Part III - Modular Forms

Lectured by Jack Thorne

Artur Avameri

Contents

1 Introduction

06 Oct 2022, Lecture 1

Definition 1.1. We define the following groups:

$$\begin{aligned}
&\mathfrak{h} = \{\tau \in \mathbb{C} \mid \mathrm{Im}(\tau) > 0\} \\
&GL_2(\mathbb{R})^+ = \{g \in GL_2(\mathbb{R}) \mid \det(g) > 0\} \\
&\Gamma(1) = SL_2(\mathbb{Z}) = \{g \in M_2(\mathbb{Z}) \mid \det(g) = 1\}.
\end{aligned}$$

Note that $\Gamma(1)$ is a subgroup of $GL_2(\mathbb{R})^+$.

Lemma 1.1. $GL_2(\mathbb{R})^+$ acts transitively on $\mathfrak{h}$ by Möbius transformations.

Proof. Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{R})^+, \tau \in \mathfrak{h}$. Then

$$\mathrm{Im}(g\tau) = \frac{1}{2i}\left(\frac{a\tau + b}{c\tau + d} - \frac{a\overline{\tau} + b}{c\overline{\tau} + d}\right) = \frac{1}{2i}\frac{(ad - bc)(\tau - \overline{\tau})}{|c\tau + d|^2} = \frac{\det(g)\mathrm{Im}(\tau)}{|c\tau + d|^2} > 0,$$

so $g\tau \in \mathfrak{h}$. This action is transitive since

$$x + iy \in \mathfrak{h} \implies \begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} i = x + iy,$$

so everything in $\mathfrak{h}$ is conjugate to i. □

Definition 1.2. If $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{R})^+$ and $\tau \in \mathfrak{h}$, then define

$$j(g, \tau) = c\tau + d.$$

This is called a **modular cocycle**. If $k \in \mathbb{Z}$ and $f : \mathfrak{h} \to \mathbb{C}$, then

$$f|_k[g] : \mathfrak{h} \to \mathbb{C}$$

is defined by

$$f|_k[g](\tau) = \det(g)^{k-1} f(g\tau) j(g, \tau)^{-k}.$$

This is the **weight k action of g on f**.

Lemma 1.2. This is a right action of $GL_2(\mathbb{R})^+$: if $g, h \in GL_2(\mathbb{R})^+$, then

$$f|_k[gh] = (f|_k[g])|_k[h].$$

Proof. We compute

$$\begin{aligned}(f|_k[g])|_k[h](\tau) = \det(h)^{k-1} f|_k[g](h\tau) j(h,\tau)^{-k} = \\ \det(h)^{k-1}\det(g)^{k-1} f(gh\tau) j(g,h\tau)^{-k} j(h,\tau)^{-k} \stackrel{?}{=} \\ \det(gh)^{k-1} f(gh\tau) j(gh,\tau)^{-k} = f|_k[gh](\tau).\end{aligned}$$

Hence we need to check that $j(gh,\tau) = j(gh,\tau)j(h,\tau)$. Note that if $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then

$$g\begin{pmatrix}\tau \\ 1\end{pmatrix} = \begin{pmatrix} a\tau + b \\ c\tau + d\end{pmatrix} = j(g,\tau)\begin{pmatrix} g\tau \\ 1\end{pmatrix}.$$

We now get

$$j(gh,\tau)\begin{pmatrix} gh\tau \\ 1\end{pmatrix} = gh\begin{pmatrix}\tau \\ 1\end{pmatrix} = g\left(j(h,\tau)\begin{pmatrix} h\tau \\ 1\end{pmatrix}\right) = j(h,\tau)j(g,h\tau)\begin{pmatrix} gh\tau \\ 1\end{pmatrix},$$

which finishes the computation and proof. □

Formulae. For $g \in GL_2(\mathbb{R})^+, \tau \in \mathfrak{h}$, we have

$$\operatorname{Im}(g\tau) = \det(g)\frac{\operatorname{Im}(\tau)}{|j(g,\tau)|^2} \text{ and } j(g,\tau)\begin{pmatrix} g\tau \\ 1\end{pmatrix} = g\begin{pmatrix}\tau \\ 1\end{pmatrix}.$$

Definition 1.3. Let $k \in \mathbb{Z}$ and $\gamma \leq \Gamma(1)$ of finite index[1]. A **weakly modular function of weight** k **and level** Γ is a meromorphic function $f : \mathfrak{h} \to \mathbb{C}$ which is invariant under the weight k action of Γ, i.e. such that

$$\forall \tau \in \mathfrak{h}, \forall \gamma \in \Gamma, f|_k(\gamma) = f.$$

We will define modular forms next time: they are weakly modular functions which are holomorphic both in $\mathfrak{h}$ and at ∞.

It is a fact that modular forms of fixed weight and level live in finite-dimensional $\mathbb{C}$-vector spaces called $M_k(\Gamma)$. These form the main objects of study in this course.

Motivation. Why study modular forms?

(1) They are related to the theory of elliptic functions. Let $E/\mathbb{C}$ be an elliptic curve and ω a holomorphic non–zero 1–form. Then there exists a unique lattice[2] $\Lambda \in \mathbb{C}$ and isomorphism $\phi : \mathbb{C}/\Lambda \to E$ such that $\phi^*(\omega) = dz$. Then

[1] In other words, γ is a (finite index) subgroup of $\Gamma(1)$.

[2] i.e. a discrete cocompact subgroup, or an abelian subgroup which is freely generated by two elements that are linearly independent over $\mathbb{R}$.

E is isomorphic to the elliptic curve $y^2 = 4x^3 - 60G_4(\Lambda)x - 140G_6(\Lambda)$ where if $k \in \mathbb{Z}$, then $G_k(\Lambda) = \sum_{\lambda \in \Lambda - \{0\}} \lambda^{-k}$. This converges absolutely for $k > 2$.

If $\tau \in \mathfrak{h}$, then $\Lambda\tau = \mathbb{Z}\tau \oplus \mathbb{Z} \subset \mathbb{C}$ is a lattice and $G_k(\tau) = G_k(\Lambda_\tau)$. This is a modular form of weight k and level $\Gamma(1)$, called an Eisenstein series.

$\mathfrak{h}/SL_2(\mathbb{Z})$ can be identified with the set of (isomorphism classes of) elliptic curves over $\mathbb{C}$.

(2) Modular forms f have Fourier expansions $\sum_{n \in \mathbb{Z}} a_n g^n$, $a_n \in \mathbb{C}$ and they often serve as a generating functions for arithmetically interesting sequences a_n.

For example, take $\theta(\tau) = \sum_{n \in \mathbb{Z}} e^{\pi i n^2 \tau}$. If $k \in 2\mathbb{N}$, then θ^k is a modular form with q–expansion $\theta^k = \sum_{n \in \mathbb{Z}} r_k(n) e^{\pi i n \tau}$, where $r_k(n)$ is the number of ways of writing n as a sum of k squares, i.e. $r_k(n) = |\{x \in \mathbb{Z}^k \mid \sum_{i=1}^k x_i^2 = n\}|$. By expressing θ^k in terms of other modular forms, we can prove formulae such as $r_4(n) = 8 \sum_{d|n, 4\nmid d} d$.

(3) The Riemann zeta function $\zeta(s)$ is an important object of study. Its pleasant features include:

- The Euler product $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$.
- It has a meromorphic continuation to $\mathbb{C}$ and has a functional equation relating $\zeta(s)$ and $\zeta(1-s)$.

A Dirichlet series $\sum_{n \geq 1} a_n n^{-s}$ which has similar properties (Euler product, meromorphic extension, some nice function equation) is called an L–function. Modular forms can be used to construct interesting examples of L–functions. In practice, we take $M_k(\Gamma)$ and decompose it under Hecke operators to get Hecke eigenforms, the nicest possible modular forms, which have the above properties.

(4) The Langlands program predicts a relation between modular forms and objects in arithmetic geometry. A special case of this is the modularity conjecture, which says that there is a bijective correspondence between elliptic curves $E/\mathbb{C}$ up to isogeny and the set of Hecke eigenforms of weight 2. This implies Fermat's last theorem. Note that this is formulated in the language of Hecke operators and L–functions.

Homework. There is a handout on Moodle called "Reminder on Complex Analysis". Have a look at it before the next lecture.

2 Modular Forms on $\Gamma(1)$

09 Oct 2022, Lecture 2

Reminder. A **meromorphic** function in an open subset $U \subset \mathbb{C}$ is a closed subset $A \subset U$ and a holomorphic function $f : U \setminus A \to \mathbb{C}$ such that $\forall a \in A$, $\exists \delta > 0$ such that $D^*(a, \delta) \subset U \setminus A$ and $\exists n \geq 0$ such that $(z-a)^n f(z)$ extends to a holomorphic function in $D(a, \delta)$.

f then has a Laurent expansion $\sum_{m \in \mathbb{Z}} a_m (z-a)^m$ valid on $D^*(a, \delta)$.

Lemma 2.1. Let f be a weakly modular function of weight k and level $\Gamma(1)$. Then there exists a meromorphic function $\tilde{f}$ in $D^*(0,1)$ (the "q-disk") such that

$$f(\tau) = \tilde{f}(e^{2\pi i \tau}).$$

Proof. f is meromorphic in $\mathfrak{h}$ by assumption. Take $\gamma = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma(1)$. Then $f|_h[\gamma](\tau) = f(\gamma\tau) = f(\tau)$, as f is invariant under the weight k action of γ. But also $f(\gamma\tau) = f(\tau + 1)$, so f is periodic.

Now map a strip of $\mathfrak{h}$ of width 1 to $D^*(0,1)$ by $\tau \mapsto e^{2\pi i \tau}$. Let $a \in D^*(0,1)$ and $\delta > 0$ be such that $D(a, \delta) \subset D^*(0,1)$. Define $\tilde{f}$ on $D(a, \delta)$ by

$$\tilde{f}(q) = f\left(\frac{1}{2\pi i} \log q\right),$$

for any branch of log defined in $D(a, \delta)$. This is meromorphic and independent of the choice of the branch of log, as f is periodic with period 1. This defines $\tilde{f}$ in $D^*(0,1)$. Finally, $\tilde{f}$ is unique since $\tau \mapsto e^{2\pi i \tau}$ is surjective. $\square$

If $\tilde{f}$ extends to a meromorphic function[3] in $D(0,1)$, then $\exists \delta > 0$ such that $\tilde{f}$ has a Laurent expansion $\tilde{f}(q) = \sum_{n \in \mathbb{Z}} a_n q^n$ valid in $D^*(0, \delta)$.

In the region $\{\tau \in \mathfrak{h} \mid \mathrm{Im}(\tau) > \frac{1}{2\pi} \log \delta\}$, we have

$$f(\tau) = \sum_{n \in \mathbb{Z}} a_n q^n,$$

where $q = e^{2\pi i \tau}$. This is called the q–**expansion** of the weakly modular function f.

Definition 2.1. Let f be a weakly modular function of weight k and level $\Gamma(1)$. We say that f is **meromorphic at** ∞ if $\tilde{f}$ extends to a meromorphic function in $D(0,1)$.

We say f is **holomorphic at** ∞ if $\tilde{f}$ is meromorphic at ∞ and has a

[3] This might not be the case if the set of poles has a limit inside the disk.

removable singularity at $q = 0$. In this case, we define

$$f(\infty) = \tilde{f}(0) = \lim_{\mathrm{Im}(\tau)\to\infty} f(\tau).$$

We say f **vanishes at** ∞ if f is holomorphic at ∞ and $f(\infty) = 0$.

Definition 2.2. A **modular function** (of weight k and level $\Gamma(1)$) is a weakly modular function (of weight k and level $\Gamma(1)$) which is meromorphic at ∞.

A **modular form** is a weakly modular function which is holomorphic in $\mathfrak{h}$ and holomorphic at ∞.

A **cuspidal modular form** is a modular form that vanishes at ∞.

Remark. We let $M_k(\Gamma(1))$ denote the set of modular forms of weight k and level $\Gamma(1)$. We write $S_k(\Gamma(1))$ for the set of cuspidal modular forms of weight k, level $\Gamma(1)$. Note $S_k(\Gamma(1)) \subset M_k(\Gamma(1))$. These are $\mathbb{C}$–vector spaces. If k is odd, then these both only contain the zero function, since taking $\gamma = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \in \Gamma(1)$ gives $f|_k[\gamma](\tau) = f(\tau)(-1)^k = f(\tau)$.

We now consider even weights only. If $k \in \mathbb{Z}$ is even, let

$$G_k(\tau) = \sum_{\lambda \in \Lambda_\tau \setminus 0} \lambda^{-k} = \sum_{(m,n)\in\mathbb{Z}^2\setminus 0} (m\tau + n)^{-k},$$

where $\Lambda_\tau = \mathbb{Z}\tau \oplus \mathbb{Z}$ for any $\tau \in \mathfrak{h}$.

If $\gamma \in \Gamma(1)$, then formally we have

$$G_k|_k[\gamma](\tau) = G_k(\gamma\tau) j(\gamma,\tau)^{-k} = \sum_{\lambda\in\Lambda_{\gamma\tau}\setminus 0} \lambda^{-k} j(\gamma,\tau)^{-k},$$

but $\Lambda_{\gamma\tau} = \mathbb{Z}\frac{a\tau+b}{c\tau+d} \oplus \mathbb{Z} = (c\tau+d)^{-1}\left(\mathbb{Z}(a\tau+b)\oplus\mathbb{Z}(c\tau+d)\right) = (c\tau+d)^{-1}\Lambda_\tau$. Hence

$$\begin{aligned} G_k|_k[g](\tau) &= \sum_{\lambda\in(c\tau+d)^{-1}\Lambda_\tau\setminus 0} \lambda^{-k}(c\tau+d)^{-k} \\ &= \sum_{\lambda\in\Lambda_\tau\setminus 0} ((c\tau+d)^{-1}\lambda)^{-k}(c\tau+d)^{-k} = G_k(\tau). \end{aligned}$$

This is justified only when the series defining $G_k(\tau)$ converges absolutely. Hence:

Proposition 2.2. Let $k > 2$ be an even integer. Then $G_k(\tau)$ converges absolutely and defines a modular form of weight k and level $\Gamma(1)$ which has

$G_k(\infty) = 2\zeta(k)$. G_k is the **weight k Eisenstein series**.

We will later see that $M_2(\Gamma(1)) = 0$.

Proof. We want to show absolute and locally uniform convergence in $\mathfrak{h}$. This will show that G_k is holomorphic by complex analysis. Let $A \geq 2$ and define $\Omega_A = \{\tau \in \mathfrak{h} \mid \operatorname{Im}(\tau) \geq \frac{1}{A}, \operatorname{Re}(\tau) \in [-A, A]\}$. We show uniform convergence in Ω_A. If $\tau \in \Omega_A, x \in \mathbb{R}$, then $|\tau + x| \geq \begin{cases} \frac{1}{A} & |x| \leq 2A \\ \frac{|x|}{2} & |x| \geq 2A. \end{cases}$ Hence

$$|\tau + x| \overset{(\dagger)}{\geq} \sup\left(\frac{1}{A}, \frac{|x|}{2A^2}\right) \geq \sup\left(\frac{1}{2A^2}, \frac{|x|}{2A^2}\right) = \frac{1}{2A^2}\sup(1, |x|).$$

$(\dagger)$ follows by drawing a diagram with the lines $y = \frac{1}{A}$ and $y = \frac{x}{2A^2}$ and marking the point $(2A, \frac{1}{A})$ on it, then noticing that out supremum always lies above the supremum of these two lines. If $(m, n) \in \mathbb{Z}^2, m \neq 0$, then

$$|m\tau + n| = |m|\left|\tau + \frac{n}{m}\right| \geq |m|\frac{1}{2A^2}\sup\left(1, \left|\frac{n}{m}\right|\right) = \frac{1}{2A^2}\sup\left(|m|, |n|\right).$$

This is also valid when $m = 0$ by inspection. If $\tau \in \Omega_A$, then

$$\begin{aligned}
&\sum_{(m,n)\in\mathbb{Z}^2\setminus 0} |m\tau + n|^{-k} \\
&\leq \left(\frac{1}{2A^2}\right)^{-k} \sum_{(m,n)\in\mathbb{Z}^2\setminus 0} \sup\left(|m|, |n|\right)^{-k} \\
&= (2A^2)^k \sum_{d\in\mathbb{N}} d^{-k} \cdot \left|\{(m,n) \in \mathbb{Z}^2 \mid \sup\left(|m|, |n|\right) = d\}\right| \\
&= (2A^2)^k \sum_{d\in\mathbb{N}} d^{-k} 8d = 8(2A^2)^k \sum_{d\in\mathbb{N}} d^{1-k} \\
&< \infty
\end{aligned}$$

whenever $k - 1 > 1$, i.e. $k > 2$. This shows absolute convergence, and uniform convergence in Ω_A by the Weierstrass M-test[4]. Hence G_k is holomorphic in $\mathfrak{h}$ and invariant under the weight k action of $\Gamma(1)$. It remains to show that G_k is holomorphic at ∞ with $G_k(\infty) = 2\zeta(k)$. For this, it suffices to check that

$$\lim_{\operatorname{Im}(\tau)\to\infty} G_k(\tau) = 2\zeta(k).$$

[4] If we have a sequence of functions $f_n : \Omega \to \mathbb{C}$ and values $M_n > 0$ with $|f_n(x)| < M_n$ and $\sum M_n < \infty$, then $\sum f_n$ converges absolutely and uniformly on Ω. Here, replace n with d and sum d over $\sum_{(m,n)\in\mathbb{Z}^2\setminus 0, \sup(|m|,|n|)=d} |m\tau + n|^{-k}$.

This follows from uniform convergence in Ω_A: we get

$$\lim_{\mathrm{Im}(\tau)\to\infty} G_k(\tau) = \sum_{(m,n)\in\mathbb{Z}^2\setminus 0} \lim_{\mathrm{Im}(\tau)\to\infty} (m\tau+n)^{-k} = \sum_{n\in\mathbb{Z}\setminus 0} n^{-k} = 2\sum_{n\geq 1} n^{-k} = 2\zeta(k).$$

□

11 Oct 2022, Lecture 3

Recap. We defined what it means for a function $f : \mathfrak{h} \to \mathbb{C}$ to be a modular form of weight k and level $\Gamma(1)$. $M_k(\Gamma(1))$ is the $\mathbb{C}$–vector space of such forms. If $f \in M_k(\Gamma(1))$, then there exists a holomorphic $\tilde{f} : D(0,1) \to \mathbb{C}$ (here we call $D(0,1)$ the q–disk) such that $\forall \tau \in \mathfrak{h}$, $f(\tau) = \tilde{f}(e^{2\pi i\tau})$. The Taylor expansion of $\tilde{f}$ gives the q–expansion

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

We have $f(\infty) = \tilde{f}(0) = a_0$. If $k > 2$ is even, then $G_k(\tau) = \sum_{\lambda\in\Lambda_\tau\setminus 0} \lambda^{-k}$ converges absolutely and defines an element of $M_k(\Gamma(1))$ with $G_k(\infty) = 2\zeta(k)$.

We define

$$E_k(\tau) = \frac{G_k(\tau)}{2\zeta(k)} = 1 + \sum_{n\geq 1} a_n q^n.$$

We will soon show that we have $a_n \in \mathbb{Q}\ \forall n \geq 1$.

We can construct more modular forms: if $f \in M_k(\Gamma(1))$ and $g \in M_l(\Gamma(1))$, then $fg \in M_{k+l}(\Gamma(1))$. To check this is a modular form, we need to check that:

- fg is holomorphic, which is true as f, g are holomorphic.
- fg is invariant under the weight $k+l$ action of $\Gamma(1)$, which is true as f, g are invariant under the weight k and l actions of $\Gamma(1)$ – this is just a computation.
- fg is holomorphic at ∞. This is true as the q–expansions multiply, so since f, g have no negative terms, the same is true for fg.

Hence we get e.g. $E_4^3, E_6^2 \in M_{12}(\Gamma(1))$ and $\frac{E_4^3 - E_6^2}{1728} \in S_{12}(\Gamma(1))$ (i.e. it is cuspidal since zero at infinity). This difference is Ramanujan's Δ–function. We will show it is nonzero later.

We now want to show that $M_k(\Gamma(1))$ is finite–dimensional. We first study $\Gamma(1)/\mathfrak{h}$. For this, introduce a fundamental set $\mathfrak{f}' \subset \mathfrak{h}$ for the $\Gamma(1)$–action. We define[5] a fundamental set to be a set that intersects each $\Gamma(1)$–orbit in exactly

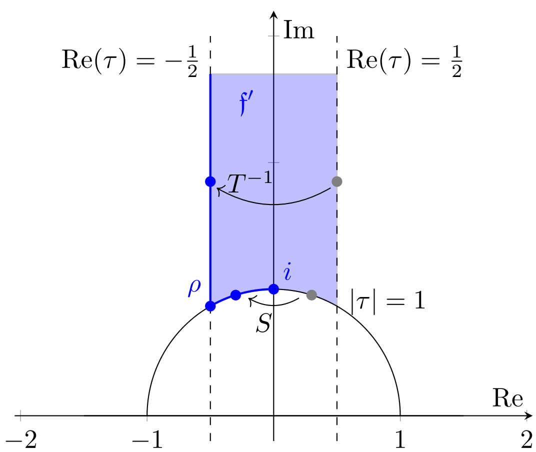
[5]Definitions in literature may vary, so we omit a formal definition.

one element. Define

$$\mathfrak{f} = \left\{ \tau \in \mathfrak{h} \mid \mathrm{Re}(\tau) \in \left[-\frac{1}{2}, \frac{1}{2}\right], |\tau| \geq 1 \right\}.$$

$$\mathfrak{f}' = \left\{ \tau \in \mathfrak{f} \mid \mathrm{Re}(\tau) \in \left[-\frac{1}{2}, \frac{1}{2}\right), |\tau| = 1 \implies \mathrm{Re}(\tau) \in \left[-\frac{1}{2}, 0\right] \right\}.$$

Introduce $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ and $S = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ in $\Gamma(1)$. We observe that every element of $\mathfrak{f}$ is conjugate under S or T^{-1} to an element of $\mathfrak{f}'$, which is true since $T(\tau) = \tau + 1$ and $S(\tau) = -\frac{1}{\tau}$.



Proposition 2.3. Let $G = \Gamma(1)/\{\pm I\}$. Then

(i) $\forall \tau \in \mathfrak{h}, \tau$ is $\Gamma(1)$–conjugate to an element of $\mathfrak{f}'$.

(ii) If $\tau, \tau' \in \mathfrak{f}'$ are $\Gamma(1)$–conjugate, then $\tau = \tau'$.

(iii) If $\tau \in \mathfrak{f}'$, then $\mathrm{Stab}_G(\tau)$ is trivial, except in the two cases $\mathrm{Stab}_G(i) = \langle S \rangle$ and $\mathrm{Stab}_G(\rho) = \langle ST \rangle$, where $\rho = e^{2\pi i/3}$.

(iv) $\Gamma(1)$ is generated by S and T.

Proof. Let H be the subgroup of G generated by S and T.

Claim. Every $\tau \in \mathfrak{h}$ is H–conjugate to an element of $\mathfrak{f}'$.

Proof. By our above observation and since $S, T \in H$, it suffices to prove that every $\tau \in \mathfrak{h}$ is H–conjugate to $\mathfrak{f}$. Take $\tau \in \mathfrak{h}$. Recall that if $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$, then $\mathrm{Im}(\gamma\tau) = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2}$.

In particular, $\forall R \geq 0$, the intersection $H\tau \cap \{\mathrm{Im}(\tau') > R\}$ is finite, since $\mathrm{Im}(\gamma\tau) > R \iff |c\tau + d|^2 < \frac{\mathrm{Im}(\tau)}{R}$, but $\Lambda_\tau = \mathbb{Z}\tau \oplus \mathbb{Z}$ is a lattice, so the set $\{(c,d) \in \mathbb{Z}^2 \mid |c\tau + d| < R'\}$ is finite.

So there exists $h \in H$ such that $\mathrm{Im}(h\tau) \geq \mathrm{Im}(h'\tau)\ \forall h' \in H$. After replacing τ by $h\tau$, we can assume $\mathrm{Im}(\tau) \geq \mathrm{Im}(h\tau)\ \forall h \in H$. Since acting by T does not change $\mathrm{Im}(\tau)$, we can also assume $\mathrm{Re}(\tau) \in \left[-\frac{1}{2}, \frac{1}{2}\right]$. We have $\mathrm{Im}(\tau) \geq \mathrm{Im}(S\tau) = \frac{\mathrm{Im}(\tau)}{|\tau|^2} \implies |\tau| \geq 1$, proving the claim and (i). $\square$

Now take $\tau, \tau' \in \mathfrak{f}'$ and suppose $\gamma\tau = \tau'$ for some $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1)$. We want to show that either $\gamma = \pm I$ or $\tau = i, \rho$.

WLOG assume $\mathrm{Im}(\tau') = \mathrm{Im}(\gamma\tau) \geq \mathrm{Im}(\tau)$, i.e. $\mathrm{Im}(\gamma\tau) = \frac{\mathrm{Im}(\tau)}{|c\tau+d|^2} \geq \mathrm{Im}(\tau)$, so $|c\tau + d| \leq 1$. However, if $\tau \in \mathfrak{f}'$, then $\mathrm{Im}(\tau) \geq \frac{\sqrt{3}}{2}$ with equality if and only if $\tau = \rho$. Hence $|c\tau + d| \geq |c|\mathrm{Im}(\tau) \geq |c|\frac{\sqrt{3}}{2} \implies |c| \leq \frac{2}{\sqrt{3}} \implies |c| = 0, 1 \implies c = 0$ or $c = \pm 1$.

- If $c = 0$, then $\gamma = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$, so $ad = 1 \implies a = d = \pm 1$, so $\gamma = \pm T^m$ for $m \in \mathbb{Z}$. However, T acts on $\mathfrak{f}'$ by shifting the real part, so it can only stay in $\mathfrak{f}'$ if $m = 0$ (as $\mathrm{Re}(\mathfrak{f}') \in \left[-\frac{1}{2}, \frac{1}{2}\right)$), so $\gamma = \pm I$ and $\tau' = \tau$.

- If $c = 1$, then $\gamma = \begin{pmatrix} a & b \\ 1 & d \end{pmatrix}$ and $|\tau + d| \leq 1$. By drawing another picture, we see that the only circles centered at integers of radius 1 which intersect $\mathfrak{f}'$ are centered at $-d = 0, -d = -1$. Hence either $d = 0$, whence $|\tau| = 1$, or $d = 1$, whence $\tau = \rho$.

  - If $c = 1, d = 0, |\tau| = 1$, then $\gamma = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & -1 \\ 1 & 0 \end{pmatrix}$ since the determinant must be 1. Then $\gamma\tau = \frac{a\tau - 1}{\tau} = a - \frac{1}{\tau} = a - \overline{\tau}$, so $\mathrm{Re}(\gamma\tau) = a - \mathrm{Re}(\tau) \in \mathrm{Re}(\mathfrak{f}' \cap \{|\tau| = 1\}) = \left[-\frac{1}{2}, 0\right]$. However, we also have $\mathrm{Re}(\gamma\tau) \in a - \left[-\frac{1}{2}, 0\right] = a + \left[0, \frac{1}{2}\right]$.

    The intersection $\left[-\frac{1}{2}, 0\right] \cap \left(a + \left[0, \frac{1}{2}\right]\right)$ can be nonempty only if either $a = 0$, whence $\mathrm{Re}(\gamma\tau) = \mathrm{Re}(\tau) = 0$, so $\tau = \gamma\tau = i$, or $a = -1$, whence $\mathrm{Re}(\tau) = \mathrm{Re}(\gamma\tau) = -\frac{1}{2}$, so $\tau = \gamma\tau = \rho$.

    If $a = 0$, then $\gamma = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -S$, which stabilizes i, and $\langle -S \rangle = \langle S \rangle$.

    If $a = -1$, then $\gamma = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} = (ST)^2$, which stabilizes ρ, and $(ST)^3 = I$, so $\langle (ST)^2 \rangle = \langle ST \rangle$.

  - If $c = 1, d = 1, \tau = \rho$, then $\gamma = \begin{pmatrix} a & b \\ 1 & 1 \end{pmatrix}$, so $\rho = \gamma\rho = \frac{a\rho+b}{\rho+1}$. We have $\rho^2 + \rho + 1 = 0$, so $\rho^2 + \rho = -1$, so $a\rho + b = \rho^2 + \rho = -1$. But $a, b \in \mathbb{Z}$ and $1, \rho$ are linearly independent over $\mathbb{R}$, so $a = 0, b = -1$, so $\gamma = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} = -ST$, which stabilizes ρ.

- If $c = -1$, we can reduce this to the case $c = 1$ by replacing γ with $-\gamma$.

We have now shown the first three parts of the proposition. It remains to show the last part, i.e. $\Gamma(1) = \langle S, T \rangle$. Since $S^2 = -I$, it is enough to show that $H = G$. Choose $\tau \in \mathrm{Int}(f)$, so $\mathrm{Stab}_G(\tau) = \{I\}$. Let $g \in G$. By our claim proving (i), $\exists h \in H$ such that $hg\tau \in \mathfrak{f}'$. We must therefore have $hg\tau = \tau$, hence $hg \in \mathrm{Stab}_G(\tau) = \{I\}$, so $g = h^{-1} \in H$. □

13 Oct 2022, Lecture 4

Notation. We write $e_\tau = |\mathrm{Stab}_G(\tau)|$.

Let f be a nonzero modular function of weight k, level $\Gamma(1)$. If $\tau \in \mathfrak{h}$, then $v_\tau(f)$ is the order of f at τ (the unique $n \in \mathbb{Z}$ such that $f(z) = (z-\tau)^n g(z)$ for some meromorphic g that is holomorphic and non–vanishing at τ). We define $v_\infty(f)$ to be the order of f at infinity, i.e. $v_\infty(f) = v_0(\tilde{f})$ for $\tilde{f}$ the meromorphic function in $D(0,1)$ with $f(\tau) = \tilde{f}(e^{2\pi i\tau})$.

Proposition 2.4. Let f be a nonzero modular function of weight k, level $\Gamma(1)$. Then

$$\sum_{\tau \in \Gamma(1)\backslash\mathfrak{h}} \frac{1}{e_\tau} v_\tau(f) + v_\infty(f) = \frac{k}{12}.$$

Proof. We first check that the sum is well–defined:

- If $\tau \in \mathfrak{h}$, then $e_\tau, v_\tau(f)$ only depend on the $\Gamma(1)$–orbit of τ. This is because if $\gamma \in \Gamma(1)$ and $\tau \in \mathfrak{h}$, then $\mathrm{Stab}_{\Gamma(1)}(\tau)$ and $\mathrm{Stab}_{\Gamma(1)}(\gamma\tau)$ are conjugate subgroups of $\Gamma(1) \implies e_\tau = e_{\gamma\tau}$. On the other hand, $f(\gamma\tau) = f(\tau)j(\gamma,\tau)^k$ and $j(\gamma,\tau)$ is holomorphic and non–vanishing on $\mathfrak{h}$, so $v_{\gamma\tau}(f) = v_{\tau(f)}$.

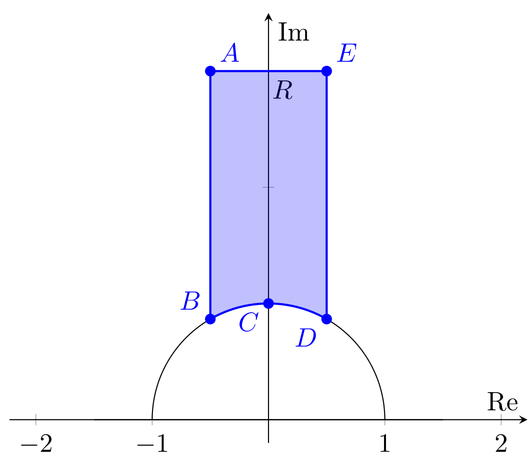
- The sum only has a finite number of nonzero terms, since if f is a modular function and $\tilde{f}$ is a meromorphic function on $D(0,1)$, then $\exists\delta > 0$ such that $\tilde{f}$ is holomorphic and non–vanishing in $D^*(0,\delta)$. Thus $\exists R > 0$ such that f is holomorphic and non–vanishing in $\{\tau \in \mathfrak{h} \mid \mathrm{Im}(\tau) > R\}$. Hence to show the sum is finite, it suffices to show that f only has a finite number of zeroes and poles in $\mathfrak{f}$ (as f intersects every $\Gamma(1)$–orbit), for which it suffices to show that f has a finite number of zeroes and poles in $\mathfrak{f} \cap \{\tau \in \mathfrak{h} \mid \mathrm{Im}(\tau) \leq R\}$, which is true as the set is compact (closed and bounded) and the zeroes and poles of f are discrete.

To prove the identity, we use contour integration. Setup: if $U \subset \mathbb{C}$ is an open subset, $f : U \to \mathbb{C}$ is holomorphic and $\gamma : [0,1] \to U$ is a path, then $\int_\gamma f(z)\mathrm{d}z = \int_{t=0}^1 f(\gamma(t))\gamma'(t)\mathrm{d}t$. We have the pullback formula: if $u : U \to V$ is a holomorphic map between open subsets of $\mathbb{C}$, $g : V \to \mathbb{C}$ is holomorphic and γ is a path in U, then $\int_{u\circ\gamma} g(z)\mathrm{d}z = \int_\gamma u^*(g(z)\mathrm{d}z) = \int_\gamma g(u(z))u'(z)\mathrm{d}z$. A particularly nice case: if $g(z) = h'(z)/h(z)$, then $g(z)\mathrm{d}z = d\log h$, so $\int_{u\circ\gamma} d\log h = \int_\gamma u^*(d\log h) = \int_\gamma d(\log h \circ u) = \int_\gamma \frac{(h\circ u)'(z)}{(h\circ u)(z)}\mathrm{d}z$.

We also have (Cauchy') argument principle: if $U \subset \mathbb{C}$ is a simply connected open subset, $\gamma \subset U$ is a simple positvely oriented closed path and g a meromorphic function in U with no zeroes or poles on γ, then

$$\frac{1}{2\pi i}\oint_\gamma d\log g = \frac{1}{2\pi i}\oint_\gamma \frac{g'(z)}{g(z)}\mathrm{d}z = \sum_{a\in\mathrm{Int}(\gamma)} v_a(g).$$

We now apply this to our modular function f. Choose $R > 0$ such that f has no zeroes or poles in $\{\tau \in \mathfrak{h} \mid \mathrm{Im}(\tau) \geq R\}$. We consider $\frac{1}{2\pi i}\oint_\gamma d\log f$, where γ is the contour $ABCDE$.



By choice of R, tere ar eno zeroes or poles of f on AE. We first consider the case where f has no zeroes or poles at all on γ. Then the argument principle gives

$$\frac{1}{2\pi i}\oint_\gamma d\log f = \frac{1}{2\pi i}\int_{AB} + \int_{BC} + \int_{CD} + \int_{DE} + \int_{EA} d\log f = \sum_{\tau\in\Gamma(1)\backslash\mathfrak{h}} \frac{1}{e_\tau} v_\tau(f)$$

(as $v_\tau(f) \neq 0$, $e_\tau = 1$ under our assumptions).

Apply the pullback formula with $u(\tau) = \tau + 1$. Then $u(AB) = ED$, $f \circ u = f$, so

$$\int_{u(AB)} d\log f = \int_{AB} d\log f \circ u = \int_{AB} d\log f = \int_{ED} d\log f = -\int_{DE} d\log f.$$

Hence $\int_{AB} + \int_{DE} d\log f = 0$.

Now take $q = e^{2\pi i\tau}$, so $f = \tilde{f} \circ q$ and $q(AE)$ is a positively oriented circle around 0 in $D(0,1)$. So

$$\frac{1}{2\pi i}\int_{q(AE)} d\log \tilde{f} = v_\infty(f) = \frac{1}{2\pi i}\int_{AE} d\log \tilde{f} \circ q = \frac{1}{2\pi i}\int_{AE} d\log f.$$

Now take $v(\tau) = S(\tau) = -\frac{1}{\tau}$. Then $v(BC) = DC$ and we know $f|_k[S](\tau) = f\left(-\frac{1}{\tau}\right)\tau^{-k} = f(\tau)$, so $f \circ v = f(\tau)\tau^k$. Hence

$$\begin{aligned}
&\int_{DC} d\log f = \int_{v(BC)} d\log f = \int_{BC} d\log(f \circ v) = \int_{BC} d\log(f(\tau)\tau^k) \\
&= \int_{BC} d\log f + kd\log\tau = \int_{BC} d\log f + k(\log C - \log B)
\end{aligned}$$

where here log is any branch of the logarithm defined on BC. But $B = \rho, C = i$, so $\log B = i\frac{2\pi}{3}$ and $\log C = i\frac{\pi}{2}$. Hence

$$\int_{CD} d\log f = -\int_{DC} d\log f + k\left(\frac{2\pi i}{3} - \frac{2\pi i}{4}\right),$$

giving

$$\int_{BC} + \int_{CD} d\log f = 2\pi i k\frac{1}{12}.$$

We have

$$\begin{aligned}
\sum_{\Gamma(1)\backslash\mathfrak{h}} \frac{1}{e^\tau} v_\tau(f) =& \frac{1}{2\pi i}\left(\int_{AB} + \int_{BC} + \int_{CD} + \int_{DE} + \int_{EA} d\log f\right) \\
=& \frac{1}{2\pi i}\left(0 + \frac{k}{12} + 0 - v_\infty(f)\right) \\
& \implies \sum_{\tau \in \Gamma(1)\backslash\mathfrak{h}} \frac{1}{e_\tau} v_\tau(f) + v_\infty(f) = \frac{k}{12}.
\end{aligned}$$

This finishes the proof in the case where there are no zeroes or poles. If there are zeroes or poles on γ, we need to modify the contour. For example, if there's a zero or a pole at a point P on AB, then consider the contour γ', which is just γ but with a small semicircle around our (discrete) pole, which satisfies the property that f has no zeroes or poles on γ'. The trickiest case is when there is a zero or pole at $B = \rho$ or $C = i$. This is Q3 on example sheet 1. $\square$

16 Oct 2022, Lecture 5

Example 2.1. Take $k = 4$, $f = E_4 \in M_4(\Gamma(1))$. Hence $\forall \tau \in \mathfrak{h}, v_\tau(E_4) \geq 0$ (as it is holomorphic in $\mathfrak{h}$). We know $E_4(\tau) = 1 + \sum_{n\geq 1} a_n q^n$, so $E_4(\infty) \neq 0$ and $v_\infty(E_4) = 0$. Hence our formula gives

$$\sum_{\tau \in \Gamma(1)\backslash\mathfrak{h}} \frac{1}{e_\tau} v(E_4) = \frac{1}{3} v_\rho(E_4) + \frac{1}{2} v_i(E_4) + \sum_{\tau \in \Gamma(1)\backslash\mathfrak{h}, \tau \not\sim \rho, i} v_\tau(E_4) = \frac{1}{3}.$$

So we have $\frac{a}{3} + \frac{b}{2} + c = \frac{1}{3}$, where $a, b, c \in \mathbb{Z}_{\geq 0}$, which gives the only solution $a = 1, b = c = 0$, so $E_4(\rho) = 0$ and $E_4(\tau) \neq 0$ if $\tau \notin \Gamma(1)\rho$.

If $k = 6$, $f = E_6$, then we get

$$\frac{1}{3} v_\rho(E_6) + \frac{1}{2} v_i(E_6) + \sum_{\tau \not\sim \rho, i} v_\tau(E_6) = \frac{6}{12} = \frac{1}{12},$$

so this forces $v_\rho(E_6) = 0$, $v_i(E_6) = 1$, $v_\tau(E_6) \neq 0$ if $\tau \not\sim \rho, i$, so $E_6(1) = 0$, $E_6(\tau) \neq 0$ if $\tau \not\sim \rho, i$.

Recall $\Delta = \frac{E_4^3 - E_6^2}{1728} \in S_{12}(\Gamma(1))$. This is nonzero since $\Delta(\rho) = \frac{E_4(\rho)^3 - E_6(\rho)^2}{1728} = -\frac{E_6(\rho)^2}{1728} \neq 0$. We also have $v_\infty(\Delta) \geq 1$ by construction, so plug in Δ to our formula to get

$$\sum_\tau \frac{1}{e_\tau} v_\tau(\Delta) + v_\infty(\Delta) = 1,$$

so $v_\infty(\Delta) = 1$, so Δ has a simple zero at ∞ and Δ is nonvanishing in $\mathfrak{h}$.

Theorem 2.5. Let $k \in 2\mathbb{Z}$. Then:

(1) If $k < 0$ or $k = 2$, then $M_k(\Gamma(1)) = 0$, $M_0(\Gamma(1)) = \mathbb{C} \cdot 1$.

(2) If $4 \leq k \leq 10$, then $M_k(\Gamma(1)) = \mathbb{C} \cdot E_k$.

(3) For any k, multiplication by Δ gives an isomorphism $M_k(\Gamma(1)) \xrightarrow{\times\Delta} S_{k+12}(\Gamma(1))$.

Proof. (1) Let $f \in M_k(\Gamma(1))$ be nonzero. Then $\sum \frac{1}{e_\tau} v_\tau(f) + v_\infty(f) = \frac{k}{12}$. Note the LHS is ≥ 0, but for $k < 0$, the RHS is < 0. If $k = 2$, then we get the equation $\frac{a}{3} + \frac{b}{2} + c = \frac{1}{6}$ for $a, b, c \in \mathbb{Z}_{\geq 0}$, which has no solutions.

Suppose $f \in M_0(\Gamma(1)) \setminus \mathbb{C} \cdot 1$. Then $f - f(\infty) \cdot 1 \in S_0(\Gamma(1))$ is a nonzero function (here 1 is the constant function 1). Then $\sum_\tau \frac{1}{e_\tau} v_\tau(f - f(\infty) \cdot 1) + \underbrace{v_\infty(f - f(\infty) \cdot 1)}_{\geq 1} = 0$, a contradiction, so $M_0(\Gamma(1)) = \mathbb{C} \cdot 1$.

(2) Let $4 \le k \le 10$ and $f \in M_k(\Gamma(1))$. Consider $f - f(\infty) \cdot E_k \in S_k(\Gamma(1))$. If this is nonzero, then

$$\sum_{\tau} \frac{1}{e_\tau} v_\tau(f - f(\infty) \cdot E_k) + \underbrace{v_\infty(f - f(\infty) \cdot E_k)}_{\ge 1} = \frac{k}{12} < 1,$$

a contradiction. So $f = f(\infty) \cdot E_k$.

(3) Our map $\times\Delta : M_k(\Gamma(1)) \to S_{k+12}(\Gamma(1))$ is a well–defined $\mathbb{C}$–linear map. It is injective, since if $\Delta f = 0$, then $f = 0$ (as Δ is nonvanishing in $\mathfrak{h}$). For surjectivity, if $f \in S_{k+12}(\Gamma(1))$, then $\frac{f}{\Delta}$ is holomorphic in $\mathfrak{h}$ and invariant under the weight k action of $\Gamma(1)$.

We need to show $\frac{f}{\Delta}$ is holomorphic at ∞, as then $\frac{f}{\Delta} \in M_k(\Gamma(1))$, so $f = \frac{f}{\Delta} f \in \mathrm{Im}(\times\Delta)$. Hence we need $v_\infty\left(\frac{f}{\Delta}\right) \ge 0$. But $v_\infty\left(\frac{f}{\Delta}\right) = \underbrace{v_\infty(f)}_{\ge 1} - \underbrace{v_\infty(\Delta)}_{=1} \ge 0$, so we're done.

□

Corollary 2.6. If $k \in 2\mathbb{Z}$, $k \ge 0$, then $M_k(\Gamma(1))$ is finite–dimensional and

$$\dim_{\mathbb{C}} M_k(\Gamma(1)) = \begin{cases} \left\lfloor \frac{k}{12} \right\rfloor + 1 & k \not\equiv 2 \pmod{12}. \\ \left\lfloor \frac{k}{12} \right\rfloor & k \equiv 2 \pmod{12}. \end{cases}$$

Proof. We proved this for $0 \le k \le 10$. In general, use induction on k: we need to show that for $k \ge 0$, $\dim_{\mathbb{C}} M_{k+12}(\Gamma(1)) = \dim_{\mathbb{C}} M_k(\Gamma(1)) + 1$.

We know $E_{k+12} \in M_{k+12}(\Gamma(1))$, so $M_{k+12}(\Gamma(1)) = \mathbb{C}E_{k+12} \oplus S_{k+12}(\Gamma(1))$. But this equals $\mathbb{C}E_{k+12} \oplus \Delta M_k(\Gamma(1))$, so $\dim_{\mathbb{C}} M_{k+12}(\Gamma(1)) = 1 + \dim_{\mathbb{C}} M_k(\Gamma(1))$. □

Example 2.2. We have $E_4^2 \in M_8(\Gamma(1)) = \mathbb{C}E_8$. So there is a relation between E_4^2 and E_8, but $E_8(\infty) = 1 = E_4(\infty)^2 \implies E_4^2 = E_8$.

Similarly, $E_4 E_6 \in M_{10}(\Gamma(1)) = \mathbb{C}E_{10}$, so the same argument gives $E_4 E_6 = E_{10}$.

Corollary 2.7. If $k \in 2\mathbb{N}$, then $M_k(\Gamma(1))$ is spanned as a $\mathbb{C}$–vector space by $\{E_4^a, E_6^b \mid a, b \in \mathbb{Z}_{\ge 0}, 4a + 6b = k\}$. In other words, if $\mathcal{M} = \oplus_{k \in \mathbb{Z}} M_k(\Gamma(1))$, then $\mathcal{M}$ is a graded $\mathbb{C}$–algebra generated by E_4 and E_6.

Proof. We proved this for $0 \le k \le 10$. If $k \ge 12$, then

$$M_k(\Gamma(1)) = \mathbb{C}E_k \oplus \Delta M_{k-12}(\Gamma(1)) = \mathbb{C}f \oplus \Delta M_{k-12}(\Gamma(1))$$

for any $f \in M_k(\Gamma(1))$ such that $f(\infty) \neq 0$ by the same argument. We can always find some $A, B \in \mathbb{Z}_{\geq 0}$ such that $4A+6B = k$, so $E_4^A E_6^B \in M_k(\Gamma(1))$ and $(E_4^A E_6^B)(\infty) \neq 0$. Now by induction, $M_{k-12}(\Gamma(1)) = \langle E_4^a E_6^b \mid 4a+6b = k-12 \rangle$, so $\Delta M_{k-12}(\Gamma(1)) = \langle \Delta E_4^a E_6^b \mid 4a+6b = k-12 \rangle$. But $\Delta \in \langle E_4^3, E_6^2 \rangle$, so

$$\Delta M_{k-12}(\Gamma(1)) = \langle E_4^a E_6^b \mid 4a + 6b = k \rangle$$

and $E_4^A E_6^B \in \langle E_4^a E_6^b \mid 4a+6b = k \rangle$, so $M_k(\Gamma(1)) = \langle E_4^a E_6^b \mid 4a+6b = k \rangle$. □

Theorem 2.8. Let $j(\tau) = \frac{E_4(t)^3}{\Delta(\tau)}$. Then j is a modular function of weight 0, level $\Gamma(1)$. It defines a bijection $\Gamma(1) \setminus \mathfrak{h} \to \mathbb{C}$ given by $\tau \to j(\tau)$. Moreover, every modular function of weight 0, level $\Gamma(1)$ is a rational function of j.[6]

The interpretation of this is that it is possible to define a Riemann surface structure on $\Gamma(1) \setminus \mathfrak{h} \sqcup \{\infty\}$ such that we get a compact Riemann surface whose meromorphic functions are exactly the modular functions of weight 0. So the theorem says that this Riemann surface, called $X(1)$, is isomorphic to the Riemann sphere, and our formula says that if $\mathcal{L}$ is an invertible sheaf on a compact Riemann surface and S is a meromorphic section, then $\sum_a v_a(S) = \deg(\mathcal{L})$. This is useful if we are also taking algebraic geometry.

[6]Remember that $\Gamma(1) \setminus \mathfrak{h}$ is the set of orbits of $\Gamma(1)$ under $\mathfrak{h}$.