Analytic primality testing.

This paper is really an attempt to learn basic analytic number theory. The thing we might want to do is clarify how the passage from a small set of modular forms with a lot of inariance to a larger set with less invariance is purely algebraic. For simplicity the weights which are considered are really what are usually called the even weights, and the levels all above 2 (but not requring congruence subgroups). This is done in sections 1 through 7.

The last section begins to apply such considerations to primality testing. It is really the elliptic modularity that is used in the last section, which isn't discussed in the earlier sections, so the two sections of this paper are currently unrelated and it should be viewed as only a working draft of a possible longer paper.

1. Modular forms

The subject of modular forms is old and has been generaliized in many directions. Therefore it is likely that the theorems which I'll state in this section are known already, and may represent a point of view only.

We'll follow Dolgachev's convention of gradings, so the space M_k of modular forms of weight k will be holomorphic entire functions $\mathbb{H} \to \mathbb{C}$ satisfying $f((az+b)/(cz+d)) = (cz+d)^{2k} f(z)$ when $ad-bc=1$ and on $\mathbb{H}$ which are holomorphic at the cusps; and we will not consider the case when k is a half-integer (although we could do so).

I should also comment, this draft likely has many errors and has not been checked.

John Atwell Moody
Coventry, July 2015

If the rule above holds only for matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ belonging to a finite index subgroup $\Gamma \subset PSl_2(\mathbb{Z})$ one says that f is 'modular of level Γ', which is weaker than being modular. For each such group Γ, from the ring $\oplus_{k=0}^{\infty} M_k(\Gamma)$, where $M_k(\Gamma)$ is the space of modular forms of weight k and level Γ, the 'Proj construction' builds a compactification $X(\Gamma)$ of the orbit space $\Gamma \setminus \mathbb{H}$. That is, if $f, g \in M_k(\Gamma)$ have the same degree, this means that the modularity will cancel when one considers the rational function f/g; it is a well-defined function the part of $\mathbb{H}$ where g is not zero, and it is invariant so defines a rational function on a variety which is a compactification of $\Gamma \setminus \mathbb{H}$.

The modular forms are much more interesting than the algebraic curve which results from this process of compactification; and athough there exist theorems of algebraic geometry (such as Riemann-Roch) which can actually construct the rational functions; the issue is, to what extent can one go all the way back, and re-introduce the modularity factor which had cancelled when one had passed to the rational function f/g.

I'll state theorems in the desired direction in this section; it is tempting to call them propositions as from the standpoint of algebraic geometry the proofs are trivial, involving no more than observations once one brings the definitions into the more general setting.

We'll begin with the modular curve $\Gamma(2)\setminus\mathbb{H}$, it is isomorphic to the complement of $C=\{0,1,\infty\}$ in $\mathbb{P}^1\cong X(2)=X(\Gamma(2))$. Consider the category of compact connected Riemann surfaces Y over $X(2)$ where the structure map $f_Y: Y\to X(2)$ is nonconstant and unbranched away from C. Let $\mathcal{M}_Y$ denote the locally free invertible sheaf $\Omega_Y(log\ f^{-1}(C))$ on Y. Then

1. Theorem. For $g: Y\to Z$ in our category (of curves over $X(2)$) there is a natural isomorphism

$$g^*\mathcal{M}_Z\to\mathcal{M}_Y.$$

Proof. This is an easy fact relating logarithmic derivatives with branched covers; the analagous theorem is also true in higher dimensions. Let $\mathcal{M}_k(Y)=\mathcal{M}_Y^{\otimes k}$.

2. Theorem. For each Y, letting $\Gamma_Y\subset\Gamma(2)$ be the Galois fundamental group of $Y\setminus f_Y^{-1}(C)$, $\Gamma(Y,\mathcal{M}_kY)$ is naturally isomorphic to the space M_k of modular forms of weight k and level Γ.

Proof. It is certainly well-known that holomorphicity at cusps on $\mathbb{H}$ is equivalent to having at most simple poles at the cusp points of the compactification. Then it remains to observe that in the one-dimensional case logarithmic poles are no different than simple poles.

Let ω_0 and ω_1 be a basis of the vector space $M_1(2)=\Gamma(X(2),\mathcal{M}_1(X(2)))$. These are two meromorphic one-forms on $\mathbb{P}^1$ with no worse than simple poles at the three points of C.

3. Theorem. For each Y and any k there is a sequence natural in Y

$$0\to\mathcal{M}_k(Y)\to\mathcal{M}_{k+1}(Y)\omega_0\oplus\mathcal{M}_{k+1}(Y)\omega_1\to\mathcal{M}_{k+2}(Y)\to 0.$$

Proof. This follows from the earlier theorems since it is true for $X(2)$. Here we could omit the symbols ω_0 and ω_1 and it would not change the truth of the statement, however with them in place we are allowed to interpret $\mathcal{M}_{k+1}\omega_i$ as two subsheaves of $\mathcal{M}_{k+2}$ for $i=0,1$

4. Corollary. For all Y over $X(2)$ and all k the coherent sheaf $\mathcal{M}_k(Y)$ is generated by global sections belonging to the vector space $M_k(2)$. In turn these have basis merely the degree k monomials in ω_0 and ω_1.

The corollary in principle actually answers the question which we stated at the beginning: how to reconstruct the modular forms from rational functions, when in the passage to rational functions the modularity coefficient has cancelled to 1? The issue is that the modularity coefficient always comes from that of the invariant differential forms ω_0, and ω_1. These are modular for the entire group $\Gamma(2)$. The issue then is the *loss* of modularity, and this is due to the fact that during the process of sheafification one multiplies ω_0 and ω_1 by functions which are invariant only for the smaller group Γ_Y.

In other words, the logarithmic forms which are invariant for various subgroups Γ actually come from the two logarithmic forms which are completely invariant for the whole of $\Gamma(2)$, but in the process of sheafification one considers linear combinations in which the coefficients have of course trivial modularity multiplier (they are invariant), but only invariant for the subgroup, not for the whole group.

Before we make this more explicit, let's consider the consequence for generating degrees of Kodaira vanishing, or, what may be simpler merely Serre duality. From the exact sequences we've considered, we obtain passing to global sections, and taking $\Gamma = \Gamma_Y$,

$$0 \to M_0(\Gamma) \to M_1(\Gamma)\omega_0 \oplus M_1(\Gamma)\omega_1 \to M_2(\Gamma) \to H^1(Y, \mathcal{O}_Y) \to 0$$

$$0 \to M_k(\Gamma) \to M_{k+1}(\Gamma)\omega_0 \oplus M_{k+1}(\Gamma)\omega_1 \to M_{k+2}(\Gamma) \to 0, \quad k \geq 1$$

Then for $g = genus(Y)$

5. Lemma. For each $\Gamma = \Gamma_Y \subset \Gamma(2)$, let $x_1, .., x_\alpha \in M_1$ span a complement of the span of ω_0, ω_1. Then the ring $\oplus_k M_k(\Gamma)$ is generated as $M(2)$ module by $x_1, ..., x_\alpha$ together with elements $y_1, ..., y_g \in M_2(\Gamma)$. The vector-space relations in M_2 are that the two subspaces $M_1\omega_0$ and $M_1\omega_1$ intersect in a one dimensional subspace of M_2. Likewise for all $k > 2$ the vector space relations are that $M_{k-1}\omega_0$ and $M_{k-2}\omega_1$ intersect along a subspace of M_k which is isomorphic to M_{k-2}.

Proof. The fact that the sequences are exact for $k \geq 1$ follows from vanishing of $H^1(Y, \Omega_Y(log\ f_y^{-1}C)^{\otimes i})$ for $i \geq 1$. The degree of the relevant divisor is $-(i-1)(2g-2) - i\ degree(f_Y^{-1}C)$. If $g > 0$ the first term is not positive and the second term negative. If $g = 0$ the second (negative) term dominates the first.

We will calculate α in a minute, and also show that these vetctor space relations are the Koszul tautologies in a free module, so the union

$$\{1\} \cup \{x_j : j = 1, ..., .\alpha\} \cup \{y_k\ k = 1, ..., g\}$$

together comprise a free basis for $M(\Gamma)$ as $M(2)$ module. A bit later we'll describe the ring structure.

Since the calculation is similar to what is known as Max Noether's construction of generators for a canonical ring, relying on vanishing theorems, while vanishing theorems for logarithmic differentials in fact of every exterior degree are also well-known, the calculation above should be viewed as an application of standard methods.

It is already included in most textbooks that the dimension of the $M_k(\Gamma)$ can be calculated by Riemann-Roch. Here we are including something about the relations using the ideas that lead into Riemann-Roch. We can double-check the dimensions by writing, just when $\Gamma \subset \Gamma(2)$, that if we write $m_k = dim(M_k(\Gamma))$ we have

$$m_2 = 2m_1 + g - 1$$

$$m_3 = 2m_2 - m_1 = 3m_1 + 2(g-1)$$

$$m_4 = 2m_3 - m_2 = 4m_1 + 3(g-1)$$

$$...$$

$$m_k = km_1 + (k-1)(g-1).$$

This is consistent with $m_k = k[\Gamma(2) : \Gamma] + 1 - g$ from Riemann-Roch if we take $m_1 = [\Gamma(2) : \Gamma] + 1 - g$.

The number and degrees of the generators of $M_k(\Gamma)$ as a free module over $M_k(\Gamma(2))$ follow from these. (They could also be deduced just from Riemann-Roch and suitable vanishing on Y a now that we know that there are module generators in just three degrees but let us proceed more directly.) Letting α, β be the number of module generators of degree 1 and 2 we have

$$dim M_k(\Gamma) = k[\Gamma(2) : \Gamma] + 1 - g$$

$$= (k+1) + \alpha k + \beta(k-1)$$

from which

$$[\Gamma(2) : \Gamma] = 1 + \alpha + \beta$$

$$1 - g = 1 - \beta.$$

Then the genus g is exactly equal to the number of module generators of degree 2, and $\alpha = [\Gamma(2) : \Gamma] - 1 - g$. Let us state this,

6. Corollary. For $\Gamma \subset \Gamma(2)$ the ring $M(\Gamma)$ of modular forms of level Γ is a free module over $M(2)$ with number of generators in each degree as follows:

degree 0:	1
degree 1:	$[\Gamma(2) : \Gamma] - (g+1)$
degree 2:	$g = genus(Y)$
degree ≥ 3:	0

2. Relation with Hodge theory, Galois theory, Poincare duality

The generators $y_1, ..., y_g$ are a basis of $M_2(Y)$ modulo its intersection with the $M(2)$ module spanned by $M_0(Y) \oplus M_1(Y)$ and this g dimensional vector space is naturally isomorphic to $H^{0,1}(Y, \mathbb{C}) = H^1(Y, \mathcal{O}_Y)$. This is also the 'anti-holomorphic part' of $H^1(Y, \mathbb{C})$.

The dual vector space, under the cup product pairing, is the subspace of M_1 consisting of the holomorphic one-forms on Y, naturally isomorphic to the holomorphic part $H^{1,0}(Y, \mathbb{C})$. We can choose our basis $x_1, ..., x_\alpha$ (which comprise a basis of $M_1(Y)$ modulo its intersection with the $M(2)$ span of 1) so that the initial sequence $x_1, ..., x_g$ comprises a dual basis of $y_1, ..., y_g$. under the cup product pairing in $H^1(Y, \mathbb{C})$. Then by degree by degree we have as $M(2)$ module

$$M(Y) \cong M(2) \otimes_{\mathbb{C}} (\mathbb{C} \oplus \mathbb{C}^{[\Gamma(2):\Gamma)]-2g-1} \oplus H^{1,0}(Y, \mathbb{C}) \oplus H^{0,1}(Y, \mathbb{C}))$$

where the first term $\mathbb{C}$ has degree zero and the last $H^{1,0}(Y, \mathbb{C})$ has degree two.

In the case $\Gamma \subset \Gamma(2)$ is normal, letting G be the quotient group, we can define finite-dimensional $\mathbb{C}G$ modules

$$A = \mathbb{C}$$

with trivial G action,

$$B = Kernel(\mathbb{C}^{cusps(Y)} \to \mathbb{C}^{cusps(X(2))}) \oplus H^{1,0}(Y, \mathbb{C}),$$

with action induced by the Galois action on cusps in the first summand and by the Galois action on the holomorphic part of $H^1(G, \mathbb{C})$ in the second summand, and

$$C = H^{0,1}(Y, \mathbb{C})$$

with the Galois action on the antiholomorphic part of cohomology.

It seems clear (proof not yet written down)

7. Theorem. $(f_Y)_*(\mathcal{O}_Y) \cong \mathcal{O}(0) \otimes A \oplus \mathcal{O}(-1) \otimes B \oplus \mathcal{O}(-2) \otimes C$ as coherent sheaf of $\mathbb{C}$G modules on $X(2) = \mathbb{P}^1$.

Also

8. Theorem. There is an equivariant pairing coming from Poincare duality

$$f_{Y*}\mathcal{O}_Y \otimes f_{Y*}\mathcal{O}_Y \to \mathcal{O}(-3)$$

which induces the perfect pairing between $H^{1,0}$ and $H^{0,1}$

4. Analytic description, first notions

The analytic construction of new generators in $M_1(\Gamma)$ and $M_2(\Gamma)$ as we mentioned, does not require finding new diferential forms with more interesting transformation rules than ω_0 and ω_1. Even the various cohomology connecting maps really formalize something elementary. On the projective line $X(2)$ interpret ω_0 and ω_1 as sections of a line bundle; there is one point where each meets the zero section, and these points are distinct, as the line bundle is isomorphic to the one whose section sheaf is $\mathcal{O}(1)$. Then the sheaf $\mathcal{M}_1$ on Y also has two sections, each with vanishing locus only the inverse image of the corresponding point of $X(2)$. The complements of the two inverse images form an open cover of Y and on each part of the open cover the sheaf $\mathcal{M}_1$ restricts to a principal sheaf. The global logarithmic one forms which are invariant for the subgroup Γ can be calculated without using any group theory, they are rational sections in any case and therefore comprise intersection of the rational sections of the two principal sheaves without poles on the open parts.

In fact the same works for any $\mathcal{M}_k$, although it is needed only for $\mathcal{M}_1$ and $\mathcal{M}_2$. It is a matter of repeating what has been said in the previous paragraph using tensor powers $\omega_0^{\otimes k}$ and $\omega_1^{\otimes k}$ in place of ω_0 and ω_1.

Here is how it will work in a little more detail: The basic elements $\omega_0, \omega_1 \in M_1(Y)$ are playing the role of homogeneous coordinates and also playing the role of forms. From an expression of degree k, if you divide by ω_0^k as a coordinate and multiply by ω_0^k as a form, this factorizes an element $M_k(Y)$ as a rational function of degree zero times a form of degree k. It appears at first like it might not be well defined where the denominator is zero, but you can also do the same with ω_1. The only issue is whether the zero locus intersect. This can happen in other situations, like in variables $[a : b : c]$ for the projective plane, a and b are both zero at $[0 : 0 : 1]$. This is what is ruled out by Theorem 3, or anyway just by the fact that $[\omega_0 : \omega_1]$ is always well defined.

The basis of $M(Y)$ has in total $1 + \alpha + g = [\Gamma(2) : \Gamma]$ elements (as many as the covering degree), and so a modular form for Γ is uniquely determined by that many homogeneous polynomials in two variables (but of degrees $k, k-1, k-2$).

In turn, the patching construction expresses each of these in terms of the two basic theta functions. Just ordinary multiplication by a rational function actually does something like the averaging that happens in Eisenstein series or theta characteristics. Since that works for every modular function it must be the most general construction.

To finally summarize what is the main lesson: that in constructing all the modular forms, it is never necessary to use any logarithmic *forms* except the original ω_0 and ω_1 which have invariance for the whole of $\Gamma(2)$. And the patching uses coefficient functions invariant for the smaller group Γ; in the process some invariance is lost. But it is never necessary to find in any other way, logarithmic forms which have any interesting transformation group, or are invariant by any but the largest finite index subgroup of $\Gamma(2)$. We will describe the patching explicitly in section 6.

Also note that if one applying these theorems in families of curves, the initial Theorem 1 will be nearly unchanged, and one will use that the restrictin of logarithmic differentials along a transverse slice are logarithmic differentials of lower dimension.

We'll give an explicit proof of this later:

9. Theorem Let λ denote the usual holomorphic λ function $\mathbb{H} \to \mathbb{C}$. Every modular form of any weight k and any level has two expressions, one as an algebraic function of $\lambda(\tau)$ times a power of $\theta(0,\tau)^4$ and one as an algebraic function of $\lambda(\tau)$ times a power of $\theta(\frac{1}{2},\tau)^4$. At every point of the modular curve one or the other of the algebraic functions is holomorphic; therefore the order of poles of the corresponding one-forms on the modular curve do not exceed those of $d\tau^{\otimes k}$ itself (which has a pole of order k at each cusp).

5. Remarks about cohomology of the $\mathcal{M}_i(Y)$

Let's explain a little more about the cohomology before proceeding on. Since $R_i f_{Y*} = 0$ for $i \geq 1$ we may calculate for i, k

$$H^i(Y, \mathcal{M}_k(Y)) = H^i(X(2), f_{Y*}\mathcal{M}_k).$$

From the previous results for $g = genus(Y)$

$$f_{Y*}\mathcal{M}_k(Y) \cong \mathcal{O}(k) \oplus \mathcal{O}(k-1)^{\oplus\alpha} \oplus \mathcal{O}(k-2)^{\oplus g}.$$

with α as before, and therefore for $k = 0, 1, 2, ...$

$$dim H^0(Y, \mathcal{M}_k(Y)) = 1,\ 2+\alpha,\ 3+2\alpha+g,\ 4+3\alpha+2g, ...$$

while

$$dim H^1(Y, \mathcal{M}_k) = g, 0, 0, ...$$

the latter also makes sense for $k = -1, -2, -3, ...$ giving $2g+\alpha,\ 3g+2\alpha+1, 4g+3\alpha+2, ...$

The direct sum $\oplus_{k=0}^{\infty}\mathcal{M}_k(Y)$ is the pushforward to Y of the sheaf of functions on the quasiprojective surface L which is the dual line bundle $\Omega_Y(\widehat{log\ f^{-1}C})$. We may assemble together the exact sequences we were considering earlier to a single exact sequence

$$0 \to \mathcal{O}_L(2Y) \to \mathcal{O}_L(Y)\omega_0 \oplus \mathcal{O}_L(Y)\omega_1 \to \mathcal{O}_L \to 0.$$

The reason we are allowing poles of degree $2, 1, 0$ on Y becomes clear if we push the sheaves down to $X(2)$ to examine them. Writing the degrees $k = 2, 1, 0$ in vertical order on the page

$$\begin{array}{ccccccccccc}
0 & \to & \begin{pmatrix} & \mathcal{O}(0) & \\ \oplus & \mathcal{O}(-1)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-2)^{\oplus g} & \end{pmatrix} & \to & \begin{pmatrix} & \mathcal{O}(1) & \\ \oplus & \mathcal{O}(0)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-1)^{\oplus g} & \end{pmatrix}\omega_0 \oplus \begin{pmatrix} & \mathcal{O}(1) & \\ \oplus & \mathcal{O}(0)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-1)^{\oplus g} & \end{pmatrix}\omega_1 & \to & \begin{pmatrix} & \mathcal{O}(2) & \\ \oplus & \mathcal{O}(1)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(0)^{\oplus g} & \end{pmatrix} & \to & 0 \\
0 & \to & \begin{pmatrix} & \mathcal{O}(-1) & \\ \oplus & \mathcal{O}(-2)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-3)^{\oplus g} & \end{pmatrix} & \to & \begin{pmatrix} & \mathcal{O}(0) & \\ \oplus & \mathcal{O}(-1)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-2)^{\oplus g} & \end{pmatrix}\omega_0 \oplus \begin{pmatrix} & \mathcal{O}(0) & \\ \oplus & \mathcal{O}(-1)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-2)^{\oplus g} & \end{pmatrix}\omega_1 & \to & \begin{pmatrix} & \mathcal{O}(1) & \\ \oplus & \mathcal{O}(0)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-1)^{\oplus g} & \end{pmatrix} & \to & 0 \\
0 & \to & \begin{pmatrix} & \mathcal{O}(-2) & \\ \oplus & \mathcal{O}(-3)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-4)^{\oplus g} & \end{pmatrix} & \to & \begin{pmatrix} & \mathcal{O}(-1) & \\ \oplus & \mathcal{O}(-2)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-3)^{\oplus g} & \end{pmatrix}\omega_0 \oplus \begin{pmatrix} & \mathcal{O}(-1) & \\ \oplus & \mathcal{O}(-2)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-3)^{\oplus g} & \end{pmatrix}\omega_1 & \to & \begin{pmatrix} & \mathcal{O}(0) & \\ \oplus & \mathcal{O}(-1)^{\oplus\alpha} & \\ \oplus & \mathcal{O}(-2)^{\oplus g} & \end{pmatrix} & \to & 0
\end{array}$$

Each column is a pushdown from L to Y and each pair of parentheses contains a pushdown from Y to $X(2)$. The fact that allowed poles have order 2,1,0 on Y reading left to right creates zeroes on $X(2)$ of the same order once the sheaves are pushed forward. The g dimensional cokernel in the top row comes from $H^1(X(2), \mathcal{O}(-2)^{\oplus g})$ on the left side.

The cokernel of the right map after taking global sections is a finite dimensional graded algebra with basis $1, x_1, ..., x_\alpha, y_1, ..., y_g$ which results when a term $\mathcal{O}(i)$ with $i \geq 0$ in the right column sits next to a term $\mathcal{O}(i)$ with $i < 0$, and otherwise the sequences are exact. The algebra $M(Y)$ is a flat deformation over $M(2)$ of this finite dimensional algebra over $\mathbb{C}$.

6. Analytic continuation from ring identities

In this section we'll show in detail how to represent each element of $M_k(Y)$ as an analytic function $\mathcal{H} \to \mathbb{C}$ using patching, assuming two things: that the structure of $M(Y)$ as a ring is known and that once the coefficients of a polynomial in one variable are known analytically so are the roots.

Abstractly, for $Y = X(\Gamma)$ and $\Gamma \subset \Gamma(2)$, once we take the numbers

$$\alpha = [\Gamma(2) : \Gamma] - (g+1),$$

$$g = genus(Y),$$

then any sequence

$$c_0;\ d_1, ..., d_\alpha;\ h_1, ..., h_g$$

consisting of polynomials in two variables u_0, u_1, with

$$degree(c_0) = k$$

$$degree(d_i) = k-1$$

$$degree(h_i) = k-2,$$

determines, bi-uniquely, an element f of $M_k(Y)$ which is given

$$f = c_0 + d_1 x_1 + ... + d_\alpha x_\alpha + h_1 y_1 + ... + y_g y_g,$$

upon replacing u_0, u_1 by ω_0, ω_1. Here $1, x_1, ..., x_\alpha, y_1, ..., y_g$ is the $M(2)$-module basis of $M(Y)$.

Let's explain how the sequence of polynomials now creates an actual entire holomorphic function

$$\mathbb{H} \to \mathbb{C}$$

which satisfies the modular identity of weight k and level Γ.

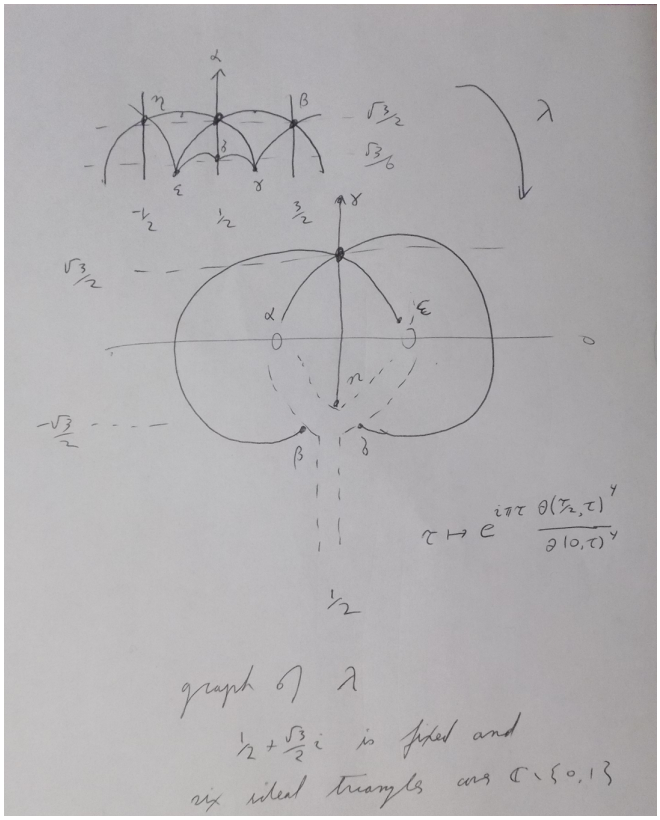
As $\Gamma(2)$-invariant forms on the upper half plane, our ω_0 and ω_1 may be taken to be

$$\omega_0 = \theta(0, \tau)^4 d\tau$$

$$\omega_1 = \theta(\frac{1}{2}, \tau)^4 d\tau$$

The quotient $\frac{\omega_1}{\omega_0}$ is a $\Gamma(2)$ invariant holomorphic function $\mathbb{H} \to \mathbb{C}$ which equals[1] $1 - \lambda(\tau)$, with λ the holomorphic λ function $\mathbb{H} \to \mathbb{C}$. Thus it descends to a meromorphic function on $X(2)$. This amounts to an isomorphism $X(2) \to \mathbb{P}^1$. If $\mathbb{P}^1$ is considered to have homogeneous coordinates $[u_0 : u_1]$ we may identify this with $[\omega_0 : \omega_1]$.

Here is a drawing of $\lambda(\tau) = 1 - \frac{\theta(1/2,\tau)^4}{\theta(0,\tau)^4}$



[1] $1 = \frac{e^{i\pi\tau}\theta(\tau/2,\tau)^4}{\theta(0,\tau)^4} + \frac{\theta(1/2,\tau)^4}{\theta(0,\tau)^4} = \lambda(\tau) + \frac{\theta(1/2,\tau)^4}{\theta(0,\tau)^4}$ by Jacobi's sum formula

For clarity, let's use the letters u_0, u_1 when we are speaking about the ring $M(2)$ algebraically, so we write $M(2) = \mathbb{C}[u_0, u_1]$ a polynomial algebra.

Theorem 3 shows that the internal sum

$$\mathcal{M}_0(Y)\omega_0 + \mathcal{M}_0(Y)\omega_1 \qquad (2)$$

is locally free. Note that $\mathcal{M}_0(Y)$ is the structure sheaf of Y, and it follows that u_0, u_1 span the locally free (in fact ample) sheaf $\mathcal{M}_1(Y)$ of rank one.

Another way of thinking about this is just to say that the ratio $[\omega_0 : \omega_1]$ is well-defined at all points of Y. That is, the inclusion $M(2) \subset M(Y)$ is unlike the inclusion $C[u, v] \subset C[u, v, w]$ representing an inclusion $\mathbb{P}^1 \subset \mathbb{P}^2$ where $[u : v]$ becomes undefined at $[0 : 0 : 1]$. Let

$$r_0 = \frac{f}{u_0^k} \in M(Y)[1/u_0]$$

$$r_1 = \frac{f}{u_1^k} \in M(Y)[1/u_1]$$

Interpret ω_0, ω_1 then as weighted homogeneous coordinates u_0, u_1 of degree one, but only within the *degree zero rational functions* r_0, r_1; and elsewhere write instead

$$\omega_0^{\otimes k} = \theta(0, z)^{4k}(d\tau)^{\otimes k}$$

$$\omega_1^{\otimes k} = \theta(1/2, z)^{4k}(d\tau)^{\otimes k}, \qquad (3)$$

so that

$$r_0\theta(0, \tau)^{4k} = r_1\theta(1/2, \tau)^{4k} \qquad (4)$$

wherever both are defined.

Removing $(d\tau)^{\otimes k}$ in passing from (3) to (4) converts invariance to modularity for the whole of the group $\Gamma(2)$. Since r_0 and r_1 are well defined meromorphic functions on Y they are invariant on $\mathbb{H}$ but only for the action of the smaller group Γ. The product functions on both sides of equation (4) therefore have modularity of weight k for the level Γ.

10. Theorem. Local freeness of the internal sum (2) implies that u_0, u_1 have no common zeroes on Y as sections of $\mathcal{M}_1(Y)$. Then the denominators u_0^k, u_1^k have no common zero in the locally free sheaf $\mathcal{M}_k(Y)$of which $r_0\omega^{\otimes k}$, $r_1\omega^{\otimes k}$ are rational sections. Starting with the left side of (4), interpeting $r_0 = \frac{f}{u_0^{degree(f)}}$ and $r_1 = \frac{f}{u_1{}^{degree(f)}}$ as algebraic functions of $\lambda(\tau)$, these have no common poles on boundary points of $\mathbb{H}$ lying over cusps of Y, and the same equation (4) then furnishes an analytic continuation to a modular function of weight k and level Γ which is holomorphic at the cusps (as is $d\tau$ itself). The corresponding k-fold one-forms $r_1^k\theta(1/2,\tau)^{4k}d\tau^{\otimes k}$ and $r_0^k\theta(0,\tau)^{4k}d\tau^{\otimes k}$ patch together to comprise well-defined meromorphic k-fold one-form on Y (now defined as a meromorphic function on all cusps) holomorphic everywhere except at the cusps, where the poles do not exceed those of order $d\tau^{\otimes k}$, namely do not exceed order k at any cusp.

The combination of constructing the ring extension $M(2) \subset M(Y)$ algebraically and then gluing in this manner must be the common generalization of special methods such as Eisenstein series and theta characteristics, in their application to constructing modular functions. A more simpe corollary not referring to analytic continuation or to $\theta(1/2,\tau)$ is this:

11. Corollary The ring $M(Y)$ is isomorphic to the ring of functions $\frac{f}{u_0^{degree(f)}}\theta(0,\tau)^{4\ degree(f)} : H \to \mathbb{C}$, where we regard $\frac{f}{u_0^{degree(f)}}$ as an 'algebraic function' of $\lambda(\tau)$. Although the $\frac{f(\lambda(\tau))}{u_0^{degree(f)}}$ can have poles points of the boundary of $\mathbb{H}$ lying over the cusps in Y these are removable in the product $\frac{f}{u_0^{degree(f)}}\theta(0,\tau)^{4\ degree(f)}$. The k-fold one-form $r_1^k\theta(0,\tau)^{4k}d\tau^{\otimes k}$ descends to a one-form on Y with poles at cusps and any of order larger than k at any cusp are 'removable.'

12. Remark. For levels which are not above level 2, one may pass to a subring by a group action. For example, the ring $M(1)$ is the invariants of the reflection group S_3, and because it is a subring all its elements already have been interpreted as analytic modular functions on $\mathbb{H}$.

It might be instructive to look at one explicit consequence of the situation where one has polarized all the Y by logarithmic forms (compatibly with the transition maps). In terms of our coordinates u_0, u_1 we might write

$$\omega_0 = \frac{u_1}{u_1 - u_0} d(\frac{u_0}{u_1})$$

$$\omega_1 = \frac{u_0}{u_0 - u_1} d(\frac{u_1}{u_0}).$$

The fact that we can use ω_0, ω_1 as homogeneous coordinates corresponds to the fact that the ratio between the right sides of these equations is the same as u_0/u_1 itself.

7. The types of Y for each g and c.

The passage from the subgroup $\Gamma \subset \Gamma$ to the over-ring $M(Y) \supset M(X(2))$ can be considered to come from the map from cohomology of a wedge of two circles to K_0 of the Riemann sphere

$$H^1(F_2, S_d) \to H^1(\mathbb{P}^1, Gl) \subset K_0(\mathbb{P}^1) = \mathbb{Z}[T, T^{-1}]$$

with S_d the permutation group and T the class of $\mathcal{O}(1)$. The map is not directly induced by functoriality of cohomology.

Although $K_0(\mathbb{P}^1)$ is not a finitely generated free abelian group, the mage of all the $H^1(F_2, S_d)$ are all totally contained in the rank 3 free abelian group

$$\mathbb{Z} + \mathbb{Z}T^{-1} + \mathbb{Z}T^{-2}$$

and a class γ of a connected Riemann surface is sent to

$$1 + \alpha(\gamma)T^{-1} + g(\gamma)T^{-2}$$

where $\alpha(\gamma) = d - 1 - g(\gamma) = c - 3 + g(\gamma)$ where $g(\gamma)$ is the genus of the associated modular curve Y and c is its number of cusps.

Thus

13. Theorem. The class in $K_0(\mathbb{P}^1)$ depends exactly on the number of cusps and the genus. Thus two classes $\gamma_1, \gamma_2 \in H^1(F_2, S_4)$ of connected Riemann surfaces $\mathbb{H}/\Gamma_1$ and $\mathbb{H}/\Gamma_2$ map to the same element of $K_0(\mathbb{P}^1)$ if and only if they are homeomorphic (=topologically isomorphic).

Now there is the issue of going back, starting from the number of cusps and the genus, to actually build the algebraic structure of all possible rings $M(Y)$.

Here is how it probably works if Γ is normal so we have a Galois group $G = \Gamma(2)/\Gamma$, and the we assume that we know how the finite group G acts on three finite dimensional vector spaces which we defined earlier

$A = \mathbb{C}$ with trivial G action,

$B = Kernel(\mathbb{C}^{cuspsY} \to \mathbb{C}^{cuspsX(2)}) \oplus H^{1,0}(Y, \mathbb{C})$, with action on the first summand induced by the permutation of cusps, action on the second induced by the inclusion of the holomorphic part of the cohomology of Y,

$C = H^{0,1}(Y, \mathbb{C})$, by the induced action on anti-holomorphic cohomology.

Then

$$A \oplus B \oplus C \cong \mathbb{C} \oplus [\mathbb{C}^{cusps-3} \oplus H^{1,0}(Y, \mathbb{C})] \oplus H^{0,1}(Y, \mathbb{C})$$

and we stated in Theorem 7 that the locally free sheaf

$$A \otimes \mathcal{O}(0) \oplus B \otimes \mathcal{O}(-1) \oplus C \otimes \mathcal{O}(-2),$$

on $X(2) \cong \mathbb{P}^1$ is isomorphic to

$$(f_Y)_* \mathcal{O}_Y$$

as a coherent sheaf with G action.

Let V be the rank d vector bundle on $\mathbb{P}^1$ with this sheaf of sections.

The dual bundle $\widehat{V} \to \mathbb{P}^1$ can be described point-by-point as follows: A point $p \in \mathbb{P}^1$ has a defining ideal sheaf $\mathcal{I}_p \subset \mathcal{O}_{\mathbb{P}^1}$; up to isomorphism $\mathcal{I}_p \cong \mathcal{O}(-1)$ though note there is not a natural unique isomorphism (as $\mathcal{O}(-1)$ depends non-functorially on $\mathbb{P}^1$ unlike its square the canonical sheaf). Once p is chosen, the fiber of V over p is the $1 + \alpha(Y) + g(Y)$ dimensional vector space

$$f_{Y*}\mathcal{O}_Y \otimes_{\mathbb{P}^1} \mathcal{O}_{\mathbb{P}^1}/\mathcal{I}.$$

A point $y \in Y$ such that $f_Y(y) = p$ gives an evaluation map to the one-dimensional vectot space $\mathcal{O}_{\mathbb{P}^1}/\mathcal{I} \cong \mathbb{C}$. Thus evaluation at y is a point of the dual vector bundle $\widehat{V}$ in the fiber over p.

14. Theorem. The vector bundle $\widehat{V} \to \mathbb{P}^1$ includes a Galois invariant multi-section of order k, which spans $\widehat{V}$ at every fiber except above $0, 1, \infty$. The (normalization of) the the multisection is isomorphic to Y.

From the multi-section we can get back the holomorphic modular functions $\mathbb{H} \to \mathbb{C}$ like this:

The algebaic curve Y has that every unramified fiber F is linearly equivalent to $K_Y + f^{-1}(C)$ for $C = 0, 1, \infty$.

It can be polarized either way (it doesn't matter) and the corresponding graded ring is $M(Y)$.

From $M(Y)$ which contains u_0, u_1 we have that by assigning u_1/u_0 to the lambda function $\lambda : \mathbb{H}- \to \mathbb{C}$ we can for each element f of M_k write

$$(\frac{f}{u_0^k})\theta(0, z)^{4k}$$

and the first term is a rational function of $\lambda(\tau)$, the second a holomorphic function $\mathbb{H} \to \mathbb{C}$, and the product is modular of weight k and level Γ, and all but order k poles at the cusps are removable as it equals

$$(\frac{f}{u_1^k})\theta(1/2, z)^{4k}$$

whenever both are defined.

The issue is then finding all the G invariant multisections of $\widehat{V} \to \mathbb{P}^1$ if there is more than one. There is likely a G invariant singular foliation of $\widehat{V}$ (the flat connection on the complement of $\{0, 1, \infty\}$) which has these as the compact (smooth) leaves.

8. Primality tests

The integer lattice points (x, y) satisfying $x, y \geq 1$, $m - \frac{1}{2} \leq xy \leq m + \frac{1}{2}$ correspond to divisors of m, for any natural number $m \geq 1$. The number of divisors is equal to $\frac{1}{2\pi i}$ times the value of the contour integral along a path surrounding the same finite set of points, of the logarithmic derivative of any holomorphic function with suitable domain of definition and which has a simple zero at each such lattice point.

Except for the choice of path of integration, the fundamental theorem of calculus indicates that the logarithmic derivative integrates to zero; choosing which points the path should wind around is identical to adding 1 for each point.

We transform such a path into a straight line by the conformal transformation of squaring a complex number. Intepret x, y as the real and imaginary coordinate in the complex plane. The divisors of a number m are bijective with the square Gaussian integers with imaginary part $2m$, and so using the principal square root function (with values in the upper half plane) write the series involving the (third) Jacobi theta function

$$\theta(\sqrt{z} + (\frac{i+1}{2}), i) = \sum_{n=-\infty}^{\infty} e^{2\pi i n(\sqrt{z} + (\frac{i+1}{2})) - \pi n^2}$$

$$= \sum_{n=-\infty}^{\infty} (-1)^n e^{-\pi n(n+1)} e^{2\pi i n \sqrt{z}} \quad (1)$$

Because $\theta(z + (\frac{i+1}{2}), i)$ has a simple zero at each Gaussian integer, we have

15. Proposition. The logarithmic derivative of (1) integrated from $-\infty$ to $-1/2$ along a horizontal line at imaginary level t has a discontinuous jump when t passes $2m$ of magnitude equal to $2\pi i$ times the number of divisors of m which are strictly less than $\sqrt{m}$. The smallest jump, by only a value of $2\pi i$, occurs if and only if m is prime or a square of a prime.

The integral can of course only be taken along the interval $[-m^2, -1/2]$, and if the sum is taken only from $-m-1$ to $m+1$ there results a finite trigonometric expression which likely has the zeroes only slightly displaced, and the change of the value of the integral between two half-integer values of t should still determine the number of divisors of m less than $\sqrt{m}$ to the nearest integer.