Nine notes on modular forms

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 invariance 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 multiplier 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 multiplier 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 $Y \setminus f_Y^{-1}(C)$, $\Gamma(Y, \mathcal{M}_k Y)$ 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 multiplier coefficient has cancelled to 1? The issue is that the multiplier 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 invariance, 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.

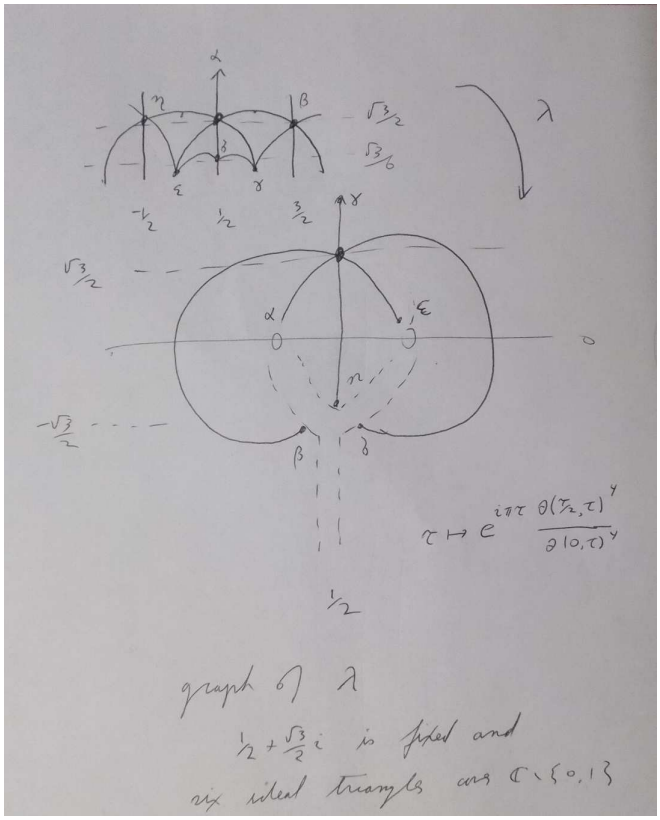
As $\Gamma(2)$-invariant forms on the upper half plane, our ω_0 and ω_1 may be taken to be

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

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

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 a rational map $\mathbb{P}^2 - \to \mathbb{P}^1$ indeterminate 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 $r_0^k\theta(0,\tau)^{4k}d\tau^{\otimes k}$ 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 $\Gamma \subset \Gamma(2)$ 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 $K_0(\mathbb{P}^1)$, using only relations from direct sums of vector bundles, 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. 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 such an integral between two nearby values of t should still determine the number of divisors of m less than $\sqrt{m}$ to the nearest integer.

Lefshetz numbers of modular curves

This note is to provide some evidence for some of the poorly proven remarks in *analytic primality testing*; it seems easiest if we organize our thinking in terms of this problem: Suppose we are given an arbitrary finite two-generator group $G = \langle g, h\rangle$ with fixed generators. This corresponds to a Galois modular[2] curve $f_Y : Y \to \mathbb{P}^1$ branched over at most $\{0, 1, \infty\}$ with $g, h, (gh)^{-1}$ the monodromy transformations corresponding to disjoint based loops about these points. Then each element $x \in G$ is an automorphism of Y and the problem is to determine the Lefshetz number of the automorphism x.

The following theorem is elementary, but we'll give a proof not relying on either the classical (transcendental) topology or the etale topology.

1. Theorem. The Lefshetz number of each such automorphism x is equal to the number of fixed cusps of x minus the number of fixed points in a general fiber of the finite covering map f_Y.

An arbitrary modular curve of level $\Gamma \subset \Gamma(2)$ is determined by such a group G and a choice of subgroup $H \subset G$. The group G arises as the reduction of $\Gamma(2)$ modulo the intersection of the conjugates of Γ while Γ is the inverse image of H.

From this cursory observation we can determine at least the genus of the modular curve $H\backslash Y$ corresponding to each choice of H, and one might expect that with finer analysis one could approach two known theorems, the Taniyama conjecture that any elliptic curve with rational j invariant has for some N a branched cover by $X_0(2N)$, the case when $G \subset SL_2(\mathbb{Z}/2N\mathbb{Z})$ is the finite group of g so that $g - 1$ has even entries and $H \subset G$ the upper triangular subgroup, and the Belyi theorem which says that any curve Y which is genuinely one-dimensional (defined over a number field) arises from some such arithmetic Γ without the requirement of containing a congruence subgroup.

[2]Let's call a curve 'modular' even if the arithmetic group Γ contains no congruence subgroup

Let's begin by comparing the residue map for $X(2) = \mathbb{P}^1$ with the same map for Y. Although we're talking about one-forms with simple poles, we'll continue our convention of referring to these as logarithmic poles since most of what we'll say has a higher dimensional analogue. Using our earlier notation, so $\mathcal{M}(Y)$ is the sheaf $\Omega_Y(log\ f_Y^{-1}C)$ with $C = \{0, 1, \infty\}$, $\mathcal{M}_k(Y) = \mathcal{M}(Y)^{\otimes k}$, and $M_k(Y) = \Gamma(Y, \mathcal{M}_k(Y))$.

The map $f_Y : Y \to \mathbb{P}^1$ induces the commutative diagram

$$\begin{array}{ccc} \Omega_{\mathbb{P}^1}(log\ C) & \to & \mathcal{O}_C \\ \downarrow & & \downarrow \\ \Omega_Y(log\ f_Y^{-1}C) & \to & \mathcal{O}_{f_Y{}^{-1}C} \end{array}$$

which we write in other notation

$$\begin{array}{ccc} \mathcal{M}(\mathbb{P}^1) & \to & \mathbb{C}^3 \\ \downarrow & & \downarrow \\ \mathcal{M}(Y) & \to & \mathbb{C}^c \end{array}$$

with c the number of cusps. The kernel of each horizontal residue map is just the holomorphic forms in each case.

Taking global sections and passing to the cokernels of the vertical maps (since there are no nonzero global holomorphic one-forms on $\mathbb{P}^1$) gives the exact sequence of finite-dimensional $\mathbb{C}G$ modules

$$0 \to H^{1,0}(Y, \mathbb{C}) \to \frac{M_1(Y)}{M_1(2)M_0(Y)} \to I_{G/\langle g\rangle} \oplus I_{G/\langle h\rangle} \oplus I_{G/\langle gh\rangle} \to 0$$

where $I_{G/\langle h\rangle}$ is the augmentation kernel of the addition map $\mathbb{C}[G/\langle h\rangle] \to \mathbb{C}$. Here we use standard notation $M_1(2)$ for $M_1(X(2))$ with $X(2)$ the modular curve we're calling $\mathbb{P}^1$. So that we can interpret the global holomorphic forms on Y with at most logarithmic poles on C to be modular forms of weight 1 (using Dolgachev's numbering convention) modulo those of level $\Gamma(2)$.

Note that this is consistent with our earlier calculation for the dimension of $\frac{M_1(Y)}{M_1(2)M_0(Y)}$ as what we called $\alpha = g + c - 3$, and we had stated without proof that the group action should be thus.

If we now add $H^{0,1}$ to the first and second term of the exact sequence we have of course by the Dolbeault decomposition and by our previous calculations involving M_2

$$0 \to H^1(Y,\mathbb{C}) \to \frac{M_2(Y)}{M_2(2)M_0(Y)+M_1(2)M_1(Y)} \oplus \frac{M_1(Y)}{M_0(Y)M_1(2)}$$

$$\to I_{G/\langle g\rangle} \oplus I_{G/\langle h\rangle} \oplus I_{G/\langle gh\rangle} \to 0.$$

If we now add a copy of $H^{0,0}(Y,\mathbb{C})\oplus H^{1,1}(Y,\mathbb{C})\oplus M_0(Y) \cong \mathbb{C}\oplus\mathbb{C}\oplus\mathbb{C}$ to the middle term and the last term, the last term becomes a permutation module and there is the exact sequence

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

$$\to \frac{M_2(Y)}{M_2(2)M_0(Y)+M_1(2)M_1(Y)} \oplus \frac{M_1(Y)}{M_0(Y)M_1(2)} \oplus M_0(Y) \oplus H^{0,0}(Y,\mathbb{C}) \oplus H^{1,1}(Y,\mathbb{C})$$

$$\to \mathbb{C}[G/\langle g\rangle] \oplus \mathbb{C}[G/\langle h\rangle] \oplus \mathbb{C}[G/\langle gh\rangle] \to 0$$

we can interpret the last term as the (global sections of) $\mathcal{O}_{f_Y^{-1}(C)}$.

If we choose any point $p \in X(2)$ which is not a cusp, we have the equivariant isomorphism of finite dimensional vector spaces, writing $\mathbb{C}$ as the reduction of $M(2)$ modulo its augmentation ideal, since the ring of modular forms $M(Y)$ is a free module over $M(2)$, with sections of number $1, \alpha, g$ in degree $0, 1, 2$, and since the group action lifts from the projective curve Y to the ring $M(Y)$ by naturality of logarithmic differentials and equivariance of cusps, we have $\mathbb{C} = \frac{M(2)}{M_1(2)\oplus M_2(2)\oplus M_3(2)\oplus\dots}$ so

$$M(Y)\otimes_{M(2)}\mathbb{C} \cong \frac{M_2(Y)}{M_2(2)M_0(Y)+M_1(2)M_1(Y)} \oplus \frac{M_1(Y)}{M_0(Y)M_1(2)} \oplus M_0(Y).$$

Thus, the three terms in the exact sequence involving the letter M aount to the same as

$$\mathcal{O}_{f_Y^{-1}(p)} = (f_{Y*}\mathcal{O}_Y)_p$$

under the free and transitive Galois action on $f_Y^{-1}p$.

(The same argument might be found more theoretically, not depending on graded rings of modular forms, by choosing finding a G invariant half-canonical divisor $\frac{1}{2}K$ and considering that the same vector space is equivariantly isomorphic also to $\Gamma(Y, \mathcal{O}_Y(f_Y^{-1}p + \frac{1}{2}K))$. action which is generically merely a free and transitive action.)

In any case now we our G equivariant exact sequence becomes

$$0 \to H^1(Y, \mathbb{C}) \to H^0(Y, \mathbb{C}) \oplus H^2(Y, \mathbb{C}) \oplus \mathcal{O}_{f_Y^{-1}(p)} \to \mathcal{O}_{f_Y^{-1}(C)} \to 0, \quad (1)$$

with $\mathcal{O}_{f_Y^{-1}(p)}$ the finite-dimensional permutation representation based on the general fiber $f_Y^{-1}(p)$ and $\mathcal{O}_{f_Y^{-1}(C)}$ the finite-dimensional permutation representation on the set of cusps of Y.

This argument currently represents a rather vague interpolation between sheaves and their global sections; nevertheless it seems to be correct in a few examples. For $X(2)$ itself it says that the Euler characteristic should be the number of cusps minus the covering degree; this is $3 - 1 = 2$.

In the course of the proof, we saw that the augmentation ideal I_G of G contains a direct sum of a copy of the holomorphic and antiholomorphic one-forms (though the embedding of antiholomorphic one-forms is as a complex vector subspace), and the quotient of I_G modulo both is the direct sum of the augmentation subspace of the three orbits of G on the cusps of Y, or if you like it is the vector space spanned by cusps modulo the three-dimensional G-invariant subspace.

From this it is easy to determine in a uniform way the genus of modular curves corresponding to subgroups $H \subset G$, as we may obtain the direct sum of holomorphic and antiholomrorphic one-forms of the corresponding modular curve as the H invariant subspace of the kernel of the map from I_G to the linear span of the cusps. The image of I_G is always anyway a complement of the three dimensional space of G invariant linear combinations of cusps.

Just to repeat this,

2. Theorem. There is a (real) vector space isomorphism between the direct sum of the holomorphic and anti-holomorphic one-forms of the modular curve corresponding to $H \subset \langle g, h \rangle$ and the complex vector subspace of the augmentation ideal $I_G \subset \mathbb{C}G$ consisting of the H invariant elements of the inverse image under the residue map from I_G to the vector space based on the cusps of the modular curve corresponding to $H = 1$, of the vector subspace of dimension three based on the three G orbit sums.

Here we have very crudely allowed ourselves to decompose and twist what should really be nicely symmetrical vector bundles in order to be able to talk about global sections. Really for an elliptic curve a basic holomorphic one-form and a basic antiholomorphic one-form correspond to basic generators for summands of $f_{Y*}\mathcal{O}_Y$ and as we have seen, the dual of the vector bundle V which has this sheaf of sections contains Y itself as a multisection (perhaps after needing to normalize).

There might be a possibility to prove the Lefschetz formula in the ordinary way, by the Lefshetz fixed point theorem. However, one would need to explain why cusps count positively while points of a G orbit count negatively.

Example. Consider a degree two cover of $X(2)$ branched at $0, 1$ and follow this by a degree two cover branched at $0, 1$ and the two copies of ∞. Let x of order two generate the Galois automorphism of the second cover. Then g has four fixed cusps and no fixed general point, so the Lefschetz trace of the action on the elliptic curve is 4. On the other hand the identity element fixes four cusps and four general points, having then Lefshetz trace 0.

We can use the Lefshetz trace calculation to identify the genus of each subgroup of our Galois group $G = C_2 \times C_2$.

In fact the theorem below is true of any finite two-generator group $\langle g, h \rangle$ and therefore for any modular curve of level higher than two

3. Theorem. For $H \subset \langle g, h\rangle$ with $\langle g, h\rangle$ finite, the genus of the corresponding modular curve is equal to 1 minus half the average number of fixed cusps plus half the average number of fixed points in a general fiber of the branched cover (where the average is taken over the elements of H).

The interesting thing about this theorem is that the exact sequence (1) provides two proofs. If we use the fact that the number of H orbits on cusps and a general fiber is in each case the average trace, the equation says that the genus of X is 1 minus half the number of cusps plus half the covering degree of $X \to \mathbb{P}^1$, which calculates the transcendental Euler characteristic of X. But if we use the fact (1) that $1 - \frac{1}{2}L(h)$ is the trace of h acting on holomorphic one-forms of Y, the average calculates the dimension of the space of H invariant holomorphic one-forms, which are the holomorphic one-forms on X.

Remark. As is easily seen from the theorem above, or directly, the genus of $H \setminus Y$ is thus determined from only G, H, g, h. It is 1 minus one-half of the average over $x \in H$ of the number of $y \in G$ such that $yxy^{-1} \in \langle g\rangle$ minus one half the average over $x \in H$ of the number of $y \in G$ such that $yxy^{-1} \in \langle h\rangle$ minus one-half the average over $x \in H$ of the number of $y \in G$ such that $yxy^{-1} \in \langle gh\rangle$ plus one-half the average over $x \in H$ of the number of $y \in G$ such that $yxy^{-1} \in \langle 1\rangle$.

This seems reminiscent of calculations involving triangle groups, though it applies to any Fgroup Γ of finite index in $\Gamma(2)$, being the inverse image of H under $\Gamma(2) \to \Gamma(2)/\Gamma(Y) = G$.

The last number calculated is just one-half of the index $[G : H]$.

Now that the value of the Lefshetz trace on Y gives the correct values of the genus of X, we have some confidence in the correctness of the previous results. The sequel will consider actions of Galois automorphisms and rational points.

Grothendieck sections and rational points of modular curves

Let's continue our convention of calling a 'modular curve' a quotient $\Gamma \setminus \mathbb{H}$ for Γ arithmetic, and we'll consider the case when $\Gamma \subset \Gamma(2)$ contained in the free group of rank two which expresses $\mathbb{H}$ as the universal cover of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. Reducing modulo the intersection of the conjugates of Γ the inclusion $\Gamma \subset \Gamma(2)$ becomes an inclusion $H \subset \langle g, h \rangle$ of a subgroup of a two-generator finite group, and correspondingly there is the Galois cover $Y \to X$ with $X = H \setminus Y$. In the previous note we made an exact sequence equivariant for the continuous H action

$$0 \to H^1(Y) \to H^0(Y) \oplus H^2(Y) \to \mathcal{O}_{f_Y{}^{-1}p} \to \mathcal{O}_{f_Y{}^{-1}C} \to 0$$

for p any chosen point of $\mathbb{P}^1$ and $C \subset \mathbb{P}^1$ the set of cusps viewed as ideal points.

Suppose now that we have chosen a number field K Galois over $\mathbb{Q}$ so that we may interpret a suitably symmetric form of Y as a scheme flat of finite type over the integers $\mathcal{O}_K$, and an embedding or 'complex place' $\mathcal{O}_K \subset \mathbb{C}$ through which we may recover the complex manifold which we previously called Y as the set of complex points $Y(\mathbb{C})$ of the scheme Y.

Now the group $Aut_{\mathbb{P}^1(\mathbb{Z})}(Y)$ of automorphisms over the *integer* projective line acts on Y and surjects onto the Galois group $Gal(K/\mathbb{Q}) = Aut(\mathcal{O}_K)$. Thus the group $Aut_{\mathbb{P}^1(\mathbb{Z})}$ mixes up geometric and arithmetic automorphisms in a single group extension. The aim now is to make a different exact sequence which is sufficiently natural that it is equivariant for $Aut_{\mathbb{P}^1(\mathbb{Z})}(Y)$.

Remark. The isomorphism types of Z forms of Y correspond to conjugacy types of liftings of the Galois action along the surjection $Y \to \mathcal{O}_K$, equivalently conjugacy classes of splittings of the group extension

$$1 \to G \to Aut_{\mathbb{P}^1(\mathbb{Z})}(Y) \to Gal(K/\mathbb{Q}) \to 1.$$

Each corresponding section group $\subset Aut_{\mathbb{P}^1(\mathbb{Z})}(Y)$ mapping isomorphically to $Gal(K/\mathbb{Q})$ determines a $\mathbb{Z}$ form of Y, however the group H need not act on the $\mathbb{Z}$-form, and we cannot recover a $\mathbb{Z}$-form of X from a single $\mathbb{Z}$ form of Y.

If a, b are coprime integers, there is an associated integer point $[a : b]$ of the projective line $\mathbb{P}^1$.

Define $\mathcal{M}(Y)$ which we'll also call $\Omega_Y(log\ {f_Y}^{-1}C)^0$ to be the subsheaf $f_Y^{-1}\Omega_{\mathbb{P}^1_{\mathbb{Z}}}(log\ C) \subset \Omega_Y(log\ {f_Y}^{-1}C)$. It is an invertible sheaf over Y which in turn is flat over $\mathcal{O}_K$. As before, we let $M_k(Y) = \Gamma(Y, \mathcal{M}^{\otimes k})$ for $k = 0, 1, 2, ...$ and in case $Y = X(2)$ we write $M_k(2) = M_k(X(2))$.

For example, when $\mathcal{O}_K = \mathbb{Z}$ we have that $M_1(2) \cong \mathbb{Z} \oplus \mathbb{Z}$, and we define the function

$$\mathbb{P}^1(\mathbb{Z}) \to M_1(2)/\{1, -1\}$$

$$[a : b] \mapsto b\omega_0 - a\omega_1$$

and we interpret the integers a, b as being elements of $\mathcal{O}_K$.

Recall that if u_0, u_1 are homogeneous coordinates on $\mathbb{P}^1$ we may write the one forms with logarithmic poles on $0, 1, \infty$ as

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

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

and these lift to the one forms

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

$$\theta(1/2, \tau)^4 d\tau$$

on the upper half plane. Thus we may also consider this map as assigning to each integer point of the projective plane a $\Gamma(2)$ invariant one-form on $\mathbb{H}$ which is holomorphic at the cusps.

This one-form, when pulled back to Y, has simple zeroes as a one-form on the fiber of $Y \to \mathbb{P}^1$ over p. More precisely, if we choose $\mathcal{O}_K$ large enough that each point of the fiber is defined over $\mathcal{O}_K$, then this one-form defines the affine subscheme whose coordinate ring modulo its $\mathbb{Z}$-torsion, normalizes to a cartesian product of one copy of $\mathcal{O}_K$ for each geometric (complex) point in the inverse image of p under $f_Y : Y \to \mathbb{P}^1$.

By attaching a sign arbitrarily to the image of p, each integer point of $\mathbb{P}^1$ determines then an element of $M_1(2) \subset M_1(Y)$ which we'll call $\gamma(p)$. Under the map

$$M_1(Y) \to Hom(\mathcal{M}_1, \mathcal{M}_2)$$

we then have

1. Theorem Associated to each rational point of $p \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$ is a natural exact of coherent sheaves on Y

$$0 \to \Omega_Y(log\ f_Y^{-1}C)^0 \stackrel{\gamma(p)}{\to} (\Omega_Y(log\ f_Y^{-1}C)^0)^{\otimes 2} \to \mathcal{O}_{f^{-1}p} \to 0,$$

equivariant for the action of $Aut_{\mathbb{P}^1(\mathbb{Z})}(Y)$.

Proof. We've defined $\mathcal{M}_k(Y)$ as a subsheaf of $\Omega_Y(log\ f_Y^{-1}C)$ such that the first three theorems of note 1 remain true. The sequence follows with the rightmost term twisted by twice the fiber, and it is natural (equivariant) for the group action. We will see later that there is a $\tau(p)$ so that the divisor of the rational function $\gamma(p)/\tau(p)$ consists of $f_Y^{-1}(p)$ plus a disjoint component, so $\mathcal{O}_{f_Y^{-1}p}$ is isomorphic to a twist by any power of $\mathcal{M}_1(Y)$. The isomorphism is equivariant for automorphisms over $\mathbb{P}^1$ since $\gamma(p)$ and $\tau(p)$ are induced from $\mathbb{P}^1$ therefore so is the sequence shown. Note that

The sheaf $\mathcal{O}(1)$ makes sense on the integer projective plane; taking global sections in the theorem and applying Leray's spectral sequence gives

2. Corollary. For each integer point $p \in \mathbb{P}^1 \setminus \{0, 1, \infty\}$ is the $Aut_{\mathbb{P}^1(\mathbb{Z})}(Y)$-equivariant exact sequence of finitely-generated $\mathcal{O}_K$ modules

$$0 \to M_1(Y) \stackrel{\gamma(p)}{\to} M_2(Y) \to \mathcal{O}_{f_Y{}^{-1}p} \to T \to 0$$

where the T is a submodule of the torsion module $H^1(\mathbb{P}^1_{\mathbb{Z}}, f_{Y*}f_Y{}^*\mathcal{O}(1))$.

3. Corollary. For each such p there is a structure on $\frac{M_2(Y)}{\gamma(p)M_1(Y)}$ of an ideal in $\mathcal{O}_{f_Y{}^{-1}p}$ such that for each section $V \subset Aut_{\mathbb{P}^1(\mathbb{Z})}(Y) \to Gal(K/\mathbb{Q})$ the integer points of the corresponding $\mathbb{Z}$ form of Y which lie over p are bijective with the V fixed points of the V action on the indedomposable idempotent elements of $\frac{M_2(Y)}{\gamma(p)M_1(Y)} \otimes_{\mathbb{Z}} \mathbb{Q}$.

To see that the ranks at least of the relevant $\mathcal{O}_K$ modules are correct, we can use that

$$rank_{\mathcal{O}_K}(M_2) = 3 \cdot 1 + 2 \cdot \alpha + 1 \cdot g$$

$$rank_{\mathcal{O}_K}(M_1) = 2 \cdot 1 + 1 \cdot \alpha + 0 \cdot g$$

with α as we defined earlier. Then the difference is

$$1 + \alpha + g$$

which indeed is the branched covering degree of Y over $\mathbb{P}^1$.

Remark. As we know, $M_2(Y)$ it contains $\omega_0 M_1(Y)$ and $\omega_1 M_1(Y)$ (which intersect along $\omega_0\omega_1 M_0(2) = \omega_0\omega_1\mathcal{O}_K$) and the quotient is isomorphic to $H^1(Y, \mathcal{O}_Y)$ which is an $\mathcal{O}_K$ module of rank g. The choice of rational point p amounts to choosing a primitive element (integer basis element) of the lattice spanned by ω_0, ω_1, and then we may assume $\gamma(p) = \omega_2$ and that we are reducing modulo the span of $M_1(Y)\omega_1$. There is then the $Aut_{\mathbb{P}^1(\mathbb{Z})}(Y)$ equivariant exact sequence

$$0 \to \mathcal{O}_K \to M_1(Y) \to \mathcal{O}_{f^{-1}p} \to H^1(Y, \mathcal{O}_Y) \to 0.$$

This can be used to re-derive our Lefshetz formula, but here the naturality is over $\mathbb{Z}$. We have seen that the relation between this natural exact sequence and the underlying ring structure of $\mathcal{O}_{f^{-1}p}$ determines in principle the integral points of Y lying over p.

Integer points of modular curves

One lesson we've learned so far is that instead of considering Galois automorphisms, Galois modules and invariants, it seems nicer to merely work naturally over $\mathbb{P}^1_{\mathbb{Z}}$ even while considering modular curves which may not be defined over $\mathbb{Z}$. That is to use naturality.

Let's start again in another attempt to describe the integer points of modular curves.

Let's begin with some generalities which were introduced in 'Easy things which number theorists know.' We'll state this a little more geometrically than before, and only describing, on a divisor on a curve flat and finite over $\mathbb{P}^1_{\mathbb{Z}}$, those integer points which happen to be *scheme theoretically isolated*

1. Theorem. Let E be a one dimensional scheme irreducible, flat and finite over $\mathbb{P}^1_{\mathbb{Z}}$. Let D be a Cartier divisor on E. Let $V \to E$ be the line bundle with section sheaf $\mathcal{O}(-D)$ (when D might have embedded components, this means the defining ideal sheaf of D) and consider $E \subset V$ to be the zero section. Let z be the global section (unique up to multiplication by units) of the line bundle with section sheaf $\mathcal{O}(D)$ whose zero-locus is D itself. Then

$$z \in \Gamma(E, \mathcal{O}_E(D)) \subset \Gamma(V, \mathcal{O}_V)$$

so we can view z as a global section of the structure sheaf of V. Note that set-theoretically, the zero locus of z in V is the union of E and the inverse image of D under the bundle projection $V \to E$. Write

$$dz \in \Gamma(V, \Omega_V(log\ E)(-E)) \subset \Gamma(V, \Omega_V).$$

Then the integer points of D which are scheme theoretically isolated from all others, are those which are disjoint from the support of $\Lambda^2(\frac{\Omega_V(log\ E)(-E))}{\mathcal{O}_V\ dz}) \subset V$.

Proof. Restricting $\Omega_V(log\ E)(-E)$ along the inclusion which we'll call $i : V \to E$ of the zero section, the pullback $i^*\Omega_V(log\ E)(-E)$ is just $\mathcal{P}_E(\mathcal{O}(D))$, and the restriction of dz is what we have called $\nabla(z)$, a global section of first principal parts of $\mathcal{O}(D)$.

It is not generally true that this section belongs to the kernel in the exact sequence

$$0 \to \Omega_E(D) \to \mathcal{P}(\mathcal{O}(D)) \to \mathcal{O}(D) \to 0$$

but it maps to an element of the kernel in the exact sequence that arises upon applying j^*, where j denotes the inclusion of the zero set of the section s of $\mathcal{O}(D)$ into E. And so we have a well-defined element which we might still call

$$\nabla(z) \in j^*(\Omega_E(D)).$$

The definition of this element involves an E^3 differential and is described in 'Easy things.' Let's repeat that here; note in the present context $Y = Spec(\mathbb{Z})$ can be ignored. We have the diagram of exact sequences

$$\begin{array}{ccccccccc} 0 & \to \Omega_{E/Y}(D) & \to & \mathcal{P}_{E/Y}(\mathcal{O}_E(D)) & \to & \mathcal{O}_E(D) & \to & 0 \\ & & & \uparrow & & \uparrow & & \\ & & 0 \to & \mathcal{O}_E\nabla(z) & \to & \mathcal{O}_E z & \to & 0 \\ & & & \uparrow & & \uparrow & & \\ & & & 0 & & 0 & & \end{array}$$

The corollary 8 of 'Easy things' says that when we take the cokernel of this upward map of rows and pull back along $j : D \to E$ we obtain in the middle place $\mathcal{P}_{D/Y}\mathcal{O}_D(D)$. Homologically speaking, when we pull back the cokernel sequence we get a non exact sequence with kernel $\mathcal{T}or_1^{\mathcal{O}_E}(\mathcal{O}_D(D), \mathcal{O}_D)$. This is is the same as $\mathcal{T}or_1^{\mathcal{O}_E}(\mathcal{O}_D, \mathcal{O}_D)$ twisted by $\mathcal{O}_E(D)$, and so it is a copy of the trivial sheaf $\mathcal{O}_D$. Thus we obtain

$$0 \to \mathcal{O}_D \to j^*\Omega_{E/Y}(D) \to \mathcal{P}_{D/Y}(\mathcal{O}_D(D)) \to \mathcal{O}_D(D) \to 0.$$

And the exact diagram

$$
\begin{array}{ccccccccccc}
 & & & & 0 & & & & & & \\
 & & & & \uparrow & & & & & & \\
 & & & & \mathcal{O}_D & \to & 0 & 0 & 0 & & \\
 & & & & \uparrow & & \uparrow & \uparrow & \uparrow & & \\
0 & \to & \mathcal{T}or_1^E(\mathcal{O}_D(D), \mathcal{O}_D) & \to & j^*\Omega_{E/Y}(D) & \to & \mathcal{P}_{D/Y}(\mathcal{O}_D(D)) & \to & \mathcal{O}_D(D) & \to & 0 \\
 & & & & \uparrow & & \uparrow & & \uparrow & & \\
 & & & & 0 & \to & j^*\Omega_{E/Y}(D) & \to & j^*\mathcal{O}_E(D) & \to & 0 \\
 & & & & & & \uparrow & & {\scriptstyle 0}\uparrow & & \\
 & & & & & & 0 & \to & j^*\mathcal{O}_E\nabla(z) & \to & j^*\mathcal{O}_{E\cdot z} & \to & 0 \\
 & & & & & & & & \uparrow & & \uparrow & & \\
 & & & & & & & & 0 & \to & \mathcal{O}_D & & \\
 & & & & & & & & & & \uparrow & & \\
 & & & & & & & & & & 0 & &
\end{array}
$$

The exact diagram implies a natural 'serpent lemma' or 'E_3 differential' which is an isomorphism from the copy of $\mathcal{O}_D$ on the lower right to the one in the upper left.

Before applying j^*, we had that $\mathcal{O}_E\nabla(z)$ embeded as a subsheaf of $\mathcal{O}_E(D)$, once we apply j^* the image of the copy of this sheaf in $j^*\mathcal{P}_{D/Y}(\mathcal{O}_D(D))$ is the same as the image $\mathcal{O}_D$ of the Tor term, which is the kernel of $j^*\Omega_{E/Y}(D) \to \mathcal{P}_{D/Y}(\mathcal{O}_D(D))$.

Here we are taking $Y = Spec(\mathbb{Z})$ and the various sheaves of differentials are the absolute differentials. We just now obtained the term $\mathcal{P}_E(\mathcal{O}_E(D))$ by restricting (=pulling back in the coherent sheaf sense) the sheaf $\Omega_V(log\ E)(-E)$ to E, and now further applying j^* means that we've further restricted to D. The diagram shows that reducing $j^*\mathcal{P}(\mathcal{O}_E(D))$ by the image of dz yields a (split) extension of $\mathcal{O}_D(D)$ by the $\mathcal{O}(D)$ tensor the cokernel of the inclusion of the conormal sheaf of D in E. Thus the second exterior power restricts to zero on precisely those components where the conormal embedding is an isomorphism. These are the components of D such that the Kahler differentials of D restrict to zero; and thus they must be both scheme-theoretically isolated and rational (=integer) points.

A more careful analysis could also incude considerations of rational points which are allowed to have scheme-theoretic intersections, but we are not considering that today.

Now let's return to the situation where we have a finite index subgroup $\Gamma \subset \Gamma(2)$ and $X = \Gamma \setminus \mathbb{H}$ the corresponding modular curve. We take $E = X$ now, and rather than relying on any Galois theory, let us make the bold assumption that X is defined over $\mathbb{Z}$ and let $f_X : X \to \mathbb{P}^1$ now denote a $\mathbb{Z}$ form which we assume is flat.

Let $p \in X(2) \setminus \{0, 1, \infty\}$ be an integer point. We choose integers a, b so that $p = [-b : a]$, or, to be more clear, so that p is the zero locus of

$$a\omega_0 + b\omega_1 \in M_1(2)$$

viewed as a global section of the locally free sheaf $\mathcal{M}_1(2)$ of one-forms on $X(2)$ with logarithmic poles at the three cusps.

We take $f_X : X \to X(2)$ the natural branched covering map, and we take $D = f_X^{-1}p$. Since f_X is flat $\mathcal{O}_D$ is a free abelian group. All irreducible components of D are one-dimensional and map onto $Spec(\mathbb{Z})$. The number of these will be less than the branched covering degree if not all are copies of $Spec\ \mathbb{Z}$.

As before we define the $\mathbb{Z}$ form $\mathcal{M}(X) = f_x^*\mathcal{M}(2)$. If we write

$$\gamma(p) = a\omega_0 + b\omega_1 \in M_1(2) \subset M_1(X)$$

then the principal ideal in the ring of modular forms $M(Y)$ is a

defining ideal for a graded homogeneous coordinate ring for the divisor D. More precisely, when we were working over $\mathbb{Q}$ there was is the module isomorphism

$$\frac{M_k(X)}{M_{k-1}(X)\gamma(p)} \to \mathcal{O}_D$$

for any $k \geq 2$. Now over $\mathbb{Z}$ the cokernel is contained in the finite abelian group $T \subset H^1(\mathbb{P}^1_{\mathbb{Z}}, f_{X*}f_X{}^*\mathcal{O}(1))$ and is still zero for $k >> 0$. let's assume this works for $k \geq 2$ for simplicity

At this point we need to deal with either second exterior powers or that E^3 differential; let's choose the latter. So we know there is an element which we call

$$\nabla(a\omega_0 + b\omega_1) \in j^*\Omega_X(D)$$

and while it is a little difficult to describe which element it is, there is the exact sequence of $\mathcal{O}_D$ modules (note D is affine)

$$0 \to \mathcal{O}_D\nabla(a\omega_0 + b\omega_1) \to j^*\Omega_X(D) \to \Omega_D \to 0.$$

Note that the rightmost term is initially $\Omega_D(D)$ however twisting has no effect (we'll see later that a divisor defined by $c\omega_0 + d\omega_1$ is equivalent to D and scheme theoretically disjoint from D).

Let's make two simplifying assumptions in the expectation that later we'll remove them with a more precise formulation. Assume that Ω_X happens to be locally free, and the inclusion $f_X^{-1}\Omega_{\mathbb{P}^1_{\mathbb{Z}}(log\ C)} \to \Omega_X(log\ f_X^{-1}C)$ happens to be an isomorphism (we called the subsheaf $\Omega_X(log\ f_X^{-1}C)^0$). Then D is linearly equivalent to $K_X + f_X{}^{-1}C$ giving $f_X{}^{-1}C$ the structure of reduced divisor (ignoring multiplicities) the sheaf $\Omega_X(D)$ is isomorphic to $\mathcal{O}_X(2K_X + f_X{}^{-1}C)$ which is also isomorphic to $\mathcal{M}_2(X)(-f^{-1}C)$; that is, we have an exact sequence of sheaves on X

$$0 \to \Omega_X(D) \to \mathcal{M}_2(X) \to \mathcal{O}^{red}_{f_X{}^{-1}C} \to 0$$

By what we've said, the sheaf on the right already has the reduced structure but we've written the superscript red to be completely clear about this.

The sheaf $\Omega_X(D)$ is acyclic (since D is effective) and this gives by passing to global sections

$$0 \to \Gamma(X, \Omega_X(D)) \to M_2(X) \to \mathcal{O}^{red}_{{f_X}^{-1}C} \to 0$$

Thus our element

$$\nabla(a\omega_0 + b\omega_1)$$

can be interpreted as an element of $M_2(X)$, a modular form of level X and weight 2, which maps to zero under what we might call the 'second residue map' from M_2 to the reduced structure sheaf of the cusps of X.

Now let's see whether the residue map

$$\Omega_X(D) \to j_* j^* \Omega_X(D)$$

is surjective on global sections. It is not since the kernel Ω_X is not acyclic, and neither is the other

We must twist our locally free sheaves by $D + {f_X}^{-1}C^{red} = 2D - K_X$.

We have

$$0 \to j^*(\mathcal{O}_X(2D-K_X))\nabla(a\omega_0+b\omega_1) \to j^*(\Omega_X(2D+{f_X}^{-1}C^{red})) \to \Omega_D \to 0.$$

Note that $\Omega_X(2D + {f_X}^{-1}C^{red}) \cong \mathcal{M}_3(X)$ while $\mathcal{O}_X(2D - K_X)$ is the kernel of $\mathcal{M}_2(X) \to \Omega_X$.

Since the twisting is high enough now we can pass to global sections; we have $\Gamma(X, \Omega_X(2D-K_X))$ is the kernel of a map $M_2(X) \to S_1(X)$ with S_1 the cusp forms of weight one. Now

$$j^*(\Omega_X(2D + {f_X}^{-1}C^{red})) \cong \frac{\Gamma(X, \mathcal{O}_X(3D))}{\Gamma(X, \mathcal{O}_X(2D)\gamma(p)}$$

$$\cong \frac{M_3(X)}{M_2(X)\gamma(p)}.$$

And

$$j^*(\mathcal{O}_X(2D - K_X)) \cong \frac{\Gamma(X, \mathcal{O}_X(2D - K_X))}{\Gamma(X, \mathcal{O}_X(f_X^{-1}C^{red}))}.$$

Recall that

$$\gamma(p) = a\omega_0 + b\omega_1 \in M_1(2) \subset M_1(X)$$

and it occurs because the relevant inclusion is the same as multiplication by $\gamma(p)$ in the ring of modular forms of level X.

In the calculations above the term $\nabla(a\omega_0 + b\omega_1)$ is not contained in the set of parentheses immediately following j^*. The main technical difficulty with this approach is that this element does not come from a global section of $\Omega_X(D)$ until *after* applying j^*.

From this, we obtain then the presentation of $\Omega_D(-K_X)$ initially as an abelian group from the exact sequence

$$\Gamma(X, \mathcal{O}_X(2D - K_X)) \to \frac{M_3}{M_2\gamma(p)} \to \Omega_D(-K_x) \to 0.$$

We have already seen that for $k \geq 2$ $\frac{M_k}{M_{k-1}\gamma(p)}$ is a copy of $\mathcal{O}_D$, its normalization is a cartesian product of rings of integers of number fields. In fact, the ring structure comes merely from introducing the following two relations into the ring of modular forms $M(X)$,

$$a\omega_0 + b\omega_1 = 0$$

$$c\omega_0 + d\omega_1 = 1$$

where c, d are chosen so that $ad - bc = 1$.

Under this map, $\frac{M_3(X)}{M_2(X)\gamma(p)}$ maps isomorphically onto $\mathcal{O}_D$, and the whole ring thus maps onto $\mathcal{O}_D$, and all the ring relations in $\mathcal{O}_D$ are the ones implied by the fact that the reduction map is a ring homomorphism.

The only slightly mysterious ingredient here is that the map

$$\Gamma(X, \mathcal{O}_X(2D - K_X)) \to \mathcal{O}_D = M_3(X)/M_2(X)\gamma(p)$$

is induced by the element $\nabla(a\omega_0 + b\omega_1)$, coming from that connecting homomorphism.

Anyway, we can state, for X a $\mathbb{Z}$-form of a level above $X(2)$ and $M(X)$ the corresponding $\mathbb{Z}$-form of the ring of modular forms,

2. Therorem. Assume $\Omega_X(log\ f_X^{-1}C)$ is equal to the locally free subsheaf $\Omega_X(log\ f_X^{-1}C))0$, and that the torsion abelian group T is zero. Let $p = [-b, a]$ be an integer point of the projective line not equal to $0, 1$, or ∞. Then

i) For $\gamma(p) = a\omega_0 + b\omega_1$ with $\omega_0, \omega_1 \in M_1(2)$ as before, letting $D \subset X$ be the inverse image of p, the abelian group $\frac{M_3(X)}{M_2(X)\gamma(p)}$ maps isomorphically onto the image $\mathcal{O}_D$ of $M(X)$ when the ring relations

$$a\omega_0 + b\omega_1 = 0$$

$$c\omega_0 + d\omega_1 = 1$$

are introduced into the ring $M(X)$ with c, d chosen so that $1 = ad - bc$.

ii) There is an element $\nabla(a\omega_0 + b\omega_1)$ (in the first principal parts of $\mathcal{O}_X(D)$) which induces a map

$$\Gamma(X, \mathcal{O}_X(2D - K_X)) \to \frac{M_3(X)}{M_2(X)\gamma(p)} \cong \mathcal{O}_D$$

whose image is an ideal. The cokernel of this map is the twisted Kahler differentials $\Omega_D(-K_X)$.

Remark. The indices k in the M_k should be multlied by 2 to obtain the more standard numbering, so what we call $M_3(X)$ is more usually called $M_6(X)$. Note well that all the above refers to a $\mathbb{Z}$ form X and a $\mathbb{Z}$ form $M(X)$.

Remark. For removing the simplifying hypotheses we may use the fact that any normalization or partial normalization map h (more generally any locally projective birational map of integral schemes) commutes with first principal parts modulo torsion in the sense $\mathcal{P}(h^*J)/torsion \cong h^*\mathcal{P}(J)/torsion$ for J rank one coherent.

Conclusion about modular forms

The conclusion of this sequence of notes about modular forms has to really only be an acknowledgement. Maybe it is now or never to write such a thing; there is no cable to charge this laptop, and the wi-fi signal is far away and weak, miles away for some reason. It is never any shame anyway if the actual content of a paper is just wrong, while also including an acknowledgement; but one should understand that this particular paper was written in five minutes, to say something that now seems easy about the Maths; but about how Maths is actually formulated, something deeper, relying also on an idea not many years ago by an undergraduate student who was here.

Three people from Queen Mary College have never had any acknowledgement from me, one is Charles Leedham-Green, we had conversations about logic, which went something like my claiming to have found the fastest-growing class of functions, and him eventually replying with something which in an obvious way could be increased, a detail left undone as idiotic as he could find.

Another is Peter Kropholler, I remember him sitting with a pint of beer with the chairman Roxburgh, when my time there was coming to an end, saying 'No one proves a great theorem, and then nothing for four or five years, and then proves another.' At the time I was privately angry, not recognizing the subtle way he had thus given me credit for his work, and a challenge to face the future.

Another is Robert Wilson, whom I hadn't met before, but had been there at Queen Mary, coming to Warwick last year to give a talk about his ideas about physics, the standard model, and recent theories in cosmology. This was intentionally incompetent (intentionally not making any connection with his work in Lie algebras), pretending to be superstitious, finding exact coincidences to many decimal places between ratios between constants in particle physics, and the number of days of the year, the effects of the moon on the earth, saying things along the lines that these relations between fundamental constants are caused by the tides. I haven't yet completely taken on board all the things which he has been saying; nor know who will.

My comments about modular forms were in answer to Cremona, Loeffler, Zerbes, Bruin, and Bartel at Warwick. The origin of an idea is a look, an expression during a conversation or passing in the hallway, in a very specific context among ambient conversations and ideas. When people say that there is a natural transformation between such and such, this is not to say that there is a pre-ordained transformation. But one can get enough confidence to assert something. The point is supposed to be that *anyone* can do that, people do it all the time, and there was once something like that, which was called the algebraic deRham theorem.

An integer form X of a modular curve has associated to it a topological space $X(\mathbb{C})$. The algebraic deRham theorem might come from Deligne wanting to solve a basic question of his advisor, or might be a common property as it is explained in Griffiths and Harris' book. Let's suppose that X is absolutely irreducible too, which I think means that $\Gamma(X, \mathcal{O}_X) = \mathbb{Z}$.

The 'ordinary' or Eilenberg-Steenrod cohomology of $X \backslash cusps$, where *cusps* is the finite set of cusps of X, is the cohomology of the sheaf of locally constant functions from $X(\mathbb{C}) \setminus cusps$ to the integers, also called the 'simplicial' or 'singular' cohomology, denoted $H^i(X(\mathbb{C}) \setminus cusps,\ \mathbb{Z})$. But it is known that this has an algebraic definition too.

If we revert to the older numbering of modular forms, this finitely-generated abelian group when i is 1 should be the same as what is known as

$$M_2(X) \oplus \frac{M_4(X)}{M_2(2)M_2(X) + M_4(2)M_0(X)}$$

where $M_k(X)$ is modular forms of weight k and level X, as long as the group or 'level' of X is a subgroup of the congruence subgroup $\Gamma(2)$.

There is also another finitely-generated abelian group associated to X, this one depends on a choice of a rational (=integral) non-cusp point $p \in X(2) = X(\Gamma(2))$, and it is the underyling abelian group of the structure sheaf of the inverse image of p under the branched covering map $f_X : X \to \mathbb{P}^1$.

And this is a free abelian group of rank one smaller. Moreover we can just say that there is a natural map of abelian groups

$$H^1(X(\mathbb{C}) \setminus cusps, \mathbb{Z}) \to \mathcal{O}_{f_X^{-1}p},$$

which has a kernel free abelian of rank one.

In the case when f_X is the identity, so $X = X(2)$, this kernel is a rank one subgroup of a rank two lattice, as $H^1(X(2) \setminus cusps,\ \mathbb{Z})$ is a free abelian group of rank two.

And the identification of the rational point p with the rank one sublattice, or perhaps with the map itself which has that rank one sublattice as its kernel, is there, and a generator of the free abelian group of rank one is a modular form of weight two and level $X(2)$, which also then has any level above 2. Under one choice of the identification, if $p = [-b : a]$ with a, b relatively prime, then this modular form is

$$\gamma(p) = a\omega_0 + b\omega_1.$$

Here we can think of ω_0 and ω_1 as basic global sections of the sheaf of one-forms on $X(2) = \mathbb{P}^1$ allowed simple poles (=logarithmic poles since this is the one dimensional situation) on the three cusps.

Somehow, in terms of homogeneous coordinates u_0, u_1 these also can be written

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

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

If we think of these as analytic functions $\mathbb{H} \to \mathbb{C}$ satisfying modularity of weight two they are

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

$$\omega_1 = \theta(1/2, \tau)^4,$$

and we append $d\tau$ if we want to think of the lifted one-forms on $\mathbb{H}$ itself.

So that a rational point of the projective line determines such an analytic function.

The class of a point in the projective line, in the cohomology of the compact complex manifold, can be identified with the isomorphism type of the torsor for the sheaf of one-forms, which consists of one-forms with a pole of residue exactly one at that point. This torsor is sometimes directly related to the $(1,1)$ form, in the notions of Kodaira, in this case describing the projective embedding which is the identity map.

That residue can be thought of as being a generator of the infinite cyclic quotient group that arises when we reduce $H^1(\mathbb{P}^1 \setminus cusps,\ \mathbb{Z})$ modulo the sublattice of rank one through $\gamma(p)$.

That is, we can think of the sheaf $\mathcal{O}_p$ as the free module of rank one consisting of all such residues, it is the image of the residue map from one forms with (logarithmic) poles at p and arbitrary poles and zeroes elsewhere, with kernel the one forms which do not have a pole at p.

When we pass to thinking about X instead of just $X(2)$, we can think of the coincidence

$$\text{covering degree} = rank\ H^1(X \setminus cusps,\ \mathbb{Z}) - 1$$

as coming from the fact it is still true on X that we have this residue map. The point here is that $M_2(X)$ modulo $\gamma(p)$ in the modular forms is the same as

$$M_2(X)\tau(p)\ modulo\ M_0(X)\gamma(p)\tau(p).$$

Here

$$\tau(p) = c\omega_0 + d\omega_1$$

with c, d chosen such that $ad - bc = 1$.

The choices of c, d are parametrized by $\mathbb{Z}$ analagous to the integers and the point at infinity on the boundary of the Poincare disk.

Now, we know that

$$M_4(X) = \frac{M_2(X)\gamma(p) \oplus M_2(X)\tau(p)}{M_0(X)\gamma(p)\tau(p)} \oplus H^1(X, \mathcal{O}_X)$$

or rather that it contains the first factor naturally with quotient being the second factor, so a splitting of the filtration describes such a direct sum decomposition.

The correspondence between modules over a graded ring and coherent sheaves can be made nice by thinking about graded modules as equivariant coherent sheaves on an affine variety, and here we are seeing that the even degree part of the ring $M(X)$ modulo the element $\gamma(p)$ is the coordiate ring of the fiber over p, this is the reduced fiber when p is not one of the three cusps.

If we think of $\gamma(p) = a\omega_0 + b\omega_1$ as a section of the sheaf of one forms on X with at most simple (=logarithmic) poles at the points of the fiber over p, and think of this as sections in the sense of sections of a vector bundle, then this section has a first principal part $\nabla(\gamma(p))$, a global section of first principal parts, and by reducing modulo $\gamma(p)$ and pulling back to the fiber we obtain a global section of first principal parts on the fiber. This happens to belong to the kernel of the map to the sheaf itself and can be interpreted as the restriction to the fiber of one forms with logarithmic poles.

Now, this is a little complicated but we can see through it! The line bundle is the one whose sections sheaf is isomrphic to one forms with at most simple (=logarithmic) poles on the *cusps*. But now we are taking sections which have simple (=logarithmic) poles on the *fiber*.

Remember that I said that the residues at the fiber, which comprise the quotient of $H^1(X \setminus cusps,\ \mathbb{Z})$ modulo that line in the lattice, are allowed to be residues of functions which can have arbitrary poles and zeroes away from the fiber. Here then we are allowed to have those poles at the cusps, because the fiber is disjoint from the cusps in $X(\mathbb{C})$ so contains no cusp in X.

The principal parts bundle is a rank two vector bundle, and $\nabla(\gamma(p))$ is not contained in the sub bundle which is one forms with poles allowed on the fiber. But when we restrict to the fiber it does end up belonging to that sub bundle for principal parts on the fiber. This is talking about how we have a principal different element for that ring of (perhaps not normal) integers.

From these considerations there is now an action, in principle determined by the structure of the ring $M(X)$, by which $\nabla(\gamma(p))$ can act on the quotient abelian group $H^1(X(\mathbb{C}) \setminus cusps,\ \mathbb{Z})/(\mathbb{Z}\gamma(p))$.

I have said other things, like about how X itself in the complex sense is a leaf of a foliated vector bundle, and so-on, and these are not really connected to what we have here. But what we have here is that there is an action L by which $L(\nabla\gamma(p))$ is a matrix acting on this quotient group.

I am purposely being adventurous now and imagining somehow, in analogy with the Weil conjectures, that we should lift of L to an endomorphism L_1 of $H^1(X(\mathbb{C}) \setminus cusps,\ \mathbb{Z})$; let L_0 be the induced action on $\mathbb{Z}\gamma(p)$ which I write as $M_0(X)\gamma(p) = H^0(X(\mathbb{C})\setminus cusps,\ \mathbb{Z})$, we have that L_i assigns an integer matrix once a cohomology basis is chosen, to each cohomology group, and we are contriving that the determinant $det(L)$ is a ratio of two determinants. Then

Theorem.

$$\prod_{i=0}^{1} det(L_i(\nabla\gamma p))^{(-1)^i} = \frac{1}{\prod_{j=1}^{m} disc(\mathcal{O}_j/\mathbb{Z})}$$

where $disc$ means discriminant and $\mathcal{O}_j$ is the structure ring of the j'th connected component of the scheme theoretic fiber of X over p, and m is the number of connected components.

The left side we've written looking like a multiplicative Lefshetz character; when the theorem implies that the points p where its absolute value is 1 are exactly those points above which the points of X are scheme-theoretically disjoint rational (=integer) points.

And

$$| \ \ [R : \mathcal{O}_{f_X^{-1}p}]^2 \ \prod_{i=0}^{1} det(L_i(\nabla\gamma p))^{(-1)^i} \ \ | = 1$$

if and only if all $\mathbb{C}$ points above p are rational, where $\mathcal{O}_{f^{-1}p} \subset R = \prod_{k=1}^{s} \mathcal{O}_{f_X^{-1}p}/P_k$, with $P_1, ..., P_s$ the minimal prime ideals and s the number of irreducible components.

In the case when X is suituably Galois, all the $\mathcal{O}_i$ are isomorphic. Then the left side as a function of p precisely determines the set of rational non-cusps of X.

To give an algebraic formula for the left side would amount to considering the graph of the map from X to $\mathbb{P}^1$, and residues on the graph. This can obviously be done if the ring structure of $M(X)$ is known. We've given a Lefshetz formula for the second factor on the left side. For the first factor, under the relations $\gamma(p) = 0, \tau(p) = 1$ the whole of $M(X)$ retracts isomorphically to $\frac{M_4(X)}{M_2(X)\gamma(p)}$ which becomes the affine coordinate ring of the fiber over p and determines this factor too.

The dependence on the level Γ may not be a Σ_0 formula; any algebraic or topological formalism is meant to be only an allegory. For example, the fact that modular forms are complex analytic entities means that the search for rational points can take place in the domain of analytic number theory. Though that was clear at the outset too.

Outline geometric proof of Mordell's conjecture

We begin with an absolutely irreducible complete (projective) curve X defined over the rationals, equivalently over the integers. Let g be the genus of X. Mordell's (proven) conjecture is the statement that if $g > 1$ then X has only finitely many rational points.

Belyi's theorem implies, even more generally just from the fact that X is legitimately one-dimensional, that we can choose a map $f : X \to \mathbb{P}^1$ branched over the three points, which we can take to be $\{0, 1, \infty\}$ under an automorphism of $\mathbb{P}^1_{\mathbb{C}}$. Therefore we can represent X as a 'modular curve' with respect to some finite index subgroup of $\Gamma(2)$. We can use this at a final stage of the proof, where it implies that $f - c = 2g - 2$ where f is the degree of the branched cover and c the number of cusps (ignoring multiplicity). We will make the assumption that f can be chosen with a number of nice properties. Our aim is not to present a useful proof, but rather to present a geometric proof with very restrictive hypotheses.

Under our very restrictive hypotheses, we'll actually show that X has no noncuspidal rational points whatsoever. That is, that the rational points are a subset of the fibers containing critical points of f. I do not pretend to have actually guessed where the rational points are, and most likely for most X there is no function f satisfying all our hypothesis. When there are are no functions f satisfying all the hypotheses, one should weaken the hypotheses while admitting finitely many noncuspidal rational points. The exceptions which are admitted in removing our very strong hypotheses until sucn an f is found to exist are known to be finite in number only because Mordell's conjecture has been previously proven.

Remark. The composite f with Belyi's maps $[x : z] \mapsto [(\frac{x}{a})^a(\frac{y}{b})^b : (\frac{z}{c})^c]$ for $\mathbb{P}^1$ itself, with $x+y = z, a+b = c$, which transform rational points over $[a : c]$ to cusps, actually shows that proving the existence of such an f is abstractly equivalent to proving Mordell' conjecture.

Let X be a Riemann surface. Choose a map $f : X \to \mathbb{P}^1_{\mathbb{C}}$ branched over three points, which we label $\{0, 1, \infty\}$ according to an automorphism of $\mathbb{P}^1$. Suppose the lifted algebraic structure on X has a $\mathbb{Z}$ form, and choose one integer structure. Here are our seven hypotheses.

i) (all critical points of f are rational) Critical points of f correspond to spectra of rings of algebraic integers, and let's just assume that each is a copy of $Spec(\mathbb{Z})$. This assumption applies for example when X is taken of level $\Gamma_0(N) \cap \Gamma(2)$ with N odd and square free. Note that each critical point of f corresponds to a cusp of X but some cusps correspond to non-critical points.

ii) (Ω_X is locally free) Let us also suppose that the sheaf of one-forms of X is locally free (even when X is interpreted as the $\mathbb{Z}$-form).

iii) (f is of Galois type) We also assume that X is Galois over $\mathbb{P}^1$, or, more generally that if there is one rational point in a fiber (of the map with cusps deleted, i.e. over a non-cusp) then all points in that fiber are rational. A fiber over a rational point more rigorously means this: we have associated to a rational point of $\mathbb{P}^1$ a map $Spec(\mathbb{Z}) \to \mathbb{P}^1$ and the scheme which is the pullback of X along this map is the fiber.

iv) (nice stabilizer actions) Assume that the stabilizer in $Aut(X)$ of each cusp component acts transitively on the fiber components which meet that cusp, or, more generally that when two fiber components meet each cusp component, the intersection subscheme of the cusp component is independent of choice of fiber component.

v) (primes behave generically) For each irreducible component C of the cusp locus and each fiber F over a point of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ assume that there is a prime number $p \in \mathbb{Z}$ such that for every rational component of F the point indexed by p is not contained in any other cusp component, and it is contained in C if and only if the fiber component meets C.

vi) (technical condition) For a cusp C and fiber F, using the prime number p from v), for any finitely-generated abelian group A, denote by $[A]$ the order of the p primary part of A, and for $A \subset B$ arbitrary abelian groups, of finite index, denote by $[B : A]$ the order of the p primary part of B/A. Let $\Phi \in \mathcal{O}_F$ be the different ideal such that $\mathcal{O}_F/\mathcal{O}_F\Phi \cong \Omega_F \otimes \omega_X^{-1}$, suppose Φ is principal, let ϕ be a generator, and let $\overline{\mathcal{O}}_F$ be the normalization. For each F such that $\overline{\mathcal{O}}_F = \mathbb{Z} \times ... \times \mathbb{Z}$, and for each inclusion τ of a fiber component into F, assume that if the fiber component meets C then $[\tau^* \frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi}] = [\frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi}]$. In other words that the kernel of the natural surjection $\frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi} \to \tau_*\tau^*(\frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi})$ has trivial p primary part. (It is onto because $\tau^*\mathcal{O}_F \cong \mathbb{Z}$). We'll show that the condition follows from iv) when the localization at p of $\overline{\mathcal{O}}_F\phi \subset \mathcal{O}_F + p\mathcal{O}_F$ happens to be the entire maximal ideal. The occurrence of τ here refers to one fiber, and is the same as what will be called τ_i later when we number the components of F.

vii) (conductor-discriminant formula). This next condition can possibly be deduced from Grothendieck duality; as I have not proved this, we have to assume it: Let $\omega \in \Omega_F \otimes \omega_X^{-1}$ be a spanning element. Assume that the pairing $[\widehat{\mathcal{O}_F}/\mathcal{O}_F] \otimes [\Omega_F \otimes \omega_X^{-1}] \to \mathbb{Q}/\mathbb{Z}$ given by $\langle r \ mod \ \mathcal{O}_F, \ \tau \rangle = trace_{\mathcal{O}_F/\mathbb{Z}}(\frac{r}{\phi}\frac{\tau}{\omega})$ is perfect (inducing a Pontryagn duality) where $\widehat{\mathcal{O}_F} = \{x \in \mathcal{O}_F \otimes \mathbb{Q} : trace(xr) \in \mathbb{Z} \text{ for all } r \in \mathcal{O}_F\}$. Note that it implies that $[\mathcal{O}_F : \mathcal{O}_F\phi] = [\widehat{\mathcal{O}_F} : \overline{\mathcal{O}_F}][\overline{\mathcal{O}_F} : \mathcal{O}_F]$ with both factors equal. Also then $[\mathcal{O}_F : \overline{\mathcal{O}_F}\phi] = [\overline{\mathcal{O}_F}\phi : \mathcal{O}_F\phi]$ since the square of the second factor is equal to the product. Also $[\overline{\mathcal{O}_F} : \overline{\mathcal{O}_F}\phi] = [\mathcal{O}_F : \mathcal{O}_F\phi]$ as one deduces directly, or because it is the index of the action of ϕ on two lattices in the same rational vector space.

Now we can state

Theorem. Suppose that $f : X \to \mathbb{P}^1$ satisfies i),...,vii). Suppose further that in the fiber F over every rational point of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ any two fiber components can be connected by an alternating sequence of intersecting fiber and cusp components. Then X has no noncuspidal rational points unless $g < 2$.

Before proving the theorem, let's give sufficient conditions for vi) to hol and discuss evidence for vii).

Discussion of vi). Consider the surjection $\frac{\mathcal{O}_F}{\mathcal{O}_F\phi} \to \frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi}$. The image is contained in $\frac{\overline{\mathcal{O}}_F}{\overline{\mathcal{O}}_F\phi}$. The ring $\overline{\mathcal{O}}_F$ is just a cartesian product of integral domains corresponding to the irreducible components of F. We only need to consider the components which meet C, so number these $\mathcal{O}_1, ..., \mathcal{O}_s$. They happen to be copies of the ring of integers $\mathbb{Z}$ but let's use slightly more general notation in case we later consider cases when fiber components may be non rational.

The p-primary component of $\frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi}$ is a subring of the p primary component of $\frac{\mathcal{O}_1}{\mathcal{O}_1\phi} \times ... \times \frac{\mathcal{O}_s}{\mathcal{O}_s\phi}$. We may say '$p$-primary component' or 'localization at p' here interchangeably. The reason that only these s components need to be considered is this: since there is no number-theoretic ramification on F, ϕ defines the self-intersection locus of F. Our choice of prime p then ensures that the p primary component of any $\frac{\mathcal{O}_i}{\mathcal{O}_i\phi}$ is zero if i is such that $Spec(\mathcal{O}_i)$ is one of the components of F not meeting C.

The p-primary component of the subring $\frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi}$ of this cartesian product, however, must be indecomposable. For, the kernel of $\frac{\mathcal{O}_F}{\mathcal{O}_F\phi} \to \frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi}$ is a nilpotent ideal. If the image localized at p contained an idempotent element besides 0 or 1 idempotent lifting would provide such an element in the localization at p of $\frac{\mathcal{O}_F}{\mathcal{O}_F\phi}$. But this is the structure sheaf of a subscheme of X supported on the intersection of the fiber over the point of $Spec(\mathbb{Z})$ corresponding to p with the subscheme of F defined by $\phi = 0$. Our hypothesis about primes implies this is a single (closed) point of our cusp component C, and so the p primary component of the is a local ring.

Condition vi) as we've chosen it requires more than indecomposability, though. If we let $I_i \subset \frac{\mathcal{O}_F}{\overline{\mathcal{O}_F\phi}}$ be the p primary component of the kernel of the i'th projection, then obviously the intersection $I_1 \cap ... \cap I_s = 0$. So $I_i = 0$ for all i if and only if $I_i = I_j$ for all i, j. If $p\mathcal{O}_{F,p} + \overline{\mathcal{O}_{F,p}\phi}$ happens to be the full maximal ideal, then $\mathcal{O}_{F,p}/\overline{\mathcal{O}_{F,p}\phi}$ will have maximal ideal just the multiples of p; as it has residue field $\mathbb{F}_p$ it must be reduced to its characteristic subring. All the $\mathcal{O}_i/\mathcal{O}_i\phi$ are isomorphic by the action iv) and within the cartesian product any subring at all contains the characteristic subring and a subring of the cartesian product, containing 1, which is reduced to its characteristic subring must equal the diagonal.

Outline proof of vii). In the first place, it is easy to verify this in the case when $\mathcal{O}_F$ is generated by one element as a ring. More generally it is not clear whether this is true abstractly about subrings of $\mathbb{Z}^n$ with principal different ideal, or whether we need to invoke more context. One might check whether coherent duality pairing

$$f_*\mathcal{O}_X \otimes f_*\Omega_X \to \Omega_{\mathbb{P}^1}$$

is perfect. Denote by j the inclusion of our integral point in $\mathbb{P}^1$, then

$$j^* f_* = f_* i^*$$

where the f_* on the left just means considering underlying abelian groups. Then the pullback of the coherent duality pairing to the copy of $Spec(\mathbb{Z})$ corresponding to a non-cusp rational point of $\mathbb{P}^1$ gives a pairing $i^*\Omega_X \otimes_{\mathbb{Z}} \mathcal{O}_F \to \mathbb{Z}$. We can write $i^*\Omega_X$ up to isomorphism as $\frac{\frac{1}{\gamma}\Omega_X}{\Omega_X}$ with $\frac{1}{\gamma} \in \mathcal{O}_F \otimes_{\mathbb{Z}} \mathbb{Q}$. Three interpretations (by trace forms, by restricting coherent duality, and by now interpreting the left factor as the image of the residue map from logarithmic forms) all appear to coincide, but it may be an exercise in duality theory to be sure that condition vii) really holds exactly as we have stated it.

To prove the theorem we'll first prove a lemma.

1. Lemma. Under conditions i) to vii), the incidence relation between cusp components which meet F and components of the fiber F itself has the structure of a combinatorial graph G with no edge-loops. That is to say, each cusp component which meets F is incident to exactly two fiber components.

Proof of Lemma 1. Retaining our choice of cusp component C and prime p number the inclusions of fiber components into the fiber $\tau_1, ...\tau_m$. For each $i = 1, 2, ..., m$, condition vi) says that

$$[\tau_i^* \frac{\mathcal{O}_F}{\overline{\mathcal{O}_F \phi}}] = \begin{cases} [\mathcal{O}_F : \overline{\mathcal{O}_F \phi}] & , \quad \text{C is incident to the } i\text{'th component of the fiber,} \\ 1 & , \quad \text{otherwise} \end{cases}$$

Suppose F meets C. If we let u be the number of fiber components meeting C we have, letting τ be the pullback to the full normalization,

$$\begin{aligned} [\mathcal{O}_F : \overline{\mathcal{O}}_F \phi]^u &= [\frac{\mathcal{O}_F}{\overline{\mathcal{O}_F \phi}}]^u \\ &= \prod_{i=1}^{m} [\tau_i^* \frac{\mathcal{O}_F}{\overline{\mathcal{O}_F \phi}}] \\ &= [\tau^* \frac{\mathcal{O}_F}{\overline{\mathcal{O}_F \phi}}] \\ &= [\frac{\overline{\mathcal{O}_F}}{\overline{\mathcal{O}_F \phi}}] \\ &= [\overline{\mathcal{O}_F} : \overline{\mathcal{O}_F \phi}] \\ &= [\mathcal{O}_F : \mathcal{O}_F \phi] \\ &= [\mathcal{O}_F : \overline{\mathcal{O}_F \phi}]\ [\overline{\mathcal{O}_F \phi} : \mathcal{O}_F \phi] \\ &= [\mathcal{O}_F : \overline{\mathcal{O}_F \phi}]^2. \end{aligned}$$

The number is a nontrivial power of or chosen prime p. Since the square of a nontrivial integer is the same as the u power, we can conclude that $u = 2$.

Proof of Theorem. Let G be now the graph with cusp components which meet F for the edges and components of the one fiber F for the vertices. We have

$$f - c = 2g - 2$$

for g the genus of X, f the order of a regular fiber (the covering degree), and c the total number of cusps. If all cusps meet F then

$$\chi(G) + \chi(X) = 0$$

where $\chi(X)$ is the Euler characteristic of the underlying topological space of the compact curve X. By the assumption of the theorem that the graph G is connected then $\chi(G) < 2$ therefore

$$2 - 2g = \chi(X) > -2$$

hence

$$g < 2.$$

If there is any cusp component not meeting F then $g = 0$ and there are one, two or three such cusp components (in the last case G being a tree); and it is not possible for more than three cusp components to be disjoint from F.

Thus, under a collection of simplifying hypotheses which actually cannot be simultaneously true except in trivial cases, we've shown that when $g > 1$ there cannot be any noncuspidal rational points; and the idea is that by weakening the hypotheses one characterises exactly what are the rational points.

Example.

Let's take as an example the elliptic curve

$$y^2 z = x(x - 3z)(x - 4z).$$

We will take as our function f the projection to the set of $[x : y]$. This branches over $[0 : 1 : 0], [0 : 0 : 1], [3 : 1 : 0], [4 : 1 : 0]$ so it does not satisfy the hypothesis of branching only over $0, 1, \infty$, and this means that the equation relating f, c, g is not true. Of course $g = 1, f = 2, c = 4$. Thus, unlike the case of a modular elliptic curve, where the incidence graph of cusps and fiber components in a fiber must have Euler characteristic zero plus the number of unused cusp components, here, if it exists, it has Euler characteristic -2 plus the number of unused cusp components.

Let's choose as F the fiber over $[1 : 1]$. The components of the fiber are indexed by $[6 : 6 : 1]$ and $[2 : 2 : 1]$; neither component of the fiber meets any of the four cusps (as one can see by considering fitting ideals of eight two by three matrices). The fiber components are glued together or order two at the point indexed in each by the prime 2, but not meeting any cusp.

We are supposed to ignore cusps which do not meet any fiber (in this case all of them), and our graph consists of two points, and does have Euler characteristic 2. The fact that the two points are not connected together allows the existence of the rational point. That is, the fact that the two fiber components do not meet at any common cusp component means we interpret them as disconnected in the graph, even while they meet each other. It would be tempting to try to find examples where the two fiber components occur as vertices of a cusp interpreted as an edge, with only three cusps uninvolved. One would think this is the more generic situation, so these should occur. The rule that each cusp (which meets any fiber component) must meet exactly two is also not difficult to satisfy in this example. In cases when the branching really is over just $\{0, 1, \infty\}$, there just not be enough cusps serve as edges to connect the graph unless $g < 2$.

In the example, the chain of rings and ideals

$$\mathcal{O}_F \subset \overline{\mathcal{O}_F}\phi \subset \mathcal{O}_F \subset \overline{\mathcal{O}}_F$$

is each as an abelian subgroup of $\mathbb{Z} \times \mathbb{Z}$ in the standard basis, the image of the respective matrix

$$\begin{pmatrix} 2 & 2 \\ 2 & 6 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

while the element ϕ is represented by the diagaonal integer 2. We see that while $\frac{\mathcal{O}_F}{\overline{\mathcal{O}}_F\phi}$ has two elements, when we tensor with either component of $\overline{\mathcal{O}}_F$ it again has two elements, consistent with condition vi). Also that the conductor-discriminant formula vii) holds, with both sides of the equation evaluating to 2. However the ramification which glues the fiber components together does not come from the cusps, rather from a ramification component that maps onto the whole of the subscheme of $\mathbb{P}^1$ defined by the ideal generated by 2.

2. Technical facts about modular curves.

In this section I'll prove some things that were stated in 'Conclusions about modular forms.' Note that the overall point of that paper having to do with interpreting the discriminant as a multiplicative Lefshetz trace on ordinary cohomology is not needed in the section above, we can interpret the Euler characteristic of S as merely the number $2-2g$ without needing to use any topology, and Mordell's conjecture is likely not related in any meaningful way with any notion of multiplicative Lefshetz characters.

We take on the hypotheses of 'Conclusions about modular forms.' In particular X is a $\mathbb{Z}$ form of level $\Gamma \subset \Gamma(2)$ an arbitrary finite-index subgroup, assumed to be absolutely irreducible. In what follows, when I say a 'fiber' I mean a regular fiber, that is, the scheme theoretic pullback of a non-cusp copy of $Spec(Z)$ in $\mathbb{P}^1$.

The $\mathcal{O}_F$-module structure.

First is the detail about how to show that the ring $\mathcal{O}_F$ acts on $M_4(X)/(M_2(X)\gamma(p))$ where $\gamma(p)$ is the modular form corresponding to the global section of $\mathcal{O}(1) = \Omega_{\mathbb{P}^1}(C)$ with C the divisor of the three cusps of $\mathbb{P}^1$, interpreted as a one form on the integer projective allowed simple poles at 3 points. If we were working over a field we would twist the residue sequence by F giving

$$0 \to \Omega_X(C) \to \Omega_X(C+F) \to \mathcal{O}_F(C) \to 0.$$

An exact sequence which is an integer form of this where i is the inclusion of the fiber (now things are indexed so $\mathcal{M}_2(X) = \Omega_X(log\ f_X^{-1}C)^0 = f_X^*\Omega_{\mathbb{P}^1_{\mathbb{Z}}}(log\ C)$)

$$0 \to \mathcal{M}_2(X) \stackrel{\gamma(p)}{\to} \mathcal{M}_4(X) \to i^*\mathcal{M}_4(X) \to 0$$

with global sections sequence still exact

$$0 \to M_2 \stackrel{\gamma(p)}{\to} M_4 \to i^*\mathcal{M}_4(X) \to 0.$$

Then $M_4(X)/(M_2(X)\gamma(p))$ is an invertible sheaf on the fiber F, and its endomorphism ring is isomorphic to $\mathcal{O}_F$.

This also gives the $\mathcal{O}_F$ action on an appropriate quotient group $H^1(X \setminus \mathbb{C},\ \mathbb{C})/H^0(X \setminus C, \mathbb{C})$ as we've mentioned elsewhere, for ordinary cohomology of the transcendental points with cusps deleted.

The action of $\nabla\gamma(p)$.

Next is the detail about how to get the matrix $L\nabla\gamma(p)$.

We get it by identifying this element with an element of $\mathcal{O}_F$ and using the $\mathcal{O}_F$ action.

The principal part $\nabla\gamma(p) \in \mathcal{P}(O_X(F))$ does not belong to the kernel term in the exact sequence

$$0 \to \Omega_X(F) \to \mathcal{P}(O_X(F)) \to \mathcal{O}_X(F) \to 0$$

but after applying i^* it does, and gives the element, let's still call it $\nabla\gamma(p)$ but now in $i^*\Omega_X(F)$. This means that from the exact sequence

$$0 \to i^*O_X\nabla\gamma(p) \to i^*\mathcal{P}(\mathcal{O}_X(F)) \to \mathcal{P}(\mathcal{O}_F(F)) \to 0$$

mapping onto

$$0 \to 0 \to \mathcal{O}_F(F) \to \mathcal{O}_F(F) \to 0$$

the result is (no need to notate twisting by F since $\mathcal{O}_F(F)$ is principal)

$$0 \to \mathcal{O}_F\nabla\gamma(p) \to i^*\Omega_X \to \Omega_F \to 0.$$

If middle term contains a principal generator ω so that

$$i^*\Omega_X = O_F \cdot \omega$$

then there is an element of $\mathcal{O}_F$ which is the quotient $\frac{\nabla\gamma(p)}{\omega}$, and the action of this via the $\mathcal{O}_F$ action on $M_4/(M_2\gamma(p))$ is what we call $\phi = L\nabla\gamma(p)$.[3]

[3]If there is no principal generator, one may look ahead to note 9; choose a $\mathbb{Z}$ basis of $\mathcal{O}_F$ and let $L\nabla\gamma(p)$ the matrix whose columns coordinatize a $\mathbb{Z}$ basis of the different ideal Φ as correctly defined on the first page of the note.

More geometric version.

Finally, about making things more geometric, one can say that the reason $\nabla\gamma(p)$ was not contained in $i^*\Omega(F)$ is only because we need to allow more poles. I am not sure if I described this correctly in 'conclusions' but certainly we can do this.

With a, b, c, d as they are there, then

$$d\ log\ \frac{au+bv}{cu+dv}$$

has poles on two disjoint fibers, one over p and one over the point q where the sectionj $\tau(p)$ is zero. When we pull back along the fiber over p the pole over q will disappear.

This is the unique spanning section (up to sign) of $\Omega(F_p + F_q)$ where F_p is the same as what we call F, that is, fiber over p and F_q is the fiber over q, if we take X to be the projective line itself.

So this same element exists when X is general, and it is uniquely characterised up to sign.

So we then use the exact sequence

$$0 \to (\mathcal{O}_F \times \mathcal{O}_{F'})\nabla(\gamma(p)\tau(p)) \to \mathcal{P}(\mathcal{O}_X(F_p + F_q)) \to \Omega_{F_p} \times \Omega_{F_q} \to 0$$

This may be slightly mistyped but anyway pulling back via i^* to just F gives what we had before, but the point is that we are taking the residue

$$Res\ d\ log \frac{au+bv}{cu+dv}$$

on F_p and F_q and ignoring the residue on F_q.

The set of all residues on F is a copy of $\mathcal{O}_F$ (still assuming $i^*\Omega_X$ principal) and we are looking at one particular residue. I think that this short-cuts that complicated way of defining $\phi = L\nabla\gamma(p)$. I think it is now simply

$$L\nabla\gamma(p) = Res\ d\ log\frac{au+bv}{cu+dv},$$

where we mean, ignore the residue on F_q and just consider the residue on $F = F_p$.

Now, if this is right, what it means for the discriminant to be 1 or -1 is that the form

$$d\ log\ \frac{au+bv}{cu+dv}$$

takes the full pole everywhere along F.

Now, this fails to have a pole at any ramified cusp, and these meet F so its residue is going to develop zeroes on ramified cusps which will ruin some of the poles, so it seems that it is rarely if at all possible except for the identity map $X = \mathbb{P}^1 \to \mathbb{P}^1$ for the discriminant to be that small.

The condition of rationality, necessary and sufficient, is that the vanishing of the residue is only as much as matches the constraints which glue together the copies of $Spec(Z)$ along the cusps in the appropriate way. This is exactly the issue which we analyzed very carefully in section 1, but under simplifying hypotheses there. (Note also, there can be finitely many whole fibers of the absolute map $X \to Spec(\mathbb{Z})$ which ramify, and one of the things condition v) did was to avoid this choice of prime.)

One way of proceeding is to begin resolving the singularities of the fiber while simultaneously changing the meromorphic form correspondingly, if the fiber is completely resolved and the meromorphic function not reduced to a holomorphic generator, the fiber must contain a nonrational point and conversely.

Example: the Fermat curves

Let's calculate the global sections of symmetric powers of one-forms on the upper half-plane which are invariant under the p commutator subgroup of $\Gamma(2)$, for a prime p, and holomorphic at cusps (upstairs). Here, we only are going to 'set out the stall' in the way of approaching another technique; that is, observing the meromorphic residue $Res\ d\ log\ \frac{au_0+bu_1}{cu_0+du_1}$ while resolving singularities in a fiber.

For the Fermat curve Y defined by $x^p + y^p = z^p$ in coordinates $[x : y : z]$ there is an obvious first choice of Belyi function, sending $[x : y : z]$ to $[x^p : y^p]$, which branches over $0, -1, \infty$ just as needed.

Then there are no further choices to make; the group is the p commutator subgroup of $\Gamma(2)$. All relations in the ring $M(Y)$ of symmetric products of one-forms invariant for the group and holomorphi at cusps, follow from one relation in a larger ring, once we embed the subring as the terms in

$$\oplus_{i=0}^{p-1}\mathbb{C}[x, y]z^i \qquad (1)$$

of total degree divisible by p. As our convention (no longer using Dolgachev's convention) we will not look at odd weights, and let's use the convention where the terms of weight $2k$ correspond to monomials of degree pk.

Just to interject a comment, 'holomorphic at the cusps' has meaning even without adjoining any boundary points to the upper half plane; if we write a form on $\mathbb{H}$ as $f(\tau)d\tau$ for f arbitrary, while the one-form may be multivalued on the algebraic curve, we can make sense of the notion that $f(\tau)$ is a multivalued holomorphic function, and that holomorphicity is what is equivalent to holomorphic at a boundary cusp on $\mathbb{H}$, but it makes sense 'downstairs,' and it means that the multivalued function $f(\tau)$ is holomorphic, and the orders of the poles of the multi-valued form $f(\tau)d\tau$ are the simple poles of $d\tau$ at cusps, minus the zeroes of f.

The ring $M(2)$ (only looking at the even part) is the subring generated by x^p and y^p, and the branched covering degree is $\alpha = p^2$. The Euler characteristic of the Fermat curve is $p^2 - 3p$ as one can see from the adjunction theorem or otherwise.

Note too that this is consistent with the rule $f - c = 2g - 2$ as we have regular fibers with p^2 elements, and p cusps over each of $0, 1, \infty$.

We can just count the monomials in (1) of degree kp, there are $kp+1-i$ of them in $\mathbb{C}[x,y]z^i$ when that number is not negative, so for k not zero there are

$$\sum_{i=0}^{p-1}((k-1)p+1+p-i) = (k-1)p^2 + \frac{(p+1)(p+2)}{2} - 1$$
$$= \frac{1}{2}(2k-1)p^2 + \frac{3}{2}p. \quad (2)$$

We can check our prediction in terms of genus and degree. We predicted that this should equal

$$(k+1)\cdot 1 + k\cdot(\alpha - 1 - g) + (k-1)\cdot g. \quad (3)$$

We have

$$\alpha = p^2$$
$$g = \frac{1}{2}(p^2 - 3p + 2)$$

and it is indeed true that substituting these values of α and g into (2) yields (3).

We also predicted then something that is at least not immediately obvious from the algebra: that when we view (1) as a ring using the Fermat equation, the components of total degree divisible by p have as a free basis over $\mathbb{C}[x^p, y^p]$ the element 1 together with the monomials of degree p (always excluding any multiple of z^p) excepting x^p, y^p and the monomials of degree $2p$ excepting any of these two times a monomial of degree p. The total number when these are added is exactly p^2 corresponding to the rule

$$1 + (\alpha - 1 - g) + g = \alpha = p^2.$$

It follows immediately from what we have already said that there is an explicit representation of the subringf based on the monomials of degree divisible by p, by holomorphic symmetric products of one-forms on the upper half plane, invariant by the p commutator subgroup of $\Gamma(2)$, and which are 'holomorphic at the cusps' at the boundary of the upper half plane.

In fact, when we look at what these turn out to be, we see that the Fermat equation simply corresponds term-by-term with the Jacobi sum formula.

That is, we are led, with no essential choices possible, to the representation

$$\begin{aligned} x^p &\mapsto \frac{15}{16}\theta(0,\tau)^4 d\tau + \frac{1}{16}\theta(1/2,\tau)^4 d\tau \\ y^p &\mapsto \frac{1}{16} e^{i\pi\tau}\theta(\tau/2,\tau)^4 d\tau \\ z^p &\mapsto \theta(0,\tau)^4 d\tau \end{aligned} \quad .$$

The Fermat equation

$$x^p + y^p = z^p$$

corresponds to the tautology known as the Jacobi equation

$$\theta(0,\tau)^4 = \theta(1/2,\tau)^4 + e^{i\pi\tau}\theta(\tau/2,\tau)^4.$$

The role of the λ function is in the fact that now

$$\frac{1}{16}\lambda(\tau) = \frac{y^p}{z^p}$$

and so

$$(\frac{x}{z})^p = 1 - \frac{1}{16}\lambda(\tau).$$

Once we have represented y^p as above, we find the so-called q expansion of $x^{p-j}y^j = (\frac{x}{y})^{p-j}y^p$ for $j = 0, 1, 2, ...p-1$ explicitly by substituting into the equation

$$x^{p-j}y^j \mapsto (1 - \frac{1}{16}\lambda(\tau))^{\frac{p-j}{p}} e^{i\pi\tau}\theta(\tau/2,\tau)^4 d\tau.$$

We can extend this to $j = p$ by the same formula without a contradiction and more generally when $a + b + c$ is divisible by p

$$x^a y^b z^c = \theta(0,\tau)^{4(a+b+c)/p}(1 - \frac{1}{16}\lambda(\tau))^{\frac{a}{p}}(\frac{1}{16}\lambda(\tau))^{\frac{b}{p}} d\tau^{\otimes \frac{(a+b+c)}{p}}$$

The leading coefficient of the q expansion $\frac{1}{16}\lambda(\tau)$ is equal to 1 and all other coefficients are integers.

We should perhaps make a comment about plane curves and Puiseux expansions, that we already know the so-called q expansion of $\lambda(\tau)$ as $\lambda(\tau) = 16q - 128q^2 + 704q^3 ....$ which results when we set $q = e^{i\pi\tau}$ in the formulas above. And the Taylor series of the $(1 - \frac{1}{16}\lambda(\tau))^{j/p}$ then yield q expansions for all the monomials of degree a multiple of p in (1) just by replacing $\lambda(\tau)$ itself by its q expansion.

That is to say, working over $\mathbb{C}$, we have given an example of the case of the p commutator subgroups of the $\Gamma(2)$, and we see that we just obtain the Fermat equation represented as the same as the Jacobi equation.

And it was not necessary to use Newton polygons, or Puiseux theory. This is because the Fermat curves are smooth plane curves.

In this sense, every curve is a plane curve; that is, if we are willing to allow compactifications where we adjoin an algebraically singular point at cusps, we can always represent our curve as a plane curve in this way; but the Puiseux theory becomes more nontrivial. Or, we can always represent a curve as a smooth curve, but it may not be a plane curve.

Now, since our equation (the Fermat equation) is an integral equation, there is already implicitly a $\mathbb{Z}$ form. Unlike the case of our assumptions in 'outline geometric proof of Mordell's conjecture,' there is not rationality at the cusps. Above each of the three cusps in $X(2)$ are a pair of cusps, one rational and one whose residue field is cyclotomic.

A non-cusp integer point of X(2) is [b:-a] distinct from [0:1], [1:0], [1:1], and choosing c,d so the determinant of

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is equal to 1, we can make sense of the rational function

$$\frac{ax^p + by^p}{cx^p + dy^p}$$

It is a ratio between two one-forms for which we've given the q expansion explicitly.

A basis of the scheme defined by $ax^p + by^p = 0$ consists of the degree $2p$ monomials modulo $ax^p + by^p$ times the degree p monomials, and setting $cx^p + dy^p$ to 1 has the same effect as just ignoring monomials of any other degree.

This makes sense even if a, b are not p'th powers; the ring of forms has no element corresponding to x or y themselves.

We consider the residue of the logarithmic deRham differential

$$d\ log \frac{ax^p + by^p}{cx^p + dy^p}$$

on the fiber over $ax^p + by^p = 0$. Over $\mathbb{Z}$, the reduced fiber over each of the three cusps has normalized coordinate ring a cartesian product of $\mathbb{Z}$ with the p'th cyclotomic integers.

When we look at the logarithmic deRham differential, we are taking the differential of an actual analytic function on the complex points of the Fermat curve. But the fact that the relations among such things are algebraic means that we can understand the residue as a section of a line bundle on a scheme finite type over the integers. More simply stated, the rational function shown above is represented by a section of the structure sheaf of the algebraic curve over the integers, and it is a rational section in that it is a section over an open subset smaller than the entirety of the scheme.

The same expression describes a complex analytic function which we have written down explicitly via its q expansion.

The specific issue here is what seems to occur in SGA7 or in notions of Neron models and elliptic curves. The idea originally by Frey, or students of Serre at this juncture says, what if we knew that a,b are p'th powers, and they add to c, another p'th power?

After all, $\lambda(\tau)$ is the series describing the cross ratio, if the elliptic curve with lattice $\mathbb{Z} + \mathbb{Z}\tau$ is described as branching over four points; and if one of these is 0 and one is ∞, such a choice of a, b is describing a Frey curve (though it's not clear where Frey lost the factor of 16).

However, it is not intrinsic to the geometry of the Fermat equation that the universal cover of $\mathbb{C} \setminus \{0, 1\}$ happens to occur when one considers elliptic curves and cross-ratios. Rather the Fermat equation describes the relation that occurs in nature among symmetric powers of forms on $\mathbb{H}$ invariant under the p commutator subgroup of $\Gamma(2)$.

To normalize the coordinate ring of a regular fiber requires separating the components. Some gluing can exist in the noncuspidal ramification at p itself, but the remainder has to take place where the fiber components meet cusps, and if $a, b, a+b$ are relatively prime the points where a fiber components meets each cusp are disjoint. Note how this is vaguely reminiscent of the a, b, c conjecture.

A rational component of the fiber is one along which the vanishing of the residue exactly matches what should be predicted by an intersection theory; at non-rational components there is further vanishing of the residue.

One of the hypotheses of the outline Mordell discussion was that the sheaf of one forms on the 'modular' curve over Z is locally free. That is not true here due to the non-cuspidal ramification at p, but it is true away from p. There we analyzed the case when Ω is locally free and all cusps are rational using a notion of duality. Here we haven't begun describing any relevant intersection theory.

The residue calculation

Earlier I was mentioning that it is not difficult to write down hypotheses which strengthen Mordell's conjecture, with the conclusion that particular modular curves have no noncuspidal rational points.

The hypotheses where this was written down most carefully included a type of Galois condition, that if a fiber contains any rational point, all must be rational.

The Fermat curves, also, as I mentioned, arise with homogeneous coordinate rings the global sections of symmetric powers of one-forms on the upper half-plane $\mathbb{H}$ holomorphic at cusps on $\mathbb{H}$, which are fixed by the p commutator subgroup of $\Gamma(2)$. The structure map to $X(2)$ is a Belyi map as always, and a corresponding integer structure on the Fermat curve coming from specifying the branch points to be $0, -1, \infty$ gives the usual integer structure of the p'th Fermat curve.

Recent news stories remind us of the depth of history which underlies the Fermat theorem, the friendship between Shimura and Taniyama in Japan in the 1950's. Wiles' proven theorem is an instance of the same strenghtened conclusion, that the Fermat curve has no noncusidal rational points except when its genus is less than two. The case of genus one also has no noncusipdal ratoinal points, leaving the pythagorean triples and the triples with an entry of zero being the only integer solutions of the Fermat equation.

In earlier cases, I was relating the different element to duality and trace forms. That would be the aim here too, and to relate it to the intersection matrix among the components of F, if a suitable intersection theory can be found.

Let's take p to be a prime, although this is not necessary.

The way to calculate integer points in a fiber over a noncuspidal point $[x^p : y^p] = [a : b]$, I've claimed, is to choose integers

$$A = -b$$
$$B = a$$
$$C$$
$$D$$

with C, D chosen so that $1 = AD - BC$.

The matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

belongs to $SL_2(\mathbb{Z})$.

The residue of the form

$$d\ log \frac{Ax^p + By^p}{Cx^p + Dy^p}$$

on the fiber F defined by the equation $Ax^p + By^p = 0$ among the set of all residues, defines a 'different element' in a twist of the restricted canonical sheaf. Let F be divisor defined by the numerator, and F' the divisor defined by the denominator. Since we've used $SL_2(\mathbb{Z})$ they are disjoint even scheme-theoretically.

The way we will calculate the residue is to use principal parts. Although there are nice ways of describing this, in terms of meromorphic connections, let's just consider it as some notation to write $\nabla(x), \nabla(y), \nabla(z)$ for the deRham differentials of x, y, z in *affine* space which can be viewed as global sections of first principal parts in various ways which are a basis rationally (in the sense of rational functions).

In the first instance, let's work over $\mathbb{Z}[1/p]$ so that we can identify, on the projective plane, the first principal parts of $\mathcal{O}(p)$ with $\mathcal{O}(p-1)$ tensor the first principal parts of $\mathcal{O}(1)$.

Then we may write

$$\mathcal{P}(\mathcal{O}(p)) = \mathcal{O}(p-1)\nabla(x) \oplus \mathcal{O}(p-1)\nabla(y) \oplus \mathcal{O}(p-1)\nabla(z).$$

This direct sum of three line bundles contains a trivial line bundle, spanned by

$$\frac{1}{p}(\nabla(x^p) + \nabla(y^p) - \nabla(z^p)).$$

Here we are allowed to use the rule of a derivation, and this is the same as

$$x^{p-1}\nabla(x) \oplus y^{p-1}\nabla(y) \oplus z^{p-1}\nabla(z).$$

Because the basic sections have no common zero anywhere on the projective plane, this is a line bundle inclusion (locally split).

Therefore the third exterior power of $\mathcal{P}(\mathcal{O}(p))$ is the same as the second exterior power of the quotient vector bundle. The quotient vector bundle restricts on the Fermat curve X to the principal parts mod torsion of $\mathcal{O}_X(D)$ where now D is the restriction to X of any hypersurface of degree p. We may take in particular our fiber F.

Putting things together a bit, the third exterior power of our rank three vector bundle corresponds to a hypersurface of degree $3p-3$ on the projective plane; and so the second exterior power of the torsion free princpal parts of the locally free sheaf $\mathcal{L} = \mathcal{O}_X(F)$ on the Fermat curve corresponds to the intersection of the Fermat curve with a degree $3p-3$ hypersurface, and we may think of this as a divisor of degree $3p^2 - 3p$ on the Fermat curve.

It is always true, because of the sequence

$$0 \to \Omega_X \otimes \mathcal{L} \to \mathcal{P}(\mathcal{L}) \to \mathcal{L} \to 0$$

which exists even if neither Ω_X nor $\mathcal{L}$ is locally free, that there is a map $\phi : \Omega \otimes \mathcal{L}^{\otimes 2} \to \Lambda^2 \mathcal{P}(\mathcal{L})$, and in cases like the case at hand, when $\mathcal{L}$ is locally free, so the exact sequence is locally split, this is an isomorphism. Here the letter $\mathcal{P}$ should be taken to be torsion free first principal parts (the reduction moduo torsion) and the kernel term should refer to the torsion-free Kahler differentials, which is the reduction of the one-forms modulo torsion.

We have said that this corresponds to a degree $3p^2 - 3p$ divisor, the degree of the canonical divisor of the Fermat curve is $p^2 - 3p$ and we see that this agrees with the degree of $\mathcal{L} \otimes \mathcal{L} \otimes \Omega$, that is, it is larger than the canonical degree by $2p^2$ which is twice the degree of F.

The inverse isomorphism, on local sections f, g of $\mathcal{L}$, is given

$$\nabla(f) \wedge \nabla(g) \mapsto fg \; d \; log(f/g).$$

1. Remark. There is a slight subtlety in that we are not allowed to factorize the expression $fg\ d\ log(f/g) = g^2 d(f/g) = f^2 d(g/f)$ as a tensor such as

$$g^{\otimes 2} \otimes d(f/g)$$

because the right factor is not a section of one forms except when $g \neq 0$. However working locally the way to see that the expression $fdg - gdf$ corresponds to a local section of $\mathcal{L} \otimes \mathcal{L} \otimes \Omega$ in a neighbournood of a point of the Fermat curve is to choose a local section s of $\mathcal{L}$ not zero at that point. Then (and the analagous thing is true for higher exterior powers in cases of higher dimension, it is a property of contracting under the Euler derivation) there is a homogeneity property of the expression so that

$$fdg - gdf = s^2(\frac{f}{s}d\frac{g}{s} - \frac{g}{s}d\frac{f}{s}).$$

The right side is an expression involving rational functions which are well-defined at our point, giving a local section of Ω, and the factor s^2 is a local section of $\mathcal{L}^{\otimes 2}$. So this can be interpreted as a tensor product then, as we've worked locally

$$s^{\otimes 2} \otimes (\frac{f}{s}d\frac{g}{s} - \frac{g}{s}d\frac{f}{s}).$$

Returning to our exposition, to obtain our 'different' element, under the isomorphism between global sections of $\mathcal{L}$ and polynomials of degree p in x, y, z modulo the Fermat relation, we shall multiply our logarithmic derivative by two terms

$$(Ax^p + By^p)(Cx^p + DY^p)d\ log\ \frac{Ax^p + By^p}{Cx^p + Dy^p}$$

and pass to the image under the isomorphism ϕ, which is

$$(A\nabla(x^p) + B\nabla(y^p)) \wedge (C\nabla(x^p) + D\nabla(y^p)).$$

To interpret this as an element of the third exterior power of principal parts of $O(p)$ on the projective plane (later restricted to F) we will wedge with the basis element of the trivial sub bunde, that is we will wedge this expression with

$$\nabla(x^p) + \nabla(y^p) - \nabla(z^p).$$

The result is

$$(xyz)^{p-1}\ determinant \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ 1 & 1 & -1 \end{pmatrix} \nabla(x) \wedge \nabla(y) \wedge \nabla(z)$$

$$= -(xyz)^{p-1}\nabla(x) \wedge \nabla(y) \wedge \nabla(z).$$

The coefficient monomial is now interpreted as a section of the line bundle $\mathcal{O}(3p-3)$ on the projective plane, and the intersection of F with the locus where this is zero is the subscheme defined by the 'different' element.

It follows that when working over $\mathbb{Z}[1/p]$, if F has a rational point then the different ideal of $\mathcal{O}_F$ is generated by $(xyz)^{p-1}$ which on each compopnent is a root of unity times an integer, and the integer does not depend on which component we are looking at.

2. Remark. In the way of explaining why we have obtained our different element as a restriction of a section of $\Omega(2F)$ when one should expect to use $\Omega(F)$, if we interpret $\mathcal{L}$ as the sheaf of rational functions with at worst simple poles on F, then the correspondence between homogeneous polynomials of degree p modulo the Fermat relations, and global sections of $\mathcal{L}$, is that a polynomial f corresponds to the rational function $\frac{f}{Ax^p+By^p}$. Thus when we multiplied our logarithmic derivative by the product of polynomials $(Ax^p + By^p)(Cx^p + Dy^p)$ the corresponding sections which we multiplied by were $(1)(\frac{Cx^p+Dy^p}{Ax^p+By^p})$. The numerator of the coefficient has divisor of zeroes F' disjoint from F and acts to multiply the residue by 1 since it restricts to 1 on F. The denominator has a simple zero on the Cartier divisor F; the product represents the same residue as the logarithmic derivative, though now on a pole of order two.

We can now state what we've proved.

3. Theorem. Let $X \subset \mathbb{P}^2$ be the Fermat curve defined by $x^p+y^p = z^p$ for p a prime number. For each $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in Sl_2(\mathbb{Z})$, let F be the fiber of $X \to \mathbb{P}^2 = \{[x^p : y^p]\}$ over the point defined by $Ax^p + By^p = 0$. Then

i) the scheme-theoretic zero locus of the residue $Res\ d\ log\ \frac{Ax^p+By^p}{Cx^p+Dy^p}$ on F agrees on the complement of the scheme $p = 0$ with the restriction to F of the scheme-theoretic zero locus of the section $(xyz)^{p-1}$ of the line bundle $\mathcal{O}_{\mathbb{P}^2}(3p-3)$ (which restricts to $\mathcal{L}^{\otimes 2} \otimes \omega_X$).

ii) If F has a rational point $[a : b : c]$ then each of x, y, z restricts on each irreducible component of F to an integer times a root of unity.

iii) Still assuming a rational point $[a : b : c]$, taking a, b, c pairwise coprime, the different ideal of $\mathcal{O}_F[1/p]$ is principal generated by the element $(abc)^{p-1}$ of the integers $\mathbb{Z}$ viewed as the characteristic subring.

iv) Still assuming a rational point $[a : b : c]$ with a, b, c coprime, if $s \in \mathbb{Z}$ is such that the scheme $Spec(\mathcal{O}_F[1/s])$ is not connected then abc is a divisor of a power of s.

Proof. We've proved all except iv) which will follow from first considerations of an intersection theory for the components of F.

It is interesting to consider the case of $p = 2$. Then the different element is precisely compatible with the intersections of the four components of F in the Pythagorean case. F is comprised of four copies of $Spec(\mathbb{Z})$ which each intersect the other three, one each, with intersection number a, b, c.

In general, when there is a rational point and $[a^p : b^p]$ is not one of the cusps, F has $(p + 2)$ irreducible components corresponding to equivalence classes of solutions $[a\omega^i : b\omega^j : c\omega^k]$, and note that adding the same constant modulo p to (i, j, k) does not affect the solution point, nor does multiplying (i, j, k) by the same nonzero number modulo p, which acts as an automorphism on an irreducible component. Thus the $p+2$ irreducible components of F, when there is a rational point, correspond bijectively with orbits of the one dimensional affine group acting on F_p^3.

One naive type of intersection theory would associate to a pair of minimal prime ideals $P, Q \subset \mathcal{O}_F$ the number of elements in the cokernel of

$$\mathcal{O}_F \to \mathcal{O}_F/P \times \mathcal{O}_F/Q.$$

Our calculation of the different element (up to p torsion) would be no different if we had started with an equation which does have a solution such as

$$-8x^3 + 7y^3 = z^3.$$

This has the solution $[x : y : z] = [2 : 3 : 5]$, it still has different element $(2.3.5)^2 \in \mathcal{O}_F[1/3]$. Because it does have a rational solution, we can look at the naive intersection numbers.

We can consider any such cubic equation, obtained by modifying the Fermat equation by including fixed rational integer coeffficients of x^p and y^p so that it does have a solution,

Writing $C_{i,j,k}$ for the irreducible component corresponding to the orbit including $[a\omega^i : b\omega^j : c\omega^k]$ we can calculate intersection numbers in a few examples. In terms of representatives for our subscript sequences, the five irreducible components of $\mathcal{O}_F$ are $C_{0,0,0}, C_{0,0,1}, C_{0,1,0}, C_{1,0,0}, C_{0,1,2}$. The first is a copy of $Spec(\mathbb{Z})$, and the others are copies of $Spec\mathbb{Z}[\omega]$ for ω a p'th root of unity.

From a few examples it seems that the naive intersection numbers in thise sense are

$$C_{0,0,0} \cdot C_{0,0,1} = 3c^3$$
$$C_{0,0,0} \cdot C_{0,1,0} = 3b^3$$
$$C_{0,0,0} \cdot C_{1,0,0} = 3a^3$$
$$C_{0,0,0} \cdot C_{0,1,2} = 3$$

$$C_{0,1,2} \cdot C_{0,0,1} = 3a^2b^2c$$
$$C_{0,1,2} \cdot C_{0,1,0} = 3a^2bc^2$$
$$C_{0,1,2} \cdot C_{1,0,0} = 3ab^2c^2$$

$$C_{0,0,1} \cdot C_{0,1,0} = 3a^2bc$$
$$C_{1,0,0} \cdot C_{0,1,0} = 3abc^2$$
$$C_{0,0,1} \cdot C_{1,0,0} = 3ab^2c.$$

To reiterate, we're considering a cubic equation of the type $qx^3 + ry^3 = z^3$ for q, r integers such that there is a solution with integers $[a : b : c]$

The fact that the different element belongs to the characteristic subring requires that for any prime s besides 3 which divides into any of a, b, c the intersection graph must remain connected when we replace all the intersection numbers by their s-adic valuation (or s primary part). The equations above verify that this is true in examples we've considered. The intersection numbers between fiber components and in fact the isomorphism type of $\mathcal{O}_F$ appear to be independent of the choice of q, r. The intersection numbers in examples have fixed expressions as monomials in a, b, c, p.

Let's restrict to the case p odd and write the equation $x^p + y^p + z^p = 0$, to make the Fermat curve more symmetrical. Also, let's use a simplified assignment of coordinates, taking $\lambda(\tau) = \frac{y^p}{z^p}$ without any factor of 16.

That is, any finite group which has a two dimensional representation has a corresponding action on the Riemann sphere, and here we consider $\mathbb{P}^1$ to be the projectivication of the two dimensional irreducible representation of the symmetric group $\mathbb{S}_3$. The group we take as acting by permuting the coordinate variables $[x : y : z]$.

The quotient of the Fermat curve X/S_3 is then 'modular' with group Γ so that

$$\Gamma(2)^{(p)} \subset \Gamma \subset \Gamma(1)$$

and it is index p^2 in $\Gamma(1)$, but not a normal subgroup. That is, the quotient $\Gamma(1)/\Gamma(2)^{(p)}$ is isomorphic to a semidirect product $F_p^2 \rtimes S_3$ and Γ is the inverse image of a subgroup copy of S_3 obtained by choosing a trivialization of the extension cocycle.

The map $X \to X/S_3$ covers the map $\mathbb{P}^1 \to \mathbb{P}^1$ which sends the λ invariant to the j invariant, at least if we parametrize things so that

$$j = \frac{(1 + \lambda + \lambda^2)^3}{\lambda^2(1 + \lambda)^2} = \frac{(x^{2p} + (xy)^p + y^{2p})^3}{(xyz)^{2p}}$$

To consider $j_0 = \frac{q}{s}$ with $q, s \in \mathbb{Z}$ we again choose A, B, C, D with $A = s, B = -q, AD - BC = 1$, as before the residue over the fiber at j_0 agrees in a neighbourhood of that fiber with

$$Res\ d\ log\ \frac{A(x^{2p} + (xy)^p + y^{2p})^2 + B(xyz)^{3p}}{C(x^{2p} + (xy)^p + y^{2p}) + D(xyz)^{3p}}.$$

In some sense, we know the answer already. If we invert 6 the fiber in X/S_3 over j_0 is the isomorphic image of the fiber in X of any one of the λ values mapping to j_0, and its different ideal must again be generated by $(abc)^{p-1}$ viewed as an element of the characteristic subring with p inverted, if there is a rational point.

But whereas this fiber has $(p+2)$ irreducible components in this case, its inverse image, the full fiber in X over j_0 now has $6(p+2)$ irreducible components, and so we expect the different element of the full fiber to have an additional factor which corresponds to intersections between fibers over different λ values.

The residue calculation is similar to what we have done already; one interesting thing is that the formula does not explicitly involve j_0 at all, meaning, there is again a polynomial expression representing a section of a line bundle, now $\mathcal{O}(13p-3)$, which defines in its restriction to the fiber over each value j_0 to the different subscheme of that fiber.

An issue is that the there is a line sub bundle of $\mathcal{P}(\mathcal{O}(6p))$ once we invert $6p$ which is 'spanned' by $\nabla(x^p+y^p+z^p)$

$$\mathcal{N} = \mathcal{O}(5p)\nabla(x^p+y^p+z^p).$$

Exactly the determinant which we already have calculated, in this case a Jacobian matrix made of the numerator and denominator of the fraction shown above along with the Fermat equation, is a generating global section of the line bundle made by wedging our global section of $\Lambda^2\mathcal{P}(\mathcal{O}(6p))$ with the whole line bundle $\mathcal{N}$, but then tensoring with the dual $\widehat{\mathcal{N}}$.

The generating section is $AD-BC$ which is 1, times

$$2p(xyz)^{2p-1}3(x^{2p}+(xy)^p+y^{2p})^2 det\begin{pmatrix} yz & xz & xy \\ 2px^{2p-1}+px^{p-1}y & 2py^{2p-1}+py^{p-1}x & 0 \\ x^{p-1} & y^{p-1} & z^{p-1}\end{pmatrix}$$

$$= 6p^2(xyz)^{2p-1}(x^{2p}+(xy)^p+y^{2p})^2(x^p-y^p)(x^p-z^p)(y^p-z^p).$$

Ignoring the factor of $6p^2$ we see that the different element for the six fibers viewed as disjoint, which was $(xyz)^{p-1}$ has now been multiplied by factors $x^py^pz^p(x^p-y^p)(x^p-z^p)(y^p-z^p)(x^{2p}+(xy)+y^{2p})^2$.

Let's show that all eight of the new factors, even with identical multiplicity (!) describe points where the six fibers intersect coming solely from congruences among values of the λ function.

4. Lemma Let a, b be coprime integers, and let $j \in \mathbb{P}^1/S_3$ be the image of $[a : b]$ in $\mathbb{P}^1$. Write $c = -a - b$. The fiber in $\mathbb{P}^1$ over j consists of the copies of $Spec(\mathbb{Z})$ indexed by $[a : b : c], [a : c : b], [b : a : c], [b : c : a], [c : a : b], [c : b : a]$, and the different element is represented on every irreducible component (after inverting 6) by the integer $6(abc)(a-b)(b-c)(c-a)(a^2+ab+b^2)^2 \in \mathbb{Z}[1/6]$.

Proof. This time the two by two determinant corresponds to the elements $(a^2 + (ab) + b^2)^3$ and $(abc)^2$. The coefficient of 6 may be removed since we are working over $\mathbb{Z}[1/6]$.

Next we will consider the diagram for X the Fermat curve

$$\begin{array}{ccc} X & \to & X/S_3 \\ \downarrow & & \downarrow \\ \mathbb{P}^1 & \to & \mathbb{P}^1/S_3 \end{array}$$

The fiber over a rational point of $\mathbb{P}^1/S_3$ is an image of a tensor product of the fibers $\mathbb{P}^1$ and of X/S_3 over that point; this describes a map from a tensor product of an algebra of rank p^2 as an abelian group, and one of rank 6 as an abelian group. The map is injective and finite. Even locally, if the discriminant of the fiber is equal to the product of the discriminants, then the coordinate ring of the fiber decomposes as the tensor product precisely with no further partial normalization.

5. Remark. Another way of thinking about the explanation in Remark 2. for why $\mathcal{L}^{\otimes 2}$ occurred, is that is that we are constructing a section of the *relative* canonical sheaf, and the factor of $\mathcal{L}^{\otimes 2}$ represents the pullback of the inverse of the canonical sheaf of $\mathbb{P}^1$. That is, when we wrote $\mathcal{L}^{\otimes 2} \otimes \Omega_X$ we might as well have written $f^*\Omega^{-1}_{\mathcal{P}^1} \otimes \Omega_X$.

This analysis will continue in 'the meaning of positive and negative.'

The meaning of positive and negative

Sometimes, in an abstract setting, when one wants to define what it means for a number to be positive, one will say that this means it is a sum of squares.

We all know that if we describe the square root of 2 algebraically, the solutions include two numbers, symmetric under an automrophism, which are negatives of each other. We were taught, if we are to think of positive numbers, one of them is extraneous.

It is a question, related to consistency of arithmetic, whether there is a purely algebraic proof that a sum of nonzero rational numbers which are positive by this definition cannot be zero.

Twisted schemes

When one is considering a subring of $\mathbb{Z}^n$, corresponding to a union of copies of $Spec(\mathbb{Z})$, there do exist nontrivial line bundles on such a union. If we consider two copies of $Spec(\mathbb{Z})$ meeting transversesly at the prime indexed by 5 in both, say, then we can construct a locally principal module by twisting the gluing identification by an element of $\mathbb{Z}/(5\mathbb{Z})^\times$. A twist by 1 or 4 would be inessential, because it lifts to an automorphism of either factor. And a twist by 2 and 3 would have the same effect, but there do exist twists of the free module, and I have not proven that the restricted canonical sheaf is not of this type.

Relation with the canonical sheaf

Let X be 'modular' corresponding to a finite index subgroup of $\Gamma(2)$, not required to be a congruence subgroup. Suppose ω_X is locally free (at non-cusp points). Let ω_0, ω_1 as usual be the basic elements (modular forms of weight to for $\Gamma(2)$, the usual coordinates on $\mathbb{P}^1$). choose numbers A, B, C, D with $AD - BC = 1$ so that the zero point of $A\omega_0 + B\omega_1$ is not a cusp, and let F be the fiber in X that point. Let $\mathcal{L} = \mathcal{M}_2(X)$, our line bundle spanned by ω_0, ω_1.

1. Lemma The residue

$$\delta = Res\ d\ log\ \frac{A\omega_0 + B\omega_1}{C\omega_0 + D\omega_1}$$

corresponds naturally to a section of the restriction to the fiber of $\omega_X \otimes \mathcal{L}^{\otimes 2}$; the ratio $\delta\omega_X^{-1}\mathcal{L}^{-2}$ is naturally an ideal in $\mathcal{O}_F$ and the module of one-forms on F, which is $\omega_X \otimes \mathcal{L}^{\otimes 2}/(\delta\mathcal{O}_F)$, is a locally prinicpal module of rank one over the ring $\mathcal{O}_F/(\delta\omega_X^{-1}\mathcal{L}^{-2})$. Finally, it happens to be prinicpal since the Picard group of $\mathcal{O}_F/(\delta\omega^{-1}\mathcal{L}^{-2})$ is trivial.

This lemma does not need proof, it is a statement of things we've said already, but before having been more vague about naturality.

That is, what I've called the different element ϕ in case the restriction of ω_X to a fiber happens to be principal (I do not know whether it always is), is the generator of the ideal in $\mathcal{O}_F$ which is defined by the formula

$$\Phi = \omega_X^{-1} \otimes \mathcal{L}^{-2} \otimes Res\ d\ log\ \frac{A\omega_0 + B\omega_1}{C\omega_0 + D\omega_1}$$

with $\mathcal{L}$ the line bundle spanned by ω_0, ω_1. It does not matter whether we use a lower case or upper case ω here, one sometimes denotes the canonical sheaf and the other the sheaf of Kahler differentials, but they are identical here as we assume Ω is locally free, an assumption which will be legitimized by inverting an appropriate integer. The fiber F' defined by $C\omega_0 + D\omega_1 = 0$ is disjoint from F and is to be disregarded, though $F + F'$ represents the pullback of the anticanonical class.

Tensor indecomposability

We mentioned that the Fermat curve X maps to the pullback of X/S_3 and $\mathbb{P}^1$ over $\mathbb{P}^1/S_3$, and it follows that the ring $\mathcal{O}_F$, a free abelian group of rank $6p^2$, of the fiber F over a rational (noncuspidal) value of j contains, as finite index within it, a tensor product of one of rank 6 with one of rank p^2.

The construction of the tensor factor of rank 6 depended on $a = -B, b = A$, and c adding to zero.

We can always tensor together the two factors abstractly, to make a ring of rank $6p^2$. However, the relation between adding and taking powers is that we may not embed the Spec of the tensor product as a closed subscheme of any curve with locally principal canonical sheaf unless the direct sum of the pulled back Kahler differentials modules is again locally principal. And, since it is supported on a discrete scheme with trivial Picard group, we may equally say 'principal.' However, we have this lemma.

2. Lemma. Let $\mathcal{F}, \mathcal{G}$ be locally principal sheaves on a Noetherian scheme. Then $\mathcal{F} \oplus \mathcal{G}$ is locally principal if and only if $\mathcal{F} \otimes \mathcal{G}$ is zero.

The proof is just to consider the specialization to one closed point, where we are speaking about one-dimensional vector spaces. Since the sum of the Kahler differentials modules from the separate factors is locally prinicipal (and even principal), it then requires that the support schemes of the two pulled back modules must be disjoint. Being on different sides of the tensor factor even requires

3. Lemma The spectrum of the tensor product of two algebras cannot be embedded as a closed subscheme any curve (over $\mathbb{Z}$) with locally principal canonical sheaf unless the discriminants of the factors are coprime.

Note that it is perfectly possible to embed a finite extension of the spectrum of such a tensor product, so this prohibition disappears after normalizing.

Duality

Let's look at the case $p = 2$. The ring of rank 4 is one which we constructed as a fiber over a λ value. It has a basis the four monomials $1, xy, xz, yz$.

The structure constants of the ring are determined by the rules $x^2 = a, y^2 = b, z^2 = c$, even while the ring does not contain any elements labelled with the letters x, y, z.

It is instructive to write the elements in rows according to the degree in which they previously had in the graded ring, even though there is not any grading anymore. We write

$$xy \quad xz \quad yz$$

$$1$$

and for instance $(xy)(yz) = bxz$.

Our element δ is represented by $(xyz)^{p-1} = xyz$, which is not in the ring. But we can construct a basis of ω_X over the fiber, using the monomials

$$x, y, z, xyz.$$

If we write all the monomials in the rows we have

$$\begin{array}{ccc} & xyz & \\ xy & xz & yz \\ x & y & z \\ & 1 & \end{array}$$

The rows of even height are sections of $\mathcal{O}_F$ while the rows of odd height are sections of ω_X on F. So for instance the equations

$$(xz)(y) = xyz$$

$$(xz)(z) = cx$$

describe ring elements acting on sections of ω_X and converting them to new sections.

Our different element, over $\mathbb{Z}[1/2]$, is $(xyz)^{p-1} = xyz$, since $p = 2$.

In any such diagram, the monomials in row $p-3$ are g in number where g is the genus of the ambient curve; they extend to a basis of the global sections of ω_X. For instance in case $p = 2$ there are none, and when $p = 3$ the monomal in the same position as 1 corresponds to the unique differential called dz on an elliptic curve.

Transpositions of negative eigenvalue

As for the ring of rank six, working over $\mathbb{Z}[1/6]$, it has different element

$$6(x^2-y^2)(x^2-z^2)(y^2-z^2)(x^2+xy+y^2)^2$$

assuming that $x+y+z=0$.

Here, we can see a picture proof of what this is saying. If we compare

$$[x:y:-x-y]$$

$$[y:x:-x-y]$$

then all three size two minor determinants are $\pm(x^2-y^2)$ In terms of the symmetric group action on $\mathbb{Z}^2$ a transposition fixes two integer lines, one each with eigenvalue $1,-1$.

But because

$$(x^2-y^2)=(x-y)(x+y)=(y-x)z$$

we see that there is an intersection, a congruence, with the cusp, which is the other type of ramification.

After pulling back, we have an equation instead

$$(x^{2p}-y^{2p})=(y^p-x^p)z^p.$$

This is geometrically explaining why the different element of the fiber in $\mathcal{L}^{\otimes 2}\otimes\omega_X[\frac{1}{6p}]$ divided by the different element of the disjoint union of its six parts was

$$6x^py^pz^p(x^p-y^p)(x^p-z^p)(y^p-z^p)(x^{2p}+(xy)^p+y^{2p})^2.$$

That is, it would have been better to write it as

$$6(x^{2p}-y^{2p})(x^{2p}-z^{2p})(y^{2p}-z^{2p})(x^{2p}+(xy)^p+y^{2p})^2.$$

Then the 6 seems to be describing the degree of the branched cover just as the factor of p^2 did on the other side (although this may be a coincidence), and the three next factors describe one each, a fixed line in $\mathbb{Z}^3$ with positive and negative eigenvalue for one of the transpositions, and finally we see the two fixed non rational lines, with eigenvalue $\pm e^{2\pi i/3}$.

'Specialization'

Before, we mentioned that if the fiber F over a λ value has any rational point, then on each of the $p+2$ components of that fiber, each of x, y, z 'specializes' to an integer times a p'th root of unity. More rigorously, each monomial in x, y, z of degree a multiple of p corresponds to an element of $\mathcal{O}_F$, and each rational monomial of degree congruent to zero modulo p also does. These must specialize to p'th roots of the corresponding rational monomials in a, b, c and so must equal the rational p'th root times a p'th root of unity. For example $(x/y)^p - (a/b)$ specializes to zero, and x/y must be one of the roots of $T^p - (a/b)$, which are a rational p'th root of a/b times a root of unity on each component. Then 'specialization' of x, y or z itself is an intersection of fractional ideals induced from $\mathbb{Z}$; the ideals are induced from $\mathbb{Z}$ so is the intersection, defining x, y, or z as an element of $\mathbb{Z}$ up to sign.

The case of a rational solution.

If the fiber has a rational point, then one of the λ values lying over its j value is rational; they are symmetric under S_3 so all six λ values are rational, and fiber over the j value is just a union of six copies of this union of $p+2$ components, intersecting various ways. We may not have the same choices of a, b, c as λ changes, but if there is a rational point in a fiber over a lambda value, then it is true for all the isomorphic fibers over the other lambda values lying over the same j value, that they too have this property, although for each there will be a different permutation.

So x on one component of one fiber may map to "a" times a root of unity and on a component of a part of the fiber lying over a different lambda value, but for the same j, will map to b times a different root of unity.

What that means, though, is that xyz always maps to abc times a root of unity on every component.

The issue about the different element not belonging to the ring is exactly that when we calculate the expression

$$(a^{2p} - b^{2p})(a^{2p} - c^{2p})(b^{2p} - c^{2p})(a^{2p} + (ab)^p + b^{2p})^2$$

of degree $10p$, each separate term, on each of a, b, c is corresponding to some integer times a root of unity, and here those roots of unity aren't even present in a^p, b^p, c^p. But we must multiply this by the inverse of the fractional ideal on the fiber which is generated by all degree $10p$ monomials. Now, they happen to generate the unit ideal; this has no significance, as it relates to our original choice of section which we call 1 which related to our choice of C and D in the matrix A, B, C, D.

But it means that the different element really is this integer, *under the assumption that the fiber contains a rational point.*

And the other factor $p^2(abc)^{p-1}$, the different element of the disjoint union of the 6 parts, likewise represents an integer times a (plus or minus) p'th root of unity as the different element, and also as an element of the ring, generating the ideal on the other side, describing the support of the ramification on the other branched cover.

So the only subtlety is that as you move around and consider different lambda values correponding to one j value, the factors like $(a^{2p} - b^{2p})$ permute among themselves. But this does not affect the fact that you can factor out $(abc)^p$ and that is constant on all components.

Now, the different elements of the algebras of rank p^2 and 6 were exactly these. This matches the product of the different elements of the two algebras, tells us also that the fiber has the same discriminant as does the tensor product.

Now, the ring of rank four which is the fiber in X/S_3 over j is contained in the ring of rank four over a λ value over j, and we have inverted 6 so that it is isomorphic to the S_3 invariants in the full fiber over j.

When we factorized the different element of the fiber over a rational j value as

$$[(abc)^{p-1}][(a^{2p}-b^{2p})(a^{2p}-c^{2p})(b^{2p}-c^{2p})(a^{2p}+(ab)^p+b^{2p})^2] \in \mathcal{O}_F[(1/6p)],$$

the factor on the left corresponded exactly with the different element of the disjoint union of six fibers over the corresponding λ values; and the factor on the right agrees with the pullback of the different element of the fiber in the λ projective plane over j.

If we actually tensor the coordinate ring of the fiber over one λ value, an algebra of rank p^2, with the coordinate ring of the fiber in the λ projective line over the corresponding j value, an algebra of rank 6, the discriminant of the tensor product would equal the actual discriminant of the whole fiber over the j value, and the two parts would come from tensor factors exactly matching this factorization. They are not coprime because the different element in the second factor can be rewritten, for instance we can factorize out z^p from $(x^{2p} - y^{2p})$, it corresponds to a transposition which interchanges two negative points in the affine cone, and fixes a point in the projective variety.

In fact this illustrates that the affine coordinate ring of a fiber is not the same as the specialization of the affine coordinate ring of a projective variety. Let's discuss this in the next section.

Modules on the fiber.

Let $R = R_0 \oplus R_1 ...$ be the homogeneous coordinate ring of a projective variety by a very ample line bundle $\mathcal{L}$, and let $f, g \in R_d$ be elements of degree d. Suppose that f, g generate $\mathcal{L}^d$.

For each matrix of integers A, B, C, D with $AD - BC = 1$ we can reduce the graded ring R modulo relations $Af+Bg = 0, Cf+Dg = 1$ for $AD - BC = 1$, and obtain the $\mathbb{Z}/d\mathbb{Z}$ graded ring

$$T = R \otimes_{\mathbb{Z}[f,g]} \mathbb{Z},$$

or, we can consider the fiber F of the map to $\mathbb{P}^1$. defined by $Af + Bg = 0$. Let $V = \mathcal{O}_{\mathcal{F}}$. We also have the locally sheaf I on F defined to be

$$I = i^* \mathcal{L},$$

which we can interpret as a locally free module over V.

Then

5. Lemma. There is an equivalence of categories between the category of $\mathbb{Z}/d\mathbb{Z}$ graded modules over T and all modules over V. There is a homomorphism

$$\psi : \mathbb{Z}/d\mathbb{Z} \to Pic(V)$$

such that for each $i \in \mathbb{Z}/d\mathbb{Z}$, the automorphism of the category of T modules which consists of shifting the grading by i corresponds to the automorphism of the category of V modules consisting of tensoring with the ideal $I^{\otimes i}$. We have $T \cong V \oplus I \oplus ... \oplus I^{d-1}$. Finally, if $\phi(1 \ mod \ d) \in Pic(F)$ is zero, there is an ideal $J \subset T$ which is complement of V.

Projective versus affine ramification

The Fermat curve has as its homogeneous coordinate ring

$$e_1 = x + y + z$$

$$e_2 = xy + xz + yz$$

$$e_3 = xyz$$

$$f_p = x^p + y^p + z^p = 0$$

$$f_{2p} = (xy)^p + (xz)^p + (yz)^p$$

$$f_{3p} = (xyz)^p.$$

This can be interpreted as an inefficient system of generators and relators, only one relator is needed. The homogeneous coordinate ring for the line bundle $\mathcal{O}(p)$ on the projective plane consists of terms of degree a muliple of p if x, y, z are given degree 1.

The homogeneous coordinate ring of the specialization at a j is the terms of degree a multiple of p in the ring with two further relations

$$Af_{2p}^3 + Bf_{3p}^2 = 0$$

$$Cf_{2p}^3 + Df_{3p}^2 = 1$$

with A, B, C, D integers with $AD - BC = 1$.

The subrings generated by x^p, y^p includes z^p of course, and it is a copy of the homogeneous coordinate ring of the specialized and unspecialized λ plane. Assume that the λ value is rational, so there is an integer point $[a^p : b^p : c^p]$.

If $\psi(1\ mod\ d) = 0$ we can ignore the grading, reduce modulo the complementary ideal J, and consider the result to be an algebra over $\mathbb{Z}$. We can construct an isomorphic copy of the rank six subalgebra by writing down as columns the ways of permutating a^p, b^p, c^p

$$\begin{pmatrix} a^p & a^p & b^p & b^p & c^p & c^p \\ b^p & c^p & a^p & c^p & a^p & b^p \\ c^p & b^p & c^p & a^p & b^p & a^p \end{pmatrix}$$

and take the subalgebra of $\mathbb{Z}^6$ generated the the three rows. In other words, sending x^p, y^p, z^p to the three rows describes an isomorphism with a rank six subalgebra of the underlying ungraded algebra of the homogenous coordinate ring of the specialized Fermat curve modulo J. The subalgebra has index $(a^p - b^p)^3(b^p - c^p)^3(c^p - a^p)^3$ in $\mathbb{Z}^6$. As an abstract un-graded algebra once reduced modulo J and tensored with $\mathbb{Z}[1/((a^p - b^p)(b^p - c^p)(c^p - a^p))]$ it decomposes into a cartesian product of six copies of that base ring, containing six different primitive idempotent elements. Let Y be the scheme in the $\mathbb{P}^1$ which parametrizes λ values, corresponding to this j The very ample line bundle for this rank six subalgebra is one of the p^2 summands of the pushforward of $\mathcal{O}(p)$ on F, giving the implication for the maps

$$\psi_F : \mathbb{Z}/(6\mathbb{Z}) \to Pic(F),$$

$$\psi_Y : \mathbb{Z}/(6\mathbb{Z}) \to Pic(Y),$$

$$\psi_F(1\ mod\ 6) = 0 \Rightarrow \psi_Y(1\ mod\ 6) = 0.$$

This in turn implies that $Y \subset \mathbb{P}^1$, once $(a^p - b^p)(b^p - c^p)(c^p - a^p)$ is inverted, will consist of six disconnected copies of the localized $Spec(\mathbb{Z})$.

Then the whole fiber F, the inverse image of Y, correspondingly decomposes into a disjoint union of the copies over the six individual λ values.

We calculated the different elements of the disjoint union, which was $p^2(abc)^{p-1}$. The different element of the whole fiber was $6p^2(abc)^{p-1}(abc)^p(a^p - b^p)(a^p - c^p)(b^p - c^p)(a^{2p} + (ab)^p + b^{2p})^2$. The factors $(a^p b^p c^p)(a^{2p} + (ab)^p + b^{2p})$ where a^p, b^p, c^p are rational integers, must then be divisors of a power of $6p(a^p - b^p)(a^p - c^p)(b^p - c^p)$. Thus

6. Theorem. For every rational value of j which lifts to a rational λ value, let F_j be the fiber over j in the Fermat curve. The element $\psi(1 \ mod \ 6) \in Pic(F_j)$ cannot zero except, possibly in special cases when $j \in \mathbb{P}^1$ is one of the three cusps or each of a^p, b^p, c^p is a power of $2, 3$, or p.

The case $p = 1$.

We are still assuming that there is a rational λ value. The fiber over j is a subscheme of $\mathbb{P}^2$ comprised of six copies of $Spec(\mathbb{Z})$, and the different element, a section of $\mathcal{O}(10)$ is perhaps best written now

$$(xy - yz)(yz - zx)(zx - xy)(xy + yz + zx)(yx + xz + zy).$$

This is a tensor product of five sections of $\mathcal{O}(2)$, and in the restriction to each irreducible component, it describes the five Cartier divisors where that component meets the five other components. The components are in two sets of three, and the decomposition is preserved by all the permutations. Note that the last two factors are equal. Two components of the same type (transformable to each other by an even element of S_3) meet each other only at the subscheme where $xy + yz + zx$ is zero, but there are two such subschemes, interchanged by any transposition. The Cartier divisor defined by this section is principal, being defined also by the rational integer $(ab - bc)(bc - ca)(ca - ab)(ab + bc + ca)(ba + ac + cb)$.

Since this section of $\mathcal{O}(10)$ or indeed the product of the first three and last four factors separately, describe a principal Cartier divisor, it seems likely that so does $\mathcal{O}(2)$ and that $\psi(2 \ mod \ 6) = 0$.

The case $p = 2$

It seems that, rather than using localizations, we can find a useful simplification by using partial normalizations which are a local isomorphism near a subscheme of interest, defining a type of etale neighbourhood (I'm not sure if I'm using the correct word here).

Since $p - 1 = 1$, the different element

$$(xyz)^{p-1}(x^2y^2 - y^2z^2)(y^2z^2 - z^2x^2)(z^2x^2 - x^2y^2)(x^2y^2 + y^2z^2 + z^2x^2)^2$$

again has an easy interpretation. The last five factors tensor together to give a section of $\mathcal{O}(20)$ which is just what we've seen already, just pulled back from the λ projective line. Any one of the 24 components belongs to a fiber over one λ value, and we already know that that fiber consists of four irreducible components, each meeting the other three transversely according to the Cartier divisors x, y, z one each. But also now we have intersections when one of the four components for one λ value meets one of the four components for another λ value.

Explicitly, the ring of rank 24 can be constructed as the subring of $\mathbb{Z}^{24}$ which is the image of the homogeneous polynomials of degree a multiple of 12 where we send $p(x, y, z)$ to a tuple

$$(p(a, b, c), p(-a, c, b), ...$$

where included is every possible permutation or assignment of signs to a, b, c which is essentially different (negating all variables is an inessential change).

The different element is not given to us as an element of this ring. Let's look at just some factors of the different element.

$$x(z^2x^2 - x^2y^2) = x \otimes x^2 \otimes (z^2 - y^2)$$

We can do a partial normalization so that we can ignore other factors of the different (and $c^2 - b^2$ is coprime to a^2) and arrive at a triple intersection point where our one component meets one other component lying over the same λ value, and two components over different λ values, and the remaining different element is $x \otimes x^2$.

I believe that the x^2 on the right is there because after a transposition, there is a meeting between the conjunction of two components in the one fiber and a symmetrically opposite conjunction of two components over the other λ value.

Now there are four components, and the coordinate ring of their union as the subring of $\mathbb{Z}^4$ spanned by all

$$(p(a,b,c), p(-a,b,c), p(a,c,b), p(-a,c,b))$$

for $p(x,y,z)$ homogeneous of degree a multiple of 12.

A neighbourhood of the relevant subscheme – the one defined by a in our originally chosen component, is isomorphic to a neighbourhood of the same scheme in the whole fiber.

The ring is the image of the invariants of the Klein four group acting on the cartesian product of four copies of $Spec(A)$ where A is the ring comprising all homogeneous polynomials of degree divisibe by twelve in $\mathbb{Z}[x,y,z]$. The group is generated by two elements τ, σ with

$$\tau(p(x,y,z), q(x,y,z), r(x,y,z), s(x,y,z)) = (q(-x,y,z), p(-x,y,z), s(-x,y,z), r(-x,y,z))$$

$$\sigma(p(x,y,z), q(x,y,z), r(x,y,z), s(x,y,z)) = (r(x,z,y), s(x,z,y), p(x,z,y), q(x,z,y)).$$

The rough intuition should be that multiple intersections cause split extensions of one forms rather than nontrivial extensions.

The Klein four group actually acts on our rank 24 ring, that action is just induced by permutations of the factors in $\mathbb{Z}^{24}$.

The images of particular types of polynomials (of various degrees in the grading and transforming according to particular characters) describe the four eigenspaces in the algebra. Without a different type of understanding of it, we cannot see a contradiction to the notion that the one forms module is principal. It seems most likely that from the direct description we would deduce that it is a copy of the augmentation ideal of the group algebra $\mathbb{F}_p K$ for K the group, whereas local principality of ω_X would force that it is a copy of $\mathbb{F}_p K$ modulo the augmentation ideal.

Let's consider this type of question without the hypothesis of a group action in the next section.

Local rings and principal differentials

7. Theorem. Let R be a local Noetherian ring containing $\mathbb{Z}$. Let m be the maximum nonunit ideal of R and suppose $\mathbb{F}_p \subset R/m$ is surjective (equality). Suppose Ω_R is a principal module. Then there is an element $\eta \in m/m^2$ so that

$$m/m^2 = \mathbb{F}_p(p \ mod \ m^2) \oplus \mathbb{F}_p\eta.$$

Proof. First we write

$$m/(m^2 + pR) \to \Omega_R \otimes_R \ R/m$$

$$m \mapsto dm.$$

This is well-defined because if $q, s \in m$

$$d(qs) = ds \otimes q + dq \otimes s = ds \otimes 0 + dq \otimes 0 = 0,$$

and for $r \in R$

$$d(pr) = dp \otimes r + dr \otimes p = 0 \otimes r + r \otimes 0$$

the last because $p \in m$.

It is also surjective because $\mathbb{Z} + m = R$, as this surjects onto $\mathbb{F}_p$ and contains the kernel of $R \to \mathbb{F}_p$.

Finally let's show the kernel of this map is zero. A linear combination of dm which maps to zero, would be reducible to zero by the Leibniz relation $d(rm) = dm \otimes r + dr \otimes m$. For r which belong to m it just says $d(rm) = 0$. In the remaining terms r can be replaced by an integer, and the relation asserts no more than that d commutes with addition. An expression reducible to zero by the relation must already be zero.

Now, if Ω_R is principal, so is $m/(m^2 + pR)$ and we may choose a single element $\eta \in m/m^2$ mapping to a generator. Then $m/m^2 = \mathbb{F}_p(p \ mod \ m) \oplus \mathbb{F}_p\eta$.

8. Corollary. Under these conditions, if in addition $R/(pR)$ has p^m elements, $R/(pR) \cong \mathbb{F}_p[T]/(T^m\mathbb{F}_p[T])$.

Proof. It is a finite local $\mathbb{F}_p$ algebra with residue field $\mathbb{F}_p$ and principal maximum ideal.

9. Corollary. Let X be a curve over $\mathbb{Z}$ and x a closed point such that that ω_X is (locally) principal in a neighbourhood of x. Suppose that there is a rational prime p so that where the residue field at x is $\mathbb{F}_p$. Suppose that R is isomorphic to the local ring at p of the subring of $\mathbb{Z}^m$ for some number m which is $\{(n_1, ..., n_m) : n_1 \equiv n_2 \equiv ... \equiv n_m \ mod\ p\}$. Then $m = 1$ if x is a nonsingular point of $Spec(R)$ and otherwise $m = 2$.

Proof. When we localize at p, the maximum ideal of such a ring is its intersection with the ideal generated by $(p, p, ..., p)$ in $\mathbb{Z}^m$, but it has minimal generating set $\{(p, 0, 0, .., 0), (0, p, 0, 0, ..., 0), ..., (0, 0, 0, 0..., p)\}$ in the smaller ring. If m is the maximal ideal, these form a basis of m/m^2, whose dimension must be at most 2.

The case $p = 2$. Beginning of the proof.

We actually construct the subring of $\mathbb{Z}^4$ which we mentioned, spanned, for all monomials $m(x, y, z)$ of degrees a multiple of $6p = 12$, by the

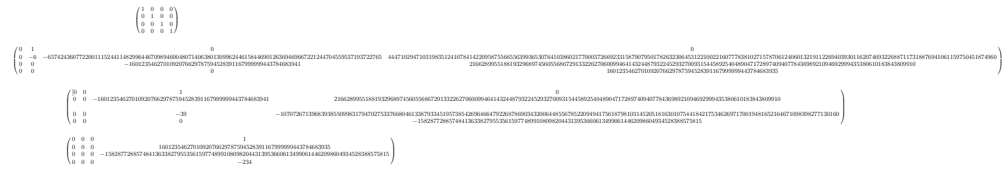
$$(m(a, b, c), m(-a, b, c), m(a, c, b), m(-a, c, b)).$$

For example for $a, b, c = 3, 5, -8$, chosen at random, we obtain[4] the ring with basis

$(1, 1, 1, 1)$
$(0, -6, -160123546270109207662978759452839116799999944378468394 1, 160123546270109207662978759452839116799999944378468393 5),$
$(0, 0, -39, -158287728857484136338279553561597748991080982044313953660613499061446209860493452838857581 5),$
$(0, 0, 0, -234)$

[4] assuming things stabilize after degree 108, and thanks to Matt Crumley http://silentmatt.com/ for putting BigInteger.js on sourceforge!

The multiplication operation is component-wise integer multiplication, and a matrix representation of the ring is given by the four matrices (shrunk so they fit on the page)



When these are reduced modulo $a = 3$ they become

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$$\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The last three are a basis of the maximal ideal, which is nilpotent of order only three, not four. Then as the conclusion of Corollary 8 does not hold, it is impossible for such a ring to exist at a closed point of the quadric where the differentials are locally free.

The numbers $3, 5, -8$ are not numbers whose squares add to zero (they happen to add to zero themselves instead). We cannot test what the construction would do if we applied it to three nonzero numbers whose squares add to zero, becuase no such triple of numbers is known.

The point about the partial normalization is that it was an isomorphism on the local ring level. Before I had been trying inverting integers. This is implicitly inverting coordinate functions somewhere ambiently, but it is just easiest to consider that we can take a partial normalizatin that does not affect a neighbourhood of the closed point we're considering.

The index of this ring in its normalization is divisible by a^4, it is $4(b-c)^2a^4$ many (but not all) examples. Such a ring is not just having the relations saying that the four integers are mutually congruent mod a. There is one more relation.

Our quick way of verifying that there is no element η in the max ideal so that $m/m^2 = F_p(p\ mod\ m^2) \oplus F_p\eta$ was to mod out by p and see if you get the only type of algebra that maps to $\mathbb{F}_p$ and has principal maximum nonunit ideal.

Those four 4×4 matrices, we'd have to make η out of a linear combination of those, yet every linear combination of those is nilpotent of order less than four.

There are examples where the index of R in $\mathbb{Z}^4$ is smaller than a^3 while the spectrum is not connected.

So, even though we can't – apparently – find 3 nonzero numbers whose squares add to zero, to test this, the only thing that is in contention is the argument which said that those four components would have to meet.

This is where we did need the different element calculation. Since the ideal in $\mathcal{O}_F$ (really the different element times the inverse of ω_X) is induced from the characteristic subring locally, and the part we are considering which is $(xyz)^{p-1}$ times $(xyz)^p$ is invariant under symmetry, this is induced from the characteristic subring globally.

Then, those two components over one λ value that meet at the subvariety defined by the integer a in both, the different element from the disjoint union of the six parts gave that ignoring other components this is an intersection defined in each by just $a^{p-1} = a$.

The other factor, if we just interpret it as we may using the specialization technique, gives us $a^p = a^2$.

This is part of a symmetric expression, so no matter which component we look at it is constant a^2.

That means this double intersection must meet components at this same subvariety of $Spec(\mathbb{Z})$ which are on other parts of the six, i.e. over other lambda values.

And because it is the subvariety defined by a, and not something like $b - c$, or $(ab + bc + ca)$ it has to come from a transposition of negative eigenvalue.

It is really hard to see numerically how this could happen.

But, the point is to make a clear proof that it cannot happen.

We know, provably, and from examples, that if that local ring R at the closed point at a prime dividing p into a has that R/pR is anything other than $\mathbb{F}_p[T]/T^m$ for some m, that this will never occur.

In examples if we put in random numbers for a, b, c in every case when the four components actually meet at a, it is not of type $F_p[T]/T^m$.

But it still needs a proof, that, to make it easier, if the subring R of $\mathbb{Z}^4$ spanned by all four tuples

$$(m(a,b,c), m(-a,b,c), m(a,c,b), m(-a,c,b))$$

for m ranging over monomials a multiple of 12, the a-primary component of the index divisible by a^3, then the order of nilpotency of the image of that maximum ideal is at most three.

Thus, to be clear we can state it as a conjecture:

Conjecture. Let a, b, c be pairwise coprime and ≥ 2. Let R be the subring of $\mathbb{Z}^4$ spanned by all

$$(m(a,b,c), m(-a,b,c), m(a,c,b), m(-a,c,b))$$

for which m is a monomial of degree a multiple of 12. Let q^e be the highest power of a prime q dividing a. Suppose that R contains no idempotent element besides 0 and 1 (which I think is equivalent to saying $[\mathbb{Z}^4 : R]$ is divisible by q^{3e}). Let m be the kernel of $R \to \mathbb{F}_q$. Then the conjecture is that the order of nilpotency of the image of m in R/qR is no larger than 3.

It is this conjecture which implies that $a^2 + b^2 + c^2$ cannot be zero.

For, if $a^2+b^2+c^2 = 0$ the different calculation would tell us that R is connected, a local ring rather than semi local. Then the conjecture tells us that the order of nilpotency is no more than 3, and this means that R/pR can't have a principal max ideal. Then the max ideal of R cannot be generated by p and one other element, and therefore the Kahler differentials the local ring cannot be principal. Then it cannot occur on the quadric $a^2 + b^2 + c^2 = 0$.

In other words, we still have to show that this always happens under the hypotheses which we've established would be true under the hypothesis of existence of a noncuspidal rational point.

Overview of a proof.

It is interesting to see that the calculation of the subring of $\mathbb{Z}^4$, in examples, when it has index divisible by a^3 the order of nilpotency modulo a prime divisor of a is never as much as four. Even for the Fermat triples one sees this. For $5, 4, 3$ with a playing the role of 5, the index in $\mathbb{Z}^4$ is divisible by a^2 only, and for $3, 4, 5$ the index is divisible by a^3 but the order of nilpotency is three.

Here is an overview of the proof which is at hand. It's just a matter of explaining how the parts fit together.

Twisted modules.

The issue here was that I can make a ring where two copies of $Spec(\mathbb{Z})$ are glued along $\mathbb{Z}/5\mathbb{Z}$. This just means the *ring* is pairs of numbers congruent mod 5. But a *module* can be pairs of numbers with one double the other mod 5.

This seems like a harmless difference but if we do not know a priori that the restriction of $\mathcal{O}(p)$ to a fiber isn't like that, we have not a method of generators and relators to describe the ring structure.

Structure of fiber

We know that only particular very restrictive combinatorial types can occur. We don't know much about the Fermat fiber combinatorial type, but the fact that the thing analytically is a pullback is suspicous because it can't be algebraically if it is to have a rational point, for the same reason 2 partial derivatives of a 2 variable function can't both be zero at a point of a smooth curve they define.

However because of the previous point, it would be beyond computer calculation to know anything at all about the whole thing.

S_3 symmetry and different element and 'Specialization.'

These are going to be used only because, in combination, they tell us which components meet which others at which points. They imply that there has to be a configuration of four components meeting at one point. Then the partial normalization which is the image of the (actual) coordinate ring in $\mathbb{Z}^4$ is a partial normalization faithfully representing the relevant local ring and we know a priori it is connected, i.e. a local ring not semilocal.

Because $\mathcal{O}(p)$ is not principal, the different element doesn't reliably tell us everything but the multiplicity of (abc) is reliable since it is constant everywhere hence induced from the characteristic subring.

Then once we know this, we can discard the S_3 symmetry, ignore the different element.

Degree of nilpotence, index in normalization.

We can get away using only limited information now. The S_3 symmetry etcetera told us that the thing is connected so we can assume the index of the subring of $\mathbb{Z}^4$ is at least divisible by the thrice the power of primes dividing a. And embeddability requires that the ring mod q for such a prime q has to be of the type $\mathbf{F}_p[T]/T^4$, i.e. the order of nilpotency of the max ideal of the ring mod p has to be the maximum value of 4.

We now can throw out the assumption of any K_4 symmetry.

Structure of argument

Now we are down to a question about subrings R of $\mathbb{Z}^4$, and for each a, b, c we have a description of it, it is the span of $(m(a,b,c), m(-a,b,c), m(a,c,b), m(-a,c,b))$ for m monomials in a, b, c of degree a multiple of 12. If we can prove that for a prime power divisor q^e of a, when its index is at least divisible by q^{3e} (and that it disconnected otherwise) the degree of nilpotency of the max ideal of R/p is less than 4, we are done with the quadric.

Even these rings for things like the Pythagorean triples $3, 4, 5$, are beyond hand calculation. But we've preserved the absurdity of having a very high intersection multiplicity geometrically, with a high order of nilpotency mod p.

We've needed to make essential use of the S_3 symmetry and the existence of a rational point, and in some ethical sense we have seen from the beginning why these contradict each other but now even in this more focussed algebraic question, it is beyond hand calculation to solve it.

But it is absurd except in weird special cases to have high nilpotency *and* high number of components. It is a matter of showing that this class of subrings of $\mathbb{Z}^4$ is far enough from generic to include any sort of weird counterexample like that. Just throwing random numbers in for a, b, c one sees basically two patterns, where either the index is divisible by a^2 only and it is disconnected, or divisible by a^3 or a^4 and the degree of nilpotency is three.

Completion of proof for general primes p.

We could simplify things by observing that the pullback of $\psi(1\ mod\ 6)$ to the partial normalization is trivial if $p = 2$ and its square is trivial in general; however it is easiest to just extend the conjecture to cover the case of all primes p and to prove that it is true. For general primes p, we still have our four components meeting at a point indexed by a, however two of them are not rational We obtain the localization R at a prime divisor q of a of the subring of

$$\mathbb{Z} \times \mathbb{Z}[\omega] \times \mathbb{Z} \times \mathbb{Z}[\omega]$$

spanned by

$$(m(a,b,c), m(\omega a,b,c), m(a,c,b), m(\omega a,c,b)$$

for m monomials of degree a multiple of $6p$, ω a primitive p'th root of unity.

The main observation is merely that any monomial

$$x^i y^j z^k$$

can be written as

$$x^i (yz)^j z^{k-j}$$

if $k \geq j$, and

$$x^i (yz)^k y^{j-k}$$

if $j \geq k$. Since bc is invertible in R the restriction $i+j+k \equiv 0\ mod\ 6p$ means the elements coming from

$$x^i y^s,\ x^i z^s$$

for $i \equiv s\ mod\ 2$ generate R as an algebra over the localized $\mathbb{Z}$. Since $b - c$ is invertible in R, being a divisor of $b^p - c^p$ which is coprime[5] to a, the difference coming from the monomials xy and xz

$$(ab, \omega ab, ac, \omega ac) - (ac, \omega ac, ab, \omega ab)$$

$$= (b-c)(a, \omega a, -a, -\omega a)$$

being in R implies that so is

$$a(1, \omega, -1, -\omega).$$

[5] label a, b, c so a is odd

From the monomials y^{2k} and z^{2k} we obtain

$$(b^{2k}, b^{2k}, c^{2k}, c^{2k}), \quad (c^{2k}, c^{2k}, b^{2k}, b^{2k})$$

and from linear combinations, both c^2 and b^2 times

$$(0, 0, b^2 - c^2, b^2 - c^2)$$

and as $b - c$ is invertible we obtain

$$(b + c)(0, 0, 1, 1).$$

From xy^{2k-1} and xz^{2k-1} we obtain

$$a(b^{2k-1}, \omega b^{2k-1}, c^{2k-1}, \omega c^{2k-1}), \ a(c^{2k-1}, \omega c^{2k-1}, b^{2k-1}, \omega b^{2k-1}),$$

and from linear combinations we obtain

$$a(b + c)(0, 0, 1, \omega).$$

A basis over localized $\mathbb{Z}$ consists of the

$(a^{2i}, (a\omega)^{2i}, a^{2i}(a\omega)^{2i})$,
$(b + c)(0, 0, a^{2i}, (a\omega)^{2i}))$,
$(a^{2i+1}, (\omega a)^{2i+1}, -a^{2i+1}, -(\omega a)^{2i+1})$,
$(b + c)(0, 0, a^{2i+1}, (\omega a)^{2i+1})$

for all cases when the superscript i or $2i + 1$ may be less than p (so for example in case $p = 2$ one takes $i = 0$ and for $p = 3$ one takes $i = 0, 1$ in the first two lines and $i = 0$ in the second two).

When $b+c$ not a multiple of q the second line with $i = 0$ shows that R is not indecomposable. The whole algebra is isomorphic to the localization of the subring of the group algebra $\mathbb{Z}C_p$ generated by $-aw$ where $C_p = \langle w \rangle$ tensored with the subring of $\mathbb{Z} \times \mathbb{Z}$ generated by $(1, 1)$ and $(0, b + c)$. Regardless of whether q is a divisor of $b + c$, $R/(Rq)$ is a tensor product of two local algebras with nontrivial nilpotent elements, and the maximum nonunit ideal cannot possibly be principal therefore.

As we've needed to assume that $q \neq 2, 3, p$ this illustrates in a special case the theorem of Fermat and Wiles, that it is impossible for $a^p + b^p + c^p = 0$ when a, b, c are nonzero and pairwise coprime, except if each of a, b, c is a power of $2, 3$, and p.

Use of scheme theory.

It was really somehow helpful to visualize things like $Spec(\mathbb{Z})$ as a subset of the Riemann sphere even though it is not. When particular closed points are considered, there they are analytically, right in the point where you had to visualize them. The issue is similar to applying a Galois automorphism to $\mathbb{Z}[\sqrt{2}]$ and worrying about the discontinuity of it, and it is tempting to work in $\mathbb{R}[T]/(T^2-2)$ and use the classical topology, things like the analytic class number formula. To use lattices and volumes, and the Euclidean norms. But there is now a second tradition which is very different from doing that.

The second tradition has to do with a type of vague attempt to make things more symmetrical under dualizing, to imagine that integers are functions with domain some type of hybrid anaytic object. It is known that it is still totally rigorous to use differential calculus on these things but it is disturbing that something seems dishonest or ghostly about the visualization of $Spec(\mathbb{Z})$.

Without wanting to be pretentious about it, I'd say that it is evidence of self-deception in the past, an un symmetrical division between discrete things and continuous things; between analysis and algebra.

Conclusion.

I wasn't going to send any more of these to anyone, hard-copy pdf's, but here is why I changed my mind about it.

One thing is, even though it's probably still wrong, I haven't actually deleted any pages, each missing part just got added...

Different things...I looked up Olga Taussky-Todd's article about sums of squares. Very responsible, thorough, hard-working it is, ... No matter how old or tired she is, how much she's had to do, you know she's going to read even Mazur, Artin, all the new things, include a really intelligent synthesis, and finish it by the time of the article deadline.

Also was thinking about how we lost a Teaching Quality Assessment point; I had been assigned to be the 'Aims and Objectives Barber of Seville,' to write aims and objectives for everyone who didn't write aims and objectives for himself.

It is actually hilariously funny, we lost the point because my own work was in the Aims and Objectives room. The inspectors focused on that. It was a Galois theory problem on my own exam, which had been self-contradictory. I'm probably putting a false spin on it; last night it was upsetting me, or depressing me about having let everyone down. It reminded me of Holden Caufield's having, supposedly, left all the fencing equipment in the subway on the way to a match. I think that book was supposed to be about something meaningful.

I was reading about Monsanto, and debating about it with people, all the history about Saccharine, PCB's, agent Orange, Dioxin. Supposedly still now GMO's are the 'only way to feed the rapidly increasing population.' I wanted to explain in some way that the argument is upside-down. It is perfectly explained in my economics book, because all I did is quote things people said that actually make sense, and that show how people can actually understand sometimes..

I included the six BCC people, now since it is the last one, I'm including as BCC's maybe non-math friends, Bill, John K, Jacob, Callan, Sam, Felix, Jamie.

Just to give an example, not to explain it all, but to give an example, when I had a summer job, Bill took all the work out of my desk drawer, when I got back to University, and discarded it without telling me. Another summer job, he tried to get me to go on the Staten Island Ferry on my lunch hour, knowing if I went with him, I'd never get back to work in time.

Jacob had seminars about – really it must have been whatever we were reading, I think they were Saturday Morning even, or maybe Thursday at 10:00. I was reading a few pages of Corps Locaux. The thing is, none of the permanent staff ever knew Jacob did this. Equally important in my mind, more important, than any seminar that happened under the eyes of the establishment, powerful weekly seminars, were these.

I don't understand how I can call it non-mathematical friendship, when it is within Mathematics that this took place, in every case.

Anyway, the reason I decided to actually post a hard-copy .pdf as an email attachment is, I had decided *not* to.

And then I fell asleep tonight, and was dreaming.

I dreamed that we had gone to get the puppy, it was a little white puppy, and we took it with us to Stratford. As we did in real life.

And that suddenly I remembered in the past, when we were with it before, having seen the puppy with its mother, drinking the warm milk.

And it occurred to me, I have no way to feed it. I just hadn't thought about it.

I mentioned my worry to Dimitra, I was desperate about it, we have to go back! And she was her usual, it might be inconvenient 'absolutely not' attitude....

I wondered, after all this time, how can it be OK?

I used to have dreams that I was supposed to be teaching a class, back at Columbia, one of those precalculus classes maybe, and I'd forgotten to go *again* and it woud be about racing around, trying to find the list of times, weeks, trying to find the room. And some students would be waiting for me to explain meaningful things, and I'd give a lecture, but then later realize, it has been four more weeks, and I have not been there, and I wonder if they are *still* waiting.

But not any more, dreaming about abandoning students or anything.

The puppy looked content enough, but I noticed its tail was loose, it is a white puppy with a brownish tail, like Skiffles has.

And it just was somehow, maybe I was trying to pet it, or maybe to comfort it somehow, or to heal it, or see if it was OK, and it was on the couch. The tail had become disconnected. The puppy was ill, from having been taken away.

And it looked ill too, and it looked like other parts might fall off, and I was feeling sick with worry about it.

I got some water from the faucet onto my hand, and let it drip onto my knee, and the puppy started licking it. It stopped for a while, and I started to panic that it wasn't right, then it started again.

My only wish was to bring it back where it always belonged.

It's hard to explain how poignant it was, that it didn't seem to be worried, from what I could determine, but I was the one who was worried, and I held it for a little while, it was talking to me.