These lecture notes are written for a course given at the GANDA conference in Johannesburg September 30 - October 4, 2019. I apologize for any typos or mathematical mistakes.

A gentle introduction to modular forms

Asbjørn Nordentoft, University of Copenhagen

September 29, 2019

1 Introduction

The idea behind this lecture is to give an introduction to modular forms, essentially from first principles. We will in our exposition follow Zagier [2] rather closely. There are many ways to approach this task and I have chosen to focus only on full level modular forms, i.e. modular forms for the full modular group. As the over-all goal of these lecture, we will give a self-contained proof of one of the congruences for partition numbers due to Ramanujan himself. So let us recall what partition numbers are.

1.1 Partition numbers

A natural combinatorial question to ask is, in how many ways one can write a number N as a sum of smaller, positive integers. If one are studying *ordered partitions*, where the order of the integers matters then the answer is easy and we leave it as an exercise.

Exercise 1. *Show that*

$$\#\left\{(n_1,\ldots,n_m) \mid n_i \in \mathbb{Z}_{\geq 1}, n_1+\ldots+n_m = N\right\} = 2^{N-1}.$$

If we however ask for the number of unordered partitions, denoted by $p(n)$, then the question gets extremely complicated. Ramanujan however succeeded in finding an asymptotic formula to the amazement of his peers (as is portrayed in the Hollywood movie "The Man Who Knew Infinity"!). Furthermore he discovered certain congruences satisfied by $p(n)$. We will in these notes present a self-contained proof of the following incredible congruence

$$p(5n+4) \equiv 0 \mod 5,$$

using modular forms. The starting point is the following product representation for the generating series (which we leave to the reader to verify):

$$\prod_{n\geq 1} \frac{1}{1-q^n} = \sum_{n\geq 1} p(n)q^n.$$

This product as a function of q turns out to posses a number of beautiful symmetries, which in essence is the subject of the theory of *modular forms*. Ramanujan found a way to ingeniously make use of these symmetries to obtain his wonderful congruences.

2 The modular group

The starting point for modular forms is the action of

$$\mathrm{SL}_2(\mathbb{R}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{R}, ad - bc = 1 \right\}$$

on the Poincare upper half-plane $\mathbb{H} = \{x + iy \in \mathbb{C} \mid y > 0\}$ given by fractional linear transformations;

$$\gamma z := \frac{az + b}{cz + d}, \quad z \in \mathbb{H}, \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Observe that the element $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ acts trivially and hence some textbooks works with the modified group $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{SL}_2(\mathbb{R})/\pm I$

Exercise 2. *Show that* $\operatorname{Im}\gamma z = \operatorname{Im} z/|cz + d|^2$ *and conclude* $\gamma z \in \mathbb{H}$. *Finally show that the above defines a group action (i.e.* $\gamma_{id} z = z$, $\gamma'(\gamma z) = (\gamma'\gamma)z$*).*

Clearly the above also defines an action of the (discrete) subgroup $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$, which we call the (full) **modular group**.
The main object of study this week will be functions $f : \mathbb{H} \to \mathbb{C}$ which transforms appropriately with respect to this action. First of all lets understand this action a bit better.

Proposition 1. *The region* $F = \{z \in \mathbb{H} \mid |z| > 1, |\operatorname{Re} x| < 1/2\}$ *is a* **fundamental domain** *for the action of* $\mathrm{SL}_2(\mathbb{Z})$ *on* $\mathbb{H}$. *This means that for all* $z \in \mathbb{H}$ *there exists* $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ *such that* $\gamma z \in \overline{F}$ *and for all* $z, z' \in F$ *with* $z \neq z'$ *we have* $\gamma z \neq z'$ *for all* $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Proof. For any $z \in \mathbb{H}$ there are only finitely many γ such that $|cz+d| \leq 1$ (why?). By the above exercise we know $\operatorname{Im} \gamma z = \operatorname{Im} z/|cz+d|^2$ and thus there exists a maximal $\gamma_0 \in \mathrm{SL}_2(\mathbb{Z})$ such that

$$\operatorname{Im}(\gamma_0 z) \geq \operatorname{Im}(\gamma z), \quad \text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}).$$

Observe that the action of $T := \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ does not affect the imaginary part. Thus we can choose a maximal γ_0 as above with $|\operatorname{Re} \gamma z| \leq 1/2$. By maximality we have for $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$

$$\operatorname{Im}(\gamma_0 z) \geq \operatorname{Im}(S\gamma_0 z) = \frac{\operatorname{Im}(\gamma_0 z)}{|\gamma_0 z|^2}.$$

Thus we conclude that $|\gamma_0 z| \geq 1$ and thus $\gamma_0 z \in \overline{F}$.
Now let $z, z' \in \overline{F}$ with $\gamma z = z'$ and assume WLOG $\operatorname{Im} z \leq \operatorname{Im} z'$. Then we have

$$\operatorname{Im} z \leq \operatorname{Im} z' = \frac{\operatorname{Im} z}{|cz+d|^2}$$

and so $|cz+d| \leq 1$. This inequality is however very restrictive since

$$|cz+d|^2 = (cx+d)^2 + (cy)^2$$

and $y \geq \sqrt{3}/2$. Thus we must have $|c| \leq 1$ and we now just have to do some case work.
If $c = 0$ then $d = \pm 1$ is the only possibility and thus $\gamma = \begin{pmatrix} \pm 1 & n \\ 0 & \pm 1 \end{pmatrix}$. Since $z, z' \in \overline{F}$, the only possibility is $|n| \leq 1$ and $|\operatorname{Re} z| = |\operatorname{Re} z'| = 1/2$.
If $c = 1$ then we have $|cz+d| = |z+d| \leq 1$ which is only possible if $|z| = 1$ and $d = 0$ or $z = \omega := e^{2\pi i/3}$, $d = 1$ or $z = \omega + 1$ and $d = -1$.
The case $c = -1$ is completely analogous to $c = 1$.
This completes the proof since all equivalent pairs $z, z' \in \overline{F}$ where shown to be on the boundary. $\square$

A few remark; from the above one can easily determine all fixed points of the $\mathrm{SL}_2(\mathbb{Z})$-action and one can show that $\mathrm{SL}_2(\mathbb{Z})$ is generated by the two matrices S, T defined above.
Now we are ready to define the main object of study

Definition 1. *A* **modular form of weight k** *is a holomorphic function* $f : \mathbb{H} \to \mathbb{C}$ *satisfying two conditions*

1. $f(\gamma z) = j(\gamma, z)^k f(z) = (cz+d)^k f(z), \quad z \in \mathbb{H}, \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z})$

2. *f is "holomorphic at ∞"*

We denote the set of all modular forms of weight k by $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$.

This last condition requires some explanation; by the modular property we see that $f(z+1) = f(z)$ and thus we can write $f(z) = \tilde{f}(q)$ where $\tilde{f} = f \circ \frac{\log}{2\pi i}$ is a holomorphic map defined on the annulus $0 < |q| < 1$. We know that $\tilde{f}$ has a Laurent expansion, which gives;

$$f(z) = \sum_{n \in \mathbb{Z}} a_n e^{2\pi i n z} = \sum_{n \in \mathbb{Z}} a_n q^n, \quad q = e^{2\pi i z}. \tag{1}$$

This is called the q-expansion of f. When $y \to \infty$ we have $e^{2\pi i z} \to 0$, thus $q = 0$ corresponds to a point at infinity (the *cusp* of $\mathrm{SL}_2(\mathbb{Z})$). We say that f is "holomorphic at ∞" if $a_n = 0$ for $n < 0$. If this is the case then $\tilde{f}$ has a removable singularity at $q = 0$.
It is an important fact that modular forms only have a finite number of zeroes and poles in $\overline{F}$. This follows since $\tilde{f}$ is a holomorphic function defined in a bounded region.

Exercise 3. *Show that if $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ for k odd then $f \equiv 0$*

Of special interest are those modular forms with $a_0 = 0$, which we call **cusp forms**. We denote all cusp forms of weight k by $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$.

Exercise 4. *Show that $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$ admit a natural structure of a $\mathbb{C}$-vector space (by pointwise addition and scalar multiplication).*

Exercise 5. *Let $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $g \in \mathcal{M}_l(\mathrm{SL}_2(\mathbb{Z}))$ be modular forms weight k and l respectively. Show that $fg \in \mathcal{M}_{l+k}(\mathrm{SL}_2(\mathbb{Z}))$. Furthermore if $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ and $g \in \mathcal{S}_l(\mathrm{SL}_2(\mathbb{Z}))$ then $fg \in \mathcal{S}_{l+k}(\mathrm{SL}_2(\mathbb{Z}))$.This exercise show that $\mathcal{M}.(\mathrm{SL}_2(\mathbb{Z})$ has a natural structure of a graded algebra over $\mathbb{C}$ and $\mathcal{S}.(\mathrm{SL}_2(\mathbb{Z}))$ is an ideal.*

2.1 Eisenstein series

Up until now we haven't seen any examples of modular forms. For $k = 0$ we have all constants functions. For $k \geq 4$ even, we can define a modular form by the method of averaging as

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

We want to argue that this series converges to a holomorphic function. So let z be contained in a compact set. The number of intergers m, n such that

$$N < |mz + n| < N + 1,$$

is uniformly bounded by a constant times the number of lattice points in the annulus of area $\pi(N+1)^2 - \pi N^2$. It should be clear that this count is $O(N)$. Thus we have the following bound

$$\sum_{\substack{m,n \\ (m,n)\neq(0,0)}} \left| \frac{1}{(mz+n)^k} \right| \leq \sum_{1 \leq N \leq \infty} \frac{N}{N^k},$$

which converges for $k \geq 4$. Since the convergence is locally uniform, we know that $G_k(z)$ defines a holomorphic function. Note that the sum in (2) does not converge absolutely for $k = 2$. We will however see later that we can define a certain Eisenstein series of weight 2 by being a bit more careful.

Exercise 6. *Show that $G_k(z)$ for $k \geq 4$ satisfy the modular transformation rule.*

In order to show that $G_k(z)$ is actually a modular form we need to calculate its q-expansion.

Lemma 1. *For $k \geq 4$ we have the following q-expansion*

$$G_k(z) = \zeta(k) + \frac{(2\pi i)^k}{(k-1)!} \sum_{n \geq 1} \sigma_{k-1}(n) q^n,$$

where $\zeta(s) = \sum_{n \geq 1} n^{-s}$ is the famous Riemann zeta-function and $\sigma_s(n) = \sum_{d|n} d^s$.
.

Proof. The starting point is the following identity due to Euler (we will omit the proof)

$$\frac{\pi}{\tan \pi z} = \frac{1}{z} + \sum_{n \geq 1} \frac{1}{z+n} + \frac{1}{z-n}, \quad z \in \mathbb{C}\backslash\mathbb{Z}. \tag{3}$$

Now the left hand-side is periodic of period one and its Fourier expansion can be calculated in the following way for $z \in \mathbb{H}$

$$\frac{\pi}{\tan \pi z} = \frac{\pi \cos \pi z}{\sin \pi z} = \frac{\pi i(e^{\pi i z} + e^{-\pi i z})}{(e^{\pi i z} - e^{-\pi i z})} = \frac{-\pi i(1+q)}{1-q} = -2\pi i \left(\frac{1}{2} + \sum_{n \geq 1} q^n \right)$$

Now differentiating (3) a total of $k-1$ times (which can be done term-wise by locally uniform convergence) and using the above we arrive, after dividing by $(-1)^{k-1}(k-1)!$, at

$$\sum_{n\in\mathbb{Z}} \frac{1}{(z+n)^k} = \frac{(-2\pi i)^k}{(k-1)!} \sum_{n\geq 1} n^{k-1} q^n \tag{4}$$

We now split the sum in (2) into those with $c=0$ and $c\neq 0$

$$\begin{aligned} G_k(z) &= \frac{1}{2}\sum_{n\neq 0}\frac{1}{n^k} + \frac{1}{2}\sum_{m\neq 0}\sum_{n\in\mathbb{Z}}\frac{1}{(mz+n)^k} \\ &= \zeta(k) + \sum_{m>0}\frac{(2\pi i)^k}{(k-1)!}\sum_{n\geq 1} n^{k-1}q^{mn} \\ &= \zeta(k) + \frac{(2\pi i)^k}{(k-1)!}\sum_{n\geq 1}\sigma_{k-1}(n)q^n \end{aligned}$$

using that for even k we have $\sum_{m\in\mathbb{Z}}\frac{1}{(mz+n)^k} = \sum_{m\in\mathbb{Z}}\frac{1}{(-mz+n)^k}$. This finishes the proof. □

A famous calculation of Euler yields for even k

$$\zeta(k) = \frac{-B_k(2\pi i)^k}{k!},$$

where B_k denotes the kth Bernoulli number defined by

$$\frac{x}{e^x-1} = \sum_{k\geq 1}\frac{B_k x^k}{k!},$$

of which the first few values are $B_2 = 1/6, B_4 = -1/30, B_6 = 1/42$.
We now define a convenient re-scaling of $G_k(z)$

$$E_k(z) := \zeta(k)^{-1}G_k(z) = 1 - \frac{2k}{B_k}\sum_{n\geq 1}\sigma_{k-1}(n)q^n,$$

which we call **the weight k (holomorphic) Eisenstein series**. Using the values for the Bernoulli numbers above we have

$$\begin{aligned} E_4(z) &= 1 + 240\sum_{n\geq 1}\sigma_3(n)q^n, \\ E_6(z) &= 1 - 504\sum_{n\geq 1}\sigma_5(n)q^n. \end{aligned}$$

Exercise 7. *Show that*

$$E_k(z) = \frac{1}{2} \sum_{\substack{m,n \in \mathbb{Z} \\ \gcd(m,n)=1}} (mz+n)^{-k}.$$

As the above q-expansion shows the Eisenstein series are *not* cusp forms. We can now however define one.

Exercise 8. *Show that* $\Delta(z) := \frac{1}{1728}(E_4(z)^3 - E_6(z)^2)$ *is a cusp form of weight 12 different from zero. (This is the famous Ramanujan* Δ*-function).*

The coefficients in the q-expansion

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

are called the Ramanujan tau-function, which has a number of very fascinating properties (note that $\tau(1) = 1$), which we will not touch upon.

2.2 The dimension formula

As we saw in the exercise $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ is a vector space. An important result is that one can actually calculate the dimension.

Recall that the zeroes and poles of a holomorphic function is a discrete set. In particular there are only finitely many zeroes and poles in any bounded region. A modular form does not quite define a function on the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$. We can however for a point in $\mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ define **the order** as the smallest n such that $a_n \neq 0$ in the Taylor expansion at *any* representative in $\mathbb{H}$, which one checks is well-defined using the modular transformation rule.

In order to calculate the dimensions of $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ we take advantage of the geometry of the fundamental region as described above. We will rely on complex analysis and do a contour integral along the boundary of F (or a modified version) with ε-neighborhoods removed around the special points $i, \omega, \omega+1, \infty$. This leads to *The Valence Formula*.

Proposition 2 (The Valence Formula)**.** *Let* $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$ *be a weight* k *modular form different from zero. Then we have*

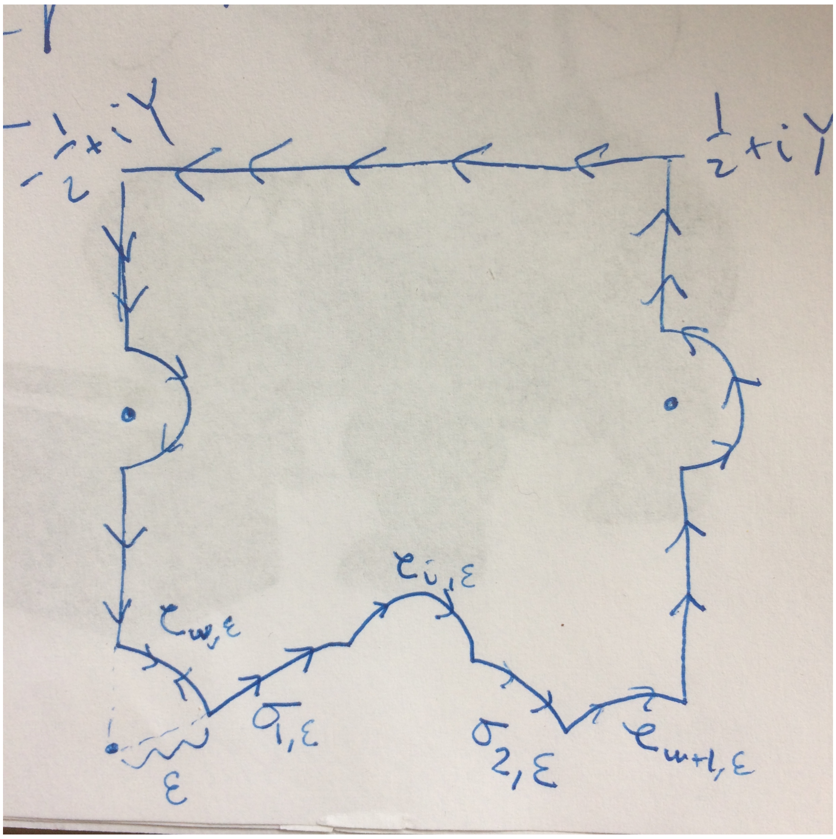
$$\sum_{P \in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}} \frac{ord_P(f)}{n_P} + ord_\infty(f) = \frac{k}{12},$$

where $n_P = 2$ *or* 3 *if* P *is equivalent to* i *or* ω *and* $n_P = 1$ *otherwise (here* $P \in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}$ *means that* P *runs through a set of representatives of the quotient space and as mentioned above the order of such a representative is well-defined).*

Proof. The main ingredient is Cauchy's formula for a meromorphic function g

$$\frac{1}{2\pi i}\int_\gamma \frac{g'(z)}{g(z)}dz = \sum_{P\in\Omega} \mathrm{ord}_P(g),$$

where γ is a simple, closed curve bounding the region Ω and $\mathrm{ord}_P(g)$ is the order of g at P (recall that $\mathrm{ord}_P(g) = 0$ for all but finitely many P since the set of zeroes and poles is discrete).
We will apply this formula to path σ (avoiding poles or zeroes of f), which is a modified version of the boundary of F. We construct σ by removing a circle segment of radius ε around $i, \omega, \omega + 1$ and a neighborhood at infinity; $\mathrm{Im}\, z > Y$ (corresponding to $0 < |q| < e^{-2\pi Y}$) and finally correcting by small semi-circles around the poles on the boundary.



As we already mentioned f only has a finite number of poles and zeroes in $\overline{F}$. Thus if we make ε small enough and Y big enough (so that σ incloses all poles and zeroes of f, except those possibly at $i, \omega, \omega+1, \infty$), we get by Cauchy's formula

$$\frac{1}{2\pi i}\int_\sigma \frac{f'(z)}{f(z)}dz = \sum_{\substack{P\in \mathrm{SL}_2(\mathbb{Z})\backslash\mathbb{H}\\ P\neq i,\omega}} \mathrm{ord}_P(f)$$

since $n_P = 1$ in this case. Now we analyze the integral on the left hand-side above.

Since $f(z) = f(z+1)$ the integral along the vertical lines $\pm 1/2$ cancel.
The integral $[-1/2+iY, 1/2+iY]$ yields after a change of variable $q = e^{2\pi i z}$

$$\frac{1}{2\pi i}\int_{[-1/2+iY,1/2+iY]} \frac{f'(z)}{f(z)}dz = \frac{1}{2\pi i}\int_{|q|=e^{-2\pi Y}} \frac{f'(q)}{f(q)}dq = \mathrm{ord}_\infty(f)$$

Now the integral along the circle segment at ω yields

$$\begin{aligned} &\frac{1}{2\pi i}\int_{\mathcal{C}_{\omega,\varepsilon}} \frac{f'(z)}{f(z)}dz \\ =&\frac{1}{2\pi i}\int_{\mathcal{C}_{\omega,\varepsilon}} \frac{\mathrm{ord}_\omega(f)}{z}dz + \frac{1}{2\pi i}\int_{\mathcal{C}_{\omega,\varepsilon}} f_0(z)dz \\ =&-\frac{\mathrm{ord}_\omega(f)}{2\pi}\int_{\pi/6+\theta(\varepsilon)}^{\pi/2} 1\, d\theta - \frac{1}{2\pi i}\int_{\mathcal{C}_{\omega,\varepsilon}} f_0(z)dz, \end{aligned}$$

which converges to $-\frac{\mathrm{ord}_\omega(f)}{6}$ as $\varepsilon \to 0$ and similarly for the integral along $\mathcal{C}_{\omega+1,\varepsilon}$. A very similar argument (do it!) yields that $\mathcal{C}_{\varepsilon,i}$ converges to $-\frac{\mathrm{ord}_i(f)}{2}$. By the modular property $f(-1/z) = z^k f(z)$, we get

$$\frac{1}{2\pi i}\int_{\sigma_{1,\varepsilon}} \frac{f'(z)}{f(z)}dz = -\frac{1}{2\pi i}\int_{\sigma_{2,\varepsilon}} \frac{f'(z)}{f(z)}dz - \frac{1}{2\pi i}\int_{\sigma_{2,\varepsilon}} \frac{k}{z}dz$$

By the above we get that

$$\frac{1}{2\pi i}\int_{\sigma_{1,\varepsilon}} \frac{f'(z)}{f(z)}dz + \frac{1}{2\pi i}\int_{\sigma_{2,\varepsilon}} \frac{f'(z)}{f(z)}dz \to \frac{k}{12}$$

as $\varepsilon \to 0$ since $\sigma_{2,\varepsilon}$ will converge to $1/12$-th of the unit circle.
Adding up all the contributions this yields the desired result in the limit $\varepsilon \to 0$. $\square$

Exercise 9. *Use the valence formula to prove that* $\Delta(z) \neq 0$ *for* $z \in \mathbb{H}$

Corollary 1.

$$\dim_{\mathbb{C}}(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) \leq \begin{cases} \lfloor \frac{k}{12} \rfloor + 1 & k \not\equiv 2 \mod 12 \\ \lfloor \frac{k}{12} \rfloor & k \equiv 2 \mod 12 \end{cases}$$

Proof. Choose $P_1, \ldots, P_{\lfloor \frac{k}{12} \rfloor + 1}$ points in the interior of F. Then for any set $f_1, \ldots, f_{\lfloor \frac{k}{12} \rfloor + 2}$ of modular forms of weight k, we can find a linear combination f, which vanishes at all P_i by solving the system of linear equations. By the valence formula this implies $f \equiv 0$. This gives the required bound for $k \not\equiv 2 \pmod{12}$.
For $k \equiv 2 \pmod{12}$ the valence formula yields that

$$2 \operatorname{ord}_\omega(f) + 3 \operatorname{ord}_i(f) \equiv 1 \mod 6$$

Since $\operatorname{ord}_\omega(f), \operatorname{ord}_i(f) \geq 0$ this means that

$$\frac{\operatorname{ord}_\omega(f)}{3} + \frac{\operatorname{ord}_i(f)}{2} \geq \frac{7}{6}$$

In particular any modular form f of weight k can have at most $\lfloor \frac{k}{12} \rfloor - 1$ zeroes in the interior of F. Now the same argument as above yields the desired bound for such k. $\square$

From the above it follows that $\dim_{\mathbb{C}}(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) = 0$ for $k \leq 0$ and we have already seen that $\dim_{\mathbb{C}}(\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))) = 0$ for k odd. It turns out that these upper bounds are sharp for $k > 0$ and k even. This follows from the following

Lemma 2. *The Eisenstein series E_4 and E_6 are algebraically independent (over $\mathbb{C}$).*

Proof. First of all it is enough to show that $E_4(z)^3$ and $E_6(z)^2$ are algebraically independent since if $P(E_4(z), E_6(z)) \equiv 0$ then a short calculation shows that we can write

$$\prod_{\substack{0 \leq j \leq 2, \\ 0 \leq j' \leq 1}} P((e^{2\pi i/3})^j X, (-1)^{j'} Y) = Q(X^3, Y^2)$$

for some $Q \in \mathbb{C}[X, Y]$ and now $Q(E_4(z)^3, E_6(z)^2) \equiv 0$.
If $Q(E_4(z)^3, E_6(z)^2) \equiv 0$ then by using the modular transformation rule, we see that all homogenous parts $Q^{(d)}(E_4(z)^3, E_6(z)^2)$ have to vanish identically. From this it follows that for all d:

$$Q^{(d)}(E_4(z)^3 / E_6(z)^2, 1) \equiv 0,$$

but since any non-zero polynomial has a finite number of roots, this implies that we can write $E_4(z)^3 = \lambda E_6(z)^2$ for some constant $\lambda \in \mathbb{C}$.
Now if we consider the function $f(z) = E_6(z)/E_4(z)$, its square is equal to $\lambda E_4(z)$ and thus f cannot have any poles (not at infinity either). This means that f is a modular form of weight 2, contradicting the inequality $\dim_{\mathbb{C}}(\mathcal{M}_2(\mathrm{SL}_2(\mathbb{Z}))) \leq 0$. This finishes the proof. $\square$

Exercise 10. *Conclude from the above lemma that the bounds in Corollary 1 are sharp.*

Putting all this together we get the following structural description of $\mathcal{M}_{\cdot}(\mathrm{SL}_2(\mathbb{Z})$.

Corollary 2. $\mathcal{M}_{\cdot}(\mathrm{SL}_2(\mathbb{Z}))$ *is freely generated by* E_4 *and* E_6 *as a graded algebra over* $\mathbb{C}$.

Now we know the dimensions of the space of modular forms, but what about the subspace of cusp forms $\mathcal{S}_k(\mathrm{SL}_2(\mathbb{Z}))$? Actually this follows quite easily.

Exercise 11. *Argue that there are no cusp forms of weight* < 12*. Show using the valence formula that* $\Delta(z) \neq 0$ *for* $z \in \mathbb{H}$*. Use this to show that multiplication by* Δ *defines an isomorphism* $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \cong \mathcal{S}_{k+12}(\mathrm{SL}_2(\mathbb{Z}))$.

3 Eisenstein series of weight 2 and the Serre derivative

As we already mentioned, the formula (2) does not converge for $k = 2$. However by using (4), if we fix the order of summation of m and n then we *do* get a convergent series:

$$G_2(z) := \frac{1}{2}\sum_{n\neq 0}\frac{1}{n^2} + \frac{1}{2}\sum_{m\neq 0}\left(\sum_{n\in\mathbb{Z}}\frac{1}{(mz+n)^2}\right),$$

which has the following Fourier expansion

$$G_2(z) = \zeta(2) + (2\pi i)^2\sum_{n\geq 1}\sigma_1(n)q^{\cdot}$$

by the above. We know from the above dimension formulas that this is not a modular form, but it still satisfies some version of the modular transformation rule.

Proposition 3. *We have*

$$G_2(\gamma z) = j(\gamma, z)^2 G(z) - \pi i\, c\, j(\gamma, z),$$

for $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

Proof. We will here present a proof due to Hecke and following Zagier [2]. Consider for $\varepsilon > 0$ the following absolutely convergent series

$$G_{2,\varepsilon}(z) = \frac{1}{2}\sum_{n\neq 0}\frac{1}{n^{2+2\varepsilon}} + \frac{1}{2}\sum_{m\neq 0}\sum_{n\in\mathbb{Z}}\frac{1}{(mz+n)^2|mz+n|^{2\varepsilon}},$$

which satisfies the transformation rule

$$G_{2,\varepsilon}(\gamma z) = j(\gamma, z)^2 |j(\gamma, z)|^{2\varepsilon} G_{2,\varepsilon}(z).$$

The point of the proof is now to show that $\lim_{\varepsilon \to 0} G_{2,\varepsilon}(\gamma z)$ exists and is equal to $G_2(z) - \pi/2y$. From this and the above, the result follows.
To do this we introduce the function

$$I_\varepsilon(z) := \int_{-\infty}^{\infty} \frac{dt}{(z+t)^2 |z+t|^{2\varepsilon}},$$

which by a change of variable is equal to

$$I_\varepsilon(z) = I_\varepsilon(x+iy) = \int_{-\infty}^{\infty} \frac{dt}{(x+iy+t)^2 |x+iy+t|^{2\varepsilon}} = \frac{I(\varepsilon)}{y^{1+2\varepsilon}}, \tag{5}$$

where

$$I(\varepsilon) = \int_{-\infty}^{\infty} \frac{dt}{(t+i)^2 |t^2+1|^{\varepsilon}}.$$

Now we consider the difference

$$\begin{aligned} &G_{2,\varepsilon}(z) - \sum_{m>0} I_\varepsilon(mz) \\ &= \sum_{n=1}^{\infty} \frac{1}{n^{2+2\varepsilon}} + \sum_{m>0} \sum_{n\in\mathbb{Z}} \left(\frac{1}{(mz+n)^2 |mz+n|^{2\varepsilon}} - \int_n^{n+1} \frac{dt}{(mz+t)^2 |mz+t|^{2\varepsilon}} \right), \end{aligned}$$

where we note that the sum over m is absolutely convergent by (5). By the mean-value theorem the summand above is bounded by $O(|mz+n|^{-3-2\varepsilon})$ and thus the sum is absolutely convergent for $\varepsilon > -1/2$. Thus we can take the limit $\varepsilon \to 0$ term by term to arrive at

$$\begin{aligned} &\lim_{\varepsilon \to 0} \left(G_{2,\varepsilon}(z) - \sum_{m>0} I_\varepsilon(mz) \right) \\ &= \sum_{m>0} \sum_{n\in\mathbb{Z}} \left(\frac{1}{(mz+n)^2} - \int_n^{n+1} \frac{dt}{(mz+t)^2} \right) \\ &= \sum_{m>0} \sum_{n\in\mathbb{Z}} \left(\frac{1}{(mz+n)^2} - \frac{1}{mz+n+1} + \frac{1}{mz+n} \right). \end{aligned}$$

Now it is easy to see that the following telescoping sum vanishes;

$$\sum_{n\in\mathbb{Z}} \left(\frac{1}{mz+n+1} - \frac{1}{mz+n} \right) = \lim_{N\to\infty} \left(\frac{1}{mz+N+1} - \frac{1}{mz-N} \right) = 0.$$

And thus the above limit equals $G_2(z)$.
Now we see easily using (5) that

$$I(0) = \int_{-\infty}^{\infty} \frac{dt}{(t+i)^2} = \left[-\frac{1}{t+i} \right]_{-\infty}^{\infty} = 0,$$

and by direct computation, we get;

$$I'(0) = \int_{-\infty}^{\infty} \frac{\log(t^2+1)}{(t+i)^2} dt = \left[\frac{1+\log(t^2+1)}{t+i} - \arctan(t) \right]_{-\infty}^{\infty} = -\pi.$$

Using Taylor expansions, the above implies;

$$\begin{aligned} &\sum_{m>0} I_\varepsilon(mz) = I(\varepsilon) \sum_{m>0} \frac{1}{(my)^{1+2\varepsilon}} \\ &= I(\varepsilon)\zeta(1+2\varepsilon) y^{-1-2\varepsilon} = \left(-\pi\varepsilon + O(\varepsilon^2) \right) \left(\frac{1}{2\varepsilon} + O(1) \right) y^{-1-2\varepsilon}, \end{aligned}$$

where we used that by partial summation

$$\zeta(1+2\varepsilon) = (1+\varepsilon) \int_1^{\infty} \frac{\sum_{n\leq x} 1}{x^{2+2\varepsilon}} dx = \frac{1}{2\varepsilon} + O(1),$$

as $\varepsilon \to 0$. Thus we conclude that

$$\sum_{m>0} I_\varepsilon(mz) \to -\pi/2y,$$

as $\varepsilon \to 0$. This finishes the proof. □

As above we consider the following renormalized version;

$$E_2(z) := \frac{6}{\pi^2} G_2(z) = 1 - 24 \sum_{n\geq 1} \sigma_1(n) q^n.$$

As a first application we will derive a product expansion for the Ramanujan Δ-function.

Corollary 3. *We have*

$$\Delta(z) := q \prod_{n\geq 1} (1-q^n)^{24}.$$

Proof. Denote by $\tilde{\Delta}$ the function defined by the above product, which we observe is indeed absolutely convergent. We get directly that $\tilde{\Delta}$ is cuspidal at ∞ and non-zero for $z \in \mathbb{H}$. We will now prove that $\tilde{\Delta}$ is a cusp form of weight 12, and

thus conclude that $\tilde{\Delta} = \Delta$ (by comparing the coefficient of q).
Since $\tilde{\Delta}(z)$ is non-zero we can consider its logarithmic derivative;

$$\frac{1}{2\pi i}\frac{\partial}{\partial z}\log \Delta(z) = 1 - 24\sum_{n\geq 1}\frac{nq^n}{1-q^n} = 1 - 24\sum_{n\geq 1}\sigma_1(n)q^n,$$

which we recognized as $E_2(z)$. Now using the proposition above we see that

$$\begin{aligned}&\frac{1}{2\pi i}\frac{\partial}{\partial z}\log\left(\frac{\Delta(\gamma z)}{j(\gamma,z)^{12}\Delta(z)}\right)\\ &= j(\gamma,z)^{-2}E_2(\gamma z) - E_2(z) - \frac{12}{2\pi i}\frac{c}{j(\gamma,z)} = 0.\end{aligned}$$

Thus we see that

$$\Delta(\gamma z) = C(\gamma)j(\gamma,z)^{12}\Delta(z),$$

for some constant $C(\gamma)$, which we would like to show is equal to 1. We only have to check this for the generators S, T. It is obvious for T since Δ is 1-periodic and it follows for S by considering

$$\Delta(Sz) = C(S)z^{12}\Delta(z)$$

at the point $z = i$. $\square$

Now we consider the following operator known as the *Serre derivative* acting on smooth maps $\mathbb{H} \to \mathbb{C}$;

$$\vartheta_k f := f' - \frac{k}{12}E_2 f,$$

where $f' := \frac{1}{2\pi i}\frac{\partial}{\partial z}f$.

Lemma 3. *The Serre derivate ϑ defines a map $\mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z})) \to \mathcal{M}_{k+2}(\mathrm{SL}_2(\mathbb{Z}))$.*

Proof. Let $f \in \mathcal{M}_k(\mathrm{SL}_2(\mathbb{Z}))$. By using chain rule, product rule and modularity we see that

$$\begin{aligned}&j(\gamma,z)^{-2}f'(\gamma z) = (f(\gamma z))'\\ &= \left(j(\gamma,z)^k f(z)\right)' = j(\gamma,z)^k f'(z) + \frac{kc}{2\pi i}j(\gamma,z)^{k-1}f(z).\end{aligned}$$

This shows that

$$f'(\gamma z) = j(\gamma,z)^{k+2}f'(z) + \frac{kc}{2\pi i}j(\gamma,z)^{k+1}f(z),$$

which combined with the proposition above, yields the wanted. $\square$

One can generalize the above to also apply to E_2 itself. We leave this as an exercise.

Exercise 12. *Show that $\vartheta_2 E_2 \in \mathcal{M}_4(\mathrm{SL}_2(\mathbb{Z}))$ and conclude that $\vartheta_2 E_2 = -E_4/6$.*

With these tools at hand we are ready to finish off with our proof of the main theorem following [1].

Theorem 1. *We have the following congruence $p(5n+4) \equiv 0 \mod 5$.*

Proof. By looking at the Fourier expansions we see that coefficient for coefficient;

$$E_4 \equiv 1, \quad E_6 \equiv E_2 \mod 5,$$

using Fermat's theorem; $n^5 \equiv n\,(5)$ in the second congruence.
This implies together with the exercise above that

$$E_4^3 - E_6^2 \equiv E_4 - E_2^2 \equiv -12E_2' \equiv 3E_2' \mod 5.$$

On the other hand the above is connected to Δ as follows

$$\begin{aligned}
&E_4^3 - E_6^2 \\
&= 1728\Delta \\
&\equiv 3q \prod_{n\geq 1}(1-q^n)^{24} \\
&\equiv 3q \frac{\prod_{n\geq 1}(1-q^n)^{25}}{\prod_{n\geq 1}(1-q^n)} \\
&\equiv 3q \frac{\prod_{n\geq 1}(1-q^{25n})}{\prod_{n\geq 1}(1-q^n)} \mod 5,
\end{aligned}$$

using the binomial theorem and the fact that $5 \mid \binom{25}{i}$ for $0 < i < 25$ (why?). From this we see that

$$E_2' \equiv q \prod_{n\geq 1}(1-q^{25n}) \left(\sum_{n\geq 1} p(n)q^n\right) \mod 5.$$

Now we compare the coefficients of q^{5n} for $n > 0$ on both sides. For E_2' all of these coefficients are $\equiv 0$ mod 5 and thus we get

$$\left(\prod_{n\geq 1}(1-q^{25n})\right)\left(\sum_{n\geq 1} p(5n+4)q^{5n+5}\right) \equiv 0 \mod 5.$$

Now the conclusion follows by an easy induction. $\square$

References

[1] Bruce C. Berndt, *Ramanujan's congruences for the partition function modulo 5, 7, and 11*, Int. J. Number Theory **3** (2007), no. 3, 349–354. MR 2352823

[2] Don Zagier, *Elliptic modular forms and their applications*, The 1-2-3 of modular forms, Universitext, Springer, Berlin, 2008, pp. 1–103. MR 2409678