4. Theta functions and the Ramanujan–Petersson conjecture

In the previous lecture, we defined a procedure for constructing a graph $G_p(\mathcal{O})$ for any prime p and order $\mathcal{O} \subset \mathbb{H}$. Actually, it makes sense to define a graph $G_n(\mathcal{O})$ for any positive integer n; it just happens that this graph tends to be Ramanujan only when n is prime. Let $A_n(\mathcal{O})$ be the adjacency matrix of $G_n(\mathcal{O})$.

Let us return to the example where $\mathcal{O} \subseteq \mathbb{H}$ is the order generated by

$$1, \quad \frac{i-\sqrt{3}k}{2}, \quad i-\sqrt{3}j, \quad \frac{1+3i+\sqrt{3}j+\sqrt{3}k}{2}.$$

In this case, the $A_n(\mathcal{O})$ are 2×2 matrices. For each n, the vectors $\begin{pmatrix}1\\1\end{pmatrix}$ and $\begin{pmatrix}1\\-1\end{pmatrix}$ are eigenvectors. The eigenvalues of the $A_n(\mathcal{O})$ acting on $\begin{pmatrix}1\\-1\end{pmatrix}$ are

$$1,\ -1,\ -1,\ -1,\ 1,\ 1,\ 0,\ 3,\ 1,\ -1,\ -4,\ \ldots .$$

Now consider the power series

$$\begin{aligned} g(q) &= q\prod_{n=1}^{\infty}(1-q^n)(1-q^{3n})(1-q^{5n})(1-q^{15n}) \\ &= q-q^2-q^3-q^4+q^5+q^6+3q^8+q^9-q^{10}-4q^{11}+\cdots . \end{aligned}$$

The coefficients of this power series are the same as the eigenvalues of $A_n(\mathcal{O})$. This is not a coincidence!

So why would these two sequences equal? It is a bit of a mystery. In number theory, sometimes you can prove something and still feel like you don't really understand why it's true. Some people might find this frustrating, but personally, I think it's really cool!

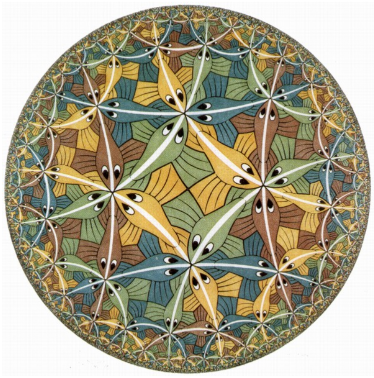
Anyway, the way that I know to prove that the above sequences are equal is to prove that they are both modular forms. To motivate the definition of a modular form, let's look at symmetries of the complex plane again.

M. C. Escher, drawing E55

This drawing has rotational and translational symmetry. If we think of this drawing as living in the complex plane, its symmetries are described by transformations of the form

$$f(z) = e^{i\theta}z + c\,.$$

It is less obvious how to describe the symmetries of the following drawing in terms of complex numbers.



M. C. Escher, Circle Limit III

But it can be done. Consider the set of transformations of the form

$$f(z) = \frac{az+b}{b^*z+a^*}, \quad a, b \in \mathbb{C}, \quad |a| > |b| .$$

These turn out to preserve the disc. We can see that they preserve its boundary, the unit circle, as follows. If $|z| = 1$, then

$$|az+b| = |azz^* + bz^*| = |a + bz^*| = |a^* + b^*z|$$

so $\frac{|az+b|}{|b^*z+a^*|} = 1$. The condition $|a| > |b|$ guarantees that the transformation preserves the interior of the circle, rather than swapping the interior and exterior.

Similarly, we can consider the following drawing.

J. Leys, after M. C. Escher, drawing E69

The drawing is meant to be infinite, filling the upper half plane. The symmetries of this drawing are of the form

$$f(z) = \frac{az+b}{cz+d}, \quad a, b, c, d \in \mathbb{R}, \quad ad - bc > 0 .$$

It is clear that such transformations preserve the real line. The condition $ad-bc > 0$ guarantees that the transformation preserves the upper half plane, rather than swapping the upper and lower half planes.

Now let's actually define modular forms.

Definition 4.1. Let k be a nonnegative integer, and let N be a positive integer. A holomorphic function f on the upper half complex plane is a *modular function* of weight k and level $\Gamma_0(N)$ if, for any integers a, b, c, d satisfying $ad - bc = 1$ and $N \mid c$,

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) .$$

We say that f is a *cuspidal modular form* if, in addition, $\lim_{z\to t} f(z) = 0$ for any $t \in \mathbb{Q} \cup \{\infty\}$.

We write $S_k(\Gamma_0(N))$ for the space of all cuspidal modular forms of weight k and level $\Gamma_0(N)$.

Proposition 4.2.

$$g(e^{2\pi i z}) \in S_2(\Gamma_0(15)) .$$

This is not at all obvious. There is a proof using Fourier analysis and a proof using complex analysis. I will give a reference below when I state a more general theorem.

We can also interpret the coefficients of f as eigenvalues of operators called *Hecke operators*. Let p be a prime number. Let T_p to be the operator defined by $T_p(\sum a_n q^n) = \sum b_n q_n$ where

$$b_n = \begin{cases} a_{pn} + p a_{n/p}, & p \mid n \\ a_{pn}, & p \nmid n . \end{cases}$$

Let's compute $T_2(g)$.

$$T_2(g) = -q + q^2 + q^3 + q^4 - q^5 + \cdots . = -g .$$

Every element of $S_k(\Gamma_0(N))$ has an expansion of the form $f(z) = \sum_{n=1}^{\infty} a_n q^n$, where $q = e^{2\pi i z}$. One can show that whenever $p \nmid N$, T_p preserves $S_k(\Gamma_0(N))$. Furthermore, $S_2(\Gamma_0(15))$ is one-dimensional. Therefore, g is an eigenvector of T_p for $p \notin \{3, 5\}$. The coefficient of q^1 in f is 1, while the coefficient of q^1 in $T_p g$ is the same as the coefficient of q^p in g. Therefore, the T_p-eigenvalue is the coefficient of q^p.

Theorem 4.3 (Ramanujan–Petersson conjecture)**.** *Let $k \geq 2$ and N be positive integers. Let p be a prime not dividing N. Then the eigenvalues of T_p acting on $S_k(\Gamma_0(N))$ satisfy*

$$|\lambda| \leq 2p^{(k-1)/2} .$$

The conjecture was proved by Eichler for $k = 2$ and by Deligne for all k. The proof uses some techniques (in particular, algebraic geometry over $\mathbb{Z}/p\mathbb{Z}$) that I will not be able to cover in this course.

In particular, when $k = 2$, we get $|\lambda| \leq 2\sqrt{p}$. To prove that the graph $G_p(\mathcal{O})$ is Ramanujan, one constructs cuspidal modular forms whose T_p-eigenvalues are the nontrivial eigenvalues of $A_p(\mathcal{O})$. The Ramanujan–Petersson conjecture then implies that $G_p(\mathcal{O})$ is Ramanujan. Let us now explain how these modular forms are constructed.

Proposition 4.4. *Let p be a prime, and let e be an eigenvector of $A_p(\mathcal{O})$. Suppose* $e \neq \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$. *For each n, let λ_n denote the eigenvalue of $A_n(\mathcal{O})$ acting on e. Define a function θ_e on the upper half complex plane by*

$$\theta_e(z) = \sum_{n=1}^{\infty} \lambda_n e^{2\pi i n z} .$$

Let

$$\mathcal{O}^* := \{x \in \mathbb{H} | \operatorname{tr} xy \in \mathbb{Z} \textit{ for all } y \in \mathcal{O}\} .$$

Let N be the smallest positive integer such that $N\mathcal{O}^* \subseteq \mathcal{O}$. Then $\theta_e(z)$ is a cuspidal modular form of weight 2 and level $\Gamma_0(N)$, and $T_p\theta_e(z) = \lambda\theta_e(z)$.

Proof. See [Shi73, Proposition 2.1]. □

The fact that $G_p(\mathcal{O})$ is Ramanujan is then a consequence of the Ramanujan–Petersson conjecture.

There is much more to the theory of modular forms than I can possibly say in one lecture. Some books that cover modular forms in more detail are [Ser73] and [DS05].

References

[DS05] F. Diamond and J. Shurman. *A first course in modular forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.

[Ser73] J.-P. Serre. *A course in arithmetic*. Springer-Verlag, New York-Heidelberg, 1973. Translated from the French, Graduate Texts in Mathematics, No. 7.

[Shi73] G. Shimura. On modular forms of half integral weight. *Ann. of Math. (2)*, 97:440–481, 1973.