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. $\square$

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

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

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

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

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 $D^*(0,1)$ 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 $\{\tau \in \mathfrak{h} \mid \operatorname{Im}(\tau) > \frac{1}{2\pi} \log \delta\}$, 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

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

$\square$