_{n\geq 1}\sigma_{k-1}(n)q^n \end{aligned}$$

In the last expression $C_k = \frac{2(-2\pi i)^k}{(k-1)!}$ and $\sigma_{k-1}(n) = \sum_{d|n} d^{k-1}$.

For any integer $k < 4$, from the dimension formula, we know that the space $\mathcal{M}_k(SL_2(\mathbb{Z})) = \{0\}$. In the case when $k \geq 4$, for any modular form $f \in \mathcal{M}_k(SL_2(\mathbb{Z}))$, we can find some complex number c such that $f - cG_k \in \mathcal{S}_k(SL_2(\mathbb{Z}))$. This implies that $\mathcal{M}_k(SL_2(\mathbb{Z})) = \mathcal{S}_k(SL_2(\mathbb{Z}))\bigoplus \mathbb{C}G_k$. In the case of level one modular forms, this decomposition is fairly easy to see.

However, for subgroup $\Gamma(N)$, the space $\mathcal{M}_k(\Gamma(N))$ splits as a direct sum of the space $\mathcal{S}_k(\Gamma(N))$ and its complement which we call the Eisenstein space and denote by $\mathcal{E}_k(\Gamma(N))$. We can explicitly find out a basis for this space. One can do the same for the subgroup $\Gamma_1(N)$ as well. The basis elements are variants of the Eisenstein series of level 1. We state this below without any proof. To see the construction of Eisenstein series for $\Gamma(N)$ and $\Gamma_1(N)$, one can refer to chapter 4 of Diamond and Shurman [2].

Theorem 3.6.2. Let $k \geq 3$ and $\overline{v} \in (\mathbb{Z}/N\mathbb{Z})^2$ be a point of order N, then the Eisenstein series for $\Gamma(N)$ are defined by the following:

$$G_k^{\overline{v}}(\tau) = \sum_{\substack{(c,d)\equiv v(N) \\ (c,d)\neq(0,0)}} \frac{1}{(c\tau + d)^k}$$

If $\overline{v} = (\overline{c_v}, \overline{d_v})$, where (c_v, d_v) is a lift of $\overline{v}$ in $\mathbb{Z}^2$, then the Fourier expansion of $G_k^{\overline{v}}(\tau)$ is computed as:

$$G_k^{\overline{v}}(\tau) = \delta(\overline{c_v})\zeta^{\overline{d_v}}(k) + \frac{C_k}{N^k}\sum_{n=1}^{\infty} \sigma_{k-1}^{\overline{v}}(n) q_N^n \ ; \ q_N^n = e^{2\pi i\tau/N},$$

where,

$$\delta(\overline{c_v}) = \begin{cases} 1 \text{ if } \overline{c_v} = 0 \\ 0 \text{ otherwise} \end{cases} \quad ; \ \zeta^{\overline{d_v}}(k) = \sum_{\substack{d\equiv d_v(N) \\ d\neq 0}} \frac{1}{d^k}$$

and

$$C_k = \frac{2(-2\pi i)^k}{(k-1)!} \ ; \ \sigma_{k-1}(n) = \sum_{\substack{m|N \\ n/m\equiv c_v(N)}} sgn(m) m^{k-1} e^{2\pi i d_v m/N}.$$

The bases of the Eisenstein space $\mathcal{E}_k(\Gamma(N))$ are constructed from the elements described in Theorem 3.6.2.

We know that the space $\mathcal{M}_k(\Gamma_1(N))$ is decomposed into χ-eigen spaces given in Theorem 3.2.19. The following theorem describes the Fourier expansion of Eisenstein series in the space $\mathcal{M}_k(N, \chi)$.

Theorem 3.6.3. Let $k \geq 3$. Let φ modulo u and ψ modulo v be any two Dirichlet characters such that $uv = N$ and φ is primitive, then the Eisenstein series for the subgroup $\Gamma_1(N)$, denoted by $E_k^{\psi,\varphi}(\tau)$ is given by the following Fourier expansion:

$$E_k^{\psi,\varphi}(\tau) = \delta(\psi)L(1-k,\varphi) + 2\sum_{n=1}^{\infty} \sigma_{k-1}^{\psi,\varphi}(n) q_N^n \ ; \ q^n = e^{2\pi i\tau/N},$$

where,

$$\sigma_{k-1}^{\psi,\varphi}(n) = \sum_{\substack{m|n \\ m>0}} \psi(n/m)\varphi(m)m^{k-1} \ ; \ L(1-k,\varphi) = \sum_{n=0}^{\infty} \frac{\varphi(n)}{n^{1-k}}$$

The bases of $\mathcal{E}_k(\Gamma_1(N))$ for $k \geq 3$ are constructed from the elements mentioned in the above theorem. For cases $k = 1$ and $k = 2$, a different process is followed to construct Eisenstein series which is outlined in chapter 4 of [2].

We conclude by remarking that the bases for the space of cusp forms are not so easy write down explicitly, as in the case of the Eisenstein space. Therefore, we develop some sophisticated machinery in order to study cusp forms in the next few sections.

3.7 Petersson inner product

The space of cusp forms can be equipped with an inner product, which will be defined as an integral. In order for the integral to make sense, we work with a measure which is invariant under the action of $SL_2(\mathbb{Z})$. Let $V \in \mathbb{C}$ be an open set. Consider the two form on $V : \omega = (dz \wedge d\bar{z})/\operatorname{Im}(z)^2$. Writing $z = x + iy$, we get $\omega = -2i(dx \wedge dy)/y^2$. We show that ω is invariant under the action of $SL_2(\mathbb{Z})$. More generally, for any $\alpha \in GL_2^+(\mathbb{R})$, we compute

$$\begin{aligned} d(\alpha z) \wedge d(\overline{\alpha z}) &= \frac{(\det \alpha)^2}{|cz + d|^4} dz \wedge d\bar{z} \\ &= \left(\frac{\operatorname{Im}(\alpha z)}{\operatorname{Im}(z)}\right)^2 dz \wedge d\bar{z} \end{aligned}$$

We work with a suitable multiple of ω given by the following expression:

$$d\mu(z) = \frac{-(dz \wedge d\bar{z})}{2i \operatorname{Im}(z)^2} = \frac{dxdy}{y^2} \;\; ; \;\; z = x + iy \tag{3.7.1}$$

The measure $d\mu(z)$ is called the *hyperbolic measure* on the upper half plane. We first study some properties of integrating functions on the upper half plane with respect to this measure.

Remark 3.7.2. Since the set $\mathbb{Q} \cup \{\infty\}$ is a countable set of measure 0, we can as well integrate over the extended upper plane $\mathbb{H}^*$.

Lemma 3.7.3. Suppose $\varphi : \mathbb{H} \to \mathbb{C}$ is a continuous and bounded function and $\alpha \in SL_2(\mathbb{Z})$. Then the integral $\int_{\mathcal{D}^*} \varphi(\alpha(\tau)) d\mu(\tau)$ converges.

Proof. Let $|\varphi(\tau)| \leq M$ for all $\tau \in \mathbb{H}$. Then,

$$\begin{aligned} \left| \int_{\mathcal{D}^*} \varphi(\alpha(\tau)) d\mu(\tau) \right| &\leq \int_{\mathcal{D}^*} \left| \varphi(\alpha(\tau)) d\mu(\tau) \right| \\ &\leq M \int_{\mathcal{D}^*} d\mu(\tau) \\ &= M \int_{-1/2}^{1/2} \int_{\sqrt{1-x^2}}^{\infty} dxdy/y^2 \end{aligned}$$

Using elementary methods from calculus, one sees that the integral in the last equation is equal to $\frac{\pi}{3}$. $\square$

Let $\Gamma \subseteq SL_2(\mathbb{Z})$ be a congruence subgroup and let $\{\alpha_j\}_j$ be the coset representatives of the space $\{\pm I\}\Gamma\backslash SL_2(\mathbb{Z})$. We can write $SL_2(\mathbb{Z}) = \bigcup_j \{\pm I\}\Gamma\alpha_j$. Because of Proposition 2.1.4, it makes sense to define the integral of a function over $X(\Gamma)$ in the following way.

Definition 3.7.4. Let φ be a continuous and bounded function on $\mathbb{H}$. Suppose φ is Γ invariant. Define the integral of φ in the modular curve $X(\Gamma)$ by the following expression:

$$\int_{X(\Gamma)} \varphi(\tau)d\mu(\tau) = \int_{\bigcup \alpha_j(\mathcal{D}^*)} \varphi(\tau)d\mu(\tau) = \sum_j \int_{\mathcal{D}^*} \varphi(\alpha_j\tau)d\mu(\tau) \tag{3.7.5}$$

Since φ is Γ invariant, the integral is independent of the coset representatives we choose.

Putting $\varphi = 1$ in definition 3.7.4, call $V_\Gamma = \int_{X(\Gamma)} d\mu(\tau)$, the *volume* of $X(\Gamma)$. Clearly, $V_\Gamma = [SL_2(\mathbb{Z}) : \Gamma]V_{SL_2(\mathbb{Z})}$, where $V_{SL_2(\mathbb{Z})} = \int_{\mathcal{D}^*} d\mu(\tau)$.

Take f and $g \in \mathcal{S}_k(\Gamma)$. In order to construct the Petersson inner product, we need to cook up a continuous, bounded, and Γ-invariant function with f and g. Consider the following function:

$$\varphi(\tau) = f(\tau)\overline{g(\tau)}(\mathrm{Im}(\tau))^k$$

The function φ is clearly continuous. It is also Γ invariant because for any $\gamma \in \Gamma$,

$$\begin{aligned}\varphi(\gamma\tau) &= f(\gamma\tau)\overline{g(\gamma\tau)}\,\mathrm{Im}(\gamma\tau)^k \\ &= (c\tau+d)^k f(\tau)\overline{(c\tau+d)}^k g(\tau)\,\mathrm{Im}(\tau)|c\tau+d|^{-k} \\ &= \varphi(\tau)\end{aligned}$$

We would finally like to establish that φ is bounded on $\mathbb{H}$. Since φ is Γ invariant, it is enough to show that φ is bounded on $\bigcup_i \alpha_i\mathcal{D}$.

Lemma 3.7.6. For any $\alpha \in SL_2(\mathbb{Z})$, the function $\varphi \circ \alpha$ is bounded on $\mathcal{D}$.

Proof. Since $\varphi \circ \alpha$ is a continuous function, it is bounded on any compact subset of $\mathcal{D}$. We only need to check that as φ is bounded as $\mathrm{Im}(\tau) \to 0$. This is seen as follows: For any $\alpha \in SL_2(\mathbb{Z})$,

$$\varphi(\alpha(\tau)) = (f[\alpha]_k g[\alpha]_k)(\tau)\,\mathrm{Im}(\tau)^k.$$

By definition, $f[\alpha]_k$ and $g[\alpha]_k$ have power series expansions in $q_h = e^{2\pi i\tau/h}$ with their a_0 term equal to 0. Therefore, $\varphi(\alpha(\tau)) = \mathcal{O}(q_h)\mathcal{O}(q_h)y^k$. Since $|q_h| \to 0$ when $\mathrm{Im}(\tau) \to \infty$ and the exponential decay dominates polynomial growth, it follows that, as $\mathrm{Im}(\tau) \to \infty, \varphi \circ \alpha \to 0$. This proves that $\varphi \circ \alpha$ is bounded on $\mathcal{D}$. $\square$

In view of the above discussion, we are ready to define the inner product on cusp forms.

Definition 3.7.7 (Petersson Inner Product). Let $\Gamma \subseteq SL_2(\mathbb{Z})$ be a congruence subgroup. The Petersson inner product on $\mathcal{S}_k(\Gamma)$ is given by:

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

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

From the definition it is easy to see that the inner product is linear in f, conjugate linear in g and Hermitian symmetric. To see that it is positive we definite compute

$$\begin{aligned}\langle f, f \rangle &= \frac{1}{V_\Gamma} \int_{X(\Gamma)} |f(\tau)|^2 \,\mathrm{Im}(\tau)^k d\mu(\tau) \\ &= \frac{1}{V_\Gamma} \sum_j \int_{\mathcal{D}^*} |f(\alpha_j \tau)|^2 \,\mathrm{Im}(\alpha_j \tau)^k d\mu(\tau)\end{aligned}$$

Since $\mathcal{D}^*$ is compact, $|f(\alpha_j(\tau)|$ and $\mathrm{Im}(\alpha_j)(\tau)$ has a minimum value, making the integral above greater than or equal to 0.

Lastly, we need to make sure that the Peterson inner product with respect to a subgroup gives us the same value when we take it with respect to the smaller group. Indeed this is true. In fact the normalizing factor $\frac{1}{V_\Gamma}$ precisely helps in ensuring that.

Lemma 3.7.8. Let $\Gamma' \subseteq \Gamma \subseteq SL_2(\mathbb{Z})$. Suppose f and $g \in \mathcal{S}_k(\Gamma)$. Then $\langle f, g \rangle_{\Gamma'} = \langle f, g \rangle_\Gamma$.

Proof. Write $SL_2(\mathbb{Z}) = \bigcup_i (\{\pm I\}\Gamma)\alpha_i$ and $\Gamma = \bigcup_j (\{\pm I\}\Gamma')\beta_j$ with $\alpha_i \in SL_2(\mathbb{Z}); 1 \le i \le [SL_2(\mathbb{Z}) : \Gamma]$ and $\beta_j \in \Gamma; 1 \le j \le [\Gamma : \Gamma']$. This gives us that $SL_2(\mathbb{Z}) = \bigcup_{i,j} (\{\pm I\}\Gamma')\beta_j \alpha_i$. Computing the inner product of f and g with respect to Γ' we get:

$$\begin{aligned}\langle f, g \rangle_{\Gamma'} &= \frac{1}{V_{\Gamma'}} \int_{X(\Gamma')} f(\tau)\overline{g(\tau)} \,\mathrm{Im}(\tau)^k d\mu(\tau) \\ &= \frac{1}{V_{\Gamma'}} \left(\sum_i \sum_j \int_{\mathcal{D}^*} f(\beta_j \alpha_i \tau)\overline{g(\beta_j \alpha_i \tau)} \,\mathrm{Im}(\beta_j \alpha_i \tau)^k d\mu(\tau) \right) \\ &= \frac{1}{V_{\Gamma'}} \left([\Gamma : \Gamma'] \sum_i \int_{\mathcal{D}^*} f(\alpha_i \tau)\overline{g(\alpha_i \tau)} \,\mathrm{Im}(\alpha_i \tau)^k d\mu(\tau) \right) \\ &= \frac{1}{V_\Gamma} \left(\sum_i \int_{\mathcal{D}^*} f(\alpha_i \tau)\overline{g(\alpha_i \tau)} \,\mathrm{Im}(\alpha_i \tau)^k d\mu(\tau) \right) \\ &= \langle f, g \rangle_\Gamma\end{aligned}$$

In the third equality we use the fact that $f(\tau)g(\tau)\,\mathrm{Im}(\tau)$ is Γ-invariant while in next one we write $1/(V_\Gamma') = 1/(V_\Gamma[\Gamma : \Gamma'])$. ∎

3.8 Adjoints of the Hecke operators

In the next few sections, we restrict our attention to $\mathcal{S}_k(\Gamma_1(N))$. We will show that the Hecke operators T_n and $\langle n\rangle$ commute with their adjoints when $(n, N) = 1$. In order to show this, we will compute the adjoints of these operators with respect to the Petersson inner product. By the Spectral Theorem from linear algebra, the space $\mathcal{S}_k(\Gamma_1(N))$ will have an orthogonal basis of simultaneous eigenvectors for these operators. We will call these eigen-vectors as *eigenforms*. We aim to establish the following theorem:

Theorem 3.8.1. The space $\mathcal{S}_k(\Gamma_1(N))$ has an orthogonal basis of simultaneous eigenforms for the Hecke operators $\{\langle n\rangle \mid (n, N) = 1\}$.

Remark 3.8.2. From now on we will use the term "operators away from the level" to refer to operators T_n or $\langle n\rangle$ with $(n, N) = 1$.

To begin with, we need some technical lemmas and definitions.

Definition 3.8.3. Suppose $\Gamma \subseteq SL_2(\mathbb{Z})$ such that $SL_2(\mathbb{Z}) = \bigcup_j \{\pm I\}\Gamma\alpha_j$ and $\alpha \in GL_2^+(\mathbb{Q})$. Then for any continuous, bounded and $\alpha^{-1}\Gamma\alpha$ – invariant function $\varphi : \mathbb{H} \to \mathbb{H}$ define,

$$\int_{\alpha^{-1}\Gamma\alpha} \varphi(\tau)d\mu(\tau) = \sum_j \int_{\mathcal{D}^*} \varphi(\alpha^{-1}\alpha_j(\tau))d\mu(\tau).$$

To see where this definition is coming from, consider the map $\alpha : \mathbb{H} \to \mathbb{H}$ such that $\tau \mapsto \alpha\tau$ for $\alpha \in GL_2^+(\mathbb{Q})$. It is easy to check that this induces a bijection from $\mathbb{H}^*/\alpha^{-1}\Gamma\alpha \to X(\Gamma)$ where the map is given by $[\tau] \mapsto [\alpha\tau]$. It is one to one and well defined because $[\tau_1] = [\tau_2]$ in $\mathbb{H}^*/\alpha^{-1}\Gamma\alpha$ if and only if $\tau_1 = \alpha^{-1}\gamma\alpha\tau_2$ for some $\gamma \in \Gamma$ if and only if $[\alpha\tau_1] = [\alpha\tau_2]$ in $X(\Gamma)$. It is clearly surjective as $[\alpha^{-1}\tau]$ maps to $[\tau]$ in $X(\Gamma)$. If $SL_2(\mathbb{Z}) = \bigcup_j \{\pm I\}\Gamma\alpha_j$, then Proposition 2.1.4 and the above bijection imply that $\mathbb{H}^*/\alpha^{-1}\Gamma\alpha = \bigcup_j \alpha^{-1}\alpha_j(\mathcal{D}^*)$ up to some boundary identifications. Therefore, definition 3.8.3 makes sense.

We will use the following three results in this section:

Lemma 3.8.4. Let $\Gamma \subseteq SL_2(\mathbb{Z})$ be a congruence subgroup and let $\alpha \in GL_2^+(\mathbb{Q})$.

1. Suppose $\varphi : \mathbb{H} \to \mathbb{C}$ is continuous, bounded and Γ-invariant. Then,

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

2. If $\alpha^{-1}\Gamma\alpha \subseteq SL_2(\mathbb{Z})$, then $V_{\alpha^{-1}\Gamma\alpha} = V_\Gamma$ and $[SL_2(\mathbb{Z}) : \alpha^{-1}\Gamma\alpha] = [SL_2(\mathbb{Z}) : \Gamma]$

3. There exists $\beta_1 \dots \beta_n \in GL_2^+(\mathbb{Q})$, where $n = [\Gamma : \alpha^{-1}\Gamma\alpha \cap \Gamma] = [\Gamma : \alpha\Gamma\alpha^{-1} \cap \Gamma]$, such that the double coset $\Gamma\alpha\Gamma$ is expressed as a disjoint union as follows:

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

Proof. Part 1 comes directly from definition 3.8.3 and 3.7.4.

To see Part 2, let $\varphi = 1$ in part 1 to get the equality of the volumes of the two subgroups. Writing $[SL_2(\mathbb{Z}) : \alpha^{-1}\Gamma\alpha]V_{SL_2(\mathbb{Z})} = V_{\alpha^{-1}\Gamma\alpha} = V_\Gamma = [SL_2(\mathbb{Z}) : \Gamma]V_{SL_2(\mathbb{Z})}$ we get the other equality.

Part 3 requires some work. By replacing Γ with the subgroup $\Gamma \cap \alpha^{-1}\Gamma\alpha$, part 2 helps us to conclude that the index of $\Gamma \cap \alpha^{-1}\Gamma\alpha$ is same as the index of $\alpha\Gamma\alpha^{-1} \cap \Gamma$ in $SL_2(\mathbb{Z})$. Using the fact that $[SL_2(\mathbb{Z}) : \Gamma][\Gamma : \Gamma \cap \alpha^{-1}\Gamma\alpha] = [SL_2(\mathbb{Z}) : \Gamma \cap \alpha^{-1}\Gamma\alpha]$ and similarly writing the relation for $\alpha\Gamma\alpha^{-1} \cap \Gamma$, we see that the index of the two subgroups in Γ is equal as well. This allows us to find $\gamma_1, \dots \gamma_n$, and $\delta_1 \dots \delta_n$ such that,

$$\Gamma = \bigcup_j (\alpha^{-1}\Gamma\alpha \cap \Gamma)\gamma_j = \bigcup_j (\alpha\Gamma\alpha^{-1} \cap \Gamma)\delta_j^{-1}$$

Setting $\Gamma_1 = \Gamma_2 = \Gamma$ in Lemma 3.3.4, we see that $\Gamma\alpha\Gamma = \bigcup_j \Gamma\alpha\gamma_j$ and $\Gamma\alpha^{-1}\Gamma = \bigcup_j \Gamma\alpha^{-1}\delta_j^{-1}$. Taking the inverse in the second relation we conclude that $\Gamma\alpha\Gamma = \bigcup_j \delta_j\alpha\Gamma$. We are almost done since we have obtained two disjoint unions of one right and one left coset space, representing the double coset,

$$\Gamma\alpha\Gamma = \bigcup_j \Gamma\alpha\gamma_j = \bigcup_j \delta_j\alpha\Gamma. \tag{3.8.5}$$

In order to locate suitable representatives β_j, $1 \leq j \leq n$, we prove that for every j, $\Gamma\alpha\gamma_j \cap \delta_j\alpha\Gamma \neq \phi$. Suppose otherwise. Then for some j, the intersection $\Gamma\alpha\gamma_j \cap \delta_j\alpha\Gamma = \phi$. It follows that $\Gamma\alpha\gamma_j \subset \bigcup_{i\neq j} \delta_i\alpha\Gamma$ and so $\Gamma\alpha\Gamma \subset \bigcup_{i\neq j} \delta_i\alpha\Gamma$, a contradiction to the decomposition of the double coset into n orbits. From the non-empty intersection $\Gamma\alpha\gamma_j \cap \delta_j\alpha\Gamma$, for each j we can choose β_j so that the relation (3.8.8) reduces to:

$$\Gamma\alpha\Gamma = \bigcup_j \Gamma\beta_j = \bigcup_j \beta_j\Gamma.$$

□

We now come to an important proposition which will help us to compute adjoint of the double coset operator $[\Gamma\alpha\Gamma]_k$. Specializing the proposition for $\Gamma = \Gamma_1(N)$ and taking a suitable α will help us to compute the adjoints of the Hecke operators.

Proposition 3.8.6. Suppose $\Gamma \subseteq SL_2(\mathbb{Z})$ be a congruence subgroup and $\alpha \in GL_2^+(\mathbb{Q})$. Let $\alpha' = \det(\alpha)\alpha^{-1}$. Then

1. If $\alpha^{-1}\Gamma\alpha \subset SL_2(\mathbb{Z})$, then for all $f \in \mathcal{S}_k(\Gamma)$ and $g \in \mathcal{S}_k(\alpha^{-1}\Gamma\alpha)$, we obtain the following:
$$\langle f[\alpha]_k, g\rangle_{\alpha^{-1}\Gamma\alpha} = \langle f, g[\alpha']_k\rangle_\Gamma$$
In particular, if $\alpha^{-1}\Gamma\alpha = \Gamma$, then the adjoint of the operator $[\alpha]_k$ denoted by $[\alpha]_k^*$, is equal to $[\alpha']_k$.

2. For all f and $g \in \mathcal{S}_k(\Gamma)$,
$$\langle f[\Gamma\alpha\Gamma]_k, g\rangle_\Gamma = \langle f, g[\Gamma\alpha'\Gamma]_k\rangle_\Gamma$$
Consequently, $[\Gamma\alpha\Gamma]_k^* = [\Gamma\alpha'\Gamma]_k$.

Proof of 1. To prove the first part, evaluate the left hand side to get,

$$\langle f[\alpha]_k, g\rangle_{\alpha^{-1}\Gamma\alpha} = 1/V_{\alpha^{-1}\Gamma\alpha}\int_{\mathbb{H}^*/\alpha^{-1}\Gamma\alpha}(f[\alpha]_k(\tau)\overline{g(\tau)}(\mathrm{Im}(\tau))^k d\mu(\tau)$$

Using part 1 and 2 of Lemma 3.8.4 the above equation equals:

$$1/V_\Gamma\int_{X(\Gamma)}\det(\alpha)^{k-1}f(\tau)\jmath(\alpha,\alpha^{-1}(\tau))^{-k}\overline{g(\alpha^{-1}(\tau))}(\mathrm{Im}(\alpha^{-1}(\tau))^k d\mu(\tau)$$

Observe that the action of α' is same as that of α^{-1}. Replacing α^{-1} with α' and noting that $\det(\alpha) = \det(\alpha')$, expanding the above expression we get,

$$1/V_\Gamma\int_{X(\Gamma)}\det(\alpha)^{k-1}f(\tau)\left(\frac{\jmath(\alpha\alpha'),\tau}{\jmath(\alpha',\tau)}\right)^{-k}\overline{g(\alpha'(\tau))}\left(\det(\alpha')\frac{\mathrm{Im}(\alpha^{-1}(\tau))}{|\jmath(\alpha',\tau)|}\right)^k d\mu(\tau)$$

Observing that $\jmath(\alpha\alpha'),\tau) = \det(\alpha)$ and writing $|\jmath(\alpha',\tau)| = \jmath(\alpha',\tau)\overline{\jmath(\alpha',\tau)}$, cancelling terms in the last expression gives us:

$$1/V_\Gamma\int_{X(\Gamma)}\det(\alpha)^{k-1}f(\tau)\overline{g[\alpha']_k(\tau)}\,\mathrm{Im}(\alpha^{-1}(\tau))^k d\mu(\tau) = \langle f, g[\alpha']_k\rangle_\Gamma.$$

□

Proof of 2. The second part requires us to compute $\langle f[\Gamma\alpha\Gamma]_k, g\rangle$. Taking β_j given in part 3 of Lemma 3.8.4, and keeping in mind Lemma 3.7.8 we compute:

$$\begin{aligned}
\langle f[\Gamma\alpha\Gamma]_k, g\rangle &= \sum_j \langle f[\beta_j]_k, g\rangle_\Gamma \\
&= \sum_j \langle f[\beta_j]_k, g\rangle_{\Gamma\cap\beta_j^{-1}\Gamma\beta_j} \\
&= \sum_j \langle f, g[\beta_j']_k\rangle_{\Gamma\cap\beta_j\Gamma\beta_j^{-1}} \\
&= \sum_j \langle f, g[\beta_j']_k\rangle_\Gamma
\end{aligned}$$

The second last equality comes from part 1 of Lemma 3.8.4. We wish to conclude that the expression in the last equality $\sum_j \langle f, g[\beta_j']_k\rangle_\Gamma$ equals $\langle f, g[\Gamma\alpha\prime\Gamma]_k\rangle$. If we show that $\Gamma\alpha'\Gamma = \bigcup_j \Gamma\beta_j'$, then we are done. This is seen as follows. Since $\Gamma\alpha\Gamma$ also equals the union of left cosets $\beta_j\Gamma$, taking inverse and multiplying by $\det(\alpha)$ in the expression $\Gamma\alpha\Gamma = \bigcup_j \beta_j\Gamma$, it follows that $\Gamma\alpha'\Gamma = \bigcup_j \Gamma(\det(\alpha)\beta_j^{-1})$. Noting that $\det(\alpha) = \det(\beta_j)$ for all j, we see that β_j' are coset representatives of $\Gamma\alpha'\Gamma$ in Γ. $\square$

With this we move on to computing the adjoints of the Hecke operators T_p and $\langle d\rangle$ away from the level.

Theorem 3.8.7. In the inner product space $\mathcal{S}_k(\Gamma_1(N))$, the Hecke operators $\langle d\rangle$ and T_p for $p \nmid N$ and $(d, N) = 1$ have adjoints given by the following operators.

$$\langle d\rangle^* = \langle d\rangle^{-1} \ ; \ T_p^* = \langle p\rangle^{-1} T_p$$

Proof. Let $f, g \in \mathcal{S}_k(\Gamma)$. Suppose $\alpha \in \Gamma_0(N)$ is the corresponding matrix for the operator $\langle d\rangle$. Since $\Gamma_1(N)$ is normal in $\Gamma_0(N)$, it follows from the first part of Proposition 3.8.6 that $\langle d\rangle^* = [\alpha]_k^* = [\alpha']_k = \langle d\rangle^{-1}$. To compute the T_p^* operator, we need to look at the operator $[\Gamma_1(N)\alpha\Gamma_1(N)]_k^*$ where $\alpha = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix}$. By the second part of Proposition 3.8.6, this amounts to finding the representation of the operator $[\Gamma_1(N)\alpha'\Gamma_1(N)]_k$. Since $p \nmid N$, we can find integers m and n such that $mp - nN = 1$. Now express α' in terms of α as follows:

$$\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n \\ N & mp \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \begin{pmatrix} p & n \\ N & m \end{pmatrix} \tag{3.8.8}$$

This helps us to find the decomposition of $\Gamma_1(N)\alpha'\Gamma_1(N)$ in terms of the decomposition of $\Gamma_1(N)\alpha\Gamma_1(N)$ into orbits of $\Gamma_1(N)$. Let $\Gamma_1(N)\alpha\Gamma_1(N) = \bigcup_j \Gamma_1(N)\beta_j$. Denote $\gamma = \begin{pmatrix} 1 & n \\ N & mp \end{pmatrix}$ and $\delta = \begin{pmatrix} p & n \\ N & m \end{pmatrix}$. Then $\gamma \in \Gamma_1(N)$ and $\delta \in \Gamma_0(N)$. The relation in (3.8.8) and the fact that $\Gamma_1(N)$ is normal in $\Gamma_0(N)$ allows us to write,

$$\Gamma_1(N)\alpha'\Gamma_1(N) = \Gamma_1(N)\alpha\Gamma_1(N)\delta \tag{3.8.9}$$

Using the decomposition of $\Gamma_1(N)\alpha\Gamma_1(N)$, we see that $\beta_j\delta$ represent the coset space $\Gamma_1(N)/\Gamma_1(N)\alpha'\Gamma_1(N)$. Finally note that $m \equiv p^{-1} \bmod N$ and so we get,

$$\begin{aligned} T_p^*(f) &= \sum_j f\left[\beta_j \begin{pmatrix} p & n \\ N & m \end{pmatrix}\right]_k \\ &= \langle p\rangle^{-1} T_p(f) \end{aligned}$$

□

Therefore the Hecke operators away from the level commute with their adjoints. This completes the proof of Theorem 3.8.1. The above computation of the adjoint of the T_p operator works only when $(p, N) = 1$. In general, if we wish to find a uniform expression for the adjoint operator which works for all Hecke operators, we will have to introduce a new operator given by $w_N = \left[\begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}\right]_k$.

Theorem 3.8.10. For any Hecke operator $T = T_n$ or $T = \langle n\rangle$, where $n \in \mathbb{Z}_+$, acting on the space $\mathcal{S}_k(\Gamma_1(N))$, the adjoint of T is the operator $T^* = w_N T w_N^{-1}$.

Before beginning the proof we note that the matrix $\delta = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}$ normalizes the subgroup $\Gamma_1(N)$. That is, $\delta^{-1}\Gamma_1(N)\delta = \Gamma_1(N)$. This allows us to view w_N as the double coset operator $[\Gamma_1(N)\delta\Gamma_1(N)]_k$.

Proof. First consider the operator $\langle n\rangle$. When $(n, N) > 1$, then $\langle n\rangle = 0$ and so its adjoint is the 0 operator as well. In the case when $(n, N) = 1$, we take the corresponding matrix γ in $\Gamma_0(N)$ with the right most entry $n \bmod N$. Observe that the matrix $w_N\gamma w_N^{-1} \in \Gamma_0(N)$ and has rightmost entry congruent to n^{-1} modulo N. We conclude that $w_N\gamma w_N^{-1} = \langle n\rangle^{-1} = \langle n\rangle^*$ by Theorem 3.8.6.

For the other type of Hecke operator we first establish the result for T_p, where p is now any prime. As before, we need to characterize the double coset $\Gamma_1(N)\alpha'\Gamma_1(N)$. We follow a similar approach in finding the decomposition of this double coset into orbits of $\Gamma_1(N)$, but this time we use a small observation that $\alpha' = \delta^{-1}\alpha\delta$. Since δ normalizes $\Gamma_1(N)$ and so does δ^{-1}, we have the equality $\Gamma_1(N)\delta = \delta\Gamma_1(N)$ and $\Gamma_1(N)\delta^{-1} = \delta^{-1}\Gamma_1(N)$.

Therefore, if $\Gamma_1(N)\alpha\Gamma_1(N) = \bigcup_j \beta_j$, then,

$$\begin{aligned}
\Gamma_1(N)\alpha'\Gamma_1(N) &= \Gamma_1(N)\delta^{-1}\alpha\delta\Gamma_1(N) \\
&= \delta^{-1}\Gamma_1(N)\alpha\Gamma_1(N)\delta \\
&= \delta^{-1}\bigcup_j \Gamma_1(N)\beta_j\delta \\
&= \bigcup_j \Gamma_1(N)\delta^{-1}\beta_j\delta
\end{aligned}$$

From the above calculation of the coset representatives we conclude that $T_p^*(f) = \sum_j f[\delta^{-1}\beta_j\delta]_k = w_N T_p w_N^{-1}(f)$. The proof works for all prime p, regardless of the fact that it divides N or not. By the definition of the T_n operator, the result extends to all $n \in \mathbb{Z}_+$. □

We will need the adjoint of the operator w_N later.

Lemma 3.8.11. Let w_N be the operator as given above acting on the space $\mathcal{S}_k(\Gamma_1(N))$. Its adjoint is given by the operator $w_N^* = (-1)^k w_N$.

Proof. Using Lemma 3.8.6, we see that $w_N^* = [\delta']_k$ where $\delta' = \det(\delta)\delta$. More precisely, $\delta' = \begin{pmatrix} 0 & 1 \\ -N & 0 \end{pmatrix}$ and so we compute:

$$\begin{aligned}
w_N^* f = f[\delta']_k(\tau) &= (-N\tau)^{-k} f(\delta'\tau) \\
&= (-1)^k (N\tau)^{-k} f(-1/N\tau) \\
&= (-1)^k f[\delta]_k(\tau) \\
&= (-1)^k w_N f
\end{aligned}$$

□

3.9 Old forms and new forms

Till now we were working with modular forms of a particular level N. In this section we would like to characterize the forms coming from lower levels. For example, if $M \mid N$, we know that $\mathcal{S}_k(\Gamma_1(M))$ is sitting inside $\mathcal{S}_k(\Gamma_1(N))$. There is yet one more way to embed $\mathcal{S}_k(\Gamma_1(M))$ in $\mathcal{S}_k(\Gamma_1(N))$. Suppose $d = N/M$. Consider the transformation:

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

Claim 3.9.1. The weight k operator $[\alpha_d]_k$ is an injective linear map from $\mathcal{S}_k(\Gamma_1(M))$ to $\mathcal{S}_k(\Gamma_1(N))$.

Proof of claim. Since $f[\alpha]_k(\tau) = d^{k-1}f(d\tau)$, it is clear that the map is injective and linear. To see that the operator lifts the level, we wish to show that $f[\alpha_d]_k \in \mathcal{S}_k(\Gamma_1(N))$. Suppose $\gamma = \begin{pmatrix} a & b \\ cN & d' \end{pmatrix} \in \Gamma_1(N)$. Write,

$$\begin{aligned} f[\alpha_d\gamma]_k(\tau) &= (cN\tau + d')^{-k}d^{k-1}f(d(\gamma\tau)). \\ &= (cM(d\tau)+d')^{-k}d^{k-1}f\left(\frac{a(d\tau)+bd}{cM(d\tau)+d'}\right) \\ &= f[\gamma'\alpha_d]_k(\tau)\ ;\ \gamma' = \begin{pmatrix} a & bd \\ cM & d' \end{pmatrix} \in \Gamma_1(M). \\ &= f[\alpha_d]_k(\tau). \end{aligned}$$

This takes care of the modularity condition. Moreover, the above calculation along with holomorphy of f gives us that $f[\alpha_d]_k$ is holomorphic at the cusps. It follows that $f[\alpha_d]_k$ is a modular form with respect to $\Gamma_1(N)$. $\square$

Given these two maps, in order to identify modular forms coming from lower levels, it makes sense to take the sum of their images at every level $M \mid N$. This will give us the following subspace of $\mathcal{S}_k(\Gamma_1(N))$.

$$\left\{ \sum_{d|N} (f_{N/d} + g_{N/d}[\alpha_d]_k) \;\middle|\; f_{N/d}, g_{N/d} \in \mathcal{S}_k(\Gamma_1(N/d)) \right\}$$

Combining the observations so far and writing everything in a fancy notation, we put together the following definition:

Definition 3.9.2 (Oldforms). For each divisor $d > 1$ of N, let i_d be the map,

$$i_d : \mathcal{S}_k(\Gamma_1(N)) \times \mathcal{S}_k(\Gamma_1(N)) \to S_k(\Gamma_1(N))$$

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

Then the subspace of oldforms at level N is:

$$\mathcal{S}_k(\Gamma_1(N))^{\text{old}} = \sum_{\substack{p|N \\ \text{prime}}} i_p(\mathcal{S}_k(\Gamma_1(N)) \times \mathcal{S}_k(\Gamma_1(N)))$$

For level $N = 1$, define the space of old forms $\mathcal{S}_k(SL_2(\mathbb{Z}))^{\text{old}} = \{0\}$.

In the above definition it is enough to sum over primes because of the following observation: Suppose that p and h are divisors of n, where p is a prime. Notice that if $f, g \in \mathcal{S}_k(\Gamma_1(N/ph))$, then $i_{ph}(f,g) \subseteq i_p(f, g[\alpha_h])$. It follows that if d is any divisor of N such that $d = ph$, then the image of $i_d \subseteq$ image of i_p. Therefore, it is enough to consider the prime divisors of N.

Definition 3.9.3. The orthogonal component of the space of old forms is known as the space of new forms with respect to the petersson inner product. That is, $\mathcal{S}_k(\Gamma_1(N))^{\text{new}} = (\mathcal{S}_k(\Gamma_1(N))^{\text{old}})^{\perp}$.

We aim to show that the Hecke operators respect this decomposition of $\mathcal{S}_k(\Gamma_1(N))$ into old and new subspaces.

Theorem 3.9.4. The subspaces of $\mathcal{S}_k(\Gamma_1(N))^{\text{old}}$ and $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ are stable under the Hecke operators T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}_+$.

Proof. We prove this case by case. Let $p \mid N$. Using the decomposition of Theorem 3.2.19, it is enough to take f and $g \in \mathcal{S}_k(N/p, \chi)$. Let χ_N modulo N denote the lift of the Dirichlet character χ modulo N/p.

1. Consider the diamond operator $\langle d\rangle$. In order to show that the operator $\langle d\rangle$ commutes with the i_p map, we show that:

$$\langle d\rangle_{N/p} f = \langle d\rangle_N f \ ; \ (\langle d\rangle_{N/p} g)[\alpha_p]_k = \langle d\rangle_N (g[\alpha_p]_k). \tag{3.9.5}$$

   Here, the subscript under the operator symbol denotes the level at which it is acting. Suppose $f \in \mathcal{S}_k(N/p, \chi)$. Then $\langle d\rangle_N f = f[\delta]_k$ where $\delta \in \Gamma_0(N)$ with rightmost entry equal to d mod N. Observe that $\delta \in \Gamma_0(N)$ implies that $\delta \in \Gamma_0(N/d)$. It follows that $f[\delta]_k = \chi(d) f = \chi_N(d) f$. We conclude that $f \in \mathcal{S}_k(N, \chi_N)$ and so $\langle d\rangle_N f = \langle d\rangle_{N/p} f$.

   Performing a similar calculation done in the proof of claim 3.9.1, it is easy to see that if $g \in \mathcal{S}_k(N/p, \chi)$, then, $g[\alpha_p]_k \in \mathcal{S}_k(N, \chi_N)$. This proves the second equality. Since $\langle n\rangle = 0$, when $(n, N) > 1$, the Hecke operator $\langle n\rangle$ stabilizes the space of old forms for all $n \in \mathbb{Z}^+$.

2. Next, we show the same equalities in 3.9.5 for the operator $T_{p'}$ where p' is a prime different from p. Let $T_{p',N/p}$ denote the operator acting at level N/p. Using the Fourier expansion of $T_{p',N/p}(f)$ from Corollary 3.4.13 and noting that $f \in \mathcal{S}_k(N, \chi_N)$, we see that $T_{p',N/p}(f) = T_{p',N}(f)$. Further, we compute the Fourier

expansion of $(T_{p',N/p}(g))[\alpha_p]_k$:

$$\begin{aligned}(T_{p',N/p}(g))[\alpha_p]_k &= p^{k-1}\sum_{n=0}^{\infty} a_{np'}(g)q^{np} + \chi(p')p'^{k-1}\sum_{n=0}^{\infty} a_n(g)q^{np'p} \\ &= T_{p,N}(g[\alpha]_k)\end{aligned}$$

The last equality comes from the fact that $p' \mid N/p$ if and only if $p' \mid N$ and so $\chi(p') = \chi_N(p')$.

3. In this case we will deal with the operator T_p. The proof is slightly different as T_p individually does not commute with the natural inclusion map and the $[\alpha_p]_k$ operator. However, for $f, g \in \mathcal{S}_k(\Gamma_1(N/p))$, we have that:

$$T_{p,N}(i_p(f,g)) = i_p(T_{p,N/p}f + p^{k-1}g, -\langle p\rangle_{N/p}f). \tag{3.9.6}$$

Indeed, if we expand the left hand side, we get, $T_{p,N}(f + g[\alpha]_k) = T_{p,N}(f) + T_{p,N}(g[\alpha_p]_k)$. Looking at the two terms separately, we see $T_{p,N}f = T_{p,N/p}f - (\langle p\rangle_{N/p}f)[\alpha_p]_k$. This follows from the explicit representation of T_p in Theorem 3.4.3 and is true regardless of whether $p \mid N/p$ or not. In the latter case we just take $\langle p\rangle_{N/p} = 0$. The expression for $T_{p,N}(g[\alpha_p]_k)$ is computed as:

$$\begin{aligned}T_{p,N}(g[\alpha_p]_k) &= \sum_j g\left[\begin{pmatrix} p & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix} 1 & j \\ 0 & p\end{pmatrix}\right]_k \\ &= \sum_j g\left[\begin{pmatrix} p & 0 \\ 0 & p\end{pmatrix}\begin{pmatrix} 1 & j \\ 0 & 1\end{pmatrix}\right]_k \\ &= p^{k-1}g\end{aligned}$$

Combining the above calculations, we get (3.9.6). Points 2 and 3 together show that the Hecke operator T_n stabilize the space of oldforms for all $n \in \mathbb{Z}^+$.

To show that the Hecke operators stabilize $\mathcal{S}_k(\Gamma_1(N))^{\mathrm{new}}$ it is enough to show that their adjoints stabilize $\mathcal{S}_k(\Gamma_1(N))^{\mathrm{old}}$. This is because $\langle f, Tg\rangle = 0$ if and only if $\langle T^*f, g\rangle = 0$ for all $f \in \mathcal{S}_k(\Gamma_1(N))^{\mathrm{old}}$ and $g \in \mathcal{S}_k(\Gamma_1(N))^{\mathrm{new}}$. By Theorem 3.8.7, when $(n, N) = 1$, $\langle n\rangle^* = \langle n\rangle^{-1}$. The discussion in 1 shows that the adjoint operator stabilizes the space of oldforms. In the case when $(n, N) > 1$, there is nothing to check.

4. To see the result for T_n^*, we use the formula for the adjoint calculated in Theorem 3.8.10. Therefore, it is enough to check that the space of oldforms is stabilized by

the operator w_N given in Theorem 3.8.10. Compute,

$$\begin{aligned}
w_N i_p(f,g) &= w_N(f + g[\alpha_p]_k) \\
&= f\left[\begin{pmatrix} 0 & 1 \\ -N & 0 \end{pmatrix}\right]_k + g\left[\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -N & 0 \end{pmatrix}\right]_k \\
&= f\left[\begin{pmatrix} 0 & 1 \\ -N & 0 \end{pmatrix}\right]_k + g\left[\begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -N/p & 0 \end{pmatrix}\right]_k \\
&= f\left[\begin{pmatrix} 0 & 1 \\ -N/p & 0 \end{pmatrix}\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\right]_k + p^{k-2}g\left[\begin{pmatrix} 0 & 1 \\ -N/p & 0 \end{pmatrix}\right]_k \\
&= i_p(p^{k-2}w_{N/p}(g), w_N(f))
\end{aligned}$$

It follows that $w_N i_p(f,g) \in \mathcal{S}_k(\Gamma_1(N))^{\text{old}}$.

□

From the discussion in Points $1-4$ we arrive at the following corollary:

Corollary 3.9.7. The spaces $\mathcal{S}_k(\Gamma_1(N))^{\text{old}}$ and $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ have orthogonal bases of eigenforms for the Hecke operators away from the level, $\{T_n, \langle n\rangle \mid (n,N) = 1\}$

3.10 The Main Lemma

The theory in the next few sections is dedicated to eliminating the condition of $(n,N) = 1$ to have an orthogonal bases for the space of newforms. To that end, in this section we prove a very important lemma which helps us to see whether a modular form is an old form, just by looking at its Fourier coefficients.

Define the map $\imath_d$ to be a sort of normalization of the α_d operator.

$$\imath_d = d^{1-k}[\alpha_d]_k : \mathcal{S}_k(\Gamma_1(M)) \to \mathcal{S}_k(\Gamma_1(N)) \ ; \ (\imath_d f)(\tau) \mapsto f(d\tau)$$

The map $\imath$ acts on the Fourier expansion as:

$$\sum_{n=1}^{\infty} a_n q^n \mapsto \sum_{n=1}^{\infty} a_n q^{dn}$$

It follows that if $f \in \mathcal{S}_k(\Gamma_1(N))$ is of the form $f = \sum_{p|N} \imath_p f_p$ with $f_p \in \mathcal{S}_k(\Gamma_1(N/p))$, then, $a_n = 0$ for all n such that $(n,N) = 1$. The Main Lemma gives us the converse of this statement.

Theorem 3.10.1 (Main Lemma). If $f \in \mathcal{S}_k(\Gamma_1(N))$ has Fourier expression $f(\tau) = \sum a_n(f)q^n$ such that $(n, N) = 1$, then f is of the form $f = \imath_p f_p$ with $f_p \in \mathcal{S}_k(\Gamma_1(N/p))$.

Using tools from previous theory and linear algebra, we will keep restating the Main Lemma into different versions, finally reducing it to a known result in representation theory.

Proof. (Sketch) As a first step, we change the subgroup we are working with. Let

$$\Gamma^1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ * & 1 \end{pmatrix} \bmod N \right\}$$

It is easy check that $\alpha_M \Gamma_1(M) \alpha_M^{-1} = \Gamma^1(N)$. By the special case in point 2, it follows that the double coset operator $[\Gamma_1(M)\alpha_M^{-1}\Gamma^1(M)]_k = [\alpha_M^{-1}]_k$ which takes $f(\tau)$ to $M^{k-1}f(\tau/M)$ is in fact an isomorphism of $\mathcal{S}_k(\Gamma_1(M))$ and $\mathcal{S}_k(\Gamma^1(M))$. After removing the constant M^{k-1}, we have the following isomorphism explicitly in terms of the Fourier series of f.

$$M^{1-k}[\alpha_M]_k : \mathcal{S}_k(\Gamma_1(M)) \to \mathcal{S}_k(\Gamma^1(M))$$

$$\sum_{n=0}^{\infty} a_n q^n \mapsto \sum_{n=0}^{\infty} a_n q_M^n$$

The above discussion leads to the following diagram:

$$\begin{array}{ccc} \mathcal{S}_k(\Gamma_1(M)) & \xrightarrow{\imath_d} & \mathcal{S}_k(\Gamma_1(N)) \\ \downarrow \wr & & \downarrow \wr \\ \mathcal{S}_k(\Gamma^1(M)) & & \mathcal{S}_k(\Gamma^1(N)) \end{array}$$

In terms of the Fourier expansion, if we draw the diagram, we see that it commutes if $M = dN$.

$$\begin{array}{ccc} \sum_{n=0}^{\infty} a_n q^n & \xmapsto{\imath_d} & \sum_{n=0}^{\infty} a_n q^{dn} \\ \downarrow & & \downarrow \\ \sum_{n=0}^{\infty} a_n q_M^n & \hookrightarrow & \sum_{n=0}^{\infty} a_n q_M^{dn} \end{array}$$

This helps us to view the map $\imath_d$ as an inclusion map and makes our life easier. More precisely, since the diagram commutes, we reformulate the Main Lemma in terms of the new subgroup as follows:

Theorem 3.10.2 (Main Lemma version 2). If $f \in \mathcal{S}_k(\Gamma^1(N))$ has Fourier expansion $f(\tau) = \sum_{n=0}^{\infty} a_n q^n$ with $a_n(f) = 0$ when $(n, N) = 1$, then $f_p = \sum_{p|N} f_p$; $f_p \in \mathcal{S}_k(\Gamma^1(N/p))$.

Next step is to translate the Main Lemma in the language of linear algebra. Recall the subgroup

$$\Gamma^0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}) \;\middle|\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & 0 \\ * & * \end{pmatrix} \mod N \right\}$$

For any divisor d of N, let $\Gamma_d = \Gamma_1(N) \cap \Gamma^0(N/d)$, a congruence subgroup of level N.

Lemma 3.10.3. The set of coset representatives for the quotient $\Gamma(N)\backslash\Gamma_d$ is

$$S = \left\{ \begin{pmatrix} 1 & bN/d \\ 0 & 1 \end{pmatrix} \;\middle|\; 0 \le b < d \right\}$$

Proof of Lemma 3.10.3. Let $\Gamma = \begin{pmatrix} a & \beta \\ c & \delta \end{pmatrix} \in \Gamma_d$. This means that a and δ are congruent to 1 modulo N; $a\delta - \beta c = 1$. Moreover, $\beta = kN/d$ for some $k \in \mathbb{Z}$. Write $k = qd + b$ with $0 \le r < d$ so that $\beta \equiv rN/d \mod N$. We claim that $\gamma{\gamma'}^{-1} \in \Gamma(N)$ where $\gamma' = \begin{pmatrix} 1 & rN/d \\ 0 & 1 \end{pmatrix} \in S$. Indeed, if we compute $\gamma{\gamma'}^{-1}$, we get:

$$\begin{pmatrix} a & \beta \\ c & \delta \end{pmatrix} \begin{pmatrix} 1 & -rN/d \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & \beta - arN/d \\ c & \delta - crN/d \end{pmatrix}$$

Call this matrix γ_1. Observe the following about the entries of γ_1. Since $N \mid (a-1)$, we have that $\beta - (arN/d) \equiv (1-a)N/d \equiv 0 \mod N$. Similarly, $\delta - crN/d \equiv 1 \mod N$. Hence, $\gamma_1 \in \Gamma(N)$. It follows that for any arbitrary matrix in $\gamma \in \Gamma_d$, we can find a matrix $\gamma' \in S$, such that $\gamma \in \Gamma(N)\gamma'$. It remains to show that the each coset space $\Gamma(N)\gamma'$ with γ' in S is distinct. This can be seen by realizing that every matrix in the set $\Gamma(N)\gamma'$ has the top right entry congruent to bN/d determined $0 \le b < d$. Therefore as coset spaces they are different. $\square$

We next introduce a suitable operator defined by averaging over the weight k operators corresponding to the representatives in S.

$$\pi_d : \mathcal{S}_k(\Gamma(N)) \to \mathcal{S}_k(\Gamma(N))$$

$$f \mapsto \sum_{b=0}^{d-1} f\left[\begin{pmatrix} 1 & bN/d \\ 0 & 1 \end{pmatrix} \right]_k$$

To find out the properties of the operator, first note that the operator acts as identity on the subspace $\mathcal{S}_k(\Gamma_d)$. Compute the following inclusion of subgroups:

$$\begin{pmatrix} 1 & bN/d \\ 0 & 1 \end{pmatrix}^{-1} \Gamma(N) \begin{pmatrix} 1 & bN/d \\ 0 & 1 \end{pmatrix} \subseteq \Gamma_d \tag{3.10.4}$$

This, and the fact that $f\left[\begin{pmatrix} 1 & bN/d \\ 0 & 1 \end{pmatrix}\right]_k$ is a modular form with respect to the subgroup given on the left hand side of the relation in 3.10.4, helps us conclude that the image of π_d lands inside the subspace $\mathcal{S}_k(\Gamma_d)$. This shows us two things. One, the map π_d is a surjection onto $\mathcal{S}_k(\Gamma_d)$. Two, it is a projection, for, $\pi_d(f) \in \mathcal{S}_k(\Gamma_d)$ and so $\pi_d(\pi_d(f)) = \pi_d(f)$ from our first observation.

The map π_d is special to us because of the way it acts on f. More explicitly, we see its action on the Fourier series of f. Let $f = \sum_{n=0}^{\infty} a_n q_N^n$.

$$\pi_d(f) = 1/d\left(\sum_{n=0}^{\infty} a_n q_N^n + \sum_{n=0}^{\infty} a_n q_N^n e^{2\pi i n/d} + \cdots + \sum_{n=0}^{\infty} a_n q_N^n e^{2\pi i (d-1)n/d}\right)$$

Collecting together the terms corresponding to $n = 0$ first, then $n = 1$ and carrying on, we get an infinite sum given by the following expression:

$$da_0 + a_1 q_N(1 + e^{2\pi i/d} + \cdots + e^{2\pi i(d-1)/d} + \cdots + a_d q_N^d(1 + e^{2\pi i} + \cdots + e^{2\pi i(d-1)}) + \ldots$$

Observe that the geometric sum corresponding to a_n with $n \nmid d$, sums up to 0, while a_n for $n \mid d$, is multiplied with the sum $(1 + e^{2\pi i} + \cdots + e^{2\pi i(d-1)})$ that adds up to d. Noting that the whole expression is multiplied by the constant $1/d$, what we finally get is the following sum:

$$\pi_d(f) = \sum_{\{n:\ d|n\}} a_n q_N^n \tag{3.10.5}$$

This shows that π_d preserves coefficients which are multiples of d and kills everything else. One more property of π_d is immediate. For d_1 and d_2 positive integers,

$$\pi_{d_1}(\pi_{d_2}(f)) = \sum_{\{n:\ d_1 d_2|n\}} a_n q_N^n = \pi_{d_1 d_2}(f) = \pi_{d_2}(\pi_{d_1}(f)). \tag{3.10.6}$$

Next, define $\pi : \mathcal{S}_k(\Gamma(N)) \to \mathcal{S}_k(\Gamma(N))$ given by $\pi = \prod_{p|N}(1 - \pi_p)$ where p denotes a prime diving N. Using the property mentioned in (3.10.6), the operator expands as:

$$\pi = \mathbb{I} - \sum_{p|N} \pi_p + \sum_{\substack{p_1|N, p_2|N \\ p_2 < p_1}} \pi_{p_1 p_2} - \ldots$$

Using the expression in (3.10.5) for π_p and the principle of exclusion and inclusion we see that,

$$\Big(\sum_{p|N}\pi_p + \sum_{\substack{p_1|N, p_2|N \\ p_2<p_1}} \pi_{p_1p_2} - \dots\Big)(f) = \sum_{\{n:\ (nN)>1\}} a_n q_N^n$$

Therefore, when π acts on f, it only keeps the coefficients a_n away from the level.

$$\pi(f) = \sum_{\{n:\ (nN)=1\}} a_n q_N^n$$

Looking back at version 2 of the Main Lemma given in 3.10.2, the hypothesis of the theorem is equivalent to assuming that $f \in \mathcal{S}_k(\Gamma^1(N)) \cap \ker(\pi)$.

Our next task is to find $\ker(\pi) = \ker\big(\prod_{p|N}(1-\pi_p)\big)$. We use the following result:

Lemma 3.10.7. Let π_p be the map defined above. Then,

$$\ker\Big(\prod_{p|N}(1-\pi_p)\Big) = \sum_{p|N}\big(\ker(1-\pi_p)\big) = \sum_{p|N}\operatorname{im}(\pi_p) \tag{3.10.8}$$

In general, the above result is true for any set of commuting projections and not necessarily π_p.

Proof. We prove the first equality first. Observe that if π is a projection, then $(1-\pi)$ is a projection. Therefore, it is enough to show that given two commuting projections π_1 and π_2, $\ker(\pi_1\pi_2) = \ker(\pi_1) + \ker(\pi_2)$. One way containment is easy to see. Suppose $x = z + y$ with $z \in \ker(\pi_1)$ and $y \in \ker(\pi_2)$. Then, $\pi_1(\pi_2(x)) = \pi_1(\pi_2(z+y)) = 0$ since π_1 and π_2 commute. This implies that $\ker(\pi_1) + \ker(\pi_2) \subseteq \ker(\pi_1\pi_2)$. For the other way containment, let $\pi_1(\pi_2(x)) = \pi_2(\pi_1(x)) = 0$. Write $\pi_2(x) = y$ and $x - \pi_2(x) = z$ so that $x = z + y$. Clearly, $y \in \ker(\pi_1)$. Compute $\pi_2(z) = \pi_2(x) - \pi_2^2(x) = 0$. This shows that $x \in \ker(\pi_1) + \ker(\pi_2)$ as required.

To establish the second equality in (3.10.8), we show that $\operatorname{im}(\pi) = \ker(1-\pi)$. Indeed, $x \in \operatorname{im}(\pi)$ if and only if $x = \pi(y)$ for some y if and only if $\pi(y) = \pi(x) = x$ if and only if $x \in \ker(1-\pi)$. □

The above lemma reduces the problem of finding the kernel of the map π to finding the image of π_p for each prime p. But we already know that π_p is a projection of $\mathcal{S}_k(\Gamma(N))$ onto $\mathcal{S}_k(\Gamma_p) = \mathcal{S}_k(\Gamma_1(N) \cap \Gamma^0(N/p))$.

The Main Lemma is now reduced to the following: If $f \in \mathcal{S}_k(\Gamma^1(N)) \cap \sum_{p|N}\mathcal{S}_k(\Gamma_1(N) \cap \Gamma^0(N/p))$, then $f \in \sum_{p|N}\mathcal{S}_k(\Gamma^1(N/p))$. We in fact have that the two spaces mentioned in the above version are equal. That is, *Theorem* 3.10.9 (Main Lemma version 3).

$$\mathcal{S}_k(\Gamma^1(N)) \cap \sum_{p|N} \mathcal{S}_k(\Gamma_1(N) \cap \Gamma^0(N/p)) = \sum_{p|N} \mathcal{S}_k(\Gamma^1(N/p)).$$

We now simplify the above expression to finally give a final version of the Main Lemma reducing it to representation theory.

Let $G = SL_2(\mathbb{Z}/N\mathbb{Z})$. We know that the group $SL_2(\mathbb{Z})$ acts on $\mathcal{S}_k(\Gamma(N))$ from the right, via the weight k operator. Clearly the group $\Gamma(N)$ acts trivially on this space and therefore by Proposition 1.4.2, the action can be considered of the group $SL_2(\mathbb{Z}/N\mathbb{Z})$. Write $N = \prod_{i=1}^n p_i^{e_i}$ as product of powers of distinct primes. By the Chinese Remainder Theorem, we can identify the group G with $\prod_{i=1}^n G_i$ where $G_i = SL_2(\mathbb{Z}/p_i^{e_i}\mathbb{Z})$ for each $1 \le i \le n$. The explicit map from G to $\prod_{i=1}^n G_i$ is given by the following assignment:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bmod p_1^{e_1}, \dots, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \bmod p_n^{e_n} \right) \tag{3.10.10}$$

Further, for each $1 \le i \le n$, Proposition 1.4.2 again helps us write $G_i \simeq SL_2(\mathbb{Z})/\Gamma(p_i^{e_i})$ so that,

$$G \simeq \prod_{i=1}^n SL_2(\mathbb{Z}/p_i^{e_i}\mathbb{Z}) \simeq \prod_{i=1}^n SL_2(\mathbb{Z})/\Gamma(p_i^{e_i}) \tag{3.10.11}$$

We now introduce two subgroups G_i for each i given by:

$$H_i = \Gamma^1(p_i^{e_i})/\Gamma(p_i^{e_i}) \ ; \ K_i = (\Gamma_1(p_i^{e_i}) \cap \Gamma^0(p_i^{e_i-1}))/\Gamma(p_i^{e_i})$$

Let $H = \prod_{i=1}^n H_i$. Denote by $\mathcal{S}_k(\Gamma(N))^H$, the subspace of $\mathcal{S}_k(\Gamma(N))$ fixed by H and similarly for any other subgroup of G. We then claim that the version 3 of the Main Lemma reduces to the following statement:

$$\mathcal{S}_k(\Gamma(N))^H \cap \sum_{i=1}^n \mathcal{S}_k(\Gamma(N))^{K_i} = \sum_{i=1}^n \mathcal{S}_k(\Gamma(N))^{\langle H, K_i \rangle} \tag{3.10.12}$$

Before moving on to prove the claim, we need a small technical lemma.

Lemma 3.10.13. For any prime p and $e \ge 1$, we have,

$$\langle \Gamma^1(p^e), \Gamma_1(p^e) \cap \Gamma^0(p^{p^{e-1}}) \rangle = \Gamma^1(p^{e-1}) \tag{3.10.14}$$

Proof. It is fairly easy to show that $\langle \Gamma^1(p^e), \Gamma_1(p^e) \cap \Gamma^0(p^{p^{e-1}}) \rangle \subseteq \Gamma^1(p^{e-1})$. We show the other way containment which is not so direct. Let Γ denote the group on the left hand side of the equality in (3.10.14). Let $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma^1(p^{e-1})$. We will show that for some matrices δ and $\delta' \in \Gamma$, $\delta\gamma\delta' \in \Gamma$. Keep in mind that a and d are congruent to 1 mod p^{e-1} while b is congruent to 0 mod p^{e-1}. The proof now goes by constantly replacing γ by matrices of the form $\gamma_1\gamma\gamma_2$ with $\gamma_1, \gamma_2 \in \Gamma$. We start by observing that if $p \mid a$, then $p \nmid b$. Otherwise, $p \mid (ad - bc) = 1$ which is not possible. Therefore, $\gamma \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ satisfies that $p \nmid (a+b)$ and so we can as well take γ such that $p \nmid a$. Similarly, we can take $p \nmid d$ by arguing the same with the product $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \gamma$.

Next, we reduce the entries of γ to have the property of p^e dividing b and c. Since $p \nmid d$, it is invertible mod p. Let $\beta \equiv -bd^{-1}$ mod p^e. Notice the following:

$$\beta \equiv 0 \bmod p^{e-1} \ ; \ b + \beta d \equiv 0 \bmod p^e \ ; \ \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \in \Gamma^1(p^e).$$

This gives us that $\begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} \gamma \in \Gamma^1(p^e) \subset \Gamma$ and has the upper right entry equal to $b + d\beta \equiv 0$ mod p^e. It is therefore enough to take $b \equiv 0$ mod p^e in γ. Similarly take $w = cd^{-1}$ mod p^e and argue with $\gamma \begin{pmatrix} 1 & 0 \\ w & 1 \end{pmatrix}$ as above to conclude that we can take the c entry in γ to be congruent to 0 mod p^e. Finally, we are reduced to the case where γ has entries with the following properties.

$$ad - bc = 1 \ ; \ a \equiv d \equiv 1 \bmod p^{e-1} \ ; \ b \equiv c \equiv 0 \bmod p^e \tag{3.10.15}$$

Consider the matrix:

$$M = \begin{pmatrix} 1 & 1-a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1-d \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$$

Using properties mentioned in (3.10.15), it is easy to see that the matrices in the product individually belong to Γ. Moreover, the matrix M is given as:

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

Observing that $ad \equiv$ mod p^e, we see that $M \equiv \gamma$ mod p^e. This implies that $M\gamma^{-1} \in \Gamma^1(p^e) \subseteq \Gamma$ and so $\gamma \in \Gamma$. ☐

Coming back to the proof of the Main Lemma, we need to prove the claim that Theorem 3.10.9 is equivalent to the statement in (3.10.12). This is seen as follows: We look at the three terms given in 3.10.9 separately.

1. First, consider the subspace $\mathcal{S}_k(\Gamma^1(N))$. This is equal to the subspace of $\mathcal{S}_k(\Gamma(N))$ fixed by the subgroup $\Gamma^1(N)/\Gamma(N)$. We wish to show that the subgroup $\Gamma_1(N)/\Gamma(N)$ is identified to H under the mapping in (3.10.10). This is easily seen by noting that any matrix in $\Gamma^1(N)/\Gamma(N)$ is of the form $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with entries satisfying the following properties that $a \equiv d \equiv 1 \bmod N$ and $b \equiv 0 \bmod N$. Under the isomorphism map in (3.10.10), this is equivalent to saying that $a \equiv d \equiv 1 \bmod p_i^{e_i}$ and $b \equiv 0 \bmod p_i^{e_i}$ for all $1 \leq i \leq n$. This is equivalent to saying that the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bmod p_i^{e_i} \in \Gamma^1(p_i)^{e_i}/\Gamma(p_i)^{e_i}$ for all $1 \leq i \leq n$. This argument essentially shows that $\Gamma^1(N)/\Gamma(N) \simeq \prod_{i=1}^{n} \Gamma^1(p_i)^{e_i}/\Gamma(p_i)^{e_i} = H$ and so $\mathcal{S}_k(\Gamma^1(N)) = \mathcal{S}_k(\Gamma(N))^H$.

2. Next, we wish to show that:

$$\sum_{p|N} \mathcal{S}_k(\Gamma_1(N) \cap \Gamma^0(N/p_i)) = \sum_{i=1}^{n} \mathcal{S}_k(\Gamma(N))^{K_i}. \tag{3.10.16}$$

For each i, write $\mathcal{S}_k(\Gamma_1(N) \cap \Gamma^0(N/p_i)) = \mathcal{S}_k(\Gamma(N))^{(\Gamma_1(N)\cap\Gamma^0(N/p_i))/\Gamma(N)}$. We follow a similar argument as in 1. Any matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ belongs to the subgroup $(\Gamma_1(N) \cap \Gamma^0(N/p_i))/\Gamma(N)$ if and only if $a \equiv d \equiv 1 \bmod N, c \equiv 0 \bmod N$ and $b \equiv 0 \bmod (N/p_i)$. This is equivalent to the fact that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bmod p_i^{e_i} \in (\Gamma_1(p_i^{e_i}) \cap \Gamma^0(p_i^{e_i-1}))/\Gamma(p_i^{e_i}) = K_i$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bmod p_j^{e_j}$ is the identity matrix for $j \neq i$. This shows that $(\Gamma_1(N) \cap \Gamma^0(N/p_i))/\Gamma(N) \simeq K_i$. This proves the relation in (3.10.16)

3. Finally, we move on to the term on the right hand side of the equality in (3.10.9) given by $\sum_{p|N} \mathcal{S}_k(\Gamma^1(N/p))$. For a fixed prime p_i, consider the term $\mathcal{S}_k(\Gamma^1(N/p_i)) = \mathcal{S}_k(\Gamma(N))^{\Gamma^1(N/p_i)/\Gamma(N)}$. Using the same argument as above, we identify the image of $\Gamma^1(N/p_i)/\Gamma(N)$ under the map in (3.10.10). The matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma^1(N/p_i)/\Gamma(N)$ if and only if $a \equiv d \equiv 1 \bmod (N/p_i)$ and $b \equiv 0 \bmod (N/p_i)$. This happens if and only if $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bmod p_i^{e_i} \in \Gamma^1(p_i^{e_i-1})/\Gamma(p_i^{e_i})$ and $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \bmod p_j^{e_j} \in \Gamma^1(p_i^{e_j})/\Gamma(p_i^{e_j})$ for all $j \neq i$ This implies that, the $\Gamma^1(N/p_i)/\Gamma(N)$ equals $\Gamma^1(p_i^{e_i-1})/\Gamma(p_i^{e_i}) \times \prod_{j\neq i} H_j$ Using Lemma 3.10.13 we see that:

$$\Gamma^1(p_i^{e_i-1})/\Gamma(p_i^{e_i}) = \langle \Gamma^1(p_i^{e_i})/\Gamma(p_i^{e_i}), (\Gamma_1(p_i^{e_i})/\cap\Gamma^0({p_i}^{e_i-1}))/\Gamma(p_i^{e_i})\rangle$$

This is exactly the subgroup $\langle H_i, K_i\rangle$. What we finally have is that $\Gamma^1(N/p_i)/\Gamma(N) = \langle H, K_i\rangle$. This completes the argument that:

$$\sum_{p|N} \mathcal{S}_k(\Gamma^1(N/p)) = \sum_{p|N} \mathcal{S}_k(\Gamma(N))^{\langle H,K_i\rangle}$$

The discussion in points 1, 2 and 3 shows that it is enough to establish statement in (3.10.12). The proof of (3.10.12) now follows from a completely group theoretic result stated below.

Theorem 3.10.17. Let V be an irreducible representation of the group $G = \prod_{i=1}^n G_i$. Let $H = H_i$ and $K = K_i$ be subgroups be subgroups. Then,

$$V^H \cap \sum_{i=1}^n V^{K_i} = \sum_i^n V^{\langle H,K_i\rangle} \tag{3.10.18}$$

In order to apply the above theorem to $\mathcal{S}(\Gamma_1(N))$, we use yet another result from representation theory: The vector space $\mathcal{S}_k(\Gamma(N))$ is a direct sum of subspaces irreducible under the G-action.

Noting that if $\mathcal{S}_k(\Gamma(N)) = \bigoplus_i W_i$ is the decomposition of the vector space into irreducible representations, then $\mathcal{S}_k(\Gamma_1(N))^H = \bigoplus_i W_i^H$, the result in 3.10.17 can now be applied to each individual irreducible component W_i of $\mathcal{S}_k(\Gamma_1(N))$. This finally finishes the proof of the Main Lemma.

□

3.11 Bases of newforms

Let $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ be an eigenform for Hecke operators away from the level. We show that f is in fact an eigenform for operators T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}^+$. When $(n, N) > 1$, $\langle n\rangle = 0$. Therefore, in this case f is an eigenform for these operators with eigen value 0. Now we only have to deal with the T_n operator. From now on, we use the terms "eigenform" and "newform" in a much general sense.

Definition 3.11.1 (Eigenform). A non-zero modular form $f \in \mathcal{M}_k(\Gamma_1(N))$ that is an eigenform for the Hecke operators T_n and $\langle n\rangle$ for all $n \in \mathbb{Z}^+$ is called a Hecke eigenform or simply an eigenform.

An eigenform $f = \sum_n a_n q^n$ is *normalised* if $a_1 = 1$.

Definition 3.11.2 (Newform). A normalized eigenform in $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ is called a newform.

In the discussion to follow, we aim to prove the following theorem.

Theorem 3.11.3. Let $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ be a non zero eigenforms for the operators T_n and $\langle n \rangle$ with $(n, N) = 1$. Then

1. The function f is a Hecke eigenform. Moreover, a suitable multiple of f is a newform. Each such newform lies in an eigenspace $\mathcal{S}_k(N, \chi)$ for some Dirichlet character χ modulo N and satisfies $T_n(f) = a_n(f)f$ for all $n \in \mathbb{Z}^+$.

2. The set of newforms in the space $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ is an orthogonal bases of the space.

3. (Multiplicity One property) If $\widetilde{f}$ satisfies the same conditions as f, and has the same T_n eigen values, then $\widetilde{f} = cf$ for some constant c.

Proof of 1. Let $f \in \mathcal{S}_k(\Gamma_1(N))$ be an eigenform for the the Hecke operators away from the level. This implies that for n such that $(n, N) = 1$, there exists eigenvalues c_n and d_n such that $T_n(f) = c_n f$ and $\langle n \rangle f = d_n f$. Define a map $\chi : (\mathbb{Z}/N\mathbb{Z})^* \to \mathbb{C}^*; n \mapsto d_n$. This clearly defines a Dirichlet character modulo N and $f \in \mathcal{S}_k(N, \chi)$. Using the formula for the Fourier coefficients of $T_n(f)$ in Proposition 3.5.4, we see that,

$$a_1(T_n f) = a_n(f) \;\text{ for all } n \in \mathbb{Z}^+ \tag{3.11.4}$$

Furthermore, f being an eigenform away from its level implies that,

$$a_1(T_n(f)) = c_n a_1(f) \;\; (n, N) = 1 \tag{3.11.5}$$

The relation in (3.11.4) and (3.11.5) give us that for n with $(n, N) = 1$, $a_n(f) = c_n a_1(f)$. Therefore if $a_1(f) = 0$, then $a_n(f) = 0$ away form the level. By the Main Lemma, f is an oldform. It follows that if $f \in \mathcal{S}_k(\Gamma_1(N))^{\text{new}}$, then $a_1(f) \neq 0$. Therefore, we can normalize f and take $a_1(f) = 1$. It remains to show that f is a Hecke eigenform. To see this, take $g_m = T_m(f) - a_m(f)f$ for any $m \in \mathbb{Z}^+$. The function g_m is clearly in $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ and g_m is an eigenform away from the level. Compute:

$$\begin{aligned} a_1(g_m) &= a_1(T_m(f)) - a_m(f)a_1(f) \\ &= a_m(f) - a_m(f) \\ &= 0 \end{aligned}$$

By the discussion above $a_n(g_m) = 0$ when $(n, N) = 1$. By the Main Lemma, $g_m \in \mathcal{S}_k(\Gamma_1(N))^{\text{old}} \cap \mathcal{S}_k(\Gamma_1(N))^{\text{new}} = 0$. This shows that $T_m(f) = a_m(f)f$ for all $m \in \mathbb{Z}^+$.

Proof of 2. We already have an orthogonal basis of eigenforms which can be normalized to give newforms. If we show that the set of all newforms are linearly independent, then by the above observation they form an orthogonal basis. Suppose there exists a non-trivial relation,

$$\sum_i c_i f_i = 0 \ ; \ c_i \in \mathbb{C} \tag{3.11.6}$$

with all the c_i non zero and as many few terms as possible, necessarily at least two. For any $m \in \mathbb{Z}^+$ applying the operator $T_m - a_m(f_1)$ to the relation in (3.11.7), we get the following relation:

$$\sum_{i>1} c_i(a_m(f_i) - a_m(f_1))f_i = 0 \ ; \ c_i \in \mathbb{C} \tag{3.11.7}$$

This has fewer terms than (3.11.7) and so it must be trivial. Since m was arbitrary, It follows that $a_m(f_i) = a_m(f)$ for all $m \in \mathbb{Z}^+$. Therefore $f_i = f_1$ for all i. This gives a contradiction to our assumption of at least two terms being in the relation in 3.11.7.

Proof of 3. From the proof of 1 we see that, $T_n(\widetilde{f}) = (a_n(\widetilde{f})/a_1(\widetilde{f}))\widetilde{f}$ and $T_n(\widetilde{f}) = (a_n(f)/a_1(f))f$. Since f and $\widetilde{f}$ have the same T_n eigenvalues, we see that $a_n(\widetilde{f})/a_1(\widetilde{f}) = a_n(f)/a_1(f)$ for all n. This shows that $a_n(\widetilde{f}) = ca_n(f)$ for all $n \in \mathbb{Z}^+$ with $c = a_1(\widetilde{f})/a_1(f)$. ☐

We do not have a basis of eigenforms for the whole space $\mathcal{S}_k(\Gamma_1(N))$, however what we do have is a bases of new forms f and functions of the form $f(n\tau)$. More precisely, we have the following theorem.

Theorem 3.11.8. The set:

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

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

Proof. (Sketch.) We first show that elements in $\mathcal{B}_k(N)$ span $\mathcal{S}_k(\Gamma_1(N))$. We will use decomposition of $\mathcal{S}_k(\Gamma_1(N))$ into old and new subspaces inductively.

$$\mathcal{S}_k(\Gamma_1(N)) = \mathcal{S}_k(\Gamma_1(N))^{\text{new}} \bigoplus \sum_{p|N} i_p(\mathcal{S}_k(\Gamma_1(N)) \times \mathcal{S}_k(\Gamma_1(N)))$$

In Theorem 3.11.3 we showed that that $\mathcal{S}_k(\Gamma_1(N))^{\text{new}}$ has a bases of newforms. So, elements in $\mathcal{B}_k(N)$ with $n = 1$ and $M = N$ span this space.

Next, Observe that the image $i_p(\mathcal{S}_k(\Gamma_1(N/p)))^2$ have elements of the form $f + g(p\tau)$ with $f, g \in \mathcal{S}_k(\Gamma_1(N/p))$. If we show that $\mathcal{S}_k(\Gamma_1(N/p))$ is spanned by elements from $\mathcal{B}_k(N/p)$, then we are done. Further decomposing it into old and new subspaces, we see that $\mathcal{S}_k(\Gamma_1(N/p))^{\text{new}}$ is spanned by newforms of level N/p. For $\mathcal{S}_k(\Gamma_1(N/p))^{\text{old}}$, by the argument above, it is enough to show that $\mathcal{S}_k(\Gamma_1(N/pp')$ is spanned by the elements of $\mathcal{B}_k(N/pp')$ for a prime p' dividing N/p. Running the same argument downwards we will eventually be left with the base case $\mathcal{S}_k(SL_2(\mathbb{Z}))$. The subspace of old forms in $\mathcal{S}_k(SL_2(\mathbb{Z}))$ is precisely $\{0\}$ since nothing is coming from lower levels here. This leaves us with $\mathcal{S}_k(SL_2(\mathbb{Z})) = \{0\}^{\perp} = \mathcal{S}_k(SL_2(\mathbb{Z}))^{\text{new}}$, which is spanned by newforms that correspond to elements with $M = 1$ and $n = 1$ in $\mathcal{B}_k(1)$.

It remains to show that the elements of the set are linearly independent. Assume that there is non-trivial relation amongst the elements in the set $\mathcal{B}_k(N)$.

$$\sum_{i,j} c_{i,j} f_i(n_{i,j}\tau) = 0 \ ; \ c_{i,j} \in \mathbb{C} \tag{3.11.9}$$

Here every f_i lies in a space $\mathcal{S}_k(M_i, \chi_i)$ with $M_i|N$ and χ_i a Dirichlet character modulo M_i. Assume that each $c_{i,j}$ in (3.11.9) is nonzero and the relation has as many few terms as possible. We make some observations first.

1. The functions f_i in the relation (3.11.9) cannot all be equal because for a given i, each function $f_i(n_{i,j}\tau)$ starts with a different power of q in its power series expansion. With only one function in the summand, the relation will never be equal to 0.

2. Each Dirichlet character χ_i lifts to a Dirichlet character $\widetilde{\chi}_i$ modulo N so that $f \in \mathcal{S}_k(N, \widetilde{\chi}_i)$. In fact, every χ_i lifts to the same Dirichlet character. To see this, assume to the contrary that at least two Dirichlet characters (say), χ_1 and χ_2 lift to different Dirichlet characters $\widetilde{\chi}_1$ and $\widetilde{\chi}_2$. Then there exists some $d \in (\mathbb{Z}/N\mathbb{Z})^*$ such that $\widetilde{\chi}_1(d) \neq \widetilde{\chi}_2(d)$. Using the same calculation as in (3.9.5) it is easy to see $\langle d \rangle_N(f_i(n_{i,j}\tau)) = (\langle d \rangle_{M_i} f_i)(n_{i,j}\tau)$. Applying $\langle d \rangle_N - \widetilde{\chi}_1(d)$ to relation (3.11.9) and using the observation above we get,

$$\begin{aligned} 0 &= \sum_{i,j} c_{i,j}(\langle d \rangle_N - \widetilde{\chi}_1(d)) f_i(n_{i,j}\tau) \\ &= \sum_{i,j} c_{i,j}((\langle d \rangle_{M_i} f_i)(n_{i,j}\tau) - \widetilde{\chi}_1(d) f_i(n_{i,j}\tau)) \\ &= \sum_{i>1,j} c_{i,j}(\chi_i(d) - \widetilde{\chi}_1(d)) f_i(n_{i,j}\tau) \end{aligned}$$

   This yields a non trivial result with fewer terms than (3.11.9), a contradiction.

3. Using a similar trick, we show that $a_p(f_i) = a_p(f_j)$ for all i, j and p, a prime not dividing N. Suppose not, then we can find a prime $p \nmid N$ and functions in the summand (3.11.9), (say) f_1 and f_2 such that $a_p(f_1) \neq a_p(f_2)$. The idea is to apply $T_{p,N} - a_p(f_1)$ to (3.11.9) and get a nontrivial relation with fewer terms than

(3.11.9), exactly as above. This will lead to a contradiction. We will need the following observation in the calculation below:

$$T_{p,N}(f_i(n_{i,j}\tau)) = (T_{p,M_i} f_i)(n_{i,j}\tau)$$

This can be seen by performing exactly the same calculation as done in point 2 of the proof of Theorem 3.9.4 for any $p \nmid N$. Now,

$$\begin{aligned} 0 &= \sum_{i,j} c_{i,j}(T_{p,N} - a_p(f_1)) f_i(n_{i,j}\tau) \\ &= \sum_{i,j} c_{i,j}((T_{p,M_i}(f_i))(n_{i,j}\tau) - a_p(f_1) f_i(n_{i,j}\tau)) \\ &= \sum_{i>1,j} c_{i,j}(a_p(f_i) - a_p(f_1)) f_i(n_{i,j}\tau) \end{aligned}$$

a contradiction as explained before.

By a result known as *Strong Multiplicity One*, points 2 and 3 imply that the functions f_i are all equal, contradicting point 1.

□

We will not state or prove *Strong Multiplicity One*. Apart from that, the proof is complete. However to give some idea, the Strong Multiplicity Theorem for new forms is a deep result which helps us to characterize new forms completely just by knowing its Fourier coefficients a_p for all but finitely many primes p.

Chapter 4

The theory of L-functions

4.1 Some basic tools from Complex analysis

We begin the chapter by introducing some tools from complex analysis which will be useful in subsequent sections. A lot of things are stated without proof. The reader is encouraged to see parts of [2] and [4] for a rigorous analysis of the topics mentioned below.

4.1.1 Gamma, zeta and L-functions

We briefly introduce the *gamma function* which will appear frequently in the text to follow. For any complex number s such that $\mathrm{Re}(s) > 0$, the gamma function is defined by the following integral.

$$\Gamma(s) = \int_0^\infty e^{-t} t^s \frac{dt}{t}$$

Notice that as $t \to 0$, the integrand $e^{-t}t^{s-1}$ is comparable to t^{s-1} and clearly $\int_0^1 t^{s-1}$ converges absolutely for $\mathrm{Re}(s) > 1$. To see that the integral converges at its upper end as well, observe that as $t \to \infty$, the integrand $e^{-t}t^{s-1}$ is comparable to $e^{-t/2}$ which makes the integral converge absolutely from 1 to ∞ as well.

We in fact have that the Gamma function defines a holomorphic function in the region where $\mathrm{Re}(s) > 1$. Moreover, it satisfies the functional equation $\Gamma(s+1) = s\Gamma(s)$. This helps in extending the domain of the function to the whole complex plane with poles at $s = 0, -1, -2,$ and so on at all negative integers. See [4] for the proof of the above statements. We will also need the *Stirling approximation* for the Gamma function which helps us conclude that the Gamma function decays rapidly along vertical lines. More precisely, for a fixed real number σ, as $t \to \pm\infty$

$$|\Gamma(\sigma + it)| \sim \sqrt{2\pi}|t|^{\sigma-1/2}e^{-\pi|t|/2} \tag{4.1.1}$$

The central object of study in this chapter is the L-function. An L-function is a Dirichlet series of the form $\sum_{n=1}^{\infty} c(n)n^{-s}$ such that $c(n) \in \mathbb{C}$ for all n. The prototype example of an L-function is the Riemann zeta function $\zeta(s)$ given by

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$$

The ζ function is a very deep object and interesting in its own right. We state some properties of the function below and finally explain the goal of this chapter.

Proposition 4.1.2. The series defining $\zeta(s)$ converges for $\mathrm{Re}(s) > 1$ and the function ζ is holomorphic in this half plane.

Proof. First observe that $|n^{-s}| = n^{-\mathrm{Re}(s)}$. For any compact subset K of the half plane $\mathrm{Re}(s) > 1$, if $\mathrm{Re}(s)$ is bounded below by σ_0 for all $s \in K$, then we get that

$$\begin{aligned} \left|\sum_{n=1}^{\infty} n^{-s} - \sum_{n=1}^{k} n^{-s}\right| &\leq \sum_{n=k}^{\infty} n^{-\mathrm{Re}(s)} \\ &\leq \sum_{n=k}^{\infty} n^{-\sigma_0} \end{aligned}$$

The series in the last expression clearly converges. Since the partial sums converge uniformly to the zeta function on any compact subset of the half plane, it defines a holomorphic function in the plane $\mathrm{Re}(s) > 1$. $\square$

Furthermore, one can show that the function given by the relation $\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$ satisfies the functional equation $\xi(s) = \xi(1-s)$ and helps to give a meromorphic continuation of the zeta function with simple poles at 0 and 1. This is non-trivial to show and the reader can refer to [4] for the precise proof.

Another interesting property of the zeta function is that it has an Euler product representation. An Euler product is an expansion of the Dirichlet series into a product over all the primes. It is easy to check that:

$$\zeta(s) = \prod_{p} (1 - p^{-s})^{-1}$$

The properties of the zeta function mentioned above give us some motivation to study L-functions and find out whether similar kind of results hold. In the sections to come, we associate an L-function to a modular form and study these similar properties of convergence, analytic continuation, and Euler product expansion.

4.1.2 The Mellin tranform and Mellin inversion

We will need some preliminary ideas in order to look at L-functions which we mention below. The *Mellin transform* of a function $f : \mathbb{R}^+ \to \mathbb{C}$ is the integral given by the following expression.

$$g(s) = \int_0^\infty f(t)t^s \frac{dt}{t} \tag{4.1.3}$$

for the values of s such that the integral is absolutely convergent. Consider the integral in two parts: First from 0 to 1, that is the integral: $\int_0^\infty f(t)t^s \frac{dt}{t}$. If the integral absolutely converges for a particular value of s, say, s_1, then it converges absolutely for all complex values s such that $\mathrm{Re}(s) \geq \mathrm{Re}(s_1)$. This is because for $0 \leq y \leq 1$,

$$\begin{aligned} \left| \int_0^\infty f(t)t^s \frac{dt}{t} \right| &\leq \int_0^\infty |f(t)t^{(\mathrm{Re}(s)-1)}|dt \\ &\leq \int_0^\infty |f(t)t^{(\mathrm{Re}(s_1)-1)}|dt \end{aligned}$$

Similarly, if the integral from 1 to ∞ absolutely converges for some s_2, then it converges for all values of s with a smaller real part than that of s_2. In sum, what we get is that the integral $g(s)$ is absolutely convergent for $\mathrm{Re}(s)$ lying in some interval (σ_1, σ_2). It may also be a half plane or all of $\mathbb{C}$.

If g is a holomorphic function of the complex variable s in some right half plane, then for $\sigma \in (\sigma_1, \sigma_2)$, its *inverse Mellin transform* is given by the integral:

$$f(t) = \frac{1}{2\pi i} \int_{s=\sigma-i\infty}^{\sigma+i\infty} g(t)t^{-s} \frac{dt}{t} \tag{4.1.4}$$

for positive t values such that the integral converges absolutely. In addition, g is required to uniformly converge to 0 as $\mathrm{Im}(s) \to \infty$. Using Cauchy's formula we can then show that the Mellin inversion is independent of the real number σ. One in fact has that, the integrals given in (4.1.3) and (4.1.4) are equivalent. This is known as the *Mellin inversion formula*.

We will also need a very helpful result from complex analysis called the Phragmén-Lindelöf principle. This is an extension of the maximum modulus principle for holomorphic functions to unbounded regions provided we assume moderate growth conditions for the function.

Proposition 4.1.5 (Phragmén-Lindelöf principle). Write $s = \sigma + it$. Let $f(s)$ be a function that is holomorphic in the the strip

$$\sigma_1 \leq \mathrm{Re}(s) \leq \sigma_2$$

such that $f(\sigma + it) = \mathcal{O}(e^{|s|^\alpha})$ for some real number $\alpha > 0$ when $\sigma_1 \leq \sigma \leq \sigma_2$. Suppose that $f(\sigma + it) = \mathcal{O}(|t|^M)$ for $\sigma = \sigma_1$ and $\sigma = \sigma_2$ and some integer M. Then $f(\sigma + it) = \mathcal{O}(|s|^M)$ uniformly in σ for all $\sigma \in [\sigma_1, \sigma_2]$.

See Lang, Complex Analysis [5] for the proof of the theorem.

4.2 Associating L-functions to modular forms

To every modular form $f \in \mathcal{M}_k(\Gamma_1(N))$, we associate a Dirichlet series known as its L-function. Let $f = \sum_n a_n q^n$ be the fourier expansion of f. Let $s \in \mathbb{C}$ be a complex variable, then the associated L-function is given by the expression:

$$L(s, f) = \sum_{n=1}^{\infty} a_n n^{-s} \tag{4.2.1}$$

It is natural to ask the about the properties of convergence of the L-function.

Proposition 4.2.2. Let $f \in \mathcal{M}_k(\Gamma_1(N))$. If f is a cusp form, then $L(s, f)$ converges absolutely for all s with $\mathrm{Re}(s) > k/2 + 1$. Otherwise, $L(s, f)$ converges absolutely for all s with $\mathrm{Re}(s) > k$.

Proof. We begin by assuming that f is a cusp form. Suppose the $\widetilde{f}(q) = \sum_{n=0}^{\infty} a_n q^n$ is the Fourier expansion of f about 0. We find out the region of convergence by estimating the Fourier coefficients. Since $\widetilde{f}$ represents a holomorphic function in the unit disc around 0, we use Cauchy's Theorem to find out the coefficients as follows.

$$a_n = 1/2\pi i \int_{C_y} \widetilde{f}(q) q^{-n-1} dq \tag{4.2.3}$$

where C_y denotes the counter clockwise circle $e^{2\pi i(x+iy)}$ for a fixed $y > 0$, as x varies from 0 to 1. Changing the variable to x we see that for a fixed y,

$$a_n = \int_{x=0}^{1} f(x+iy)e^{2\pi in(x+iy)}dx \tag{4.2.4}$$

We wish to find a bound for the function $f(\tau)$. Consider the function $\varphi(\tau) = f(\tau)y^{k/2}$. It is easy to check that $\varphi(\gamma\tau) = \varphi(\tau)$ for all $\gamma \in \Gamma_1(N)$. By running the same argument as done in Lemma 3.7.6, we see that φ is bounded on $\mathbb{H}$. Hence $|\phi(\tau)| = |f(\tau)y^{k/2}| \leq M$ for some positive integer M. This gives a bound for the function $f(\tau)$ dependent on the imaginary part of τ. What we finally have is,

$$\begin{aligned} |a_n| &\leq \int_{x=0}^{1} My^{-k/2}|e^{-2\pi in(x+iy)}|dx \\ &= My^{-k/2}e^{2\pi yn} \end{aligned}$$

This expression is valid for all $y > 0$. In particular when we put $y = 1/n$, we see that $|a_n| \leq Me^{2\pi}n^{k/2}$ for all $n \in \mathbb{Z}^+$. This implies that $|a_n n^{-s}| = \mathcal{O}(n^{k/2} - \mathrm{Re}(s))$. The function $L(s,f)$ converges in the region $k/2 - \mathrm{Re}(s) < -1$, that is, $k/2 + 1 < \mathrm{Re}(s)$.

When f is not a cusp form, we write it as a sum of an Eisenstein series in $\mathcal{M}_k(\Gamma_1(N))$ and a cusp form as discussed in section 3.6. We estimate the Fourier coefficients of the Eisenstein series first. From the discussion in section 3.6, we know that $|a_n(E_k)| \leq |C_k|\sigma_k(n) = |C_k|\sum_{d|N} d^{k-1}$. Furthermore,

$$n^{k-1} \leq \sum_{d|n} d^{k-1} = \sum_{d|n} (n/d)^{k-1} < (n^{k-1})\zeta(k-1) \tag{4.2.5}$$

This shows that $|a_n| = \mathcal{O}(n^{k-1})$. Since n^{k-1} grows faster than $n^{k/2}$, the growth of the Fourier coefficients of f is of order n^{k-1}. It follows that when f is not a cusp form, then $|a_n(f)n^{-s}| = \mathcal{O}(n^{k-1} - \mathrm{Re}(s))$. This gives us that $L(s,f)$ converges for $k-1-\mathrm{Re}(s) > -1$, that is, $\mathrm{Re}(s) > k$. □

A similar argument as done for the ζ function shows that the L-function defines an analytic function in the region where it converges. In the next section, we will show that for a cusp form f, there exists an analytic continuation of $L(s,f)$ to the whole complex plane. A standard tool to analytically continue functions is to derive a functional equation. As we will see in the coming sections, the Mellin transform of the cusp form will be used to find a suitable functional equation of its L-function.

Proposition 4.2.6. The Mellin transform of $f(it)$ is the function:

$$g(s) = (2\pi)^{-s}\Gamma(s)L(s,f) \ ; \ \mathrm{Re}(s) > k/2 + 1$$

Proof. From the formula of the Mellin transform of f, write,

$$g(s) = \int_{t=0}^{\infty} \sum_{n=1}^{\infty} a_n e^{2\pi i n(it)} t^s \frac{dt}{t} \tag{4.2.7}$$

We need to make sure that this integral is absolutely convergent at both the ends. We split the integral into two parts. First consider the integral from 1 to ∞. Notice that for $t \geq 1$, the sum $\sum_{n=1}^{\infty} a_n e^{-2\pi nt} = \mathcal{O}(e^{-2\pi t})$. This implies that for some $C \in \mathbb{C}$,

$$\left| \int_{t=1}^{\infty} \sum_{n=1}^{\infty} a_n e^{-2\pi nt} t^{s-1} dt \right| \leq \int_{t=1}^{\infty} |Ce^{-2\pi t} t^{\mathrm{Re}(s)-1} dt|$$

and the last integral clearly converges for all s. To analyze the the integral from 0 to 1, notice that, as $t \to 0$,

$$\begin{aligned} \left| \sum_{n=1}^{\infty} a_n e^{-2\pi nt} \right| &\leq \sum_{n=1}^{\infty} |a_n e^{-2\pi nt}| \\ &\leq \sum_{n=1}^{\infty} e^{-2\pi nt} \\ &= \frac{1}{e^{2\pi t} - 1} \\ &\leq C' 1/t \end{aligned}$$

for some constant $C' \in \mathbb{C}$. This allows us to write

$$\begin{aligned} \left| \int_{t=0}^{1} \sum_{n=1}^{\infty} a_n e^{-2\pi nt} t^{s-1} dt \right| &\leq \int_{t=0}^{1} |C' t^{\mathrm{Re}(s)-2} dt| \\ &= \left. \frac{t^{\mathrm{Re}(s)-1}}{\mathrm{Re}(s) - 1} \right|_{t=0}^{1} \end{aligned}$$

This clearly gives a finite value for $\mathrm{Re}(s) > 1$, in particular for $\mathrm{Re}(s) > k/2 + 1$.

Therefore it makes sense to interchange the sum and the integral. Changing the variable by substituting $2\pi nt = x$, we get

$$\begin{aligned} g(s) &= \sum_{n=1}^{\infty} \int_{t=0}^{\infty} a_n e^{-2\pi nt} t^s \frac{dt}{t} \\ &= \sum_{n=1}^{\infty} \int_{t=0}^{\infty} \frac{a_n}{(2\pi n)^s} e^{-x} x^s \frac{dx}{x} \\ &= (2\pi)^{-s} L(s, f) \Gamma(s) \end{aligned}$$

□

We digress a bit to normalized eigenforms and derive a property of the L-function that characterizes them completely. To that end, we first prove a nice result which helps us identify (normalised) eigenforms, based on some properties of its Fourier coefficients.

Proposition 4.2.8. Let $f \in \mathcal{M}_k(N, \chi)$. Then f is a normalized eigenform if and only if its Fourier coefficients satisfy the following conditions:

1. $a_1(f) = 1$
2. $a_{p^r}(f) = a_p(f)a_{p^{r-1}}(f) - \chi(p)p^{k-1}a_{p^{r-2}}(f)$
3. $a_{mn}(f) = a_m(f)a_n(f)$ when $(m, n) = 1$.

Proof. Suppose f is a normalized eigenform. Then $a_1 = 1$ by definition. Using the definition in (3.5.1) and the fact that f is an eigenform, we see that,

$$a_{p^r}(f)f = T_{p^r}(f) = T_pT_{p^{r-1}}(f) - \chi(p)p^{k-1}T_{p^{r-2}}(f) \tag{4.2.9}$$

$$= (a_p(f)a_{p^{r-1}}(f) - \chi(p)p^{k-1}a_{p^{r-2}}(f))f \tag{4.2.10}$$

This gives us 2. The property in 3 similarly follows because of (3.5.3). Conversely, suppose the Fourier coefficients of f satisfies the three conditions. First condition implies that f is normalized. It is enough to prove that f is an eigenform for the T_p operator. For if we assume that f is an eigenform for the operator T_{p^n} for all $n \leq r$ then, running same calculation in (4.2.9) backwards with condition 2, we see that f is an eigenform for the operator T_{p^r}. This helps us to conclude that f is in fact an eigenform for T_n for all $n \in \mathbb{Z}^+$. Indeed, if $n = p_1^{r_1} \dots p_m^{r_m}$ is the prime factorization of n, then,

$$\begin{aligned} T_n(f) &= T_{p_1^{r_1}} \dots T_{{p_m}^{r_m}}(f) \\ &= (a_{p_1^{r_1}}(f) \dots a_{{p_m}^{r_m}}(f))f \\ &= a_{p_1^{r_1} \dots {p_m}^{r_m}}(f)f \\ &= a_n(f)f \end{aligned}$$

To get the second last equality, we have used the third condition. Therefore, what remains to prove is the base case of T_p and then by induction, the argument is complete. We proceed by computing the m^{th} Fourier coefficient of $T_p(f)$. When $p \nmid N$, it is easy to see from the formula in (3.4.13) that $a_m(T_p(f)) = a_{mp}(f)$ which is same as $a_p(f)a_m(f)$

since $(p, m) = 1$. When $p \mid m$ with $m = p^r m'$ and $(m', p) = 1$, we compute:

$$\begin{aligned}
a_m(T_p(f)) &= \sum_{d|p} \chi(d) d^{k-1} a_{mp/d^2}(f) \\
&= \chi(p) p^{k-1} a_{m/p}(f) + a_{mp}(f) \\
&= \chi(p) p^{k-1} a_{p^{r-1}m'}(f) + a_{m'p^{r+1}}(f) \\
&= \chi(p) p^{k-1} a_{p^{r-1}}(f) a_{m'}(f) + a_{m'}(f) a_{p^{r+1}}(f) \quad \text{by condition 3} \\
&= a_{m'}(f)(\chi(p) p^{k-1} a_{p^{r-1}}(f) + a_{p^{r+1}}(f)) \\
&= a_{m'}(f) a_p(f) a_{p^r} \text{ by condition 2}
\end{aligned}$$

Condition 3 again gives us that the last expression equals $a_p(f)a_m(f)$ in this case as well. Finally, this shows that $T_p(f) = a_p(f)f$ for any prime p. □

Theorem 4.2.11. Let $f \in \mathcal{M}_k(N, \chi)$ such that $f(\tau) = \sum_{n=0}^{\infty} a_n q^n$. Then the following are equivalent:

1. f is a normalised eigenform.
2. The associated L-function of f has an Euler Product expansion:

$$L(s, f) = \prod_p (1 - a_p p^{-s} + \chi(p) p^{k-1-2s}) \tag{4.2.12}$$

The proof of the theorem requires a small useful result to be proven.

Lemma 4.2.13. Given the Fourier expansion of f as above with $a_1 = 1$, we have the following equality of sums:

$$\sum_{n=1}^{\infty} \prod_{p||n} a_{p^r} p^{-r} = \prod_p \sum_{r=0}^{\infty} a_{p^r} p^{-r} \tag{4.2.14}$$

where the product is taken over all primes and the symbol $p^r||n$ signifies that p^r is the highest power of prime dividing n.

More generally, if g is a function on prime powers such that $g(1) = 1$, one has,

$$\sum_{n=1}^{\infty} \prod_{p||n} g(p^r) = \prod_p \sum_{r=0}^{\infty} g(p^r)$$

provided that the sum in above expression makes sense.

To stay with the main theorem, we assume this result for now and prove it at the end.

Proof. We will show that the three conditions given in Proposition 4.2.8 for the coefficients of f are necessary and sufficient to ensure that $L(s, f)$ has an Euler product expansion. First, suppose that $f \in \mathcal{M}_k(N, \chi)$ and satisfies the three conditions in 4.2.8. Let p denote an arbitrary prime number. Multiplying condition 3 with p^{-rs} we get

$$a_{p^r} p^{-rs} = p^{-rs} a_p a_{p^{r-1}} - p^{-rs} \chi(p) p^{k-1} a_{p^{r-2}} \tag{4.2.15}$$

Summing over $r \geq 2$ in (4.2.15) we see that

$$\begin{aligned} \sum_{r=2}^{\infty} a_{p^r} p^{-rs} &= \sum_{r=2}^{\infty} p^{-rs} a_p a_{p^{r-1}} - \sum_{r=2}^{\infty} p^{-rs} \chi(p) p^{k-1} a_{p^{r-2}} \\ &= a_p p^{-s} \sum_{r=1}^{\infty} p^{-rs} a_{p^r} - p^{-2s} \chi(p) p^{k-1} \sum_{r=0}^{\infty} p^{-rs} a_{p^r} \end{aligned}$$

This gives us the equality:

$$\sum_{r=0}^{\infty} a_{p^r} p^{-rs} (1 - a_p p^{-s} + p^{-2s} \chi(p) p^{k-1}) = a_1 + a_p p^{-s} (1 - a_1) \tag{4.2.16}$$

Putting the value of $a_1 = 1$ in (4.2.16), we get:

$$\sum_{r=0}^{\infty} a_{p^r} p^{-rs} = \frac{1}{1 - a_p p^{-s} + p^{-2s} \chi(p) p^{k-1}} \tag{4.2.17}$$

Now calculating the expression of the L-function gives that,

$$\begin{aligned} L(s, f) = \sum_{n=1}^{\infty} a_n n^{-s} &= \sum_{n=1}^{\infty} \left(\prod_{p||n} a_{p^r} \right) n^{-s} \\ &= \sum_{n=1}^{\infty} \prod_{p||n} a_{p^r} p^{-rs} \text{ ; by condition 3} \\ &= \prod_{p} \sum_{r=0}^{\infty} a_{p^r} p^{-rs} \text{ ; by Lemma 4.2.13} \end{aligned}$$

Finally, (4.2.17) helps us arrive at the Euler expansion of the L-function Conversely, given the Euler product expansion in (4.2.12), expanding it in the form of a geometric series and using Lemma 4.2.13 we compute:

$$\begin{aligned} L(s,f) &= \prod_p (1 - a_p p^{-s} + p^{-2s}\chi(p)p^{k-1}) \\ &= \prod_p \sum_{r=0}^{\infty} b_{p,r} p^{-rs} \; ; \text{ for some } \{b_{p,r}\} \\ &= \sum_{n=1}^{\infty} \prod_{p^r || n} b_{p,r} p^{-rs} \\ &= \sum_{n=1}^{\infty} \Big(\prod_{p^r||n} b_{p,r} \Big) n^{-s} \end{aligned}$$

We would like to conclude that $a_n = \prod_{p||n} b_{p,r}$. However there is no uniqueness theorem for L-functions. What we do is the following. Since both the series define the same L-function, they will give the same Mellin transform. From the Mellin inversion formula we will get back the Fourier series with coefficients a_n which will be equal to the Fourier series with coefficients $\prod_{p^r||n} b_{p,r}$. By the uniqueness theorem of the Fourier series, for each n, the coefficient a_n will be equal to $\prod_{p^r||n} b_{p,r}$. More precisely, from the discussion in section 4.2.6 and Proposition 4.2.6 we can write

$$\sum_{n=1}^{\infty} a_n e^{-2\pi n t} = \sum_{n=1}^{\infty} \Big(\prod_{p||n} b_{p,r} \Big) e^{-2\pi n t} = \int_{\sigma - i\infty}^{\sigma + i\infty} (2\pi)^{-s} L(s,f)\Gamma(s) ds$$

for sufficiently large σ in the right half plane where the Mellin transform exists. We just have to make sure that the integral mentioned in the right hand side converges. Here we will use two results two approximate the integral. First is the Stirling approximation for the Gamma function in (4.1.1) and second is the fact that the L function is bounded for sufficiently large σ. This helps us to analyze the growth of the integral as follows:

$$\begin{aligned} \lim_{t\to\infty} \left| \int_{\sigma-it}^{\sigma+it} (2\pi)^{-s} L(s,f)\Gamma(s) ds \right| &\leq \lim_{t\to\infty} \int_{-t}^{t} |(2\pi)^{-\sigma - iy} L(\sigma+iy, f)\Gamma(\sigma+iy)| dy \\ &\leq \lim_{t\to\infty} \int_{-t}^{t} M(2\pi)^{-\sigma}\sqrt{2\pi}|y|^{\sigma-1/2} e^{-\pi|y|/2} dy \\ &\leq \lim_{t\to\infty} \int_{0}^{t} M'|y|^{\sigma-1/2} e^{-\pi|y|/2} dy \end{aligned}$$

where M and M' define suitable constants. The integral in the last expression clearly converges and so our claim follows. Furthermore, for $n = p^r$, we see that $a_{p^r} = b_{p,r}$. Therefore, the calculation done above for the L-function helps us conclude the result in (4.2.17). Taking $\mathrm{Re}(s) \to \infty$ in (4.2.17) we get $a_1 = 1$. Now that the first condition holds, we do the calculation leading to (4.2.16) backwards to get (4.2.15). This gives us condition 2. To realize condition 3, notice that for positive integers m and n such that

$(m,n) = 1$,

$$\begin{aligned} a_n a_m &= \prod_{p||n} b_{p,r} \prod_{p||m} b_{p,r} \\ &= \prod_{p||nm} b_{p,r} \\ &= a_{mn} \end{aligned}$$

This completes the proof of the theorem. It only remains to show Lemma 4.2.13. □

Proof of Lemma 4.2.13. Instead of the infinite product of primes in (4.2.14), we first consider the product over a finite number of primes(say) $p_1, \ldots p_n$. Then,

$$\prod_{i=1}^{n} \sum_{r=0}^{\infty} a_{p_i^r} p_i^{-r} = \sum_{\substack{e_i=0 \\ 1 \leq i \leq n}}^{\infty} \prod_{i=1}^{n} a_{p_i^{e_i}} p_i^{-e_i} \tag{4.2.18}$$

As e_i runs from 0 to infinity for each $1 \leq i \leq n$, corresponding to a unique tuple of powers $(e_1, \ldots, e_n)$ we get a unique product term $\prod_{i=1}^{n} a_{p_i^{e_i}} p_i^{-e_i}$ in the summand. By the Fundamental theorem of arithmetic, the above expression reduces to summing over all the positive integers m which are only divisible by p_i for any $1 \leq i \leq n$ and for a such fixed m, taking the product over the highest power of prime dividing it. That is, writing the set $S = \{m \in \mathbb{Z}^+ : p_i \mid m \text{ for } 1 \leq i \leq n \ ; \ p \nmid m \text{ if } p \neq p_1, \ldots p_n\}$, we get the following expression:

$$\prod_{i=1}^{n} \sum_{r=0}^{\infty} a_{p_i^r} p_i^{-r} = \sum_{m \in S} \prod_{p^r || m} a_{p^r} p^{-r} \tag{4.2.19}$$

Therefore, taking the product over all primes in the left hand side of the expression in (4.2.19), instead of summing over the elements in S, we will have to sum over all the positive integers on the right hand side. The proof for any function g of prime powers with $g(1) = 1$ is the same. □

4.3 Analytic continuation of the L-function of a cusp form

Let f be a cusp form of weight k and $L(s,f) = \sum_n a_n n^{-s}$ be its associated L-function. Via the means of the machinery developed in the previous section, we show that $L(s,f)$ has an analytic continuation to the whole complex plane.

Recall the Mellin transform $g(s)$ of a cusp form f calculated in Proposition 4.2.6. In order to show that $L(s,f)$ has an analytical continuation to the whole s plane, we will derive a functional equation for the function $\Lambda_N(s) = N^{s/2} g(s)$. We will use the following operator on the space of cusp forms.

$$W_N : \mathcal{S}_k(\Gamma_1(N)) \to \mathcal{S}_k(\Gamma_1(N))$$

$$f \mapsto i^k N^{1-k/2} f\left[\begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix}\right]_k \tag{4.3.1}$$

Observe that W_N is a suitable multiple of the operator w_N introduced in section 3.8. We calculated the adjoint of w_N in Lemma 3.8.11, which helps us conclude the following result about the above operator.

Lemma 4.3.2. The operator W_N is self adjoint.

Writing $W_N = i^k N^{1-k/2} w_N$ and using Lemma 3.8.11, we see that

$$\begin{aligned} W_N^* &= \overline{i^k} N^{1-k/2} w_N^* \\ &= i^k (-1)^k N^{1-k/2} (-1)^k w_N \\ &= W_N \end{aligned}$$

The operator W_N is special in the sense that it is a unitary operator. Indeed, if we calculate

$$\begin{aligned} W_n^2(f) &= W_N(i^k N^{-k/2} \tau^{-k} f(-1/N\tau)) \\ &= i^{2k} N^{-k} \tau^{-k} (-1/N\tau)^{-k} f\left(\frac{-1}{N(-1/N\tau)}\right) \\ &= f(\tau) \end{aligned}$$

The calculation above and Lemma 4.3.2 show that $W_N W_N^* = W_N^* W_N = I$. From the Spectral Theorem in linear algebra we have an orthogonal decomposition of $\mathcal{S}_k(\Gamma_1(N))$ into the eigenspaces of W_N. Since the only possible eigenvalues for a unitary operator is ± 1, let $\mathcal{S}_k(\Gamma_1(N))^+$ denote the eigenspace corresponding to the eigenvalue 1, and similarly for the eigenvalue -1, we get:

$$\mathcal{S}_k(\Gamma_1(N)) = \mathcal{S}_k(\Gamma_1(N))^+ \bigoplus \mathcal{S}_k(\Gamma_1(N))^-.$$

With the help of the orthogonal decomposition and the W_N operator we will now prove the analytic continuation of the L function associated to any cusp form.

Theorem 4.3.3. Suppose $f \in \mathcal{S}_k(\Gamma_1(N))^{\pm 1}$. Then the function $\Lambda_N(s) = N^{s/2} g(s)$ where $g(s)$ is the Mellin transform of f extends to an entire function satisfying the functional equation:

$$\Lambda_N(s) = \pm \Lambda_N(k-s).$$

Consequently, $L(s, f)$ extends to an analytic function to the full s plane.

Proof. Take $f \in \mathcal{S}_k(\Gamma_1(N))^{\pm 1}$. Then

$$\Gamma_N(s) = N^{s/2} \int_{t=0}^{\infty} f(it) t^s \frac{dt}{t} \tag{4.3.4}$$

Changing the variables from t to $t/\sqrt{N}$ in (4.3.4) we get,

$$\begin{aligned} \Gamma_N(s) &= \int_{t=0}^{\infty} f(it/\sqrt{N}) t^s \frac{dt}{t} \\ &= \int_{t=1}^{\infty} f(it/\sqrt{N}) t^s \frac{dt}{t} + \int_{t=0}^{1} f(it/\sqrt{N}) t^s \frac{dt}{t} \end{aligned}$$

We will deal with these two integrals separately and show that they converge to an entire function. Consider the first summand, $\int_{t=1}^{\infty} f(it/\sqrt{N}) t^s \frac{dt}{t}$. Observe that as $t \to \infty$, $f(it/\sqrt{N})$ is of order $e^{-2\pi t/\sqrt{N}}$. To show that the integral from 1 to ∞ converges to an entire function we will use the following theorem from complex analysis.

Theorem 4.3.5. Let $F(s, z)$ be defined for $(s, z) \in \Omega \times [0, 1]$ where Ω is an open set in $\mathbb{C}$. Suppose F satisfies the following properties.

1. It is holomorphic in s for each $z \in [0, 1]$.
2. It is continuous on $\Omega \times [0, 1]$.

Then the function f on Ω given by the following integral is holomorphic.

$$f(s) = \int_0^1 F(s, z) ds$$

Coming back to our integral, we take $f_n(s) = \int_1^n f(it/\sqrt{N}) t^s \frac{dt}{t}$. This clearly satisfies the properties in the Theorem 4.3.5 and so for each $n \in \mathbb{Z}^+$, the function f_n is holomorphic. It remains to show that the sequence $\{f_n\}_n$ converge uniformly on compact subsets of Ω. Let K be a compact subset of Ω such that the $\mathrm{Re}(s)$ is bounded above by some positive constant σ_0. Then

$$\begin{aligned} \left| \int_1^{\infty} f(it/\sqrt{N}) t^s \frac{dt}{t} - \int_1^{\infty} f(it/\sqrt{N}) t^s \frac{dt}{t} \right| &\leq \int_n^{\infty} |f(it/\sqrt{N}) t^s| \frac{dt}{t} \\ &\leq \int_n^{\infty} C e^{-2\pi t/\sqrt{N}} t^{\sigma_0} \frac{dt}{t} \end{aligned}$$

The last expression is independent of s and is finite. It follows that the integral $\int_1^{\infty} f(it/\sqrt{N}) t^s \frac{dt}{t}$ defines a holomorphic function.

For the second integral, first observe that

$$W_N f(i/\sqrt{N} t) = t^k f\left(\frac{it}{\sqrt{N}}\right) \tag{4.3.6}$$

Now write $\int_{t=0}^{1} f(it/\sqrt{N})t^s \frac{dt}{t} = \int_{t=0}^{1}(W_N f)(i/\sqrt{N}t)t^{s-k}\frac{dt}{t}$. Changing the variable by substituting $t = 1/x$, we see that,

$$\begin{aligned}\int_{t=0}^{1} f(it/\sqrt{N})t^s \frac{dt}{t} &= -\int_{t=\infty}^{1} (W_N f)(ix/\sqrt{N})x^{k-s}\frac{dx}{x} \\ &= \int_{t=1}^{\infty} \pm f(ix/\sqrt{N})x^{k-s}\frac{dx}{x}\end{aligned}$$

By the same argument as done for the first integral we see that this also defines an analytic function on the whole s-plane. This shows that $\Lambda_N(s)$ extends to an entire function to the s-plane. Combining the two integrals, what we have is the following:

$$\Lambda_N(s) = \int_{t=1}^{\infty} \left(f(it/\sqrt{N})t^s \pm f(it/\sqrt{N})t^{k-s} \right) \frac{dt}{t} = \pm\Lambda_N(k-s).$$

□

Observe that the crucial step in the proof was to use the result in (4.3.6) which allowed us to show that the integral from 0 to 1 from converges. For this purpose it was important to work with the W_N operator. In particular, if we were to work with level 1 modular forms, things would have been much simpler as the W_N operator reduces to just one of the generators of $SL_2(\mathbb{Z})$. That is, if f is a cusp form of level 1 then, $f(z) = (-1)^k z^{-k} f(1/z)$. For $z = iy$ with $y > 0$, the equation is equivalent to saying that,

$$f(iy) = i^k y^{-k} f(i/y) \tag{4.3.7}$$

The same argument as done in the previous proof would result in the following functional equation for L-function associated to cusp forms of level 1:

$$\Lambda(s, f) = i^k \Lambda(k-s, f) \tag{4.3.8}$$

Assume $f \in \mathcal{S}_k(N, \psi)$ where ψ is a Dirichlet character $\mod N$. We work in a much more general setting now to find out the functional equation for L-functions "twisted" by a primitive Dirichlet character. Let $f \in S_k(N, \psi)$. Consider the function, $g = w_N f$. We know that if $w_N = [\delta]_k$, then the matrix δ normalizes $\Gamma_1(N)$. Moreover, if $\gamma = \begin{pmatrix} a & b \\ cN & d \end{pmatrix} \in \Gamma_0(N)$ then, $\delta\gamma\delta^{-1} \in \Gamma_0(N)$ with lower right entry a. Also, observe that since $ad - bNc = 1$, we have $\psi(a) = \overline{\psi(d)}$. This helps us conclude that $g[\gamma]_k = f[\delta\gamma]_k = f[\delta\gamma\delta^{-1}]_k[\delta]_k = \psi(a)f[\delta] = \overline{\psi(d)}g$. This implies that $g \in \mathcal{S}_k(N, \overline{\psi})$.

Definition 4.3.9. (Dirichlet L-function) Suppose χ is a Dirichlet character $\bmod N$. Then the expression for Dirichlet L-functions is given by:

$$L(s,\chi) = \sum_{n=1}^{\infty} \chi(n) n^{-s}$$

Remark 4.3.10. For notational convenience, we drop the notation $w_N = [\delta]_k$, and refer to the weight k operator as simply $[w_N]_k$.

Theorem 4.3.11. Suppose $f \in \mathcal{S}_k(N,\psi)$. and $g = f[w_N]_k \in \mathcal{S}_k(N,\overline{\psi})$. Let χ be a primitive character modulo D. Assume D and N are co-prime. Let $A(n)$ and $B(n)$ be the Fourier coefficients of f and g. Suppose the "twisted" L-functions for the primitive character χ associated to f and g and their respective Mellin transforms are given as follows.

$$L(s,f,\chi) = \sum_{n=1}^{\infty} \chi(n)A(n)n^{-s} \ ; \ L(s,g,\chi) = \sum_{n=1}^{\infty} \chi(n)B(n)n^{-s}$$

$$\Lambda(s,f,\chi) = (2\pi)^{-s}\Gamma(s)L(s,f,\chi) \ ; \ \Lambda(s,g,\chi) = (2\pi)^{-s}\Gamma(s)L(s,g,\chi)$$

Then,

$$\Lambda(s,f,\chi) = i^k \chi(N)\psi(D)\frac{\tau(\chi)^2}{D}(D^2N)^{-s+k/2}\Lambda(k-s,g,\overline{\chi}) \tag{4.3.12}$$

Proof. We set the following notations first:

$$f_\chi(\tau) = \sum_{n=1}^{\infty} \chi(n)A(n)q^n \ ; \ g_{\overline{\chi}}(\tau) = \sum_{n=1}^{\infty} \overline{\chi}(n)B(n)q^n.$$

We use the formula in (3.2.14) to write:

$$\chi(n)q^n = \frac{\chi(-1)(\tau(\chi)}{D} \sum_{\substack{m \bmod D \\ (m,D)=1}} \overline{\chi}(m)e^{2\pi i(m/D+z)} \tag{4.3.13}$$

Multiplying (4.3.13) by $A(n)$ and taking the sum over n from one to infinity we get,

$$\begin{aligned} f_\chi(\tau) &= \frac{\chi(-1)(\tau(\chi)}{D} \sum_{\substack{m \bmod D \\ (m,D)=1}} \overline{\chi(m)} \sum_{n=1}^{\infty} A(n)e^{2\pi i(\frac{zD+m}{D})} \\ &= \frac{\chi(-1)\tau(\chi)}{D} \sum_{\substack{m \bmod D \\ (m,D)=1}} \overline{\chi(m)} f\left[\begin{pmatrix} D & m \\ 0 & D \end{pmatrix}\right]_k \end{aligned}$$

This, and Lemma 3.3.2 helps us to conclude that $f\chi$ is a cusp form as well with respect to some congruence subgroup. Keeping in mind that $f = g[w_N]_k$ Compute that,

$$
\begin{aligned}
& f_\chi\left[\begin{pmatrix} 0 & -1 \\ D^2N & 0 \end{pmatrix}\right]_k = f_\chi\left[\begin{pmatrix} 0 & -1/ND \\ D & 0 \end{pmatrix}\right]_k \\
= \; & \frac{\chi(-1)\tau(\chi)}{D} \sum_{\substack{m \bmod D \\ (m,D)=1}} \overline{\chi(m)} g\left[\begin{pmatrix} 0 & -1 \\ D^2N & 0 \end{pmatrix}\begin{pmatrix} D & m \\ 0 & D \end{pmatrix}\begin{pmatrix} 0 & -1/ND \\ D & 0 \end{pmatrix}\right]_k \\
= \; & \frac{\chi(-1)\tau(\chi)}{D} \sum_{\substack{m \bmod D \\ (m,D)=1}} \overline{\chi(m)} g\left[\begin{pmatrix} D^2 & 0 \\ -NmD & 1 \end{pmatrix}\right]_k \\
= \; & \frac{\chi(-1)\tau(\chi)}{D} \sum_{\substack{m \bmod D \\ (m,D)=1}} \overline{\chi(m)} g\left[\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix}\begin{pmatrix} D & r \\ 0 & D \end{pmatrix}\right]_k
\end{aligned}
$$

In the last equality, for each m, the integers r and s chosen such that $Ds - rNm = 1$. Now, observe that, $\overline{\chi}(m) = \chi(-N)\chi(r)$. This helps us to write the last expression as:

$$
f_\chi\left[\begin{pmatrix} 0 & -1 \\ D^2N & 0 \end{pmatrix}\right]_k = \frac{\chi(N)\tau(\chi)}{D} \sum_{\substack{r \bmod D \\ (r,D)=1}} \chi(r) g\left[\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix}\begin{pmatrix} D & r \\ 0 & D \end{pmatrix}\right]_k \tag{4.3.14}
$$

Since $g \in \mathcal{S}_k(N, \overline{\psi})$, we have,

$$
g\left[\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix}\right]_k = \overline{\psi(s)} g = \psi(D) g \tag{4.3.15}
$$

By following a similar argument with $g_{\overline{\chi}}$ instead of f_χ we land up with the expression:

$$
g_{\overline{\chi}} = \frac{\chi(-1)\tau(\overline{\chi})}{D} \sum_{\substack{m \bmod D \\ (r,D)=1}} \overline{\chi(r)} g\left[\begin{pmatrix} D & r \\ 0 & D \end{pmatrix}\right]_k \tag{4.3.16}
$$

Combining the observations in (4.4.13) and (4.4.12), with the results in equation (3.2.13) and Proposition 3.2.11, we further write the expression in (4.4.11) as

$$
\begin{aligned}
& \frac{\chi(N)\tau(\chi)}{D} \sum_{\substack{r \bmod D \\ (r,D)=1}} \chi(r) g\left[\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix}\begin{pmatrix} D & r \\ 0 & D \end{pmatrix}\right]_k \\
= \; & \frac{\chi(N)\tau(\chi)}{D} \sum_{\substack{r \bmod D \\ (r,D)=1}} \chi(r)\psi(D) g\left[\begin{pmatrix} D & r \\ 0 & D \end{pmatrix}\right]_k \\
= \; & \frac{\chi(N)\psi(D)\tau(\chi)^2}{D} g_{\overline{\chi}}
\end{aligned}
$$

This helps us to write,

$$f_\chi = \frac{\chi(N)\psi(D)\tau(\chi)^2}{D} g_{\overline{\chi}} \Bigg[\begin{pmatrix} 0 & -1 \\ D^2N & 0 \end{pmatrix}^{-1} \Bigg]_k \tag{4.3.17}$$

Evaluating both the sides at $\tau = iy$, we get,

$$f_\chi(iy) = \frac{\chi(N)\psi(D)\tau(\chi)^2}{D} i^k y^{-k} N^{-k/2} D^{-k} g_{\overline{\chi}} \left(\frac{i}{D^2Ny} \right) \tag{4.3.18}$$

Now running the same argument as in the proof of 4.3.3, we get the relations between the functional equations:

$$\Lambda(s, f, \chi) = i^k \chi(N)\psi(D) \frac{\tau(\chi)^2}{D} (D^2N)^{-s+k/2} \Lambda(k-s, g, \overline{\chi}) \tag{4.3.19}$$

□

4.4 Converse Theorem of Weil

In this section, we aim to establish the converse of the theorems we proved in the previous section regarding the analytic continuation of the L-function associated to cusp forms. More precisely, we answer the following question: Given a sequence of Fourier coefficients such that the L-function associated to the coefficients can be extended to the whole complex plane and satisfies the desired functional equation, can we say that the Fourier coefficients are coming from a modular form?

It turns out that this is indeed true with some additional assumptions. The converse theorem for level 1 cusp forms is pretty straightforward and we prove it so as to lay out the essential idea. We will go backwards in the proof of Theorem 4.3.3 to arrive at the equation (4.3.7) using the Mellin inversion formula.

Theorem 4.4.1. Let $A(n) > 0$ be a sequence of complex numbers such that $|A(n)| = \mathcal{O}(n^K)$ for some sufficiently large real number K. Let $L(s)$ be defined by the series $\sum_{n=1}^{\infty} A(n)n^{-s}$ convergent for sufficiently large s. Assume that $\Lambda(s)$ has analytic continuation for all s, is bounded in every vertical strip strip $\sigma_1 \leq \mathrm{Re}(s) \leq \sigma_2$, and satisfies the functional equation given in 4.3.3. Then the function $f(\tau) = \sum_{n=1}^{\infty} A(n)q^n$ is an element of $\mathcal{S}_k(\Gamma(1))$.

.

Proof. Observe that since $|A(n)|$ are of polynomial orders for all $n \geq 1$, the Fourier series defining the function $f(\tau)$ converges. Also, this implies that for sufficiently large s, the L series converges. Clearly the function $f(z)$ is periodic and therefore it satisfies the modularity condition for the matrix $T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. We next show that f satisfies the modularity condition for the matrix $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Consequently, by Proposition 1.1, it will be cusp form of level 1. In order to establish the above statement, we need to check that for all $\tau \in \mathbb{H}$,

$$f(\tau) - \tau^{-k} f(-1/\tau) = 0. \tag{4.4.2}$$

It is enough to show that $f(iy) = i^k y^{-k} f(i/y)$ for positive real values of y, because the function on the left hand side of the equation (4.4.2) will then be a holomorphic function vanishing on the positive imaginary axis. By the Identity Theorem in complex analysis, the zeroes of any non-zero function cannot have an accumulation point in the interior of the domain and so the function $f(\tau) - \tau^{-k} f(-1/\tau)$ must vanish at all of the upper half plane. We will obtain the desired result using the Mellin inversion formula. From Theorem 4.2.6, we know that,

$$\int_0^\infty f(iy) y^s \frac{dy}{y} = \Lambda(s, f)$$

For sufficiently large σ where the above equation holds, we can use the Mellin inversion formula as discussed in section 4.1.3 to write,

$$f(iy) = \frac{1}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \Lambda(s, f) y^{-s} ds = \frac{i^k}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \Lambda(k - s, f) y^{-s} ds \tag{4.4.3}$$

To the get the last equality, we have used the functional equation. We make a change of variable from $k - s$ to s to get that:

$$f(iy) = \frac{i^k y^{-k}}{2\pi i} \int_{k - \sigma - i\infty}^{k - \sigma + i\infty} \Lambda(s, f) y^{-s} ds \tag{4.4.4}$$

We would like to shift the line of integral back to the original line. In order to apply Cauchy's Theorem, we first estimate the integrand. Notice that for a sufficiently large $\sigma_1 > 0$ where the L-function is bounded, using the Stirling approximation in (4.1.1), we get a constant C such that $|\Lambda(s, f) y^{-s}| \leq C|t|^{\sigma - 1/2} e^{-|t|/2}$. This goes to 0 rapidly as t goes to infinity at both ends. In order to use the Phragmén Lindelöf principle, we further compare $e^{|t|/2} \geq C_1 t^M$ for some positive constants C_1 and M such that $M > \sigma_1 - 1/2$. This implies that $|\Lambda(s, f) y^{-s}| \leq C_2 |t|^{\sigma_1 - 1/2 - M}$. Similarly for sufficiently negative $\sigma_2 < 0$, we use the functional equation to relate $\Lambda(\sigma_2 + it, f)$ to $\Lambda(k - \sigma_2 - it, f)$

so that the L function makes sense for $\mathrm{Re}(s) = k - \sigma_2$. Running the same argument as above, we see that $|\Lambda(\sigma_2 + it)y^{-s}| = \mathcal{O}|t|^{k-\sigma_2-1/2-M'}$ for some constant M' such that $k-\sigma_2-1/2-M' < 0$. So, even at the other end for sufficiently negative σ, the Λ function decays to 0 as t goes to infinity. By assumption $\Lambda(s,f)$ is bounded in all vertical strips and so $\Lambda(s,f)$ is trivially of order $e^{|s|}$. By the Phragmén-Lindelöf principal stated in 4.1.5, the function $\Lambda(s,f)$ is of order $|s|^A$ uniformly in the strip $\sigma \in [\sigma_1, \sigma_2]$ for some constant $A < 0$. This allows us to conclude that $\Lambda(s,f)$ decays uniformly to 0 as t goes to infinity for σ in any arbitrary compact set.

Now we integrate the function $|\Lambda(s,f)|y^{-k}$ over any rectangular contour where the vertical lines of the contour are the lines (say) $\sigma + it$ and $k - \sigma + it$. Uniform convergence of the integrand to 0 as $t \to \infty$ will make the integral along the horizontal lines go to zero. By Cauchy's Theorem we can shift the line of integral in (4.4.5) to conclude that,

$$f(iy) = \frac{i^k y^{-k}}{2\pi i} \int_{\sigma - i\infty}^{\sigma + i\infty} \Lambda(s,f) y^{-s} ds = i^k y^{-k} f(i/y) \tag{4.4.5}$$

□

Since we explicitly knew the generators of $SL_2(\mathbb{Z})$, it made our task easier to go backwards. However, while working with modular forms of level N, the fact that we don't know the generators of the subgroup has to accounted for by assuming sufficiently many functional equations for L-functions twisted by a primitive character.

Theorem 4.4.6. (Weil) Suppose N is a positive integer and ψ is a Dirichlet character modulo N, not assumed to be primitive. Suppose that $A(n)$ and $B(n)$ are sequences of complex numbers such that $|A(n)|$ and $|B(n)|$ are of order n^K for some sufficiently large real number K. Let D be relatively prime to N and χ be a primitive Dirichlet character modulo D. Let $L_1(s,\chi) = \sum_{n>0} \chi(n)A(n)n^{-s}$ and let $L_2(s,\overline{\chi}) = \sum_{n>0} \chi(n)B(n)n^{-s}$. Write $\Lambda_1(s,\chi) = (2\pi)^{-s}\Gamma(s)L_1(s,\chi)$ and $\Lambda_2(s,\overline{\chi}) = (2\pi)^{-s}\Gamma(s)L_2(s,\overline{\chi})$. Let S be the set of finitely primes including those dividing N. Assume that whenever the conductor D of χ is either 1 or a prime, then $D \notin S$. Furthermore, assume that $\Lambda_1(s,\chi)$ and $\Lambda_1(s,\overline{\chi})$ have analytic continuation to all of s, are bounded in every vertical strip $\sigma_1 \leq \mathrm{Re}(s) \leq \sigma_2$ and satisfy the functional equation:

$$\Lambda_1(s,\chi) = i^k \chi(N)\psi(D)\frac{\tau(\chi)^2}{D}(D^2N)^{-s+k/2}\Lambda_2(k-s,\overline{\chi}) \tag{4.4.7}$$

Then $f(\tau) = \sum_n A(n)q^n$ is a modular form in $\mathcal{S}_k(N,\psi)$.

Remark 4.4.8. Observe that the matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ lies in the group $\Gamma_1(N)$. This implies that modular forms with respect to $\Gamma_1(N)$ have period 1 and so it makes sense to define the Fourier series of f in Theorem 4.4.6 with $q = e^{2\pi i t}$.

We begin by remarking that the since f and g are defined as Fourier series, they are clearly holomorphic. The growth conditions on the coefficients also ensures that the series converges. Moreover, by the result mentioned in Proposition 1.4.10, it also takes care of the fact that f is holomorphic at the cusps. It only remains to show that f satisfies the modularity condition and lies in the ψ-space of $\mathcal{M}_k(\Gamma_1(N))$. Before moving on to proving this, we need a small result. Suppose $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then γ has two fixed points if and only if $|tr(\gamma)| = |a+d| < 2$. This can be shown along similar lines as in the beginning of section 2.3.0.1. We call such matrices as *elliptic*.

Lemma 4.4.9. Let f be a holomorphic function on $\mathbb{H}$ and let $M \in SL_2(\mathbb{R})$ be elliptic of infinite order such that $f[M]_k = f$. Then $f = 0$.

Proof. Let $a \in \mathbb{H}$ be the fixed point of the elliptic element M. Consider the transformation $C = \begin{pmatrix} 1 & -a \\ 1 & -\overline{a} \end{pmatrix}$. As seen in section 2.4, this takes a to 0 and $\overline{a}$ to ∞ so that CMC^{-1} fixes 0 and ∞. This implies that $CMC^{-1} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}$. Since M is of infinite order, α is not a root of unity. Let $g = f[C]_k$. Observe that $g[CMC^{-1}]_k = g$ and so $g(z) = \alpha^k g(\alpha^2 z)$. Since f is holomorphic on $\mathbb{H}$, it has a power series expansion about a. The transformation C takes a to 0 and consequently g has a power series expansion about 0. It follows that:

$$g(z) = \sum_{n=0}^{\infty} b_n z^n = \sum_{n=0}^{\infty} \alpha^{2n+k} b_n z^n.$$

This implies that $b_n = \alpha^{2n_k} b_n$. Now, $\alpha^{2n+k} \neq 1$ for any n, k. It follows that $b_n = 0$ for all $n \in \mathbb{Z}^+$. Finally, we get $g = 0$ and hence $f = g[C^{-1}]_k = 0$. $\square$

Proof of the Theorem 4.4.6. We set some notation first. Let

$$f_\chi(\tau) = \sum_{n=1}^{\infty} A(n)\chi(n)q^n \;;\; g_{\overline{\chi}}(\tau) = \sum_{n=1}^{\infty} B(n)\overline{\chi(n)}q^n$$

where χ is a primitive Dirichlet character $\mod D$ and $D \neq 1$ or $D \notin S$. The growth conditions on $A(n)$ and and $B(n)$ ensure that the series converge. Our first task is go backwards in the proof of Theorem 4.3.11 from (4.3.19) to (4.4.10) using the Mellin inversion formula. Again, by the Identity Theorem from complex analysis, it is enough to show that the equation in (4.3.18) holds for f_χ and $g_{\overline{\chi}}$ given above.

Assume that either $D = 1$ or D is a prime not in S. For sufficiently large σ, we use the Mellin inversion formula as follows. Since $\Lambda_1(s, \chi)$ has analytic continuation to the whole s plane and is given by the expression

$$\int_0^\infty f_\chi(iy)y^s\frac{dy}{y} = \Lambda_1(s,\chi)$$

we can use the Mellin inversion formula and write:

$$f_\chi(iy) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty}\Lambda_1(s,\chi)y^{-s}ds$$

The functional equation in (4.4.7) holds for prime $D \notin S$ or when $D=1$. We use it to to write the above expression as:

$$f_\chi(iy) = \frac{i^k\chi(N)\psi(D)\tau(\chi)^2}{2\pi iD}\int_{\sigma-i\infty}^{\sigma+i\infty}(D^2N)^{-s+k/2}\Lambda_2(k-s,\overline{\chi})y^{-s}ds$$

Arguing the same way to estimate the integrand as done in the previous proof, the Phragmén-Lindelöf principal helps us to shift the line of integral as follows:

$$\begin{aligned} f_\chi(iy) &= \frac{i^k\chi(N)\psi(D)\tau(\chi)^2}{2\pi iD}\int_{(k-\sigma)-i\infty}^{(k-\sigma)+i\infty}(D^2N)^{s-k/2}\Lambda_2(s,\overline{\chi})y^{s-k}ds \\ &= \frac{i^k\chi(N)\psi(D)\tau(\chi)^2}{2\pi iD}\int_{\sigma-i\infty}^{\sigma+i\infty}(D^2N)^{s-k/2}\Lambda_2(s,\overline{\chi})y^{s-k}ds \end{aligned}$$

This implies that,

$$f_\chi(iy) = \frac{\chi(N)\psi(D)\tau(\chi)^2}{D}i^k(D^2N)^{-k/2}y^{-k}g_{\overline{\chi}}\left(\frac{i}{D^2Ny}\right)$$

This establishes (4.4.10). Recall that in the proof of Theorem 4.3.11, in order to establish the expression in (4.4.11), we only require the fact that $f[w_N]_k = g$ and the rest is all manipulative computation. But this is achieved exactly when we put $D=1$ in (4.4.10). Even the relation in (4.3.19) is true regardless of the fact that f and g are modular forms. In summary, what we have in hand are the following three expressions.

$$f_\chi\left[\begin{pmatrix}0 & -1\\ D^2N & 0\end{pmatrix}\right]_k = \frac{\chi(N)\psi(D)\tau(\chi)^2}{D}g_{\overline{\chi}} \tag{4.4.10}$$

$$f_\chi\left[\begin{pmatrix}0 & -1\\ D^2N & 0\end{pmatrix}\right]_k = \frac{\chi(N)\tau(\chi)}{D}\sum_{\substack{r \bmod D\\ (r,D)=1}}\chi(r)g\left[\begin{pmatrix}D & -r\\ -Nm & s\end{pmatrix}\begin{pmatrix}D & r\\ 0 & D\end{pmatrix}\right]_k \tag{4.4.11}$$

$$g_{\overline{\chi}} = \frac{\chi(-1)\tau(\overline{\chi})}{D}\sum_{\substack{m \bmod D\\ (r,D)=1}}\overline{\chi(r)}g\left[\begin{pmatrix}D & r\\ 0 & D\end{pmatrix}\right]_k \tag{4.4.12}$$

From these three equations, we would now like to establish the expression in (4.4.13), that is,

$$g\left[\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix}\right]_k = \overline{\psi(s)}g = \psi(D)g \tag{4.4.13}$$

where $Ds - rNm = 1$. This is a non-trivial task. Let $D \notin S$ be a prime.

Claim 4.4.14. If $c(r)$ is any function on $\mathbb{Z}/D\mathbb{Z}$ such that $\sum\limits_{r \bmod D} c(r) = 0$, then,

$$\sum_{\substack{r \bmod D \\ (r,D)=1}} c(r)g\left[\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix}\begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix}\right]_k = \sum_{\substack{r \bmod D \\ (r,D)=1}} c(r)\psi(D)g\left[\begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix}\right]_k \tag{4.4.15}$$

Here, for a given r we choose m and s, both dependent on r, so that $Ds - Nmr = 1$.

Proof of claim 4.4.14. Plugging the expression in (4.4.12) for g_χ in (4.4.10), we get

$$f_\chi\left[\begin{pmatrix} 0 & -1 \\ D^2N & 0 \end{pmatrix}\right]_k = \frac{\chi(N)\psi(D)\tau(\chi)^2\chi(-1)\tau(\overline{\chi})}{D^2} \sum_{\substack{m \bmod D \\ (r,D)=1}} \overline{\chi(r)}g\left[\begin{pmatrix} D & r \\ 0 & D \end{pmatrix}\right]_k$$

Equating the above expression on the right with the expression in (4.4.13) and simplifying using the result in 3.2.13, we arrive at the following equality:

$$\sum_{\substack{r \bmod D \\ (r,D)=1}} \chi(r)g\left[\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix}\begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix}\right]_k = \sum_{\substack{r \bmod D \\ (r,D)=1}} \chi(r)\psi(D)g\left[\begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix}\right]_k$$

We note that the Dirichlet characters mod D form a $D-1$ dimensional space. When we exclude the identity operator, the rest span a $D-2$ dimensional space such that the functions in the space satisfy the property that sum of the evaluation on residue classes mod D is zero. To conclude the above result for any function $c(r)$ in claim 4.4.14, we use the fact that space of such functions form a $D-2$ dimensional vector space as well. Therefore, the space is spanned by the Dirichlet characters modulo D and the result now follows for any arbitrary function with the desired properties. □

Now that (4.4.15) is true, we express it in a slightly different way. We first extend the right action of $GL_2(\mathbb{Q})^+$ on holomorphic functions on $\mathbb{H}$ via the weight k operator to a right action of the group algebra $\mathbb{C}[GL_2(\mathbb{Q})^+]$. More precisely, if $\alpha \in \mathbb{C}[GL_2(\mathbb{Q})^+]$ is an element such that $\alpha = \alpha_1\gamma_1 + \cdots + \alpha_n\gamma_n$, with $\gamma_i \in GL_2(\mathbb{Q})^+$, then define the action of α on f naturally as

$$f[\alpha]_k = \sum_i \alpha_i f[\gamma_i]_k$$

Let Ω denote the annihilator of the element $g = \sum_{n\geq 1} B(n)q^n$ in $\mathbb{C}[GL_2(\mathbb{Q})^+]$. This whole set up allows us to express the equation in (4.4.15) as:

$$\sum_{\substack{r \bmod D \\ (r,D)=1}} c(r) \begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix} \begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix} \equiv \sum_{\substack{r \bmod D \\ (r,D)=1}} c(r)\psi(D) \begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix} \bmod \Omega$$

In particular, when D is an odd prime, then for a fixed r, we take c to be the function that is 1 on the residue classes r and -1 on the residue classes $-r$ and zero elsewhere. Rewriting the expression with the particular $c(r)$ as described above we get,

$$\left(\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix} - \psi(D)\right) \begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix} \equiv \left(\begin{pmatrix} D & r \\ Nm & s \end{pmatrix} - \psi(D)\right) \begin{pmatrix} 1 & -r/D \\ 0 & 1 \end{pmatrix} \bmod \Omega \tag{4.4.16}$$

Let D and s be distinct odd primes not in S such that $Ds \equiv 1 \bmod N$.

Choose r and m such that $Ds - Nmr = 1$. Then replacing s with D and D with s in (4.4.16), we land up with the expression:

$$\left(\begin{pmatrix} s & -r \\ -Nm & D \end{pmatrix} - \psi(s)\right) \begin{pmatrix} 1 & r/s \\ 0 & 1 \end{pmatrix} \equiv \left(\begin{pmatrix} s & r \\ Nm & D \end{pmatrix} - \psi(s)\right) \begin{pmatrix} 1 & -r/s \\ 0 & 1 \end{pmatrix} \bmod \Omega \tag{4.4.17}$$

Multiplying (4.4.16) with the matrix $\begin{pmatrix} 1 & r/s \\ 0 & 1 \end{pmatrix}$ we get:

$$\left(\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix} - \psi(D)\right) \begin{pmatrix} 1 & 2r/D \\ 0 & 1 \end{pmatrix} \equiv \left(\begin{pmatrix} D & r \\ Nm & s \end{pmatrix} - \psi(D)\right) \bmod \Omega \tag{4.4.18}$$

Further, we multiply the expression in (4.4.17) with

$$-\psi(D) \begin{pmatrix} 1 & r/D \\ 0 & 1 \end{pmatrix} \begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix} \begin{pmatrix} 1 & 2r/D \\ 0 & 1 \end{pmatrix}$$

Regroup the matrices using the fact that $\psi(sD) = 1$ to finally get that

$$\begin{aligned} \left(\begin{pmatrix} D & r \\ Nm & s \end{pmatrix} - \psi(D)\right) \begin{pmatrix} s & -r \\ -Nm & D \end{pmatrix} \begin{pmatrix} 1 & 2r/s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix} \begin{pmatrix} 1 & 2r/s \\ 0 & 1 \end{pmatrix} \\ \equiv \left(\begin{pmatrix} D & -r \\ -Nm & s \end{pmatrix} - \psi(D)\right) \begin{pmatrix} 1 & 2r/D \\ 0 & 1 \end{pmatrix} \bmod \Omega \end{aligned} \tag{4.4.19}$$

We write the big factor in the left hand side of the expression (4.4.16):

$$\begin{pmatrix} s & -r \\ -Nm & D \end{pmatrix} \begin{pmatrix} 1 & 2r/s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} D & -r-Nm & s \end{pmatrix} \begin{pmatrix} 1 & 2r/s \\ 0 & 1 \end{pmatrix} = M$$

Then equations in (4.4.18) and (4.4.19) combine to show that,

$$\left(\begin{pmatrix} D & r \\ Nm & s \end{pmatrix} - \psi(D) \right) (1-M) \in \Omega \tag{4.4.20}$$

Multiplying the matrices which define M we compute that,

$$M = \begin{pmatrix} 1 & 2r/D \\ -2Nm/s & -3+\frac{4}{DS} \end{pmatrix}$$

Let $g_1 = \left(g \left[\begin{pmatrix} D & r \\ Nm & s \end{pmatrix} \right]_k - \psi(D)g \right)$ Since g is a holomorphic function on $\mathbb{H}$, g_1 is holomorphic on $\mathbb{H}$. Because of (4.4.20), we have that $g_1 - g_1[M]_k = 0$. This implies that $g_1 = g_1[M]_k$. Now notice that $|tr(M)|$ equals $|1-3+\frac{4}{Ds}| < 2$ since D and s are odd primes. This gives us that the matrix M is elliptic. Also, notice that the trace of M is a rational number which is not an integer and so not an algebraic integer. It follows that the eigenvalues are not roots of unity and so M has infinite order. We can now use Lemma 4.4.9 proved in the beginning to conclude that $g_1 = 0$. So far whatever computation we did proves that, for D and s odd primes not in S,

$$g \left[\begin{pmatrix} D & r \\ Nm & s \end{pmatrix} \right]_k = \psi(D)g \; ; \; Ds - rNm = 1. \tag{4.4.21}$$

In general, let $\begin{pmatrix} a & b \\ cN & d \end{pmatrix}$ be any matrix in $\Gamma_0(N)$. Then $ad - bcN = 1 \bmod N$ and so $(a, cN) = (d, cN) = 1$. We now use Dirichlet's Theorem on primes in an arithmetic progression which states that for any integer a' such that $(a', N') = 1$, there exists infinitely many primes p such that $p \equiv a' \bmod N'$. Dirichlet's Theorem allows us to find primes D and s and integers u and v such that $D = a - uNc$ and $s = d - vNc$. Observe that $sD \equiv 1 \bmod N$ and so the result in (4.4.21) holds. Now compute that,

$$g \left[\begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \begin{pmatrix} D & r \\ Nc & s \end{pmatrix} \begin{pmatrix} 1 & v \\ 0 & 1 \end{pmatrix} \right]_k = \begin{pmatrix} a & av+ud-uvNc+r \\ Nc & d \end{pmatrix}$$

We would like to put $r = b + uvNc - av - ud$ such that the upper right entry equals b. One can also easily check that with $\det \begin{pmatrix} D & r \\ Nc & s \end{pmatrix} = 1$.

Using the fact that f and g are periodic since they are defined by a Fourier series and the result in (4.4.21) for D and s odd primes, we finally conclude that

$$g\left[\begin{pmatrix}1 & u\\0 & 1\end{pmatrix}\begin{pmatrix}D & r\\Nc & s\end{pmatrix}\begin{pmatrix}1 & v\\0 & 1\end{pmatrix}\right]_k = \psi(D) = \psi(a)g = \overline{\psi}(d)g = g\left[\begin{pmatrix}a & b\\cN & d\end{pmatrix}\right]_k$$

In particular when $a \equiv d \equiv 1 \bmod N$, we see that $g \in \mathcal{S}_k(\Gamma_1(N))$ and in fact the above result helps us conclude that $g \in \mathcal{S}_k(N, \overline{\psi})$. Moreover the relation $g[w_N^{-1}]_k = f$ helps us conclude that $f \in \mathcal{S}_k(N, \psi)$. This completes the proof of the Converse Theorem of Weil. $\square$

Appendix A

A quick introduction to the analytic theory of elliptic curves

A.1 Lattices and the complex torus

In this Appendix we give a brief introduction to complex torus, Weierstrass $\wp$ function and elliptic curves. For most of the proofs, a brief sketch is provided. For a detailed study, one can refer to [3].

Definition A.1.1 (Lattice). A lattice in $\mathbb{C}$ is the set $\Lambda = \omega_1\mathbb{Z} \bigoplus \omega_2\mathbb{Z}$, where ω_1 and ω_2 belong to $\mathbb{C}$ and are linearly independent over $\mathbb{R}$.

The complex numbers ω_1 and ω_2 form a "basis" of the lattice Λ. Since we are concerned with the upper half plane, by convention we set $\omega_1/\omega_2 \in \mathbb{H}$.

Definition A.1.2 (Complex torus). Let Λ be the lattice generated by ω_1 and ω_2. The complex torus is the set $\mathbb{C}/\Lambda = \{z + \Lambda \mid z \in \mathbb{C}\}$.

When we visualize this set in $\mathbb{C}$, we get a parallelogram with opposite sides identified and hence the name. The next lemma helps us to characterize the basis of a lattice.

Lemma A.1.3. Let $\Lambda = \omega_1\mathbb{Z} \oplus \omega_2\mathbb{Z}$ and $\Lambda' = \omega_1'\mathbb{Z} \oplus \omega_2'\mathbb{Z}$ be two lattices such that ω_1/ω_2 and ω_1'/ω_2' belong to $\mathbb{H}$. Then $\Lambda' = \Lambda$ if and only if

$$\begin{pmatrix} \omega_1' \\ \omega_2' \end{pmatrix} = \gamma \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \text{ for some } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}).$$

Proof. Suppose that $\begin{pmatrix}\omega_1' \\ \omega_2'\end{pmatrix} = \gamma \begin{pmatrix}\omega_1 \\ \omega_2\end{pmatrix}$ for some $\gamma = \begin{pmatrix}a & b \\ c & d\end{pmatrix} \in SL_2(\mathbb{Z})$. Then clearly $\Lambda' \subseteq \Lambda$. For the other way round containment we multiply the equality with the matrix γ^{-1}.

Conversely, suppose $\Lambda = \Lambda'$, then $\begin{pmatrix}\omega_1' \\ \omega_2'\end{pmatrix} = A \begin{pmatrix}\omega_1 \\ \omega_2\end{pmatrix}$ and $\begin{pmatrix}\omega_1 \\ \omega_2\end{pmatrix} = B \begin{pmatrix}\omega_1' \\ \omega_2'\end{pmatrix}$. This implies that $AB = I$. Moreover, since ω_1/ω_2 and ω_1'/ω_2' belong to $\mathbb{H}$, the determinant of A is 1 and so $A \in SL_2(\mathbb{Z})$.

□

We know that any map between compact Riemann surfaces is either constant or a surjection. We next classify all non-constant maps between the complex tori.

Proposition A.1.4. Suppose $\varphi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ is a holomorphic map between the complex tori. Then there exists complex numbers a and b with $a\Lambda \subseteq \Lambda'$ such that $\varphi(z + \Lambda) = az + b + \Lambda'$. The map is invertible if and only if $a\Lambda = \Lambda'$.

Proof. (Sketch) We recall that $\mathbb{C}$ is the universal covering for the tori $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$. Suppose that p and p' are the corresponding covering maps. We can lift the map φ to a map $\widetilde{\varphi}$ between the coverings such that $\varphi \circ p = p' \circ \widetilde{\varphi}$. This implies that for a fixed $\omega \in \Lambda$, $\widetilde{\varphi}(z + \omega) = \widetilde{\varphi}(z) + \omega_z'$ with $\omega_z' \in \Lambda'$. Consider the function $z \mapsto \widetilde{\varphi}(z + \omega) - \widetilde{\varphi}(z)$. This is a continuous map from $\mathbb{C}$ to a discrete subgroup Λ of $\mathbb{C}$. This implies that the map is constant and so ω_z' does not depend on z. This gives us that for all $z \in \mathbb{C}$, $\widetilde{\varphi}(z + \omega) = \widetilde{\varphi}(z) + \omega'$ for some ω' in Λ'. Differentiating with respect to z we arrive at the following result:

$$\widetilde{\varphi}'(z + \omega) = \widetilde{\varphi}'(z)$$

This shows that $\widetilde{\varphi}'$ is doubly periodic with respect to Λ and hence a well defined function from $\mathbb{C}/\Lambda$ to $\mathbb{C}$. It follows that $\widetilde{\varphi}'$ is an entire and a bounded function and so constant. This characterizes $\widetilde{\varphi}$ showing that there exists constants a and b such that $\widetilde{\varphi}(z) = az + b$ which helps us conclude the result for φ as well. From this it is easy to check that $a\Lambda \subseteq \Lambda'$. If $a\Lambda \neq \Lambda'$ such that $a\Lambda \subsetneq \Lambda'$, then there exists some $z \in \Lambda'$ such that $z/a \notin \Lambda$. Consequently the map φ will not be an injection. Other way round, one can check that if $a\Lambda = \Lambda'$ then the map $z + \Lambda' \mapsto (z - b)/a + \Lambda$ inverts φ. □

Corollary A.1.5. Suppose $\varphi : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$ is a holomorphic map between the complex tori such that $\varphi(z + \Lambda) = az + b + \Lambda'$ with $a\Lambda \subseteq \Lambda'$. Then the map φ is a group homomorphism if and only if $b \in \Lambda'$ and so $\varphi(z + \Omega) = az + \Lambda'$.

In particular, the above corollary implies that there exists a non zero holomorphic group isomorphism between the complex tori $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ if and only if there exists $m \in \mathbb{C}$ such that $m\Lambda = \Lambda'$.

Lemma A.1.6. Every complex torus is isomorphic to a complex torus whose lattice is generated by 1 and a complex number $\tau \in \mathbb{H}$.

Proof. Let $\Lambda = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ be an arbitrary lattice with $\omega_1/\omega_2 \in \mathbb{H}$. Let $\tau = \omega_1/\omega_2$ and $\Lambda_\tau = \tau\mathbb{Z} + \mathbb{Z}$. We have that $(1/\omega_2)\Lambda = \Lambda_\tau$. By Corollary A.1.5, the map $\varphi_\tau : \mathbb{C}/\Lambda \to \mathbb{C}/\Lambda_\tau$ where $z + \Lambda$ maps to $(1/\omega_2)z + \Lambda_\tau$ is an isomorphism. □

Remark A.1.7. The complex number τ is not unique! If τ' is another such complex number, then $\tau' = \gamma\tau$ for some $\gamma \in SL_2(\mathbb{Z})$ by Lemma A.1.3.

A.2 The Weierstrass $\wp$ function

We know that doubly periodic functions with respect to a lattice can be considered to be functions on the torus $\mathbb{C}/\Lambda$. Since the image of such functions is bounded, all holomorphic double periodic functions are constant. The first simple example of a doubly periodic function is given by the following expression:

$$\sum_{\omega\in\Lambda}(z-\omega)^{-N} \ ; \ N \geq 3$$

For $N = 2$, the above series does not converge absolutely. The most famous example is obtained by adding a few "error terms" to the series.

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{\omega\in\Lambda \\ \omega\neq 0}}\left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right) \text{ for } z \in \mathbb{C}, z \neq \omega.$$

One can check that the series converges absolutely and uniformly on compact subsets of $\mathbb{C}/\Lambda$. This is the Weierstrass $\wp$ function which will help us relate the complex tori to elliptic curves.

Replacing z with $-z$, it is easy to see that $\wp(z)$ is an even function.

Proposition A.2.1. The Wierstrass $\wp$ function is periodic.

Proof. We differentiate term by term to see that $\wp'$ is periodic with respect to Λ.

$$\wp'(z) = \sum_{\substack{\omega\in\Lambda \\ \omega\neq 0}} \frac{2}{(z-\omega)^3}$$

This implies that for each $z \in \omega$, the funtion $\wp'(z+\omega) - \wp'(z) = 0$ for all $z \in \mathbb{C}$. It follows that $\wp(\omega + z) - \wp(z)$ is a constant function, say c_ω. Substituting $z = -\omega/2$ and using the fact that $\wp$ is an even function, the result follows. □

Remark A.2.2. By construction the Weierstrass $\wp$ function and its derivative have poles of order 2 and 3 respectively precisely at the points of the lattice.

A.2.1 Eisenstein series for a lattice

Let Λ be a lattice. The Eisenstein series for Λ denoted as $G_k(\Lambda)$ is given by the following expression:

$$G_k(\Lambda) = \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega_k}$$

By Lemma A.1.6, the Eisenstein series for any lattice reduces to the first examples of modular forms we saw.

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

Very surprisingly, they come up as coefficients in the Laurant series expansion of the $\wp$ function.

Proposition A.2.3. The Laurant series expansion of the Weierstrass $\wp$ function is given by the series

$$\wp(z) = \frac{1}{z^2} + \sum_{n=2}^{\infty} (n+1) G_{n+1}(\Lambda) z^n$$

in the neighborhood $0 < |z| < \{\inf |\omega| \mid \omega \in \Lambda\}$.

Proof. This is a direct calculation. We write

$$\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} = \frac{1}{\omega^2}\left(\sum_{n=1}^{\infty}(n+1)(z/\omega)^n\right)$$

This allows us to express the Weierstrass $\wp$ function as:

$$\wp(z) \;\; = \;\; \frac{1}{z^2} + \sum_{\substack{\omega \in \Lambda \\ \omega \neq 0}} \frac{1}{\omega^2}\left(\sum_{n=1}^{\infty}(n+1)(z/\omega)^n\right)$$

The series is absolutely convergent and so can change the sums to get the desired expression. □

Proposition A.2.4. The functions $\wp$ and $\wp'$ satisfy the differential equation

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

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

Proof. From Proposition A.2.3, we conclude that

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

where $\varphi(z)$ defines an analytic function. Write $z^2\varphi(z) = f(z)$. Now the left hand side is Λ – periodic and therefore, so is $f(z)$. This implies that $f(z)$ is bounded and so constant. Observing the fact that $f(0) = 0$, the result now follows. □

The differential equation in Proposition A.2.4 gives us first hints of the Weierstrass $\wp$ function being associated to a cubic equation.

Definition A.2.5 (Elliptic curves). Equations of the form $y^2 = 4x^3 - a_2x - a_3$; $a_2^3 - 27a_3^2 \neq 0$ are called elliptic curves.

The Weierstrass $\wp$ function satisfies the elliptic curve with coefficients as the Eisenstein series.

Proposition A.2.6. Let $\Lambda = \omega_1\mathbb{Z} + \omega_2\mathbb{Z}$ and $\omega_3 = \omega_1 + \omega_2$. Then the cubic equation satisfied by $\wp$ and $\wp'$ is factorized as:

$$y^2 = 4(x - e_1)(x - e_2)(x - e_3) \ ; \ e_i = \wp(\omega_i/2) \text{ for } i = 1, 2, 3.$$

Moreover, the right hand side has distinct roots.

Proof. (Sketch) Since $\wp'$ is an odd function and $\omega_i/2 + \Lambda = -\omega_i/2 + \Lambda$, we see that $\wp'$ takes zero at these points. For a doubly periodic function, the number of zeroes equals the number of poles counting multiplicity and so these are precisely the zeroes of $\wp'$. From the relation satisfied by $\wp$ and $\wp'$ in Proposition A.2.4, we see that the equation factors as mentioned above. Moreover the three roots are distinct since $\wp$ has degree 2 and takes each value twice counting multiplicity. □

Corollary A.2.7. The Δ function given by the expression

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

is non vanishing at all values of $\mathbb{H}$.

Proof. Change the basis of the lattice Λ in Proposition A.2.6 and write it as Λ_τ for any $\tau \in \mathbb{H}$. Observe that the Δ function is in fact the discriminant of the polynomial $p_\tau(x) = 4x^3 - g_2(\tau)x - g_3(\tau)$. By Proposition A.2.6, this has distinct roots and so the discriminant is non zero. This is true for any $\tau \in \mathbb{H}$. □

Remark A.2.8. Observe that the delta function being non zero for all $\tau \in \mathbb{H}$ for the particular case of the elliptic curve with Eisenstein series as its coefficients, is exactly the condition on the coefficients in definition A.2.5.

A.3 Elliptic curves as complex tori

Let Λ be a lattice and $\wp_\Lambda$ be the corresponding Weierstrass function. Consider the map $z \mapsto (\wp_\Lambda(z), \wp'_\Lambda(z))$. Via this map, Proposition A.2.4 helps us conclude that every point $z \in \mathbb{C}/\Lambda$ determines a point on the graph of the curve $p(x) = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$ representing an elliptic curve E. The function $\wp(z)$ takes any value twice (counting multiplicity) and so the equation $\wp(z) = x$ has two solutions, (say) $\pm z_0$. Since $\wp'$ is odd, the point $(\wp_\Lambda(z_0), \wp'_\Lambda(z_0)) \neq (\wp_\Lambda(-z_0), \wp'_\Lambda(-z_0))$ in E whenever z_0 is not a zero of $\wp'_\Lambda$. To take care of the values for which $\wp'_\Lambda(z)$ is zero, observe that $e_i = \wp(w_i/2 + \Lambda)$ for $i = 1, 2, 3$ is a double value for $\wp_\Lambda$, that is, it is of multiplicity 2. We can further extend the map to all of $\mathbb{C}$ by mapping the lattice points to a suitable point "infinity" of E. In sum, we arrive the following result:

Proposition A.3.1. For any lattice Λ, the Weierstrass $\wp$ function and its derivative give a bijection

$$(\wp(z), \wp'(z)) : \mathbb{C}/\Lambda \to E$$

where E is the elliptic curve $y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda)$.

Not only does every complex number $\mathbb{C}/\Lambda$ via the Weierstrass $\wp$ function lead to an elliptic curve, but even the converse holds.

Proposition A.3.2. Suppose E is an elliptic given in definition A.2.5, there exists a lattice Λ such that $a_2 = g_2(\Lambda)$ and $a_3 = g_3(\Lambda)$.

We give a brief sketch of the proof. We use the following fact about the j function defined in section 1.5: the j given by $j(\tau) = 1728\frac{(g_2(\tau))^3}{\Delta(\tau)}$ surjects from $\mathbb{H}$ onto $\mathbb{C}$.

Proof. (Sketch) We handle the case when a_2 and a_3 are not equal to zero first. There exists some $\tau \in \mathbb{H}$ such that $j(\tau) = 1728a_2^3/(a_2^3 - 27a_3^2)$. This implies that

$$\frac{(g_2(\tau))^3}{g_2^3(\tau) - 27g_3^2(\tau)} = \frac{a_2^3}{a_2^3 - 27a_3^2}$$

After taking reciprocals and cancelling terms we see that,

$$\frac{a_2^3}{g_2(\tau)^3} = \frac{a_3^2}{g_3(\tau)^2} \tag{A.3.3}$$

Let $\Lambda = \omega_1\mathbb{Z}\oplus\omega_2\mathbb{Z}$ such that $\omega_1/\omega_2 = \tau$. Then $g_2(\Lambda) = \omega_2^{-4}g_2(\tau)$ and $g_3(\Lambda) = \omega_2^{-6}g_3(\tau)$. We want $a_2 = g_2(\Lambda)$ and $a_3 = g_3(\Lambda)$. Therefore, if we choose ω_2 such that $\omega_2^{-4} = a_2/g_2(\tau)$, then the condition for g_2 is straightaway satisfied. Moreover the equality in A.3.3 helps us conclude the condition for g_3 as well. The case when a_2 or a_3 is zero is dealt with separately. For that we specifically consider the lattices Λ_{μ_3} where $\mu_3 = e^{2\pi i/3}$ and Λ_i respectively. We use the fact that $g_2(\mu_3) = 0$, while in the other case $g_3(i) = 0$. Once we have this, the proof is quite similar to the above case. $\square$

This gives us a one to one correspondence between the complex tori and elliptic curves. Furthermore, one in fact has a bijection up to isomorphism as well. Thus complex tori and elliptic curves are interchangeable objects.

Bibliography

[1] Daniel Bump. Automorphic Forms and Representations. Cambridge University Press, 1998.

[2] Fred Diamond and Jerry Shurman. A First Course in Modular Forms. Graduate Texts in Mathematics 228, 2005.

[3] Gareth A. Jones and David Singerman. Complex Functions, an algebraic and geometric viewpoint. Cambridge University Press 1987.

[4] Serge Lang. Complex Analysis. Graduate Texts in Mathematics 103, 1993.

[5] Elias M. Stein and Rami Sakarchi. Complex Analysis. Princeton Lectures in Analysis, 2006.