An Introduction to Classical Modular Forms and Harmonic Maass forms

Michael H. Mertens and Larry Rolen

Contents

Introduction

Modular forms play a role in so many important branches of not only number theory, but also mathematics in general (e.g. (algebraic) geometry, topology, Lie theory, just to name a few) and also mathematical physics, that they are an indispensable tool. More recently, the theory of new modular objects, such as harmonic Maass forms, has begun to play a similar role, and has a similarly long list of fields where they have found applications. In particular, combinatorics and q-series have provided the subject with many key questions. In fact, a lot of the best applications in the theory of harmonic Maass forms and motivating examples which historically motivated important results in harmonic Maass forms have come from its interactions with these other subjects, in particular combinatorics and physics. For example, combinatorialists have a different set of tools and intuitions, and a knack for discovering interesting new examples of modular-type objects. Thus, collaboration and communication between the two areas is very fruitful in both directions. These notes aim to provide a short introduction to the subject of modular forms and harmonic Maass forms for anyone who wants to familiarize themselves with the basic concepts.

These notes can of course give but a glimpse of the theory, let alone applications, of modular forms and harmonic Maass forms. Most of the subjects mentioned here are treated only superficially, as doing them full justice would require too much space, and a great many other topics cannot be mentioned at all. To lie a proper foundation, the authors have kept the first part of these notes on classical modular forms relatively formal and rigorous, providing some proof sketches of the most important facts or at least precise references where a more detailed treatment can be found in existing textbooks on the subject. Some sections in this part follow parts of existing textbooks rather closely, in which case this is indicated at the beginning of the section. The second part on harmonic Maass forms is then intentionally less formal, in order to provide a feeling for the new developments that have come from this relatively recent concept (most of the results mentioned are less than 15 years old). The interested reader will find multiple references to original publications in this part, most of which should be readable with the background provided in these notes and perhaps some further study. In the references, a couple of texts (textbooks, surveys, book chapters) are listed that are not necessarily explicitly cited in these notes, but which the authors would particularly like to recommend for further reading. Such references are marked with an asterisk next to the title.

We will begin with a brief overview of some basic facts and definitions of modular forms, focusing on modular forms for the full modular group $\mathrm{SL}_2(\mathbb{Z})$ at first and then on the more general setting of modular forms for (congruence) subgroups, including some examples such as Eisenstein series, the Δ-function, theta series and eta quotients. We will also briefly discuss the theory Jacobi forms and even more briefly describe the theory of singular moduli. The third section will cover some particular applications of the theory of modular forms to partitions, namely a sketch of the proof of Rademacher's series representation of the partition numbers and finally a short outline on a general framework for congruences of the partition function. Following this, we will outline some of the background and historical motivation for the newer theory of harmonic Maass forms, with an eye towards applications in related fields such as combinatorics. Of course, this general survey will not describe all the details of these varied topics, and so the original texts and more extensive books should be consulted by the interested reader. However, the intention of the authors is to provide a practical and concise guide to these topics for students and experts in related fields to spark interest and interdisciplinary collaboration.

1 Basics

1.1 The upper half-plane

Throughout, we denote by $\mathfrak{H}$ the *complex upper half-plane*,

$$\mathfrak{H} = \{\tau = u + iv \in \mathbb{C} \,:\, \operatorname{Im}(\tau) = v > 0\}.$$

This is a model for the hyperbolic plane.

Its group of holomorphic isometries (with respect to the hyperbolic metric) or biholomorphic automorphisms is well-known to be isomorphic to $\mathrm{PSL}_2(\mathbb{R})$. It is however more convenient for our purposes to work with $\mathrm{SL}_2(\mathbb{R})$ instead of the projective group. This group acts on $\mathfrak{H}$ via *Möbius transformations*,

$$\left(\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tau\right) \mapsto \gamma.\tau := \frac{a\tau + b}{c\tau + d}.$$

To see this, note for instance that

$$\operatorname{Im}\left(\frac{a\tau + b}{c\tau + d}\right) = \frac{\operatorname{Im}(\tau)}{|c\tau + d|^2},$$

hence the action is well-defined. That this actually is a group action whose core is generated by $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ is an easy direct verification.

It is now a natural question to ask whether there exist holomorphic functions $f\colon \mathfrak{H} \to \mathbb{C}$ which are invariant under the action of discrete subgroups of $\mathrm{SL}_2(\mathbb{R})$, the most elementary example of which is $\mathrm{SL}_2(\mathbb{Z})$. Unfortunately, it turns out that the answer to this question is *no*, except for constant functions, but it is *yes* if we relax the invariance requirement slightly.

Before discussing this point further, we shall require a description of a fundamental domain of the action of $\mathrm{SL}_2(\mathbb{Z})$. It can be seen in an abstract way that such a fundamental domain containing inner points must exist through Baire's category theorem, but it can be verified in a much more elementary way. Specifically, an exact fundamental domain of the action of $\mathrm{SL}_2(\mathbb{Z})$, i.e. a connected domain[1] $\mathcal{F} \subseteq \mathfrak{H}$ such that for any $\tau \in \mathfrak{H}$ there exists a $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ such that $\gamma.\tau \in \mathcal{F}$ and for any two distinct points $\tau \neq \tau' \in \mathfrak{H}$ there is no $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ with $\gamma.\tau = \tau'$, is given by

$$\mathcal{F} := \left\{\tau \in \mathfrak{H} \,:\, -\frac{1}{2} < \operatorname{Re}(\tau) \leq \frac{1}{2}, |\tau| \geq 1, \text{ and } |\tau| > 1 \text{ for } \operatorname{Re}(\tau) < 0\right\},$$

whose picture is often given in many places. The picture in Figure 1 was generated using SAGE version 8.1 [Sage]. An important fact on its own and also important in proving that

---

[1] Technically, the fundamental domain is not open, but its interior is a domain

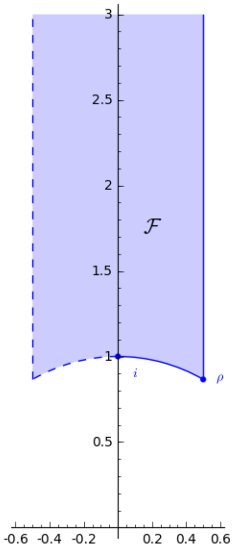

Figure 1: Standard fundamental domain of $\mathrm{SL}_2(\mathbb{Z})$

$\mathcal{F}$ is indeed a fundamental domain is the fact that $\mathrm{SL}_2(\mathbb{Z})$ is generated by two elements,

$$\mathrm{SL}_2(\mathbb{Z}) = \left\langle T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle .$$

The matrices T and S are often referred to as *translation* and *(modular) inversion* respectively, which is why in some places in the literature the letter J instead of S is used.

There are two special points in $\mathcal{F}$, as will be important later, namely the points i and $\rho = \frac{1+i\sqrt{3}}{2}$, which are the only points in $\mathcal{F}$ with a non-trivial stabilizer in $\mathrm{SL}_2(\mathbb{Z})$: i is fixed by S, which has order 4, and ρ is fixed by $U = TS = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, which has order 6. We call such points *elliptic fixed points*. Note that every other point in $\mathcal{F}$ is only fixed (trivially) by $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. For proofs and more details on the facts mentioned here, we refer the reader to [CS, Chapter 4].

1.2 Basic definitions and first results on modular forms

Returning to the question whether there are functions (essentially) invariant under $\mathrm{SL}_2(\mathbb{Z})$, we define the notion of a modular form.

Definition 1.1. *A function* $f\colon \mathfrak{H} \to \mathbb{C}$ *is called a* modular form *of weight* $k \in \mathbb{Z}$ *for* $\mathrm{SL}_2(\mathbb{Z})$ *if*

(i) f *is holomorphic on* $\mathfrak{H}$*,*

(ii) f is invariant under the weight k Petersson slash operator*, that is we have*

$$(f|_k\gamma)(\tau) := (c\tau + d)^{-k} f\left(\frac{a\tau + b}{c\tau + d}\right) = f(\tau)$$

for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and $\tau \in \mathfrak{H}$,

(iii) f is holomorphic at ∞*, i.e. $f(iv)$ is bounded as $v \to \infty$.*

If we even have that $f(iv) \to 0$ as $v \to \infty$, we call f a cusp form*.*

In the following remarks we record several elementary observations on modular forms.

Remark 1.2. *Modular forms (resp. cusp forms) of weight k form a vector space over $\mathbb{C}$ which we denote by M_k (resp. S_k). It follows directly from the definition that products of modular forms are again modular forms and products of modular forms and cusp forms are again cusp forms. That is, we have*

$$M_k \cdot M_\ell \subseteq M_{k+\ell} \quad \textit{and} \quad S_k \cdot M_\ell \subseteq S_{k+\ell}.$$

Remark 1.3. *Since $\mathrm{SL}_2(\mathbb{Z})$ is generated by the two matrices T and S which map $\tau \in \mathfrak{H}$ to $\tau + 1$ and $-1/\tau$ resp., a function $f\colon \mathfrak{H} \to \mathbb{C}$ satisfies the transformation law in (ii) of Definition 1.1 if and only if we have*

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

for all $\tau \in \mathfrak{H}$. In particular, the invariance $f(\tau+1) = f(\tau)$ implies through a standard fact of complex analysis that a modular form $f \in M_k$ has a Fourier expansion *of the form*

$$f(\tau) = \sum_{n\in\mathbb{Z}} \alpha_f(n) e^{2\pi i n\tau}.$$

The growth condition in (iii) of Definition 1.1 yields that $\alpha_f(n) = 0$ for $n < 0$ for $f \in M_k$ and and that we additionally have $\alpha_f(0) = 0$ if and only if f is a cusp form.

Remark 1.4. *Since $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ we find for $f \in M_k$ that*

$$f = f|_k \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = (-1)^k f,$$

hence there are no non-zero modular forms of odd weight.

The following requires a little more work.

Lemma 1.5. *There are no non-zero modular forms of negative weight.*

Proof. Let $k < 0$ and $f \in M_k$. It is not hard to see that the non-holomorphic function $\tilde{f}(\tau) := v^{k/2}|f(\tau)|$ satisfies $\tilde{f}(\gamma.\tau) = \tilde{f}(\tau)$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. If $k < 0$, $\tilde{f}$ is bounded as $v \to \infty$ because f is, so in particular it is bounded for all τ with $\mathrm{Im}(\tau) > \varepsilon > 0$. Thus we find for the Fourier coefficients of f that

$$|\alpha_f(m)| = \left| e^{2\pi mv} \int_0^1 f(u+iv)e^{-2\pi imu}du \right| \leq v^{-k/2}e^{2\pi mv} \int_0^1 \tilde{f}(u+iv) \leq Cv^{-k/2}e^{2\pi mv}$$

for some constant C which doesn't depend on v. Since the left-hand side of this inequality doesn't depend on v, we can take the limit $v \to 0$ which yields $\alpha_f(m) = 0$ for all m, hence $f \equiv 0$, so the claim follows. ☐

One is now inclined to ask whether there actually are any non-trivial examples of modular forms at all. In fact one can construct very explicit and important examples, which are motivated by the theory of *elliptic functions*, on which we cannot touch here.

Example 1.6. *We define the* Eisenstein series *of weight k by*

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

One can show that this series is absolutely and locally uniformly convergent on $\mathfrak{H}$ as soon as $k > 2$ (see for instance [Kil, Lemma 2.7]). Assuming this from now on, we can verify that these functions are indeed modular. Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. Then we have

$$\begin{aligned} G_k\left(\frac{a\tau+b}{c\tau+d}\right) &= \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0\}} \left(m \cdot \frac{a\tau+b}{c\tau+d} + n\right)^{-k} \\ &= (c\tau+d)^k \sum_{(m,n)\in\mathbb{Z}^2\setminus\{0\}} ((ma+nc)\tau + (mb+nd))^{-k}. \end{aligned}$$

Since clearly with (m,n) also $(m,n)\cdot\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = (ma+nc, mb+nd)$ runs through all of $\mathbb{Z}^2\setminus\{0\}$, this yields, as claimed,

$$G_k\left(\frac{a\tau+b}{c\tau+d}\right) = (c\tau+d)^k G_k(\tau).$$

It remains to verify that Eisenstein series don't always vanish identically (which by Remark 1.4 they do if k is odd). For this we can compute their Fourier expansion. Before stating the result, let us recall the definition of the *Riemann zeta function*,

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s}, \quad \mathrm{Re}(s) > 1$$

and of the *Bernoulli numbers* B_n,

$$\sum_{n=0}^{\infty} B_n \frac{t^n}{n!} = \frac{t}{e^t - 1}, \quad |t| < 2\pi.$$

Note that B_n is always rational and non-zero if n is even.

Theorem 1.7. *For even $k \geq 4$ we have that*

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

where $\sigma_k(n) := \sum_{d|n} d^k$ denotes the usual divisor power sum and we use the standard abbreviation $q := e^{2\pi i \tau}$.

The normalized *Eisenstein series are given by*

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

Proof. This is a direct consequence of the so-called *Lipschitz summation formula*

$$\sum_{n\in\mathbb{Z}} (\tau + n)^{-k} = \frac{(-2\pi i)^k}{(k-1)!}\sum_{m=1}^{\infty} m^{k-1} e^{2\pi i m \tau},$$

which holds for $\tau \in \mathfrak{H}$ and integers $k \geq 2$. For a full proof, we refer for example to [Kil, Proposition 2.8]. □

Example 1.8. *The first few normalized Eisenstein series are given explicitly by*

$$\begin{aligned}
E_4(\tau) &= 1 + 240\sum_{n=1}^{\infty} \sigma_3(n) q^n, \\
E_6(\tau) &= 1 - 504\sum_{n=1}^{\infty} \sigma_5(n) q^n, \\
E_8(\tau) &= 1 + 480\sum_{n=1}^{\infty} \sigma_7(n) q^n, \\
E_{10}(\tau) &= 1 - 264\sum_{n=1}^{\infty} \sigma_9(n) q^n, \\
E_{12}(\tau) &= 1 + \frac{65520}{691}\sum_{n=1}^{\infty} \sigma_{11}(n) q^n.
\end{aligned}$$

We now want to turn a little more towards the structure of modular forms. For this, we need the following very important theorem known as the *valence formula* for $\mathrm{SL}_2(\mathbb{Z})$. We mention that it can be thought of as an instance of the Riemann-Roch theorem, but we won't take this viewpoint here.

Theorem 1.9. *Let $f : \mathfrak{H} \to \mathbb{C}$, $f \not\equiv 0$, be a meromorphic function satisfying $f|_k\gamma \equiv f$ for some $k \in \mathbb{Z}$ (note that in this setup, k might be negative) and suppose that f has a Fourier expansion $f(\tau) = \sum_{n=n_0}^{\infty} \alpha_f(n)q^n$ for some $n_0 \in \mathbb{Z}$ with $\alpha_f(n_0) \neq 0$. Further define* $\mathrm{ord}_\infty(f) := n_0$. *Then we have that*

$$\mathrm{ord}_\infty(f) + \frac{1}{2}\,\mathrm{ord}_i(f) + \frac{1}{3}\,\mathrm{ord}_\rho(f) + \sum_{w \in \mathcal{F}\setminus\{i,\rho\}} \mathrm{ord}_w(f) = \frac{k}{12}.$$

Proof. One computes the integral

$$\frac{1}{2\pi i}\int_C \frac{f'(\tau)}{f(\tau)}d\tau$$

where C is the positively oriented boundary curve of a truncated version of the fundamental domain $\mathcal{F}$, but deformed to avoid the points i, ρ, and $-\overline{\rho}$, as well as any zeros or poles of f on the boundary of $\mathcal{F}$. By the residue theorem, this integral essentially yields the left-hand side of the valence formula, and evaluating the integral one finds the value $k/12$, letting the deformations go to 0. For a detailed proof, see for instance [CS, Theorem 5.6.1]. $\square$

It should be remarked that in fact only finitely many terms in the sum over points in the fundamental domain can be non-zero due to the identity theorem (a non-zero meromorphic modular form can only have finitely many zeros or poles in the fundamental domain). Furthermore, the factors $\frac{1}{2}$ and $\frac{1}{3}$ in front of $\mathrm{ord}_i(f)$ and $\mathrm{ord}_\rho(f)$ reflect the fact that i and ρ are elliptic fixed points whose stabilizers have orders $2 \cdot 2$ and $2 \cdot 3$.

An easily computable but non-obvious corollary of the valence formula yields formula for the dimension of the spaces M_k.

Theorem 1.10. *We have*

$$\dim_{\mathbb{C}} M_k = \begin{cases} \lfloor k/12 \rfloor & \text{if } k \equiv 2 \pmod{12}, \\ \lfloor k/12 \rfloor + 1 & \text{if } k \not\equiv 2 \pmod{12}. \end{cases}.$$

Proof. See for example [Kil, Theorem 3.5]. $\square$

In particular, one finds that $M_2 = \{0\}$ and $M_k = \mathbb{C}E_k$ precisely for the values $k = 4, 6, 8, 10, 14$.

Remark 1.11. *The finite-dimensionality of* M_k *is the source of numerous sometimes surprising identities. For instance it follows from Remark 1.2 that* $E_4^2 \in M_8 = \mathbb{C}E_8$*, so, since* E_4^2 *and* E_8 *have the same leading Fourier coefficient, we must have* $E_4^2 = E_8$*. Comparing the coefficients, one finds the so-called* Hurwitz identity *for divisor sums,*

$$\sigma_7(n) = \sigma_3(n) + 120 \sum_{r=1}^{n-1} \sigma_3(r)\sigma(n-r).$$

It is interesting to note that this identity can be formulated entirely in terms of elementary number theory without appealing to modular forms or other advanced mathematics, though real effort is needed to give a combinatorial proof. A proof using only (formal) power series was found by Zagier and Skoruppa in 1978 (published in [KK, Bemerkung I.4.2]). A different elementary proof relying on a series of papers by Liouville is outlined in a math stackexchange post by user Philimathmuse under the link below[2]*, but it doesn't seem to have appeared in print.*

The following important result which also can be derived essentially from the valence formula in Theorem 1.9 is that it is possible to generate all modular forms very easily.

Theorem 1.12. *The space*

$$M_* := \bigoplus_{k=0}^{\infty} M_k$$

is an infinite-dimensional $\mathbb{C}$*-algebra. More precisely, we have that*

$$M_* = \mathbb{C}[E_4, E_6]$$

is the free polynomial algebra generated by E_4 *and* E_6 *(which are easily seen to be algebraically independent).*

Proof. See for instance [Ste, Theorem 2.17]. □

Remark 1.13. *In view of the results in Section 2 where modular forms for other groups than* $\mathrm{SL}_2(\mathbb{Z})$ *are considered it is important to point out that it is* **not** *always the case that the algebra of modular forms is isomorphic to a free polynomial algebra. In fact this is almost never true except for finitely many exceptions. It is always true however that the algebra of modular forms is a finitely generated* $\mathbb{C}$*-algebra, but there usually are non-trivial relations among the generators.*

We conclude this subsection by discussing one further example of modular forms, which is the first non-trivial example of a cusp form.

---

[2] `https://math.stackexchange.com/questions/663919`

Example 1.14. *Consider the so-called* Δ*-function defined by*

$$\Delta(\tau) := \frac{E_4(\tau)^3 - E_6(\tau)^2}{1728} =: \sum_{n=1}^{\infty} \tau(n) q^n = q - 24q^2 + 252q^3 - 1472q^4 + 4830q^5 + O(q^6) \in S_{12}$$

This is a cusp form of weight 12*. The coefficients* $\tau(n)$ *are referred to as the Ramanujan* τ*-function. We clearly have* $\operatorname{ord}_\infty(\Delta) = 1$*, so the valence formula applied to* Δ *yields*

$$1 + \frac{1}{2}\operatorname{ord}_i(\Delta) + \frac{1}{3}\operatorname{ord}(\Delta) + \sum_{w \in \mathcal{F} \setminus \{i, \rho\}} \operatorname{ord}_w(\Delta) = 1.$$

Since Δ *is holomorphic on* $\mathfrak{H}$ *by definition, we must have* $\operatorname{ord}_\tau(\Delta) \geq 0$ *for all* $\tau \in \mathfrak{H}$*, hence it must be that* $\Delta(\tau) \neq 0$ *for all* $\tau \in \mathfrak{H}$*. Through a standard fact in complex analysis, this implies that* Δ *must admit an representation as an infinite product. Indeed, one can show for example using the theory of elliptic functions or properties of the Dedekind eta function (see Section 2.2.2) that*

$$\Delta(\tau) = q \prod_{n=1}^{\infty} (1 - q^n)^{24}.$$

1.3 Operators

In the previous section, we saw various examples of modular forms. We can construct new modular forms from these in a fairly general manner. First we extend the definition of the slash operator to the larger group $\mathrm{GL}_2^+(\mathbb{Q})$ of 2×2-matrices with rational entries and positive determinant by setting

$$(f|_k\gamma)(\tau) := (\det\gamma)^{k/2}(c\tau + d)^{-k} f\left(\frac{a\tau + b}{c\tau + d}\right).$$

Then for $f \in M_k$ the function $f|_k\gamma$ for any $\gamma \in \mathrm{GL}_2^+(\mathbb{Q})$ is again a modular form of weight k, although usually not for the full modular group $\mathrm{SL}_2(\mathbb{Z})$, but for the conjugated group $\gamma^{-1}\,\mathrm{SL}_2(\mathbb{Z})\gamma \leq \mathrm{SL}_2(\mathbb{Q})$. Since one usually considers modular forms for subgroups of $\mathrm{SL}_2(\mathbb{Z})$, one can also view $f|_k\gamma$ as a modular form for the subgroup $\Gamma = (\gamma^{-1}\,\mathrm{SL}_2(\mathbb{Z})\gamma) \cap \mathrm{SL}_2(\mathbb{Z})$, which always has finite index in $\mathrm{SL}_2(\mathbb{Z})$.

Some standard operators on modular forms can be expressed in terms of this generalized slash operator.

Definition 1.15. *For* $f \in M_k$ *and* $m, N, r \in \mathbb{N}$ *and* χ *a Dirichlet character mod* N *we define the operators*

$$f|V_m := m^{-k/2} f|_k \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix},$$

$$f|U_m := m^{k/2-1}\sum_{j=0}^{m-1} f|_k \begin{pmatrix}1 & j\\ 0 & m\end{pmatrix},$$

$$f|S_{N,r} := \frac{1}{N}\sum_{j=0}^{N-1} e^{-2\pi i r j/N} f|_k \begin{pmatrix}1 & j/N\\ 0 & 1\end{pmatrix},$$

$$f\otimes\chi := \sum_{r=0}^{N}\chi(r)(f|S_{N,r}).$$

For all these operators it is usually very convenient to consider their action on the Fourier expansion of a modular form.

Lemma 1.16. *Let $f \in M_k$ be a modular form with Fourier expansion $f(\tau) = \sum_{n=0}^{\infty}\alpha_f(n)q^n$. Then we have in the notation of Definition 1.15*

$$(f|V_m)(\tau) = f(m\tau) = \sum_{n=0}^{\infty}\alpha_f(n)q^{mn},$$

$$(f|U_m)(\tau) = \sum_{n=0}^{\infty}\alpha_f(mn)q^n,$$

$$(f|S_{N,r})(\tau) = \sum_{n\equiv r\ (N)}\alpha_f(n)q^n,$$

$$(f\otimes\chi)(\tau) = \sum_{n=0}^{\infty}\alpha_f(n)\chi(n)q^n.$$

Proof. In all cases, this is a straightforward computation. Note that for all operators, we are acting by upper triangular matrices. □

Some of the most important operators on modular forms are *Hecke operators.* In order to define these, we consider the set

$$\mathcal{M}_m := \{M \in \mathbb{Z}^{2\times 2} \;:\; \det M = m\}$$

for $m \in \mathbb{N}$. The group $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathcal{M}_m$ by left and right multiplication on $\mathcal{M}_m$,

$$\mathrm{SL}_2(\mathbb{Z})\cdot\mathcal{M}_m = \mathcal{M}_m = \mathcal{M}_m\cdot\mathrm{SL}_2(\mathbb{Z}).$$

As one can show without too much difficulty, the set

$$\mathrm{SL}_2(\mathbb{Z}) : \mathcal{M} := \left\{\begin{pmatrix}a & b\\ 0 & d\end{pmatrix} \in \mathcal{M}_m \;:\; ad = m,\ d>0,\ b\in\{0,...,d-1\}\right\}$$

is a full set of representatives of $\mathrm{SL}_2(\mathbb{Z}) \setminus \mathcal{M}_m$ (for a proof, see for instance [KK, Satz IV.1.2]), which is easily seen to have cardinality

$$\#(\mathrm{SL}_2(\mathbb{Z}) : \mathcal{M}_m) = \sigma_1(m),$$

so in particular, it is finite. With this we can define the following.

Definition 1.17. *For $m \in \mathbb{N}$ and $f \in M_k$ we define the mth* Hecke operator *acting on f by*

$$f|T_m^{(k)} := m^{k/2-1} \sum_{M \in \mathrm{SL}_2(\mathbb{Z}):\mathcal{M}} f|_k M.$$

Note that since the weight k is usually clear from context, we often write T_m instead of $T_m^{(k)}$. We now record the action of Hecke operators on Fourier expansions.

Lemma 1.18. *Let $f \in M_k$ with Fourier expansion $f(\tau) = \sum_{n=0}^{\infty} \alpha_f(n) q^n$ and let $g = f|T_m$. Then g has a Fourier expansion with coefficients*

$$\alpha_g(n) = \sum_{d|\gcd(m,n)} d^{k-1} \alpha_f(mn/d^2) \text{ mit } n \geq \begin{cases} 0 & \text{if } \alpha_f(0) \neq 0 \\ 1 & \text{if } \alpha_f(0) = 0. \end{cases}.$$

Proof. With the specific choice of representatives $\mathrm{SL}_2(\mathbb{Z}) : \mathcal{M}_m$, we find that

$$f|T_m(\tau) = m^{k-1} \sum_{\substack{a,d>0 \\ ad=m}} \sum_{b=0}^{d-1} d^{-k} f\left(\frac{a\tau + b}{d}\right).$$

From here, the result is a computation analogous to that in Lemma 1.16. For a detailed proof, see for instance [Ser, Proposition VII.5.12]. □

From this we can deduce one of the most important properties of Hecke operators.

Theorem 1.19. *For $f \in M_k$ and $m \in \mathbb{N}$ we have $f|T_m \in M_k$. If $f \in S_k$, then so is $f|T_m$.*

Proof. We first verify the transformation property. Since $f|_k\gamma = f$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, the definition of T_m is independent of the choice of representatives and since for $M \in \mathcal{M}_m$ we also have $M\gamma \in \mathcal{M}_m$, we have that

$$\begin{aligned} f|T_m|\gamma &= \sum_{M \in \mathrm{SL}_2(\mathbb{Z}):\mathcal{M}_m} (f|_k M)|_k \gamma = \sum_{M \in \mathrm{SL}_2(\mathbb{Z}):\mathcal{M}_m} f|_k(M\gamma) \\ &= \sum_{M \in (\mathrm{SL}_2(\mathbb{Z}):\mathcal{M}_m)\gamma} f|_k M = f|T_m \end{aligned}$$

since $(\mathrm{SL}_2(\mathbb{Z}) : \mathcal{M}_m)\gamma$ is simply another set of representatives of $\mathrm{SL}_2(\mathbb{Z}) \setminus \mathcal{M}_m$. The claim on cusp forms follows immediately from Lemma 1.18. □

We record some important facts about the Hecke operators themselves in the following theorem, whose proof we omit as it can be found in many standard sources.

Theorem 1.20. *Let $m, n \in \mathbb{N}$ be coprime, p a prime, $r \in \mathbb{N}$, and $f \in M_k$. Then the following are true.*

(i) $(f|T_m)|T_n = (f|T_{mn})$*, so the Hecke operators are multiplicative in their indices and in particular, they commute.*

(ii) $(f|T_{p^r})|T_p = f|T_{p^{r+1}} + p^{k-1}f|T_{p^{r-1}}$.

Proof. See for instance [Ste, Proposition 2.29]. ☐

Since the Hecke operators are endomorphisms of the vector space M_k and they all commute, it makes sense to ask for simultaneous eigenforms under all of them. These are usually referred to as *Hecke eigenforms* and play a central role in the theory of modular forms. Although we will not have space to go into much more detail on the theory of eigenforms, we record the following key result.

Theorem 1.21. *Let $f \in M_k$ be not a constant with Fourier expansion $f(\tau) = \sum_{n=0}^{\infty} \alpha_f(n)q^n$. Then the following are equivalent.*

(i) *f is a simultaneous eigenform for all T_m, $m \in \mathbb{N}$.*

(ii) *f is a simultaneous eigenform for all T_p, where p is prime.*

(iii) *$\alpha_f(1) \neq 0$ and for any coprime $m, n \in \mathbb{N}$ we have $\alpha_f(m)\alpha_f(n) = \alpha_f(1)\alpha_f(mn)$, so the Fourier coefficients of Hecke eigenforms are (essentially) multiplicative functions.*

Proof. The equivalence of (i) and (ii) follows from Theorem 1.20. For the implication $(i) \Rightarrow (ii)$, see for instance [Ser, Corollary VII.5.2], the implication $(ii) \Rightarrow (i)$ follows because up to a constant non-zero factor $\alpha_f(1)$, the eigenvalue of f under T_m equals the Fourier coefficient $\alpha_f(m)$ by Lemma 1.18. ☐

Remark 1.22. *The multiplicativity of the Fourier coefficients of Hecke eigenform is an extremely important property, for example when one talks about their L-*functions *(naively speaking, one replaces q^n in the Fourier expansion of a modular form by n^{-s} for some $s \in \mathbb{C}$ with* $\mathrm{Re}(s)$ *sufficiently large). The multiplicativity then translates to the fact that the associated L-function of a Hecke eigenform has a so-called Euler product. A toy example of this is the Riemann zeta function which for* $\mathrm{Re}(s) > 1$ *can be written as*

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p \text{ prime}} (1 - p^{-s})^{-1}.$$

L-functions of modular forms have played a central role in the past century of number theory, for example in the proof of Fermat's Last Theorem or the Birch and Swinnerton-Dyer Conjecture. Moreover, Euler products are required to make connections to L-functions arising from arithemetic-geometric sources such as the theory of elliptic curves as in the Modularity Theorem, formerly known as the Taniyama-Shimura-Weil Conjecture. For a thorough treatment of this, we refer to [DS].

We note one important corollary of Theorem 1.21, which had been conjectured by Ramanujan after he had computed about 30 values of the Ramanujan τ-function (i.e. Fourier coefficients of the cusp form Δ) and first proven by Mordell, some 15 years before Hecke set up his general theory of operators.

Corollary 1.23. *The Ramanujan τ-function is multiplicative.*

Proof. The space S_{12} is one dimensional, hence, since Hecke operators map cusp forms to cusp forms, Δ must be a simultaneous Hecke eigenform, so by Theorem 1.21 its Fourier coefficients are multiplicative. $\square$

2 More advanced theory

2.1 More general settings for modular forms

As we have already seen above, one can produce modular forms for subgroups of $\mathrm{SL}_2(\mathbb{Z})$ from those for the full modular group by slashing with matrices in $\mathrm{GL}_2^+(\mathbb{Q})$. We now give a general definition of such forms.

Definition 2.1. *Let $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ be a finite-index subgroup of $\mathrm{SL}_2(\mathbb{Z})$. Then a function $f : \mathfrak{H} \to \mathbb{C}$ is called a* modular form *of weight $k \in \mathbb{Z}$ for Γ if*

(i) *f is holomorphic on $\mathfrak{H}$,*

(ii) *$f|_k\gamma \equiv f$ for all $\gamma \in \Gamma$,*

(iii) *f is* holomorphic at the cusps*, which means that for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ we have $(f|_k\gamma)(iv)$ is bounded as $v \to \infty$.*

If f satisfies instead of (iii) the stronger condition that $(f|_k\gamma)(iv) \to 0$ as $v \to \infty$, we call f a cusp form for Γ. We denote the spaces of weight k modular forms resp. cusp forms by $M_k(\Gamma)$ resp. $S_k(\Gamma)$.

Remark 2.2. *It should be remarked that since Γ has finite index in $\mathrm{SL}_2(\mathbb{Z})$, one just needs to check condition (iii) in Definition 2.1 for a finite set of representatives of $\mathrm{SL}_2(\mathbb{Z})/\Gamma$. Note however that this is not always trivial to do in practice. Also, one cannot see in general if f is a cusp form just by looking at its Fourier expansion at ∞.*

Remark 2.3. *Even though we have first encountered modular forms for subgroups of* $\mathrm{SL}_2(\mathbb{Z})$ *as originating from modular forms for the full group* $\mathrm{SL}_2(\mathbb{Z})$*, it is not generally true that all such modular forms can be obtained in this way. Those that can't be are (essentially) what are referred to as* newforms*.*

It often happens in applications that one wants to relax Definition 2.1 even further by requiring instead of (*ii*) the weaker condition that

$$f|_k\gamma = \varepsilon(\gamma)f$$

for a certain type of function $\varepsilon : \Gamma \to \mathbb{C}$ called a *multiplier system*. While we won't define this term in full generality, one should essentially think of a one-dimensional representation of Γ which (for convenience) should satisfy that there is some $N \in \mathbb{N}$ with $\varepsilon(\gamma)^N = 1$ for all $\gamma \in \Gamma$. The corresponding spaces are denoted by $M_k(\Gamma, \varepsilon)$ resp. $S_k(\Gamma, \varepsilon)$. For a quick overview on multiplier systems, the reader is referred to [Iwa, Sections 2.6 and 2.7]. A more detailed account on multiplier systems can be found in [Ran, Chapter 3].

Remark 2.4. *Note that if* $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \notin \Gamma$*, there might be modular forms of odd weight for* Γ *(but none for negative weight). If* $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) \in \Gamma$*, then* $M_k(\Gamma, \varepsilon)$ *can only be non-trivial if* $\varepsilon(\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)) = (-1)^k$*.*

One has the following important subgroups of $\mathrm{SL}_2(\mathbb{Z})$. In what follows, let $N \in \mathbb{N}$ be a positive integer.

$$\begin{aligned}
\Gamma_0(N) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod N \right\}, \\
\Gamma_1(N) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod N \right\}, \\
\Gamma(N) &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \ : \ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod N \right\}.
\end{aligned}$$

The group $\Gamma(N)$ is called the *principal congruence subgroup* of level N. We also give the indices of these groups here (see for instance [CS, Section 6.2]).

$$\begin{aligned}
[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(N)] &= N \prod_{p|N} \left(1 + \frac{1}{p}\right), \\
[\mathrm{SL}_2(\mathbb{Z}) : \Gamma_1(N)] &= N^2 \prod_{p|N} \left(1 - \frac{1}{p^2}\right), \\
[\mathrm{SL}_2(\mathbb{Z}) : \Gamma(N)] &= N^3 \prod_{p|N} \left(1 - \frac{1}{p^2}\right).
\end{aligned}$$

Note that most of the results for modular forms on the full modular group discussed in Section 1 have more or less direct analogues for modular forms for subgroups. For example, a generalization of the valence formula in Theorem 1.9 to arbitrary finite index subgroups can be found in [CS, Theorem 5.6.11], a general dimension formula is given in [CS, Theorem 5.6.18] and Hecke operators can also be considered for (congruence) subgroups, see for instance [CS, Chapter 10]. Some care must be taken though for Hecke operators of index dividing the level.

Sometimes one also wants to consider modular forms of half-integral weight. Essentially because the complex square-root function is inherently multi-valued, this cannot be achieved by simply allowing half-integer powers in the automorphy factors. Instead one is required to consider a different group under which these functions are invariant, the *metaplectic group*. One can however circumvent this using the following definition which goes back to Shimura [Shi].

Definition 2.5. *A function $f : \mathfrak{H} \to \mathbb{C}$ is called a modular form of weight $k \in \frac{1}{2} + \mathbb{Z}$ for $\Gamma_0(4N)$ if*

(i) f is holomorphic on $\mathfrak{H}$,

(ii) we have

$$(f|_k\gamma)(\tau) := \left(\frac{c}{d}\right)\left(\frac{-4}{d}\right)^k (c\tau+d)^{-k} f\left(\frac{a\tau+b}{c\tau+d}\right) = f(\tau)$$

for all $\gamma \in \Gamma_0(4N)$ and all $\tau \in \mathfrak{H}$, where here $\left(\frac{a}{b}\right)$ denotes the Jacobi symbol and we choose the principal branch of the square-root, which is positive for positive real arguments,

(iii) f is holomorphic at the cusps.

Note that by choosing an appropriate multiplier system it is possible to have different kinds of half-integer weight modular forms, as we see for instance in the context of the Dedekind eta function in Section 2.2.2.

Even though we have now relaxed the notion of modular form quite considerably, one still can obtain the following result (see [Iwa, Equation (2.68)]).

Theorem 2.6. *For any subgroup $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ of finite index (indeed for any so-called Fuchsian group of the first kind), multiplier system ε and weight $k \in \frac{1}{2}\mathbb{Z}$, the space $M_k(\Gamma, \varepsilon)$ is finite-dimensional, more precisely we have*

$$\dim_{\mathbb{C}} M_k(\Gamma, \varepsilon) \leq C(k+1)(\mathrm{vol}(\Gamma \setminus \mathfrak{H}) + 1),$$

for some constant $C > 0$, where $\mathrm{vol}(\Gamma\backslash\mathfrak{H})$ denotes the (hyperbolic) volume of a fundamental domain of Γ.

One of the important appeals of the theory of modular forms is that it is a source of sometimes surprising identities, which arise by comparing two modular forms. A first example of this is the Hurwitz identity in Remark 1.11. To do this in general, the following result often referred to as the *Sturm bound*, although it is in the presented form essentially already due to Hecke, is invaluable.

Theorem 2.7. *Let $f, g \in M_k(\Gamma)$ for some $\Gamma \leq \mathrm{SL}_2(\mathbb{Z})$ of finite index with Fourier expansions $f(\tau) = \sum_{n=0}^{\infty} \alpha_f(n) q^n$ and $g(\tau) = \sum_{n=0}^{\infty} \alpha_g(n)$. Then we have $f(\tau) = g(\tau)$ for all $\tau \in \mathfrak{H}$ if and only if*

$$\alpha_f(n) = \alpha_g(n) \text{ for all } n \leq \left(\left\lfloor \frac{k}{12} \right\rfloor + 1 \right) [\mathrm{SL}_2(\mathbb{Z}) : \Gamma].$$

Proof. See [Kil, Theorem 3.13]. □

Thus, comparing two modular forms boils down to checking finitely many of their Fourier coefficients, even though in many applications, the indices of the groups involved are so large that one has to do more or less clever tricks before it is feasible to actually do this in practice. It should also be remarked that Sturm's original paper [Stu] is actually about congruences between modular forms.

2.2 Further examples

In this section, we will discuss some further examples of sources of modular forms.

2.2.1 Theta series

In this section let $Q \in \mathbb{Z}_{>0}^{m \times m}$ be a positive definite integer matrix whose diagonal entries are even (one should think of the term *even quadratic form* here). Then we define the *theta series* of Q by

$$\Theta_Q(\tau) := \sum_{\ell \in \mathbb{Z}^m} q^{\ell^{tr} Q \ell / 2} = \sum_{n=0} r_Q(n) q^n,$$

where we set

$$r_Q(n) := \#\{\ell \in \mathbb{Z}^m \ : \ \ell^{tr} Q \ell = 2n\}.$$

These numbers $r_Q(n)$ have played an important role in number theory for centuries, dating back at least to Fermat, possibly even Diophantus, who asked (and partly answered) the question which numbers can be represented for instance as the sum of 2 or 4 squares. In those cases, Q would just be twice the 2×2 or 4×4 identity matrix, so they asked when $r_Q(n) \neq 0$. It is even more interesting to ask for an (elementary) formula for $r_Q(n)$. This can often be provided through the theory of modular forms by means of the following important theorem due to Schoeneberg [Sch].

Theorem 2.8. *The function* Θ_Q *is a modular form of weight* $m/2$ *for some* $\Gamma_0(N)$ *and some multiplier system which can be determined explicitly. In particular, if* $\det Q = 1$*, we have* $\Theta_Q \in M_{m/2}(\mathrm{SL}_2(\mathbb{Z}))$*.*

The basic idea of the proof is essentially to employ *Poisson summation* which in its simplest form states that for a function $f : \mathbb{R} \to \mathbb{R}$ with sufficiently rapid decay at ±∞ one has

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{m\in\mathbb{Z}} \widehat{f}(m),$$

where $\widehat{f}$ denotes the *Fourier transform* of f. As it turns out, the Fourier transform of $\exp(-\ell^{tr}Q\ell\pi v)$ is closely related to $\exp(-\ell^{tr}Q\ell\pi/v)$, from where Schoeneberg's theorem can be inferred. See for instance [Zag4, Proposition 9] for an account of the proof where Q has one variable. A full proof can be found in [KK, Abschnitt V.4.5].

Example 2.9. *It can be deduced from Theorem 2.8 that if* $\det Q = 1$*, then we must have* $8 \mid m$*, so the lowest dimension where the second part of that theorem applies is* $m = 8$*. Indeed, there is a matrix in* $\mathbb{Z}^{8\times 8}_{>0}$ *with even entries on the diagonal and determinant* 1*,*

$$\mathbb{E}_8 = \begin{pmatrix} 4 & -2 & 0 & 0 & 0 & 0 & 0 & 1 \\ -2 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

So by Schoeneberg's theorem, we know that $\Theta_{\mathbb{E}_8} \in M_4(\mathrm{SL}_2(\mathbb{Z}))$*, which implies immediately, because this space is one-dimensional, that* $\Theta_{\mathbb{E}_8} = E_4$*, so by comparing coefficients we find*

$$r_{\mathbb{E}_8}(n) = 240\sigma_3(n)$$

for all n*.*

2.2.2 The Dedekind eta function

One extremely important modular form that can be used to build a surprisingly large proportion of modular forms in general is the *Dedekind eta function.* It is defined by

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty} (1 - q^n).$$

Recall that we already saw that

$$\Delta(\tau) = q \prod_{n=1}^{\infty} (1-q^n)^{24} = \eta(\tau)^{24}.$$

So one may think of the eta function as the 24th root of Δ, which somewhat suggests that it should be a modular form of weight $12/24 = 1/2$ in some sense. Figuring this out in detail however is not so easy. One has the obvious transformation behaviour

$$\eta(\tau+1) = e^{\pi i/12}\eta(\tau)$$

under translation. It is a bit less obvious to see that

$$\eta(-1/\tau) = \sqrt{\frac{\tau}{i}}\eta(\tau),$$

where, as always we choose the principal branch of the square-root. So there is a reasonable transformation behaviour under both generators of $\mathrm{SL}_2(\mathbb{Z})$ which confirms that η is a modular form of weight 1/2 with some multiplier system. It is however not at all easy to figure out the general transformation behaviour of η under a general matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. This was first done by Dedekind who proved the following result.

Theorem 2.10. *For* $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ *with* $c > 0$ *we have that*

$$\eta\left(\frac{a\tau+b}{c\tau+d}\right) = \exp\left(\pi i \left(\frac{a+d}{12c} + s(-d,c) - \frac{1}{4}\right)\right)\sqrt{\frac{c\tau+d}{i}}\eta(\tau),$$

where for coprime integers h, k, $k > 0$, *we denote by*

$$s(h,k) = \sum_{r=1}^{k-1} \left(\frac{r}{k} - \frac{1}{2}\right)\left(\frac{rh}{k} - \left\lfloor \frac{rh}{k} \right\rfloor - \frac{1}{2}\right)$$

the Dedekind sum.

Proof. See [Apo, Chapter 3] $\square$

Even though this multiplier system seems rather complicated, it is relatively well-controllable. For example we can use this (in principle) to show the following version of Euler's famous Pentagonal Number Theorem.

Corollary 2.11. *We have*

$$\eta(24\tau) = \frac{1}{2}\sum_{n\in\mathbb{Z}} \chi_{12}(n) q^{n^2} \in S_{1/2}(576, \chi_{12}),$$

where $\chi_{12}(n) = \left(\frac{12}{n}\right)$.

One important feature of the eta function is that it can be used to build new examples of modular forms in a systematic manner.

Definition 2.12. *For $N \in \mathbb{N}$ and integers $r_d \in \mathbb{Z}$, $d \mid N$, we call the expression*

$$\prod_{d|N} \eta(d\tau)^{r_d}$$

an eta quotient *of level N.*

By investigating the arithmetic properties of Dedekind sums, one can give explicit criteria when an eta quotient of level N defines e.g. a modular form in $M_k(\Gamma_0(N))$. In some cases, the algebra of modular forms for $\Gamma_0(N)$, $M_*(\Gamma_0(N)) = \bigoplus_{k=0}^{\infty} M_k(\Gamma_0(N))$, is even generated by such eta quotients (see [RW]), which can be very useful for explicit computations.

2.2.3 Poincaré series

An important general way to construct modular forms in $M_k(\Gamma, \varepsilon)$ in a very general setting is through the method of *Poincaré series*, which we briefly want to mention in this section. As we shall see, these have played a central role in the theory of harmonic Maass forms as they can be used as convenient bases, and they are important in applications of the subject to asymptotic analysis of many interesting arithmetic sequences. The idea of Poincaré series is to average some function over the group Γ. More precisely, let $\varphi : \mathfrak{H} \to \mathbb{C}$ be a holomorphic function invariant under the group $\Gamma_\infty = \mathrm{Stab}_\Gamma(\infty)$. In the case of $\Gamma = \Gamma_0(N)$ for instance, one has $\Gamma_\infty = \langle \pm T \rangle$, so that here, φ should be one-periodic. With this, we formally define the series

$$\mathbb{P}(\tau) := \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \varepsilon(\gamma)(\varphi|_k \gamma)(\tau).$$

It is easy to see that $\mathbb{P}$ is holomorphic on $\mathfrak{H}$ and transforms like a modular form in $M_k(\Gamma, \varepsilon)$, provided that it converges absolutely and locally uniformly. If φ has moderate growth approaching 0, this is the case as soon as the weight k is sufficiently large. For example, choosing $\Gamma = \Gamma_0(N)$, $\varepsilon = \chi$ a Dirichlet character modulo N, and $\phi(\tau) = e^{2\pi i m\tau}$ for a positive integer m, the Poincaré series $\mathbb{P}$ above converges for $k > 2$ and by varying m, one obtains a basis for the full space $S_k(\Gamma_0(N), \chi)$ of cusp forms (see e.g. [Iwa, Corollary 3.5]). Of course, since there are infintely many such Poincaré series (one for each m), there must be linear relations among them. However, a basis consisting of Poincaré series is usually not very useful when investigating arithmetic properties of Fourier coefficients since the Fourier coefficients of Poincaré series generically have (at least conjecturally) transcendental Fourier coefficients.

In many cases, one can however give explicit formulas for these Fourier coefficients, often in terms of infinite sums of so-called *Kloosterman sums* and *Bessel functions*, but we refrain from giving these expressions here. They can be found for example in [Iwa, Section 3.2]. The interested reader should note that an early example of such an expansion, which we now recognize as an example of the theory of Poincaré series, is the Rademacher expansion of the partition function $p(n)$, which build on the work of Hardy and Ramanujan using the Circle Method. This idea goes back to Petersson [Pet], but has only received greater interest during the past few years. For a more modern treatment of this idea, specifically applied to the partition function, we refer the reader to [Pri].

2.3 Jacobi forms

Jacobi forms are in a sense an amalgam of elliptic functions (i.e. doubly-periodic functions) and modular forms. They have deep connections to many important types of modular forms. Even though there are examples dating back to Jacobi (hence the name), their systematic study originated from certain *Siegel modular forms*, which are a certain kind of multivariable modular forms. Their study and the perspective that follows in these notes was originally laid out systematically by Eichler and Zagier [EZ].

Definition 2.13. *A function* $\phi\colon \mathbb{C}\times\mathfrak{H}\to\mathbb{C}$ *is called a* Jacobi form *of weight* k *and index* m *if*

(i) ϕ *is holomorphic on* $\mathbb{C}\times\mathfrak{H}$*,*

(ii) we have

$$\phi(z+\lambda\tau+\mu,\tau)=e^{-2\pi i m(\lambda^2\tau+2\lambda z)}\phi(z,\tau)$$

for all $\lambda,\mu\in\mathbb{Z}$*,* $z\in\mathbb{C}$*,* $\tau\in\mathfrak{H}$ *(*elliptic transformation law*),*

(iii) we have

$$\phi\left(\frac{z}{c\tau+d},\frac{a\tau+b}{c\tau+d}\right)=(c\tau+d)^k e^{2\pi i m\frac{cz^2}{c\tau+d}}\phi(z,\tau)$$

for all $\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)\in\mathrm{SL}_2(\mathbb{Z})$*,* $z\in\mathbb{C}$*,* $\tau\in\mathfrak{H}$ *(*modular transformation law*),*

(iv) ϕ *has a Fourier expansion*

$$\phi(z,\tau)=\sum_{n,r\in\mathbb{Z}}c(n,r)q^n\zeta^r,$$

where $\zeta=e^{2\pi i z}$ *with* $c(n,r)=0$ *whenever* $n<r^2/4m$*.*

The space of Jacobi forms of weight k *and index* m *is denoted by* $J_{k,m}$*.*

The prototypical examples of Jacobi forms are the *Jacobi theta functions*

$$\theta_{m,\ell}(z,\tau)=\sum_{\substack{r\in\mathbb{Z}\\ r\equiv\ell\ (2m)}}q^{r^2/4m}\zeta^r,$$

which are Jacobi forms of weight $1/2$ and index m for some subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

Here, we introduce the three most direct ways in which a Jacobi form can capture modular forms. The first such way is through evaluation at so-called *torsion points.*

Theorem 2.14. *Let $\phi \in J_{k,m}$ and $\alpha, \beta \in \mathbb{Q}$. Then $e^{2\pi i m \lambda^2 \tau}\phi(\alpha\tau + \beta, \tau)$ is a modular form of weight k for some explicitly known congruence subgroup depending on α, β and m.*

Proof. See [EZ, Theorem 1.3]. ☐

In addition to its Fourier expansion, a Jacobi form has at least two other important expansions. The first one is afforded by the following theorem.

Theorem 2.15. *Let $\phi \in J_{k,m}$ be a Jacobi form. Then we can write*

$$\phi(z,\tau) = \sum_{\ell \pmod{2m}} h_\ell(\tau)\theta_{m,\ell}(z,\tau),$$

which we call the theta expansion *of ϕ. The functions h_ℓ are then modular forms of weight $k - 1/2$ for some specific subgroup of $\mathrm{SL}_2(\mathbb{Z})$.*

Proof. See [EZ, Theorem 5.1]. ☐

Remark 2.16. *If one considers the vector $\vec{h}(\tau) = (h_0(\tau), ..., h_{2m-1}(\tau))$, one finds that this vector transforms indeed like a modular form of weight $k - 1/2$ for the full modular group $\mathrm{SL}_2(\mathbb{Z})$ for a certain representation*[3]*, called the* Weil representation, *meaning that there is a map $\rho : \mathrm{SL}_2(\mathbb{Z}) \to \mathrm{GL}_{2m}(\mathbb{C})$, such that*

$$\vec{h}\left(\frac{a\tau + b}{c\tau + d}\right) = \rho\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right)\left(\frac{c\tau + d}{i}\right)^{k-1/2}\vec{h}(\tau).$$

The third important expansion of a Jacobi form that yields modular forms is its Taylor expansion in $z = 0$.

Theorem 2.17. *Define the space of* weak *Jacobi forms $\widetilde{J}_{k,m}$ by relaxing condition (iv) in Definition 2.13 to $c(n,r) = 0$ whenever $n < 0$. Then the Taylor coefficients of $\phi \in \widetilde{J}_{k,m}$ in $z = 0$ are essentially modular forms of even weight for $\mathrm{SL}_2(\mathbb{Z})$. More precisely, there is an explicit isomorphism*

$$\widetilde{J}_{k,m} \to \begin{cases} M_k \oplus M_{k+2} \oplus ... \oplus M_{k+2m} & k \text{ even,} \\ M_{k+1} \oplus M_{k+3} \oplus ... \oplus M_{k+2m-3} & k \text{ odd.} \end{cases}$$

---

[3] In the case of half-integral weight, it is actually only a projective representation of $\mathrm{SL}_2(\mathbb{Z})$, i.e. a representation for a certain covering group called the *metaplectic* group $\mathrm{Mp}_2(\mathbb{Z})$.

For even weight, the isomorphism is given by

$$\phi(z,\tau) \mapsto \left(\sum_{\ell\,(2m)} [h_\ell(\tau), \theta_{m,\ell}(0,\tau)]_\nu \right)_{\nu=0}^{m},$$

where $[\bullet,\bullet]_\nu$ *denotes the* ν*th* Rankin-Cohen bracket*, a bilinear differential operator on modular forms that acts essentially like a product [Coh]. A similar formula holds also in the case of odd weight.*

Proof. See [EZ, Theorem 9.2]. □

2.4 Singular moduli

This sections follows closely the exposition on the subject in Section 6 of [Zag4].

The subject of singular moduli is one of the most important in the arithmetic theory of modular forms. Let us motivate it by the following observation that

$$e^{\pi\sqrt{163}} = 262537412640768743.99999999999250072...$$

is surprisingly close to an integer. Here we would like to briefly outline an explanation for this phenomenon.

For this we look at *Klein's modular invariant*

$$j(\tau) := \frac{E_4(\tau)^3}{\Delta(\tau)} = q^{-1} + 744 + 196\,884q + 21\,493\,760q^2 + O(q^3).$$

This clearly transforms like a modular form of weight 0, so we have indeed

$$j\left(\frac{a\tau+b}{c\tau+d}\right) = j(\tau),$$

but it has a pole at infinity and is therefore not a modular form in our sense, but rather a *modular function*, i.e. a meromorphic function transforming like a modular form of weight 0. The function j is one of the earliest, and perhaps the most important, examples of a modular function, as the following theorem illustrates.

Theorem 2.18. *(i) Every modular function for* $\mathrm{SL}_2(\mathbb{Z})$ *is a rational function in* j*.*

(ii) Every modular function for $\mathrm{SL}_2(\mathbb{Z})$ *which is holomorphic in* $\mathfrak{H}$ *(but may have a pole at* ∞*) is a polynomial in* j*.*

(iii) Every modular function for any finite index subgroup of $\mathrm{SL}_2(\mathbb{Z})$ *is an algebraic function in* j*.*

Proof. Claim (i) is proven for instance in [CS, Theorem 5.7.3] (ii) is then an easy consequence of the valence formula. The proof of (iii) is more involved. $\square$

One can compute several special values of j, e.g.

$$j(\rho) = 0, \quad j(i) = 1728, \quad j(i\sqrt{2}) = 8000, \quad j\left(\frac{1+i\sqrt{7}}{2}\right) = -3375,$$

$$j\left(\frac{1+i\sqrt{15}}{2}\right) = -\frac{191025 + 85995\sqrt{5}}{2}, \ldots$$

It is a rather striking phenomenon, when a transcendental function evaluated at algebraic arguments again gives algebraic values. In this case, the reason lies in the following theorem.

Theorem 2.19. *For every $m \in \mathbb{N}$ there is a polynomial $\Psi_m(X,Y) \in \mathbb{Z}[X,Y]$ of degree $\sigma_1(m)$ in both variables such that*

$$\Psi_m(j(M.\tau), j(\tau)) \equiv 0$$

for all $M \in \mathcal{M}_m$ with $\mathcal{M}_m$ as in Section 1.3.

Sketch of proof. Consider

$$\prod_{M \in \mathrm{SL}_2(\mathbb{Z}):\mathcal{M}_m} (X - j(M.\tau)) =: \Psi_m(X, j(\tau)).$$

Then this is well-defined because $j(\gamma.\tau) = j(\tau)$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ and we have

$$\Psi_m(X, j(\gamma.\tau)) = \Psi_m(X, j(\tau))$$

because $\mathcal{M}_m \cdot \gamma = \mathcal{M}_m$ (recall that we used a similar argument in the context of Hecke operators). This means that the coefficient of X^r in $\Psi_m(X, j(\tau))$ is a modular function in τ for $\mathrm{SL}_2(\mathbb{Z})$ which has no poles in $\mathfrak{H}$ (because none of the $j(M.\tau)$ do), so each coefficient is a polynomial in $j(\tau)$, which yields the definition of a polynomial $\Psi_m(X,Y) \in \mathbb{C}[X,Y]$. For the sake of brevity, we omit the proof that $\Psi_m(X,Y)$ has indeed integral coefficients and the same degree in both variables. $\square$

This yields the following theorem.

Theorem 2.20. *Let $\mathfrak{z} \in \mathfrak{H}$ be a* CM point*, i.e. there are $A, B, C \in \mathbb{Z}$ such that $A\mathfrak{z}^2 + B\mathfrak{z} + C = 0$. Then the* singular modulus *$j(\mathfrak{z})$ is an algebraic integer.*

Sketch of proof. The matrix $M = \left(\begin{smallmatrix} B & C \\ -A & 0 \end{smallmatrix}\right)$ satisfies $M.\mathfrak{z} = \mathfrak{z}$ as one immediately checks. Since $\det M = AC$, we have

$$\Psi_{AC}(j(M.\mathfrak{z}), j(\mathfrak{z})) = \Psi_{AC}(j(\mathfrak{z}), j(\mathfrak{z})) = 0.$$

$\square$

Remark 2.21. *One should say about the above proof sketch that it is not necessarily clear that* $\Psi_m(X,X)$ *is not just the zero polynomial (in fact this can happen). But since one can redefine* $\Psi_m(X,Y)$ *by dividing out all factors* $(X-Y)$ *without changing the important properties in Theorem 2.19, this can be reconciled.*

Remark 2.22. *One can in fact show, e.g. by using (iii) of Theorem 2.18, that* $f(\mathfrak{z})$ *is an algebraic number (not necessarily integral) for every modular function* f *for any finite index subgroup of* $\mathrm{SL}_2(\mathbb{Z})$ *with algebraic Fourier coefficients. This explains essetially the remarkable identity*

$$\cfrac{e^{2\pi/5}}{1+\cfrac{e^{-2\pi}}{1+\cfrac{e^{-4\pi}}{1+\ddots}}} = \sqrt{\frac{5+\sqrt{5}}{2}} - \frac{1+\sqrt{5}}{2}$$

that Ramanujan recorded in his first letter to Hardy, because the left-hand side can be identified as the value at $\mathfrak{z} = i$ *of a modular function for the group* $\Gamma(5)$*. This function is closely related to the famous Rogers-Ramanujan identities.*

The values $j(\mathfrak{z})$ for CM points $\mathfrak{z}$ play an important role in class field theory, as we shall explain now. Let $Q(X,Y) = AX^2 + BXY + CY^2 =: [A,B,C] \in \mathbb{Z}[X,Y]$ be a binary quadratic form of discriminant $D = B^2 - 4AC < 0$ and $A > 0$ and define $\mathfrak{z}_Q \in \mathfrak{H}$ by $Q(\mathfrak{z}_Q, 1) = 0$. Also define $\mathcal{Q}_D = \{[A,B,C] : B^2 - 4AC = D\}$ as the set of quadratic forms of discriminant D.

Theorem 2.23. *Let* $D < 0$ *be a* fundamental discriminant*, i.e. the discriminant of a quadratic field, and set*

$$H_D(X) = \prod_{Q \in \mathrm{SL}_2(\mathbb{Z}) \backslash \mathcal{Q}_D} (X - j(\mathfrak{z}_Q)).$$

The $H_D(X) \in \mathbb{Z}[X]$ *and* $\deg H_D(X) = h(D)$*, the* class number *of* D*. The splitting field of* $H_D(X)$ *is the so-called* Hilbert class field *of* $\mathbb{Q}(\sqrt{D})$*, i.e. a Galois extension of* $\mathbb{Q}(\sqrt{D})$ *whose Galois group is isomorphic to the ideal class group of* $\mathbb{Q}(\sqrt{D})$*.*

Proof. See [Cox, §11]. ☐

3 Applications to partitions

In this section, we explore some applications of the theory of modular forms to *partitions*. It is well-known that the generating function of the *partition function*

$$p(n) := \#\{(\lambda_1, ..., \lambda_\ell) : \lambda_1 \geq ... \geq \lambda_\ell > 0, \lambda_1 + ... + \lambda_\ell = n\}$$

has a very simple description as an infinite product,

$$P(q) = \sum_{n=0}^{\infty} p(n)q^n = \prod_{n=1}^{\infty} (1-q^n)^{-1} = \frac{q^{1/24}}{\eta(\tau)}.$$

Hence, $P(q)$ is essentially a weakly holomorphic modular form (in particular, it has poles at cusps) of weight $-1/2$ with some multiplier system, namely the inverse of that of the Dedekind eta function (cf. Theorem 2.10). We want to exploit this fact to say something about the numbers $p(n)$.

3.1 Asymptotics

This part of the exposition is essentially a short summary of Chapter 5 of [Apo]. The procedure described here, pioneered by Hardy and Ramanujan and applicable to the study of many important number-theoretic phenomena even beyond the realm of modular forms, is called the *Cirlce Method*.

The basic idea is simple. By Cauchy's Theorem, we have that

$$p(n) = \frac{1}{2\pi i} \int_C \frac{P(q)}{q^{n+1}} dq,$$

where C is a simple, closed contour inside the unit disk looping around $q = 0$ exactly once. From the product expansion we can see that $P(q)$ has a singularity whenever q approaches a root of unity. The idea now is to choose a special contour C and let it approach the unit circle form inside in a certain way to be able to replace $P(q)$ in Cauchy's Theorem above by a more elementary function plus some (small) error. For this we divide C into parts $C_{h,k}$ for coprime h, k with $k \leq N$ for some previously chosen $N \in \mathbb{N}$, which are close to the roots of unity $e^{2\pi i h/k}$ of order up to N,

$$p(n) = \frac{1}{2\pi i} \int_C \frac{P(q)}{q^{n+1}} dq = \sum_{k=1}^{N} \sum_{h\ (k)^*} \int_{C_{h,k}} \frac{P(q)}{q^{n+1}} dq = \sum_{k=1}^{N} \sum_{h\ (k)^*} \int_{C_{h,k}} \frac{\psi_{h,k}(q)}{q^{n+1}} dq + \text{“Error”},$$

where here and from now on, $\sum_{a\ (b)^*}$ means a sum over all $a = 0, ..., b-1$ coprime to b.

In order to do this, we require the following lemma which is a convenient reformulation of Dedekind's Theorem 2.10.

Lemma 3.1. *Let* $q = \exp\left(\frac{2\pi i h}{k} - \frac{2\pi z}{k^2}\right)$, $q' = \exp\left(\frac{2\pi i H}{k} - \frac{2\pi}{z}\right)$ *with* $\mathrm{Re}(z) > 0$, $\gcd(h,k) = 1$ *and* $hH \equiv -1 \pmod{k}$. *Further define* $\omega(h,k) = e^{\pi i s(h,k)}$ *and*

$$\Psi_k(z) = z^{1/2} \exp\left(\frac{\pi}{12z} - \frac{\pi z}{12k^2}\right).$$

Then we have

$$P(q) = \omega(h,k) \frac{\Psi_k(z)}{k^{1/2}} P(q').$$

Next, following Rademacher, we choose a special path of integration. Suppose C is a circular contour inside the unit disk (in the q-plane). Then this essentially corresponds to a line-integral from i to $i+1$ in the τ-plane. Now we replace this line integral by an integral along the upper arcs of so-called *Ford circles* (see Figure 2, the picture was again generated using SAGE version 8.1 [Sage]). Denote this so-called Rademacher path by $\mathcal{P}(N)$.

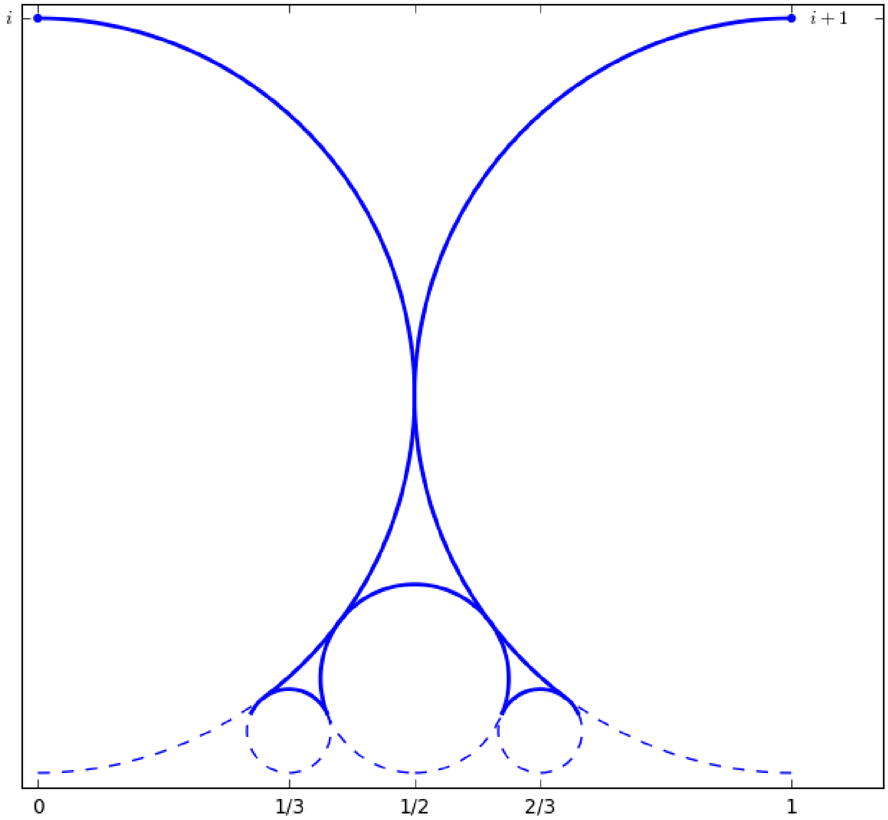

Figure 2: Rademacher path $\mathcal{P}(3)$.

By making an appropriate change of variables we can write

$$\begin{aligned} p(n) &= \frac{1}{2\pi i}\int_C \frac{P(q)}{q^{n+1}}dq \\ &= \int_i^{i+1} P(e^{2\pi i\tau})e^{-2\pi in\tau}d\tau \\ &= \int_{\mathcal{P}(N)} P(e^{2\pi i\tau})e^{-2\pi in\tau}d\tau \end{aligned}$$

$$= \sum_{k=1}^{N} \sum_{h\ (k)^*} ik^{-2} e^{-2\pi inh/k} \int_{z_1(h,k)}^{z_2(h,k)} e^{2\pi nz/k^2} P\left(\exp\left(\frac{2\pi ih}{k} - \frac{2\pi z}{k^2}\right)\right) dz,$$

where $z_1(h,k)$ and $z_2(h,k)$ are the images of points where the Ford circles in the path $\mathcal{P}(N)$ touch under the change of variables $\tau = \frac{h}{k} + \frac{iz}{k^2}$ which occurred in the last step. The contour of integration becomes the arc of a circle around 1/2 of radius 1/2 joining these points. Using Lemma 3.1, one can then write

$$p(n) = \sum_{h,k} ik^{-5/2} \omega(h,k) e^{-2\pi inh/k} (I_1(h,k) + I_2(h,k)),$$

where $\sum_{h,k}$ for the double sum over h and k and we define

$$I_1(h,k) = \int_{z_1(h,k)}^{z_2(h,k)} \Psi_k(z) e^{2\pi hz/k^2} dz$$

and

$$I_2(h,k) = \int_{z_1(h,k)}^{z_2(h,k)} \Psi_k(z) \left[P\left(\exp\left(\frac{2\pi iH}{k} - \frac{2\pi}{z}\right)\right) - 1\right] e^{2\pi nz/k^2} dz.$$

One can show now that $|I_2(h,k)| \leq Ck^{3/2}N^{-3/2}$ for some constant $C > 0$, whence

$$\left| \sum_{h,k} ik^{-5/2} \omega(h,k) e^{-2\pi inh/k} I_2(h,k) \right| \leq CN^{-1/2}.$$

Therefore we find that

$$p(n) = \sum_{h,k} ik^{-5/2} \omega(h,k) e^{-2\pi inh/k} I_1(h,k) + O(N^{-1/2}).$$

Now letting $N \to \infty$ and evaluating the integral $I_1(h,k)$ explicitly one arrives at the following important theorem due to Rademacher [Rad].

Theorem 3.2. *For $n \in \mathbb{N}$, we have the following formula for $p(n)$,*

$$p(n) = \frac{1}{\pi\sqrt{2}} \sum_{k=1}^{\infty} A_k(n) k^{1/2} \frac{d}{dn} \left(\frac{\sinh\left(\frac{\pi}{k}\sqrt{\frac{2}{3}\left(n - \frac{1}{24}\right)}\right)}{\sqrt{n - \frac{1}{24}}} \right)$$

with

$$A_k(n) = \sum_{h\ (k)^*} e^{\pi is(h,k) - 2\pi inh/k}.$$

Essentially by specializing to the term $k = 1$ in the above series for $p(n)$, one recovers an older result due to Hardy and Ramanujan.

Theorem 3.3. *As $n \to \infty$, we have the asymptotic equality*

$$p(n) \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{2n/3}}.$$

Remark 3.4. *Hardy and Ramanujan in fact had a full asymptotic expansion for $p(n)$, using essentially the argument outlined here, but their expansion does not give a convergent series. Some 20 years after their work was published, Rademacher realized that the proof by Hardy-Ramanujan could be modified to yield the convergent series expression in Theorem 3.2.*

It should also be pointed out that there are many much easier ways to obtain just the main term of the asymptotic expansion given in Section 3.1, one very important one being to use a so-called Tauberian theorem due to Ingham [Ing].

3.2 Congruences

This section is essentially a short summary of parts of Chapter 5 of [Ono2].

Many people know the famous *Ramanujan congruences*,

$$\begin{aligned} p(5n+4) &\equiv 0 \pmod{5}, \\ p(7n+5) &\equiv 0 \pmod{7}, \\ p(11n+6) &\equiv 0 \pmod{11}. \end{aligned}$$

Over the last 100 years, many proofs of these congruences, both using formal or q-series, but also modular forms techniques (see for instance [Ber] for a rather elementary, but modular forms related proof), have been published, also generalizations for congruences modulo powers of 5, 7, and 11 have been found (some of which have been conjectured by Ramanujan). For a statement of these and further references, see [Ono2, Theorem 5.7].

In this section, we want to construct further congruences of the form

$$p(An+B) \equiv 0 \pmod{M}.$$

In fact, for any M coprime to 6, there are infinitely many non-nested arithmetic sequences $An+B$, such that such a congruence holds. Here are some examples modulo larger primes,

$$\begin{aligned} p(17 \cdot 41^4 n + 1122838) &\equiv 0 \pmod{17}, \\ p(19 \cdot 101^4 n + 815655) &\equiv 0 \pmod{19}, \\ p(23 \cdot 5^4 n + 3474) &\equiv 0 \pmod{23}, \\ &\vdots \end{aligned}$$

In this last section, we would like to explain the outline the proof of the following theorem due to Ahlgren and Ono [Ahl, AO, Ono1] underlying these and in fact all (known) congruences for $p(n)$. For this we define for a prime $\ell \geq 5$ the set

$$\mathcal{S}_\ell := \left\{ \beta \in \{0, ..., \ell - 1\} \ : \ \left(\frac{\beta - \delta_\ell}{\ell} \right) \in \{0, \varepsilon_\ell\} \right\}$$

where $\delta_\ell := \frac{\ell^2 - 1}{24}$ and $\varepsilon_\ell := \left(\frac{-6}{\ell} \right)$.

Theorem 3.5. *Let $\ell \geq 5$ prime, $m \in \mathbb{N}$, $\beta \in \mathcal{S}_\ell$, then a positive proportion of primes $Q \equiv -1 \pmod{24\ell}$ satisfy*

$$p\left(\frac{Q^3 n + 1}{24} \right) \equiv 0 \pmod{\ell^m}$$

for all $n \equiv 1 - 24\beta \pmod{24\ell}$ and $\gcd(Q, n) = 1$.

The proof of this result hinges on the following proposition.

Proposition 3.6. *Given a prime $\ell \geq 5$, $m \in \mathbb{N}$, $\beta \in \mathcal{S}_\ell$, there exists $\lambda_{\ell,m} \in \mathbb{Z}$ and $F_{\ell,m,\beta} \in S_{\lambda_{\ell,m}+1/2}(\Gamma_0(576\ell^5)) \cap \mathbb{Z}[[q]]$ such that*

$$F_{\ell,m,\beta} \equiv \sum_{n=0}^{\infty} p(\ell n + \beta) q^{24\ell n + 24\ell\beta - 1} \pmod{\ell^m}.$$

Sketch of proof. The function

$$E_{\ell,t}(\tau) := \frac{\eta(\tau)^{\ell^t}}{\eta(\ell^t \tau)}$$

is a modular form of weight $(\ell^t - 1)/2$ for the group $\Gamma_0(\ell^t)$ with multiplier system $\chi_{\ell,t} = \left(\frac{(-1)^{(\ell^t - 1)/2} \ell^t}{\cdot} \right)$, which follows from Dedekind's Theorem 2.10. It is almost a cusp form in the sense that it vanishes at all cusps except ∞, i.e. $(E_{\ell,t}|\gamma)(iv) \to 0$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z}) \setminus \Gamma_0(\ell^t)$ and we have $E_{\ell,t}^{\ell^{m-1}} \equiv 1 \pmod{\ell^m}$ for all $m \in \mathbb{N}$.

Next we define

$$f_\ell(\tau) := \frac{\eta(\ell\tau)^\ell}{\eta(\tau)} =: \sum_{n=0}^{\infty} a_\ell(n) q^n,$$

which defines a modular form in $M_{(\ell-1)/2}(\Gamma_0(\ell), \left(\frac{\cdot}{\ell} \right))$. We can also write this as

$$f_\ell(\tau) = \sum_{n=0}^{\infty} p(n) q^{n + \delta_\ell} \prod_{n=1}^{\infty} (1 - q^{\ell n})^\ell.$$

We require the following twisted version of f_ℓ,

$$\tilde{f}_\ell(\tau) := \sum_{n=1}^{\infty} \left(1 - \varepsilon_\ell \left(\frac{n}{\ell}\right)\right) a_\ell(n) q^n,$$

which through standard theory of modular operators as outlined in Section 1.3 defines a modular form in $M_{(\ell-1)/2}(\Gamma_0(\ell^3), (\frac{\cdot}{\ell}))$. Since $\tilde{f}_\ell$ vanishes at the cusp ∞ and the function $E_{\ell,t}$ defined above vanishes at all other cusps, the function

$$f_{\ell,m'}(\tau) := E_{\ell,t}^{\ell^{m'}}(\tau)\tilde{f}_\ell(\tau)$$

is in fact a cusp form on $\Gamma_0(\ell^3)$ with multiplier system $\chi_{\ell,t}\left(\frac{\cdot}{\ell}\right)$ of weight $k = m'(\ell^t - 1)/2 + (\ell - 1)/2$, provided that m' is sufficiently large. We also have

$$f_{\ell,m'} \equiv \tilde{f}_\ell \pmod{\ell^{m'+1}}$$

since $E_{\ell,t} \equiv 1 \pmod{\ell^m}$ and $\operatorname{ord}_\infty(f_{\ell,m'}) \geq \delta_\ell + 1$ because $\operatorname{ord}_\infty(f_\ell) \geq \delta_\ell$ and by construction, the leading term in $f_\ell(\tau)$ disappears in $\tilde{F}_\ell(\tau)$. Thus the function

$$F_{\ell,m'}(\tau) := \frac{f_{\ell,m'}(\tau)}{\eta(\ell\tau)^\ell}$$

vanishes at infinity and hence is a cusp form if m' is sufficiently large. Now we observe that

$$\tilde{F}_{\ell,m'}(\tau) \equiv \sum_{n \equiv 0\ (\ell)} p(n - \delta_\ell) q^{n - \ell^2/24} + 2 \sum_{\left(\frac{n}{\ell}\right) = -\varepsilon_\ell} p(n - \delta_\ell) q^{n - \ell^2/24} \pmod{\ell^m}.$$

Again through application of standard operators, we find that $\tilde{F}_{\ell,m'}(24\tau) \in S_{k'}(\Gamma_0(576\ell^3))$ for m' sufficiently large for suitable k'. The result now follows by applying an appropriate sieve operator. □

To complete the proof of Theorem 3.5, we need the following general lemma on Hecke operators for half-integer weight modular forms. For prime index p coprime to $4N$, one can define the Hecke operator T_{p^2} acting on $S_{\lambda+1/2}(\Gamma_0(4N))$ in terms of the operators defined in Section 1.3 by

$$f|T_{p^2} := f|U_{p^2} + \left(\frac{(-1)^\lambda}{p}\right) p^{\lambda-1} f \otimes \left(\frac{\cdot}{p}\right) + p^{2\lambda-1} f|V_{p^2}.$$

Lemma 3.7. *Let $N \in \mathbb{N}$, $\lambda \in \mathbb{N}$ and let $f \in S_{\lambda+1/2}(\Gamma_0(4N))$ with integer Fourier coefficients. Then for any $M \in \mathbb{N}$, a positive proportion of primes $p \equiv -1 \pmod{4MN}$ satisfies*

$$f|T_{p^2} \equiv 0 \pmod{0} \pmod{M}.$$

If we now apply the above lemma to the cusp form $F_{\ell,m,\beta}$ from Proposition 3.6, we find that for a positive proportion of primes $Q \equiv -1 \pmod{24\ell}$ satisfies

$$F_{\ell,m,\beta}|T_{Q^2} \equiv 0 \pmod{\ell^m}.$$

But

$$F_{\ell,m,\beta}(\tau) \equiv \sum_{n=0}^{\infty} a_{\ell,m,\beta}(n)q^n \equiv \sum_{n\equiv 24\beta-1\ (24\ell)} p((n+1)/24)q^n \pmod{\ell^m}.$$

By looking at the Qnth coefficient of $F|T_{Q^2}$, one finds that

$$0 \equiv a_{\ell,m,\beta}(Q^3n) \equiv p\left(\frac{Q^3n+1}{24}\right) \pmod{\ell^m},$$

which is the claim from Theorem 3.5.

4 History and motivation of Harmonic Maass Forms

Having discussed a few of the basic facets and examples of the theory of modular forms, we now switch gears to talk about their generalization to the space of harmonic Maass forms. There are numerous beautiful examples of combinatorial generating functions which are both q-hypergeometric series and modular forms. As a fundamental example, consider the partition function $p(n)$. As we have seen above, the generating function $P(q) := \sum_n p(n)q^n$ is essentially a modular form of weight $-1/2$, and this can be used to prove congruences and even exact formulas for the numbers $p(n)$. We saw this by using the classical product formula for P. There is another formula for P, which arises from a counting of partitions while keeping track of sizes of Durfee squares in the Ferrers diagrams. Namely,

$$P(q) = \sum_{n\geq 0} \frac{q^{n^2}}{(q)_n^2},$$

where $(a;q)_n = (a)_n := \prod_{j=0}^{n-1}(1-aq^j)$. This form is beautiful and convenient, but hard to deduce modularity from. Being able to do so would be incredibly useful, and is a problem frequently encountered in this area, either explicitly or implicitly. This was summarized neatly in the following problem, popularized by Andrews [And3].

Problem (Andrews)**.** *How can one prove the modularity properties of $P(q)$ directly from the q-hypergeometric expression above, i.e., without first proving a $\sum = \prod$ identity?*

This problem is hard (in particular, unsolved). However, while modularity is difficult to pin down exactly on the level of q-hypergeometric series, modular forms do leave some fingerprints behind. This can be seen via the following elementary lemma, which follows directly from the transformations in the definition of a modular form. This version of the lemma and its proof can be found in Vlasenko and Zwegers' work [VZ] on Nahm's conjecture.

Lemma 4.1. *Suppose $f \in M_k(\Gamma)$ where $\Gamma \subseteq \mathrm{SL}_2(\mathbb{Z})$ is a subgroup of finite index. Then there are constants a, b such that as $\varepsilon \searrow 0$,*

$$e^{a/\varepsilon} f\left(\frac{i\varepsilon}{2\pi}\right) \sim be^{-k} + o(\varepsilon^N)$$

for all $N \geq 0$. That is, there is only one term in the asymptotic expansion **to all orders***.*

The point is that we are letting the modular variable τ tend to 0 from above along the imaginary axis, or letting q tend to 1 radially from within the unit disk, and this can be related to the behavior at $i\infty$ by modularity. But this is easy to determine using the Fourier expansion.

However, among q-hypergeometric series, this type of asymptotic expansion is very rare. Almost all q-hypergeometric series have complicated expansions, usually of an infinite form, which are very far from this elementary asymptotic behavior. In fact, in his original letter where Ramanujan describes his mock theta functions, he begins by giving a "generic" example with an infinite type expansion which doesn't look related to modular forms.

Example 4.2. *A famous conjecture of Nahm claims that certain q-hypergeometric series are modular if and only if certain elements of the so-called Bloch group (an object from K-theory) are torsion. The first case of this conjecture claims to predict exactly those $A, B, C \in \mathbb{Q}$ for which*

$$F_{A,B,C}(q) := \sum_{n \geq 0} \frac{q^{\frac{An^2}{2}+Bn+C}}{(q)_n}$$

is modular. Zagier solved this [Zag3] by proving that $F_{A,B,C}$ is modular for exactly 7 (explicitly given) triples $(A, B, C) \in \mathbb{Q}^3$. For example, these examples include the famous Rogers-Ramanujan function:

$$\sum_n \frac{q^{n^2}}{(q)_n} = \frac{1}{(q;q^5)_\infty (q^4;q^5)_\infty},$$

where the right-hand "product-side" can be directly shown to be modular. Zagier's proof works essentially by combining the type of reasoning encoded in the elementary lemma above with a brilliant method for computing the asymptotic expansions near $q = 1$ of each of the functions $F_{A,B,C}$, he reduced the candidates for modularity to finitely many cases. In each of these cases, modularity is established by applying a $\sum = \prod$ identity, such as that for the Rogers-Ramanujan function above.

5 A surprising discovery

Ramanujan was of course well aware of examples such as the Rogers-Ramanujan identities. Every example of such an identity, and correspondingly of a q-hypergeometric modular form, is highly interesting. As mentioned above, he computed (at least one) examples of q-hypergeometric asymptotic expansions near roots of unity like $q = 1$. Shortly before his untimely death, while isolated in India, he wrote a brief letter to Hardy describing a startling set of functions he discovered which he knew are not quite modular (in his language, are not "theta functions," or quotients, products, and linear combinations thereof), but which are so uncannily close that something must be very special about them. For instance, one of his functions was the series

$$f(q) := \sum_n \frac{q^{n^2}}{(-q)_n^2},$$

which looks similar to the expression for P above, but with a minus sign inserted. In particular, it doesn't have an asymptotic expansion as in Lemma 4.1 (which can be modified to approach other roots of unity, or, in τ, for τ approaching rational numbers vertically from above), but it doesn't have the kind of infinite, wild expansion which is typical for non-modular series. In particular, Ramanujan considered the (up to a rational power of q) weight 1/2 weakly holomorphic modular form

$$b(q) := (q; q^2)_\infty \sum_{n\in\mathbb{Z}} (-1)^n q^{n^2}.$$

Ramanujan then described a shocking numerical near cancellation between these two functions near $q = -1$, where they both explode. For example, we can compute with the help of a modern computer that $f(-0.998) \approx -6 \cdot 10^{90}$. However,

$$f(-0.998) + b(-0.998) \approx 3.992 \approx 4.$$

(We briefly note that the number 4 on the right hand side of this formlua also has an important meaning, which we will discuss shortly.)

This near miss example is what Ramanujan called a *mock theta function*, as it is "imitating" or pretending to be the "theta function" $b(q)$ near $q = -1$, and also has predictable behavior near other roots of unity. Ramanujan's letter goes on to give a definition of mock theta functions as he conceived of them, where he attempts to analytically describe the strangeness of what he observed, and to give a number of further examples he claims satisfy similar properties. His examples gave tantalizing hints of a new structure of modular-type objects, and this feeling served as a final challenge of Ramanujan which fascinated number theorists for decades to come.

6 A framework for Ramanujan's mock theta functions

In the decades following Ramanujan's letter to Hardy, Andrews, Berndt, Hickerson, Watson, and a number of others undertook important work expanding Ramanujan's examples and shedding light on their transformation properties, combinatorial implications, and q-series representations. These provided clues of a broader theory, including expressions for mock theta functions which would later be recognized by Zwegers as part of what we will refer to below as "Zwegers' three-fold path." A full framework for the class of functions Ramanujan was studying finally emerged about 80 years after his original letter, in the early 2000's in Zwegers' thesis [Zwe]. This can be explained using the theory of harmonic Maass forms, which was independently developed in a seminal paper of Bruinier and Funke [BF]. Their definition of a harmonic Maass form, can be roughly summarized as follows. A key point is that harmonic Maass forms are no longer required to be holomorphic, but rather to satisfy a second order differential equation instead (and, being in the kernel of a "Laplacian," they are termed "harmonic").

"Definition". *A harmonic Maass form (HMF) is a function which satisfies the following.*

1. It transforms like a modular form.

2. It is in the kernel of

$$\Delta_k := -\xi_{2-k}\xi_k,$$

where

$$\xi_k := 2iv^k \overline{\frac{\partial}{\partial \overline{\tau}}}.$$

3. It grows like at most as quickly as a weakly holomorphic modular form at the cusps (i.e., it has a principal part, and once this is subtracted, it decays rapidly).

The space of HMFs of weight k is denoted by H_k (with the subgroup of $\mathrm{SL}_2(\mathbb{Z})$ often suppressed here for simplicity).

Note that classical modular forms satisfy the conditions of the definition of a HMF, as the operator ξ_k kills holomorphic functions by definition, and so Δ_k does too. The operator ξ_k plays a central role in the theory of harmonic Maass forms. It shuttles between "dual weights" k and $2-k$, which in many ways reflect dual worlds of existence. In particular, ξ_k applied to a HMF of weight k yields a modular object of weight $2-k$. Since Δ_k splits as a composition of two ξ_k operators (up to a sign, which is there for historic reasons), the image under ξ_k must be in the kernel of ξ_{2-k}. But this precisely means that the image is a holomorphic function, and so one must obtain a holomorphic modular form. By considering the growth condition at cusps, it turns out that one always obtains a *cusp form*. Thus, one obtains a map

$$\xi_k \colon H_k \to S_{2-k}.$$

Bruinier and Funke gave a general proof that this map is *surjective* (one can think of this as a sort of "Serre duality"), and, following Zagier [Zag5], we commonly refer to the image of a HMF under it as the *shadow* of the HMF. Moreover, just as holomorphicity and translation invariance imply the existence of Fourier expansions of modular forms, harmonicity and translation invariance prove that HMFs have a Fourier expansion which splits into two pieces (since its a second order differential equation):

$$f \in H_k \implies f = f^+ + f^-.$$

Here, f^+ is called the *holomorphic part* and f^- is called the *non-holomorphic part*. Fittingly, f^+ is holomorphic, and f^- is not (assuming its non-zero, i.e., that we don't have a classical modular form). In particular, f^+ is an ordinary q-series, whose coefficients often encode some important object in whatever application we may be studying, and the coefficients of f^-, while decorated by certain non-holomorphic functions, are the same numbers which occur in the Fourier expansion of the shadow. Thus, the non-holomorphic part "comes from" classical modular forms and understanding it is equivalent to knowing about the shadow of f.

Definition 6.1. *A mock modular form is the holomoprhic part of a harmonic Maass form. A mock theta function is a mock modular form whose shadow is a unary theta series.*

Around the same time as Bruinier and Funke developed this theory, Zwegers proved that these same properties can be applied to modifications of Ramanujan's mock theta functions, and so, in this language, he proved the following in his famous Ph.D. thesis [Zwe].

Theorem 6.2 (Zwegers)**.** *Ramanujan's mock theta functions are mock theta functions according to the definition above (up to multiplying by a rational power of q and a rescaling of q by a rational power).*

Thus, Zwegers proved that Ramanujan's functions are missing an extra piece, which, when added, "corrects" their modularity transformations. Moreover, this piece is "simpler" than the original mock theta function, as it contains the same information as a unary theta function, and so this completion is particularly useful for understanding the original mock theta function. This realization led to an explosion of applications across mathematics and physics, including to representation theory, combinatorics, black holes, and arithmetic geometry.

7 The three-fold path of Zwegers

Zwegers gave three related ways to realize Ramanujan's mock theta functions as examples of broad classes of functions. Firstly, he studied properties of *Appell-Lerch series*, which have a shape such as

$$\sum_{n\in\mathbb{Z}} \frac{q^{Q(n)}}{1-aq^n},$$

where the sum over $\mathbb{Z}$ can be replaced with other lattices, and $Q(n)$ can be a quadratic form, for instance. If one expands the denominator as a geometric series in a certain range, one can find expressions for *indefinite theta functions*. As an example of the shape of such an indefinite theta function, we have the following identity of Andrews [And1] for Ramanujan's fifth order mock theta function $f_0(q)$:

$$f_0(q) = \frac{1}{(q)_\infty}\left(\sum_{\substack{n+j\geq 0\\ n-j\geq 0}} - \sum_{\substack{n+j<0\\ n-j<0}}\right)(-1)^j q^{n(5n+1)/2-j^2}.$$

This is called an indefinite theta function as the quadratic form in the exponents of q is indefinite (has both positive and negative eigenvalues). As mentioned above, this means that a sum over all choices of (n,j) would diverge, which is why we must do something more complicated such as summing only in a certain range. (More detailed notes on the theory of indefinite theta functions along with exercises can be found in [Rol])

Finally, Zwegers showed mock modularity properties of coefficients of meromorphic Jacobi forms. Important work of Dabholkar, Murthy, and Zagier [DMZ] follows up on this, and the reader (including readers with interest in q-series) are strongly encouraged to read their work as well. It should also be pointed out that coefficients of Jacobi forms, for example of infinite products in two variables, have often been important in combinatorics, for example, in using the constant term method. For instance, in his important memoir on generalized Frobenius partitions [And2], Andrews proves formulas which can be used to connect combinatorial generating functions to both coefficients of Jacobi forms and mock modular forms, as they fit into the examples along the lines of what is briefly discussed here.

8 A few important applications

Here, we will survey just a few of the many topics where mock modular forms play a role.

8.1 Combinatorics

As a starting point, let us consider the Ramanujan congruences modulo 5 and 7 for $p(n)$:

$$p(5n+4) \equiv 0 \pmod 5, \qquad p(7n+5) \equiv 0 \pmod 7.$$

Dyson conjectured [D], and Atkin and Swinnerton-Dyer proved [ASD], that these congruences are combinatorially explained by the *rank* statistic of a partition π, $r(\pi)$, which is the largest part minus the number of parts. If we form the generating function of partitions refined to keep track of ranks:

$$\mathcal{R}(\zeta; q) := \sum_n \sum_{\pi \text{ is a partition of } n} q^n \zeta^{r(\pi)},$$

where $\zeta := e^{2\pi i z}$ and $z \in \mathbb{C}$, then we observe that $\mathcal{R}(1; q) = P(q)$ is modular, and $\mathcal{R}(-1; q) = f(q)$ is mock modular. Bringmann and Ono proved in [BriO] that these two observations are instances of the following fact.

Theorem 8.1 (Bringmann-Ono)**.** *For any root of unity ζ, $\mathcal{R}(\zeta; q)$ is a mock modular form (up to minor considerations such as multiplying by a rational power of q).*

This theorem immediately implies many congruences, and also allows one to study asymptotics of ranks very efficiently. We will see how these two properties, congruences and asymptotics, can be studied for mock modular forms in generality below.

8.2 Arithmetic Geometry

As an example of the applications to geometry, we consider the picture for elliptic curves. Suppose that E is an elliptic curve over $\mathbb{Q}$. Then by the Modularity Theorem, there is a weight 2 newform associated to the curve, which we will denote by f_E. We denote by Λ_E the lattice such that $E \cong_{\mathbb{C}} \mathbb{C}/\Lambda_E$. We will sketch an idea due to Guerzhoy [Gue1, Gue2], which allows us to give a "canonical" lifting under the (∞-to-1) map $\xi_0 \colon H_0 \to S_2$ (recall that Bruinier and Funke proved that this is surjective). The crucial point of this particular lift is that it is compatible with "arithmetic," such as p-adic properties. Other methods for lifting, such as following Bruinier and Funke's proof or using the method of Poincaré series sketched below, either aren't as explicit (the following method relies only on classical modular forms theory and classical elliptic functions and is direct), or destroy arithmetic (see Section 2.2.3; this method is better-suited to existence proofs and asymptotic analysis).

Consider the Weierstrass $\wp$-function for the lattice Λ_E, an elliptic function, i.e. $\wp(z+\omega) = \wp(z)$ for all $z \in \mathbb{C} \setminus \Lambda_E$ and $\omega \in \Lambda_E$. The negative of its antiderivative is the *Weierstrass zeta function*:

$$\wp(\Lambda_E; z) = -Z'(\Lambda_E; z).$$

Now Z is not elliptic, but only nearly so. Returning to the modular form $f_E =: \sum_{n\geq 1} a_n q^n$ again, we consider the formal antiderivative (known as the *Eichler integral*)

$$\mathcal{E}_E(z) := \sum_{n\geq 1} \frac{a_n}{n} q^n,$$

so that

$$q\frac{d}{dq}\left(\mathcal{E}_E\right) = f_E.$$

Now this function is nearly modular of weight 0, but not quite. This is because slashing in weight 0 and 2 (and only in this case!) commutes with differentiation, in the sense that for any function g on the upper half-plane and any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$,

$$D\left(g|_0\gamma\right) = (Dg)|_2\gamma$$

where $D := q\frac{d}{dq}$. Thus, for any γ in the group Γ under which f_E transforms, the function $\mathcal{E}_E|_0(1-\gamma)$ has derivative equal to zero, and so is a constant. Thus, we obtain a map

$$\Psi \colon \Gamma \to \mathbb{C}$$

by setting $\Psi(\gamma)$ to be the constant obtained by evaluating $\mathcal{E}_E|_0(1-\gamma)$ at any point in $\mathbb{H}$. By classical theory of Eichler and Shimura, the image of Ψ is exactly the lattice Λ_E. To phrase this another way, the "errors to modularity" of $\mathcal{E}_E$ all live in this lattice. Thus, if we consider the function $Z(\Lambda_E; \mathcal{E}_E)$, it would actually be modular of weight 0 if Z were elliptic, as errors to modularity would be swallowed up by lattice invariance. However, since

Z is not quite an elliptic function, but it can be corrected with a minor non-holomorphic correction term. Working this out, one finds that $Z(\Lambda_E; \mathcal{E}_E)$ is really a *mock modular form*. Moreover, the non-holomorphic part is especially simple, so computing the shadow of the mock modular form is essentially trivial, and easily returns the original cusp form f_E. Thus, we have answered our question of how to find explicit lifts under ξ_0. In a paper by Alfes, Griffin, Ono, and the second author [AGOR], the implications for elliptic curves are explored. By applying work of Bruinier and Ono [BruO], as well as Waldspurger [Wal] and Kohnen-Zagier [KZ], it turns out that we can directly construct a single HMF whose holomorphic part encodes the vanishing of L-derivatives of quadratic twists of E, and whose non-holomorphic part has coefficients encoding the vanishing of the central L-values of the twists. By the Birch and Swinnerton-Dyer conjecture, these coefficients should thus determine the rank of the quadratic twists of E up to rank 2, that is, they "know" whether the rank is 0, 1, or greater than 1.

More generally, these types of lifting problems are central in many questions on harmonic Maass forms, and as we now know, in classical modular forms. Returning to q-series, recall that *Lehmer's conjecture* states that all coefficients $\tau(n)$ of the Ramanujan Delta function $\Delta(\tau)$ are non-zero. Again by Bruinier-Funke, we can show that "nice" lifts of Δ to weight $2 - 12 = -10$ exist. It turns out that Lehmer's conjecture is equivalent to the coefficients of the corresponding mock modular form being transcendental (see [BOR]). Thus, an explicit understanding of the lifting problem in a similar manner which is compatible with arithmetic for higher weights would be highly interesting.

9 Quantum Modular forms

Following Zagier (see [Zag6] for his introduction of the term along with a number of motivating examples), a *quantum modular form* of weight k, roughly speaking, is a function $f\colon \mathbb{Q} \cup \infty \to \mathbb{C}$ such that the "errors to modularity" $f|_k(1-\gamma)$ are "nice functions" for all γ in some subgroup of $\mathrm{SL}_2(\mathbb{Z})$ (as this sounds similar to the previous section, it is worth pointing out that Eichler integrals can be used to define quantum modular forms). What nice means depends on the context, and the point is that this definition includes many different examples with a different flavor which Zagier surveyed in his original paper on the subject and which have been developed since. A typical behavior might be a function which is defined on $\mathbb{Q}$, with no way or no obvious way to extend it to $\mathbb{R}$, but whose cocycles become not only defined on $\mathbb{R}$ but also become continuous, differentiable, smooth, analytic, etc.

As one example, recall that above we saw that Ramanujan's mock theta function $f(q)$ is almost cancelled out by the (essentially) modular form $b(q)$ near $q = -1$:

$$f(q) + b(q) \approx 4.$$

Ramanujan conjectured this value 4 in his letter, as well as similar almost cancelling values of $f(q) \pm b(q)$ at other roots of unity. It turns out that these constants which Ramanujan studied encode a hidden structure, namely, a quantum modular form. This is a good example of the fact that going back to the original sources in a subject and really reading them closely can be very useful in any field. The hidden structure behind Ramanujan's calculations is revealed in the following elegant result of Folsom, Ono, and Rhoades [FOR1, FOR2].

Theorem 9.1 (Folsom-Ono-Rhoades)**.** *For every even order $2k$ root of unity ζ, we have the following radial limit formula:*

$$\lim_{q\to\zeta}\Big(f(q)-(-1)^k b(q)\Big) = -4\sum_n (1+\zeta)^2(1+\zeta^2)^2\cdots(1+\zeta^n)^2\zeta^{n+1}.$$

Note the right hand side is actually a *finite sum* for any root of unity ζ. As an exercise, consider what happens for odd order roots of unity. What is the meaning of the right hand side? We already know that $f(q)$ is a specialization of the rank generating function $\mathcal{R}$. It turns out that there is an elegant radial limit theorem for any other specialization of $\mathcal{R}$, which are mock modular as discussed above. To explain this, we need the *crank generating function* $\mathcal{C}(\zeta;q)$, which is defined in the same way as $\mathcal{R}$ but with "rank" replaced by "crank," another partition statistic which simultaneously explains Ramanujan's congruences modulo 5, 7, *and* 11. This was conjectured by Dyson and found and proven to work by Andrews and Garvan [AG], following up on work of Garvan [Gar] on a crank for vector partitions. The generating function enjoys the infinite product formula:

$$\mathcal{C} = \frac{(q)_\infty}{(\zeta q)_\infty(\zeta^{-1}q)_\infty}.$$

This is essentially the inverse of the Jacobi theta function times η^2, and so it is a meromorphic Jacobi form of weight 1/2. Thus, its specializations, including $b(q)$, are all weakly holomorphic modular forms of weight 1/2. Furthermore, if we consider a generating function $\mathcal{U}(\zeta;q)$ which encodes strongly unimodal sequences by size and a statistic on them which is also called rank, then it turns out that

$$\mathcal{U}(\zeta;q) = \sum_n (-\zeta q)_n(-\zeta^{-1}q)_n \cdot q^{n+1}.$$

Folsom, Ono, and Rhoades [FOR1, FOR2] then proved the following general limiting formula, which pleasingly connects three types of combinatorial generating functions in one and generalizes the above formula for $f(q)$ and $b(q)$.

Theorem 9.2 (Folsom-Ono-Rhoades)**.** *Under certain conditions on the roots of unity ζ, ζ', and for certain elementary factors $*$, we have that*

$$\lim_{q\to\zeta'}\left(\mathcal{R}(\zeta;q) - *\mathcal{C}(\zeta;q)\right) = \mathcal{U}(-\zeta;\zeta').$$

This can be proved using asymptotic calculations, which require significant work. The reader is also referred to later, simpler proofs for such formulas by Zudilin [Zud], and later by Folsom, Ono, and Rhoades [FOR2], and the authors of this survey highly recommend these works especially for a combinatorial audience. Since specializations of $\mathcal{R}$ are mock, and specializations of $\mathcal{C}$ are modular, it is natural to ask what kind of function $\mathcal{U}$ is. We noted above that it becomes a finite sum for roots of unity ζ, so it is well defined as a function on $z \in \mathbb{Q}$. As an example, $\mathcal{U}(-1;q) = \sum_n (q)_n^2 q^{n+1}$, Bryson, Pitman, Ono, and Rhoades proved [BPOR] that this function is actually (at roots of unity)

$$\mathcal{U}(-1;q) = F(q^{-1}),$$

where $F(q)$ is Kontsevich's function

$$F(q) = \sum_n (q)_n.$$

This function is known to be a quantum modular form, as Zagier proved [Zag2] by establishing his "strange identity"

$$F(q) \text{``} = \text{''} -\frac{1}{2}\sum_{n\geq 1} n\chi_{12}(n) q^{\frac{n^2-1}{24}}.$$

Quotation marks are used here as the left hand side now really only makes sense at roots of unity, while the *partial theta function* (it is summed over only half of a lattice) is only defined for $|q| < 1$. Note that this expansion is very closely related to the theta function representation for $\eta(\tau)$ in Corollary 2.11. The world where $F(q)$ lives, the *Habiro ring* which is known to encode invariants in knot theory, can be thought of as a set of "analytic functions at the roots of unity." In particular, we can consider the expansion of this function around roots of unity, such as the following one around $q = 1$:

$$\sum_n (1-q;1-q)_n =: \sum_n \xi_n q^n.$$

It is an instructive exercise to prove that the coefficients on the right hand side converge, as well as that similar expansions make sense at other roots of unity. These coefficients ξ_n are important numbers in knot theory. They were used by Zagier in the study of *Vassiliev invariants.* Just as modular and mock modular (as we will see) forms are very useful in studying asymptotics, Zagier used quantum modularity (in current language) to study the growth properties of these numbers. It also turns out that the same numbers count a number of interesting things in combinatorics, and these were further studied by Andrews and Sellers [AS]. They observed and proved that they satisfy infinitely many Ramanujan-type congruences. Further work by Guerzhoy, Kent, and the second author explained [GKR] this from a general quantum modular perspective and for general types of partial theta functions. Thus, in contrast to Ramanujan's statement that "false" (essentially the same theory as partial) theta functions "do not enter mathematics as beautifully" as his mock theta functions, they do enter into mathematics, in fact Ramanujan's mathematics, quite beautifully. Though, to be fair, it is difficult for most things to match the beauty of the mock theta functions.

10 General philosophy of applications to combinatorics

We conclude with a general discussion of how the theory of harmonic Maass forms gives powerful tools for studying two types of properties of sequences of numbers when you are lucky enough to have mock modular forms lurking: congruences and asymptotics.

10.1 Congruences

We have seen that modular forms are a powerful tool for proving congruences of arithmetic sequences. Here, we briefly sketch an approach for obtaining congruences for mock modular forms. The key idea is to reduce to the classical modular case. For instance, if the shadow is a linear combination of unary theta functions (for example, for a mock theta function; this property is also almost required for most combinatorial sequences as mock modular forms don't usually have integral Fourier coefficients otherwise). Then we are lucky, and we can twist away the shadow. Now we may apply the sieving operators from the classical theory of modular forms, which simply restrict the Fourier coefficients of a modular form by throwing away all that don't lie in a fixed arithmetic progression, and which preserve modularity. The proof of this fact only uses the transformation properties, i.e., it is a statement on the level of slash operators. Thus, sieving the coefficients of a HMF also yields a new HMF. But if the shadow is supported on finitely many square classes, then we can choose a progression where the coefficients of the shadow vanish. This then reduces us to the world of classical modular forms, and we can generically check in many cases that this procedure doesn't also annihilate the holomorphic part. By combining with the classical theory of modular form congruences, we thus obtain many congruences for coefficients of the mock modular form we may want to study. For instance, this already tells us, by well-known work of Treneer [Tre], that mock theta functions satisfy infinitely many linear congruences.

10.2 Asymptotics

We have seen that modularity can be used to prove asymptotics, and indeed exact formulas, for the partition function $p(n)$ by means of a Tauberian Theorem (for the asymptotic) or the Circle Method (for more precision, depending on which variant of the Circle Method is used). Rademacher's formula turns out to be a hint of a more general structure of modular forms. We want a procedure to take any modular form (or mock modular form; as we shall see, answering the first question naturally leads us to the second) and give a procedure for computing the asymptotics of the coefficients of the modular form quickly. In the case of the partition function, we saw that the Circle Method consists of a careful study of the generating function near its singularities, and so its principal parts at (all equivalent in this case) the cusps dictate the asymptotic growth properties. The theory of harmonic Maass forms gives a way to diagonalize spaces based on principal parts. As is well-known, for general congruence subgroups, weakly holomorphic modular forms with arbitrary principal parts don't exist. However, *harmonic Maass* forms are allowed to have arbitrary principal part. Considering just the cusp ∞ for simplicity, for any $k \leq 0$, any $m \geq 1$, and any congruence subgroup Γ, we have a *Maass-Poincaré series* $F_{k,m,\Gamma} \in H_k(\Gamma)$ such that

$$F_{k,m,\Gamma} = q^{-m} + O(1)$$

at ∞, and $F_{k,m,\Gamma}$ is bounded at all inequivalent cusps. This series is a lift of the classical Poincaré cusp forms under ξ_k, and it is defined by a similar procedure as we saw for the Eisenstein series and other Poincaré series before; namely, by averaging a certain function hit with slash operators to force modularity while retaining good convergence and growth properties. Since there are no (non-constant) holomorphic modular forms of weight $k \leq 0$, this means that we can take the appropriate linear combination of these Poincaré series which matches the principal part of any weakly holomorphic, and more generally any mock modular form in these weights, and the difference is simply a constant. The advantage here is that the series $F_{k,m,\Gamma}$ have completely explicit formulas for their Fourier coefficients, in terms of Kloosterman sums and Bessel functions, just as Rademacher found for $p(n)$. Thus, we can quickly determine an *exact formula* for the coefficients of any of these forms with a routine, finite check. We conclude by briefly considering other weights. For weights between 0 and 2, convergence of the Poincaré series is an issue. However, general work of Goldfeld and Sarnak [GS] can be used to analytically continue the Poincaré series in weights $1/2$ and $3/2$ (which, as described above, are often the ones of combinatorial interest), and the same formulas for Fourier expansions still hold. In these weights, as well as for methods of Poincaré series which could be applied in weights larger that $3/2$ (note that there are no cusp forms of non-positive weight, so modifications have to be made for large weights), note that due to the existence of holomorphic modular forms, the method of matching principal parts won't yield exact formulas. However, the obstruction to obtaining exact formulas will then only have polynomial growth, and if you started with a function with a pole at some cusp, so that there will be rapid growth of the Fourier coefficients, the method of Poincaré series still yields a strong estimate, and, in particular, an asymptotic.

References

[Ahl] S. Ahlgren, *The partition function modulo composite integers M*, Math. Ann. **318** (2000), 795–803.

[AO] S. Ahlgren and K. Ono, *Congruence properties of the partition function*, Proc. Natl. Acad. Sci., USA **98** (23) (2001), 12882–12884.

[AGOR] C. Alfes, M. Griffin, K. Ono, and L. Rolen, *Weierstrass mock modular forms and elliptic curves*, Res. Number Theory **1** (2015), Art. 24, 31 pp.

[And1] G. E. Andrews, *The fifth and seventh order mock theta functions*, Trans. Amer. Math. Soc. **293** (1986), no. 1, 113–134.

[And2] G. E. Andrews, *Generalized Frobenius Partitions*, Memoirs of the American Mathematical Society, **301**, Providence, RI, May 1984.

[And3] G. E. Andrews, *Partitions: At the Interface of q-Series and Modular Forms*, Ramanujan J. (2003) 7: 385.

[AG] G. E. Andrews and F. G. Garvan, *Dyson's crank of a partition*, Bull. Amer. Math. Soc. (N.S.) **18** (1988), no. 2, 167–171.

[AS] G. E. Andrews and J. A. Sellers, *Congruences for the Fishburn numbers*, J. Number Theory **161** (2016), 298–310.

[Apo] T. M. Apostol, *Modular Forms and Dirichlet Series in Number Theory*, 2nd edition, Springer-Verlag, 1990.

[ASD] A. O. L. Atkin and P. Swinnerton-Dyer, *Some properties of partitions*, Proc. London Math. Soc. (3) **4** (1954), 84–106.

[Ber] B. C. Berndt, *Ramanujan's congruences for the partition function modulo* 5, 7, *and* 11, Int. J. Number Thy. **3** (2007), 349–354.

[BFOR] K. Bringmann, A. Folsom, and K. Ono, and L. Rolen *Harmonic Maass forms and mock modular forms: theory and applications**, AMS Colloquium Series, 2017.

[BriO] K. Bringmann and K. Ono, *Dyson's ranks and Maass forms*, Ann. of Math. (2) **171** (2010), no. 1, 419–449.

[BF] J. H. Bruinier and J. Funke, *On two geometric theta lifts*, Duke Math. J. **125** (2004), 45–90.

[BruO] J. H. Bruinier and K. Ono, *Heegner divisors, L-functions, and harmonic weak Maass forms*, Ann. Math. **172** (2010), 2135–2181.

[BOR] J. H. Bruinier, R. C. Rhoades, and K. Ono, *Differential operators for harmonic weak Maass forms and the vanishing of Hecke eigenvalues*, Math. Ann. (2008) 342: 673.

[BPOR] J. Bryson, K. Ono, S. Pitman, and R. C. Rhoades, *Unimodal sequences and quantum and mock modular forms*, Proc. Natl. Acad. Sci. USA **109** (2012), no. 40, 16063–16067.

[Coh] H. Cohen, *Sums Involving the Values at Negative Integers of L-Functions of Quadratic Characters*, Math. Ann. **217** (1975), 271–285.

[CS] H. Cohen and F. Strömberg, *A Classical Introduction to Modular Forms**, Graduate Studies in Mathematics **179**, 2017.

[Cox] D. A. Cox, *Primes of the Form* $x^2 + ny^2$, 2nd edition, Wiley, 2013.

[DMZ] A. Dabholkar, S. Murthy, and D. Zagier, *Quantum Black Holes, Wall Crossing, and Mock Modular Forms**, to appear in Cambridge Monographs in Mathematical Physics.

[DS] F. Diamond and J. Shurman, *A First Course in Modular Forms*, Springer-Verlag, 2005.

[D] F. J. Dyson, *Some guesses in the theory of partitions*, Eureka (1944), no. 8, 10–15.

[EZ] M. Eichler and D. Zagier, *The theory of Jacobi forms**, Birkhäuser-Verlag, 1985.

[FOR1] A. Folsom, K. Ono, and R. C. Rhoades, *Mock theta functions and quantum modular forms*, Forum Math. Pi **1** (2013), e2, 27.

[FOR2] A. Folsom, K. Ono, and R. C. Rhoades, *Ramanujan's radial limits*, Ramanujan 125, Contemp. Math., vol. 627, Amer. Math. Soc., Providence, RI, 2014, 91–102.

[Gar] F. G. Garvan, *New combinatorial interpretations of Ramanujan's partition congruences mod* 5*,* 7 *and* 11, Trans. Amer. Math. Soc. **305** (1988), no. 1, 47–77.

[GS] D. Goldfeld and P. Sarnak, *Sums of Kloosterman sums*, Invent. Math. **71** (1983), no. 2, 243–250.

[Gue1] P. Guerzhoy, *A mixed mock modular solution of the Kaneko-Zagier equation*, Ramanujan J., **36** (2015), no. 1-2, 149–164.

[Gue2] P. Guerzhoy, *On Zagiers adele*, Res. Math. Sci., 2014, 1:7.

[GKR] P. Guerzhoy, Z. A. Kent, and L. Rolen, *Congruences for Taylor expansions of quantum modular forms*, Res. Math. Sci. **1** (2014), Art. 17, 17.

[HR] G. H. Hardy and S. Ramanujan, *Asymptotic Formulae in Combinatory Analysis*, Proc. London Math. Soc. **17** (1918), 75–115.

[Ing] A. E. Ingham, *A Tauberian Theorem for Partitions*, Ann. Math. **42** (5) (1941), 1075–1090.

[Kil] L. J. P. Kilford, *Modular Forms: A Classical and Computational Introduction**, Imperial College Press, 2008.

[KZ] W. Kohnen and D. Zagier, *Values of L-series of modular forms at the center of the critical strip*, Invent. Math. **64** (1981), no. 2, 175–198.

[Iwa] H. Iwaniec. *Topics in Classical Automorphic Forms*, AMS Graduate Studies in Mathematics **17**, 1997.

[KK] M. Koecher and A. Krieg, *Elliptische Funktionen und Modulformen*, 2nd edition, Springer-Verlag, 2007.

[Ono1] K. Ono, *The distribution of the partition function modulo m*, Ann. Math. **151** (2000), 293–307.

[Ono2] K. Ono, *The Web of Modularity: Arithmetic of the Coefficients of Modular Forms and q-series**, CBMS **102**, 2004.

[Ono3] K. Ono, *Unearthing the visions of a master: harmonic Maass forms and number theory**, Current developments in mathematics, 2008, 347454, Int. Press, Somerville, MA, 2009.

[Pet] H. Petersson, *Konstruktion der Modulformen und der zu gewissen Grenzkreisgruppen gehrigen automorphen Formen von positiver reeller Dimension und die vollstndige Bestimmung ihrer Fourierkoeffizienten*, S.-B. Heidelberger Akad. Wiss. Math.-Nat. Kl. **1950** (1950), 417–494.

[Pri] W. de A. Pribitkin, *Revisiting Rademacher's Formula for the Partition Function $p(n)$*, Ramanujan J. **4** (4) (2000), 455–467.

[Rad] H. Rademacher, *On the expansion of the partition function in a series*, Ann. Math. **44** (1943), 416–422.

[Ran] R. A. Rankin, *Modular forms and functions*, Cambridge University Press, 1977.

[Rol] L. Rolen, *Notes from: School on mock modular forms and related topics"*, Kyushu University. November 21-25, 2016, `https://math.vanderbilt.edu/rolenl/Notes.html`

[RW] J. Rouse and J. J. Webb, *On spaces of modular forms spanned by eta-quotients*, Adv. Math. 272 (2015), 200224.

[Sage] SageMath, the Sage Mathematics Software System (Version 8.1), The Sage Developers, 2017, https://www.sagemath.org.

[Sch] B. Schoeneberg, *Das Verhalten von mehrfachen Thetareihen bei Modulsubstitutionen*, Math. Ann. **116** (1) (1939), 511–523.

[Ser] J.-P. Serre, *A Course in Arithmetic**, Springer-Verlag, 1973.

[Shi] G. Shimura, *Modular Forms of Half-Integral Weight*, Ann. Math. **97** (1973), 440–481.

[Ste] W. Stein, *Modular Forms: A Computational Approach*, available at https://wstein.org/books/modform/stein-modform.pdf

[Stu] J. Sturm, *On the congruence of modular forms*, in *Number Theory (New York, 1984–1985)*, Lecture Notes in Mathematics **1240**, Springer-Verlag, 1987, 275–280.

[Tre] S. Treneer, *Congruences for the coefficients of weakly holomorphic modular forms*, Proc. London Math. Soc. (3) **93** (2006), no. 2, 304–324.

[VZ] M. Vlasenko and S. Zwegers, *Nahm's conjecture: asymptotic computations and counterexamples*, Communications in Number Theory and Physics **5** (2011), 617–642.

[Wal] J.-L. Waldspurger, *Sur les coefficients de Fourier des formes modulaires de poids demi-entier*, J. Math. Pures Appl. (9) **60** (1981), no. 4, 375–484.

[Zag1] D. Zagier, *Modular Forms**, in *From Number Theory to Physics*, Springer-Verlag, Springer-Verlag, 1992, available at http://people.mpim-bonn.mpg.de/zagier/

[Zag2] D. Zagier, *Vassiliev invariants and a strange identity related to the Dedekind eta-function*, Topology **40** (2001), no. 5, 945–960.

[Zag3] D. Zagier, *The dilogarithm function*, in *Frontiers in Number Theory*, Physics and Geometry II, P. Cartier, B. Julia, P. Moussa, P. Vanhove (eds.), Springer-Verlag, Berlin-Heidelberg-New York (2006), 3–65.

[Zag4] D. Zagier, *Elliptic Modular Forms and Their Applications**, in *The 1-2-3 of Modular Forms*, Springer-Verlag, 2008.

[Zag5] D. Zagier, *Ramanujan's mock theta functions and their applications (after Zwegers and Bringmann–Ono)**, Sḿinaire Bourbaki. Vol. 2007/2008. Astérisque No. 326 (2009), Exp. No. 986, viiviii, 143164 (2010).

[Zag6] D. Zagier, *Quantum modular forms*, Quanta of maths, Clay Math. Proc., vol. 11, Amer. Math. Soc., Providence, RI, 2010, 659–675.

[Zud] W. Zudilin, *On three theorems of Folsom, Ono, and Rhoades*, Proc. Amer. Math. Soc. **143** (2015), 1471–1476.

[Zwe] S. Zwegers, *Mock theta functions*, Ph.D. thesis, Universiteit Utrecht, 2002.