(*October 9, 2011*)

Modular curves, raindrops through kaleidoscopes

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/~garrett/

- $\mathfrak{H}$ as homogeneous space for $SL_2(\mathbb{R})$
- The simplest non-abelian quotient $SL_2(\mathbb{Z})\backslash\mathfrak{H}$
- Non-abelian solenoid: raindrop through a kaleidoscope
- Enabling the action of $SL_2(\mathbb{R})$

Solenoids are limits of one-dimensional things, circles. As complicated as the solenoids might be, all the limitands are *the same*, just circles.

Hoping for a richer supply of phenomena, we should wonder about limits of *two*-dimensional things (surfaces). It was convenient that all the circles in the solenoids were *uniformized* compatibly, that is, were compatibly expressed as *quotients* of $\mathbb{R}$. For ease of description, we choose to look at surfaces X which are quotients of the upper half-plane $\mathfrak{H}$ in $\mathbb{C}$, with quotient mappings $\mathfrak{H} \to X$ that fit together compatibly. [1]

The real line $\mathbb{R}$, being a *group*, is a *homogeneous space* in the sense that (of course) it acts *transitively* on itself. The upper half-plane $\mathfrak{H}$ is not reasonably a group itself, but *is* acted upon by $SL_2(\mathbb{R})$ (2-by-2 real matrices with determinant 1) acting by so-called *linear fractional transformations*

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}(z) = \frac{az+b}{cz+d}$$

We verify that this action is transitive, making $\mathfrak{H}$ a homogeneous space for the group $SL_2(\mathbb{R})$. The oddity of the action *can* be put in a larger context, which we will do a bit later.

Modular curves are quotients $\mathfrak{H} \to \Gamma\backslash\mathfrak{H}$ where Γ varies among certain finite-index subgroups of $SL_2(\mathbb{Z})$, the group of 2-by-2 integer matrices with determinant 1. They are called *modular*[2] for historical reasons discussed earlier. They are *curves* in the sense that they are *complex* one-dimensional (though *real* two-dimensional).

We will see that $SL_2(\mathbb{Z})\backslash\mathfrak{H}$, the simplest explicit case, is *topologically* S^2 with a point missing. However, when we keep track of geometry consistent with the relevant group action, the missing point is *infinitely far away*, so the shape is not a *round* sphere, but stretched out like a *raindrop*. The *hyperbolic geometry* appropriate to the upper half-plane will be discussed briefly in the next chapter.

Some of the discussion here will seem peculiarly specific, or peculiarly idiosyncratic, especially by comparison to the ease with which we manipulate $\mathbb{R}$ and $\mathbb{Z}$. The contrast is partly explained by the fact that two-dimensional objects with non-abelian group actions are genuinely more complicated. Indeed, at some future point, we may decide (with hindsight) that as groups or geometric objects $\mathbb{R}$ and $\mathbb{R}/\mathbb{Z}$ are misleading.

And why not look at quotients of $\mathbb{C}$ (or $\mathbb{R}^2$) to make surfaces? Indeed, the quotients of $\mathbb{C}$ by lattices are elliptic curves, certainly worthy of study. We choose a different course with its own interest.

[1] The *Uniformization Theorem* in complex analysis asserts that a compact, connected, Riemann surface (that is, a compact, connected, one-dimensional complex manifold) is either complex projective 1-space $\mathbb{P}^1$ (a.k.a. *the Riemann sphere*), or is $\mathbb{C}/\Lambda$ for a lattice Λ, or is a quotient $\Gamma\backslash\mathcal{D}$ of the unit disk $\mathcal{D}$ by a suitable group $\Gamma \subset GL_2(\mathbb{C})$. (It is easy to check that the map $z \to (z+i)/(iz+1)$ is an isomorphism of the disk to the upper half-plane, so uniformizing by one is equivalent to uniformizing by the other.) In this last (and very interesting) case of uniformization by the disk or half-plane, however, the groups Γ rarely can be described sufficiently tangibly for our present purposes.

[2] For the same historical reasons, $SL_2(\mathbb{Z})$ is sometimes called *the modular group*. By now this is an anachronism.

1. $\mathfrak{H}$ *as homogeneous space for* $SL_2(\mathbb{R})$

Before trying to uniformize surfaces, we must explain the structure of the half-plane $\mathfrak{H}$ as *homogeneous space*, that is, as a space acted upon *transitively* by a *group*.

When uniformizing circles as $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$, the group $\mathbb{R}$ acting transitively on such a circle is not very different from the circle itself. Indeed, circles $\mathbb{R}/\mathbb{Z}$ are groups themselves, since $\mathbb{Z}$ is normal in $\mathbb{R}$, which is unavoidable since $\mathbb{R}$ is *abelian*. By contrast, the upper half-plane

$$\mathfrak{H} = \{z = x + iy \ : \ y > 0\} \subset \mathbb{C}$$

does *not* have any reasonable group structure itself. Luckily, the group

$$G = SL_2(\mathbb{R}) = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \ : \ \text{real matrices with } \ ad - bc = 1\}$$

acts on $\mathfrak{H}$ with the **linear fractional transformation action** [3]

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}(z) = \frac{az + b}{cz + d}$$

and we have:

[1.0.1] Claim: The group $SL_2(\mathbb{R})$ stabilizes $\mathfrak{H}$ and acts transitively on it. [4] In particular,

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{y} & 0 \\ 0 & \frac{1}{\sqrt{y}} \end{bmatrix}(i) = x + iy \qquad (\text{for } x \in \mathbb{R},\ y > 0)$$

Further, for $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{R})$ and $z \in \mathfrak{H}$

$$\operatorname{Im} g(z) = \frac{\operatorname{Im} z}{|cz + d|^2}$$

Proof: The first formula is clear. The second formula would imply that the upper half-plane is stabilized. Compute directly:

$$2i \cdot \operatorname{Im}(\begin{bmatrix} a & b \\ c & d \end{bmatrix}(z)) = \frac{az + b}{cz + d} - \frac{a\overline{z} + b}{c\overline{z} + d} = \frac{(az + b)(c\overline{z} + d) - (a\overline{z} + b)(cz + d)}{|cz + d|^2}$$

$$= \frac{adz - bc\overline{z} - bcz + ad\overline{z}}{|cz + d|^2} = \frac{z - \overline{z}}{|cz + d|^2}$$

since $ad - bc = 1$. ///

[3] This action can easily (if awkwardly) be discussed in a completely *ad hoc* fashion, but, in fact, arises as an artifact (in coordinates) of a natural action of $GL_2(\mathbb{C})$ on complex projective space $\mathbb{P}^1$. This situation itself is a very special case of actions of $GL(n + 1, \mathbb{C})$ (complex invertible matrices of size $n + 1$) on projective n-space $\mathbb{P}^n$. This broader context, as well as action of subgroups stabilizing complex n-balls, will be discussed just a little later.

[4] In fact, *every* holomorphic automorphism of $\mathfrak{H}$ is given by an element of $SL_2(\mathbb{R})$. This follows from *Schwarz' lemma* (on the disk), which allows us to deduce that an automorphism of the disk fixing 0 is a rotation.

[1.0.2] Remark: The extra information about how the imaginary part transforms will be useful in determining a *fundamental domain* just below.

Since $SL_2(\mathbb{R})$ acts transitively on $\mathfrak{H}$, we can express $\mathfrak{H}$ as a *quotient* of $SL_2(\mathbb{R})$. For example,

[1.0.3] Claim: The isotropy group in $SL_2(\mathbb{R})$ of the point $i \in H$ is the **special orthogonal group** [5]

$$SO(2) = \{g \in SL_2(\mathbb{R}) : \ g^\top \cdot g = 1_2\} = \{\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} : \ \theta \in \mathbb{R}\}$$

Proof: For real a, b, c, d, the equation $(ai + b)/(ci + d) = i$ gives $ai + b = -c + id$, so $a = d$ and $c = -b$. The determinant condition $ad - bc = 1$ gives $a^2 + b^2 = 1$, which we can reparametrized via trigonometric functions as indicated. ///

[1.0.4] Corollary: We have an isomorphism of $SL_2(\mathbb{R})$-spaces

$$SL_2(\mathbb{R})/SO(2) \approx \mathfrak{H} \quad \text{via} \quad g \cdot SO(2) \to g(i)$$

[1.0.5] Remark: The last claim and its corollary had only a boring analogue in the case of the transitive action of $\mathbb{R}$ on itself, since all isotropy groups were trivial.

[1.0.6] Remark: Yes, quotients such as $SL_2(\mathbb{R})/SO(2)$ have more structure than just topological, and these will be relevant.

[1.0.7] Remark: It will eventually become clear that the effect of taking the quotient of $SL_2(\mathbb{R})$ by $SO(2)$ is a hindrance, and that we should prefer to consider the *three-dimensional* group $SL_2(\mathbb{R})$ and its quotients $\Gamma\backslash SL_2(\mathbb{R})$, rather than $SL_2(\mathbb{R})/SO(2)$ and its quotients $\Gamma\backslash SL_2(\mathbb{R})/SO(2) \approx \Gamma\backslash\mathfrak{H}$ for several reasons. However, in the short term, and to connect with (and exploit) historical artifacts, we do treat $\mathfrak{H}$ and *its* quotients $\Gamma\backslash\mathfrak{H}$.

---

2. *The simplest non-abelian quotient* $SL_2(\mathbb{Z})\backslash\mathfrak{H}$

We will make surfaces as quotients $\Gamma\backslash\mathfrak{H}$ of the half-plane $\mathfrak{H}$ by subgroups Γ [6] of $G = SL_2(\mathbb{R})$ such that the quotient $\Gamma\backslash\mathfrak{H}$ is reasonably small. [7] The simplest beginning choice is

$$\Gamma = SL_2(\mathbb{Z}) = \{\text{2-by-2 integer matrices with determinant 1}\}$$

Both for use just below and to show that $SL_2(\mathbb{Z})$ is a large group, we note:

---

[5] The choice of corner in which to put the $-\sin\theta$ does not matter much in the larger scheme of things, and often the opposite choice is made, but there are some reasons one might make the present choice. Still, it doesn't really matter.

[6] As in the case of the solenoids, the local topology of quotients is simplest for quotients by *discrete* subgroups. See the appendix where configurations $\Gamma\backslash G/K$ are considered, for Γ discrete in a topological group G, and K a *compact* subgroup of G.

[7] The *Uniformization Theorem* makes *compact* (connected) surfaces, but our explicitly-constructed surfaces will *not* be compact. This creates non-trivial issues in many regards, but the relative simplicity of the description of the groups Γ makes these other complications acceptable.

[2.0.1] Claim: Given *relatively prime* integers c, d, there are integers a, b such that $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma$.

Proof: From basic number theory we know that there are integers m, n such that

$$\text{greatest common divisor } c, d = m \cdot c + n \cdot d$$

Here the greatest common divisor is 1, and take $a = n$, $b = -m$, so $ad - bc = 1$. ///

To be able to draw a picture of the quotient, we take an archaic [8] approach which nevertheless succeeds in this case. First, we find a **fundamental domain** for Γ on $\mathfrak{H}$, meaning to find a *nice* set of representatives for the quotient. Second, see how the edges of the fundamental domain are glued together when mapped to the quotient $\Gamma\backslash\mathfrak{H}$.

[2.0.2] Claim: Every Γ-orbit in $\mathfrak{H}$ has a representative in

$$\overline{F} = \{z \in \mathfrak{H} : |z| \geq 1,\ |\text{Re}(z)| \leq \frac{1}{2}\}$$

More precisely, each orbit has a *unique* representative in

$$F = \{z \in \mathfrak{H} : |z| > 1,\ -\frac{1}{2} \leq \text{Re}(z) < \frac{1}{2}\} \cup \{z \in \mathfrak{H} : |z| = 1, \text{Re}(z) \leq 0\}$$

[2.0.3] Remark: The fundamental domain is illustrated in the picture *[... iou ...]*...

Proof: From above, for $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma$

$$\text{Im} \begin{bmatrix} a & b \\ c & d \end{bmatrix} (z) = \frac{\text{Im } z}{|cz+d|^2}$$

The set of complex numbers $cz + d$ is a subset of the lattice $\mathbb{Z} \cdot z + \mathbb{Z}$, which (by its discreteness in $\mathbb{C}$) has (at least one) smallest (in absolute value) non-zero element. Thus, $\inf |cz + d| = \min |cz + d| > 0$, taking the infimum or minimum over *relatively prime* c, d, which we have observed are exactly the lower rows of elements of Γ. Then

$$\sup \frac{1}{|cz+d|} = \max \frac{1}{|cz+d|} < \infty$$

Thus, for fixed $z \in \mathfrak{H}$,

$$\sup \text{Im} \begin{bmatrix} a & b \\ c & d \end{bmatrix} (z) = \sup \frac{\text{Im}\, z}{|cz+d|^2} = \max \frac{\text{Im}\, z}{|cz+d|^2} < \infty$$

Thus, in each Γ-orbit there is (at least one) point z assuming the maximum value of $\text{Im}\, z$ on that orbit.

Since $\text{Im} \begin{bmatrix} a & b \\ c & d \end{bmatrix} (z) = \text{Im}\, z/|cz + d|^2$, for z giving maximal $\text{Im}\, z$ in its orbit, it must be that

$$|cz + d| \geq 1$$

[8] The approach which succeeds in general is a distant descendant of the present argument, and (as usual) is successful because it discards many (adroitly chosen) details, which turn out to to have been inessential.

for all c,d relatively prime. Thus, for example, for $d=0$ there is the *inversion*

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}(z) = -1/z$$

Thus, $|1\cdot z+0|\geq 1$, so for $\operatorname{Im} z$ maximal in its Γ-orbit, $|z|\geq 1$.

We can adjust any $z\in\mathfrak{H}$ by

$$\begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}(z) = z+n \quad (\text{for } n\in\mathbb{Z})$$

to normalize $-1/2\leq \operatorname{Re}(z) < 1/2$.

So take $|z|\geq 1$ and $|\operatorname{Re}(z)|\leq 1/2$ and show that $|cz+d|\geq 1$ for *all* c,d. Break z into its real and imaginary parts $z=x+iy$. Then

$$|cz+d|^2 = (cx+d)^2 + c^2y^2 = c^2(x^2+y^2) + 2cdx + d^2 \geq c^2(x^2+y^2) - |cd| + d^2$$

$$\geq c^2(|z|^2 - \frac{1}{4}) + \frac{c^2}{4} - |cd| + d^2 \geq c^2(|z|^2 - \frac{1}{4})$$

Thus, for $|c|\geq 2$, we have $|cz+d|>1$ when $|z|\geq 1$ and $|x|\leq 1/2$.

For $c=0$, necessarily $d=\pm 1$, and the only corresponding elements of Γ are

$$\begin{bmatrix} \pm 1 & n \\ 0 & \pm 1 \end{bmatrix}$$

The only z's with $|z|\geq 1$ and $|x|\leq 1/2$ that can be mapped to each other by such group elements are $-\frac{1}{2}+iy$ and $\frac{1}{2}+iy$. We whimsically keep the former as our chosen representative.

For $c=\pm 1$,

$$|cz+d|^2 = 2xd + d^2 + |z|^2 \geq -|d| + d^2 + 1 \geq 1 \quad (\text{for } d\in\mathbb{Z})$$

In fact, for $|x|<1/2$, there is a *strict* inequality

$$2xd + d^2 + |z|^2 > -|d| + d^2 + 1 \geq 1$$

so $|cz+d|>1$. When $|x|=1/2$, still $-|d|+d^2+1>1$, *except* for $d=0,\pm 1$.

Thus, first without worrying about strictness of the inequalities, $|cz+d|\geq 1$ for $|z|\geq 1$ and $|x|\leq 1/2$, and the set $\overline{F}$ contains (at least one) representative for every orbit. What remains is to eliminate duplicates.

We have already observed that the only duplicates for $|z|>1$ have $|x|=1/2$, and $z\to z+1$ maps the $x=-1/2$ line to the $x=1/2$ line.

Now consider $|z|=1$. For $|x|<1/2$, the only cases where $|cz+d|=1$ are with $c=\pm 1$ and $d=0$, which correspondes to matrices

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} * & \pm 1 \\ \mp 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}\cdot\begin{bmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{bmatrix} \quad (\text{for some } n\in\mathbb{Z})$$

For $|z|=1$, the inversion $z\to -1/z$ maps $z=x+iy$ to

$$-\frac{1}{z} = -\overline{z}/|z|^2 = -\overline{z} = -x+iy$$

Thus, for $|x|<1/2$, the only one among these products that maps z back to the fundamental domain is exactly the inversion $z\to -1/z$. This inversion identifies the two arcs

$$\{|z|=1 \text{ and } -\frac{1}{2}\leq x\leq 0\} \qquad \{|z|=1 \text{ and } 0\leq x\leq \frac{1}{2}\}$$

Thus, we should include only one or the other of these two arcs in the strict fundamental domain.

Last, with $|z| = 1$ and $|x| = 1/2$, there are exactly four group elements modulo $\pm 1_2$ (the center $\{\pm 1_2\}$ acts trivially) that map z to the closure of the fundamental region. These are: the identity, one of the translations $z \to z \pm 1$, the inversion $z \to -1/z$, and the *composite* of the translation and the inversion. That is, in addition to the identity,

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{map } -\frac{1}{2} + \frac{i\sqrt{3}}{2} \text{ to the boundary of } \overline{F}$$

and

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{map } \frac{1}{2} + \frac{i\sqrt{3}}{2} \text{ to the boundary of } \overline{F}$$

Thus, in the quotient $\Gamma\backslash\mathfrak{H}$, the identification of the sides $x = \pm 1$ creates a (topological) cyclinder, and the identification of the two arcs on the bottom closes the bottom of the cylinder. Thus, topologically, we have a cylinder closed at one end, which is a disk. But the non-euclidean geometry [9] (if we were to pay more attention to details) suggests that the *top* of the cylinder is infinitely far away, and the radius of the cylinder goes to 0 as one goes toward the open top end, so it is more accurate to think of the quotient $\Gamma\backslash\mathfrak{H}$ as a raindrop shape. ///

[... iou ...] pictures

---

3. *Non-abelian solenoid: raindrop through a kaleidoscope*

Here we make a simple but non-*abelian* solenoid, that is, projective limit of finite-to-one surjections of geometric objects. Since the *bottom* object is the raindrop $SL_2(\mathbb{Z})\backslash\mathfrak{H}$, the family of coverings arguably is visualizable as an effect comparable to that of a kaleidoscope.

[... iou ...] pictures!

There are several subgroups of $SL_2(\mathbb{Z})$ of traditional interest and with traditional notations, but we only need one explicit type, the **principal congruence subgroup of level** N: for positive integral N, let

$$\Gamma_N = \Gamma(N) = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z}) \ : \ a = 1 \bmod N, b = 0 \bmod N, c = 0 \bmod N, d = 1 \bmod N\}$$

$$= \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z}) \ : \ \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \bmod N\}$$

Observe that Γ_N is the *kernel* of the *reduction-mod-N* homomorphism [10]

$$GL_2(\mathbb{Z}) = \mathrm{Aut}_{\mathbb{Z}}(\mathbb{Z}^2) \to \mathrm{Aut}_{\mathbb{Z}/N}((\mathbb{Z}/N)^2) = GL_2(\mathbb{Z}/N)$$

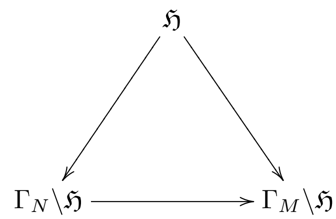
Thus, Γ_N is *normal* in $\Gamma_1 = SL_2(\mathbb{Z})$. For $M|N$ we have $\Gamma_N \subset \Gamma_M$, and a natural map

$$\Gamma_N\backslash\mathfrak{H} \to \Gamma_M\backslash\mathfrak{H} \quad \text{by} \quad \Gamma_N \cdot z \to \Gamma_M \cdot z$$

[9] We will look at the non-euclidean geometry of the upper half-plane and other examples a little later.

[10] This reduction homomorphism is *surjective*. This is not entirely trivial to prove, but is a feasible and wholesome exercise.

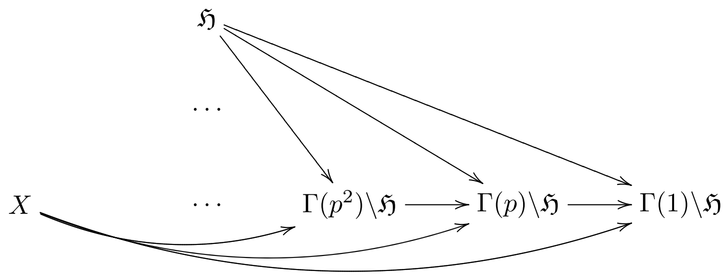
that fits into a commutative diagram with the respective quotient maps from $\mathfrak{H}$



As a warm-up, we'll not form the largest projective limit of these quotients, but only the p-power limit for a fixed prime p, much as we did in forming the classical solenoid earlier. Fix a prime p. Without necessarily being concerned about the geometric details of the quotients involved, we form the projective limit

$$X = \lim_n \Gamma(p^n)\backslash\mathfrak{H}$$

This has the picture



As proved useful in the earlier study of the classical solenoids, it will often be useful to identify elements of the projective limit with (compatible) sequences

$$\ldots \to z_2 \to z_1 \to z_0$$

of points in $z_n \in \Gamma(p^n)\backslash\mathfrak{H}$ that are *compatible* in the sense that the projection $\Gamma(p^n)\cdot z_{n+1}$ of z_{n+1} to the n^{th} limitand is z_n.

Just as we initially found $\mathbb{Z}_2$ in the automorphisms of the 2-solenoid from examination of the diagram for that projective limit, the fact that all the groups $\Gamma(p^n)$ are *normal* in $\Gamma(1) = SL_2(\mathbb{Z})$ exhibits some automorphisms of the limit.

[3.0.1] Claim:

$$\lim_n \Gamma(1)/\Gamma(p^n) \approx \lim_n SL_2(\mathbb{Z}/p^n) \approx SL_2(\mathbb{Z}_p)$$

acts on $\lim_n \Gamma(p^n)\backslash\mathfrak{H}$ in a natural fashion.

Proof: First, note that the quotient group $\Gamma(1)/\Gamma(p^n)$ acts on $\Gamma(p^n)\backslash\mathfrak{H}$ by

$$\gamma\cdot\Gamma(p^n)\cdot z = \gamma\Gamma(p^n)\gamma^{-1}\cdot\gamma z = \Gamma(p^n)\cdot\gamma z$$

from the normality of $\Gamma(p^n)$. For $\gamma \in \Gamma(p^n)$, the action is trivial since γ is absorbed: $\gamma\cdot\Gamma(p^n) = \Gamma(p^n)$. Thus, a (compatible) family of group elements

$$\ldots \to \gamma_2 \to \gamma_1 \to \gamma_0$$

with $\gamma_n \in \Gamma(1)/\Gamma(p^n)$ with the compatibility condition

$$\gamma_{n+1} \cdot \Gamma(p^n) = \gamma_n \cdot \Gamma(p^n)$$

gives an automorphism of the limit by

$$(\ldots \to \gamma_2 \to \gamma_1 \to \gamma_0) \cdot (\ldots \to z_2 \to z_1 \to z_0) = (\ldots \to \gamma_2 z_2 \to \gamma_1 z_1 \to \gamma_0 z_0)$$

as claimed. It remains to check that

$$\lim_n \Gamma(1)/\Gamma(p^n) \approx \lim_n SL_2(\mathbb{Z}/p^n) \approx SL_2(\mathbb{Z}_p)$$

The surjectivity of $\Gamma(1) \to SL_2(\mathbb{Z}/p^n)$ is left as an exercise. The kernel of this homomorphism is certainly $\Gamma(p^n)$. The elements of the projective limit are compatible families

$$\cdots \xrightarrow{\text{mod } p^3} \begin{bmatrix} a_3 & b_3 \\ c_3 & d_3 \end{bmatrix} \xrightarrow{\text{mod } p^2} \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \xrightarrow{\text{mod } p} \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix}$$

This means that each of the four sequences of entries is a compatible family of elements in the projective limit

$$\mathbb{Z}_p \approx \lim \left(\cdots \xrightarrow{\text{mod } p^3} \mathbb{Z}/p^3 \xrightarrow{\text{mod } p^2} \mathbb{Z}/p^2 \xrightarrow{\text{mod } p} \mathbb{Z}/p \right)$$

That is, we have the isomorphism $\lim_n SL_2(\mathbb{Z}/p^n) \approx SL_2(\mathbb{Z}_p)$ as claimed. ///

[3.0.2] Remark: The topological group $SL_2(\mathbb{Z}_p)$ is quite non-abelian. However, being a limit of finite groups, it is *compact*. [11] Compactness of a topological group is a good feature, in the sense that this compactness makes the group tractable in many regards. However, in the present situation it would be a serious mistake to overlook the presence of a much larger (non-compact) group $SL_2(\mathbb{Q}_p)$ of automorphisms of $\lim_n \Gamma(p^n)\backslash\mathfrak{H}$.

The next task is to find a larger but cofinal diagram in order to make more automorphisms easily visible. To do so, we must allow movement *outside* the group $SL_2(\mathbb{Z})$, although not too far, in a sense that will be made clear.

A **p-power congruence subgroup** is a subgroup Γ of $SL_2(\mathbb{Q})$ which contains some $\Gamma(p^n)$ with finite index. That is, for some $0 \le n \in \mathbb{Z}$

$$\Gamma \supset \Gamma(p^n) \qquad [\Gamma : \Gamma(p^n)] < \infty$$

[3.0.3] Claim: Let $g \in SL_2(\mathbb{Z}[\frac{1}{p}])$. Then the action $\Gamma \to g\Gamma g^{-1}$ stabilizes the set of p-power congruence subgroups of $SL_2(\mathbb{Q})$.

Proof: Let Γ be a p-power congruence subgroup. Given $g \in SL_2(\mathbb{Z}[\frac{1}{p}])$, we must show that $g\Gamma g^{-1}$ contains some $\Gamma(p^\ell)$, and with finite index. That is, we want to show that Γ contains some $g^{-1}\Gamma(p^\ell)g$ with finite index. Since the subgroups $\Gamma(p^\ell)$ are of finite index in each other, to verify the finite-index condition it suffices to verify it for any sufficiently *small* $\Gamma(p^\ell)$.

Let m be large enough such that there are integral matrices A, B such that we can write

$$g = 1_2 + p^{-m}A \qquad g^{-1} = 1_2 + p^{-m}B$$

[11] The compactness of an automorphism group created in this fashion is not surprising: there is a *bottom* limitand, the automorphism group of each limitand over the bottom one is *finite*, and the projective limit of finit groups is compact.

Let $\gamma = 1_2 + p^n N$ be in $\Gamma(p^n)$, where N is an integral matrix. Then

$$g^{-1}\gamma g \;=\; (1 + p^{-m}B)\,(1 + p^n N)\,(1 + p^{-m}A)$$

$$= 1 + p^{-m}B + p^{-m}A + p^{-2m}BA + p^n N + p^{n-m}BN + p^{n-m}NA + p^{n-2m}BNA$$

The first four summands sum to 1, since $g\,g^{-1} = 1$, so this is

$$1 + p^n N + p^{n-m}NA + p^{n-m}BN + p^{n-2m}BNA$$

Thus, for $n > 2m$, we have $g^{-1}\gamma g \in \Gamma(p^{n-2m})$, so

$$g^{-1}\,\Gamma(p^n)\,g \subset \Gamma(p^{n-2m}) \subset \Gamma \quad \text{(for } n \text{ large enough such that } \Gamma(p^{n-2m}) \subset \Gamma)$$

That is, for large enough n such that $\Gamma(p^{n-2m}) \subset \Gamma$, we do have the desired containment

$$\Gamma(p^n) \subset g\Gamma(p^{n-2m})g^{-1} \subset g\Gamma g^{-1}$$

To verify the finite-index condition,

$$[g\Gamma g^{-1} : \Gamma(p^n)] = [\Gamma : g^{-1}\Gamma(p^n)g] = [\Gamma : \Gamma(p^{n-2m})] \cdot [\Gamma(p^{n-2m}) : g^{-1}\Gamma(p^n)g]$$

since the indices are not altered by conjugating inside a larger group. And then

$$[\Gamma(p^{n-2m}) : g^{-1}\Gamma(p^n)g] = [g\Gamma(p^{n-2m})g^{-1} : \Gamma(p^n)] \le [g\Gamma(p^{n-4m})g^{-1} : \Gamma(p^n)] < \infty$$

when $n > 4m$, by the same computation as above. This proves that conjugation by elements of $SL_2(\mathbb{Z}[\frac{1}{p}])$ stabilizes the set of p-power congruence subgroups. ///

Thus, consider the larger family of limitands $\Gamma\backslash\mathfrak{H}$ where Γ is a p-power congruence subgroup, with the natural maps

$$\Gamma\backslash\mathfrak{H} \to \Gamma'\backslash\mathfrak{H} \quad \text{(for } \Gamma \subset \Gamma')$$

[3.0.4] Claim: The family $\Gamma\backslash\mathfrak{H}$ of quotients with p-power congruence subgroup Γ has *cofinal* subfamily consisting of the quotients $\Gamma(p^n)\backslash\mathfrak{H}$ by principal congruence subgroups $\Gamma(p^n)$, giving a natural isomorphism

$$\lim_{\Gamma = p \text{ power}} \Gamma\backslash\mathfrak{H} \quad\approx\quad \lim_n \Gamma(p^n)\backslash\mathfrak{H}$$

Proof: Since each such Γ contains some $\Gamma(p^n)$, for each Γ there is a surjection

$$\Gamma(p^n)\backslash\mathfrak{H} \to \Gamma\backslash\mathfrak{H}$$

That is, essentially by the definition of p-power congruence subgroups, the collection of quotients by principal congruence subgroups is *cofinal*. Cofinal limits are naturally isomorphic. ///

[3.0.5] Corollary: An element $g \in SL_2(\mathbb{Z}[\frac{1}{p}])$ acts on the limit of the p-power congruence quotients $\lim_\Gamma \Gamma\backslash\mathfrak{H}$ by an action induced from the compatible family of *isomorphisms*

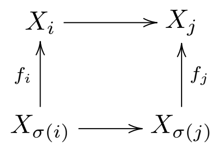
$$g : \Gamma\backslash\mathfrak{H} \to g\Gamma g^{-1}\backslash\mathfrak{H} \qquad \text{given by} \qquad g\cdot(\Gamma\cdot z) = g\Gamma g^{-1}\cdot(g\cdot z)$$

Proof: As usual, a map of a limit $X = \lim_{i\in I} X_i$ to itself is given by a compatible family of maps $X \to X_i$ to the limitands. One way to give such a family is as follows. Let $p_j : X \to X_j$ be the projection to the j^{th}

limitand. Let σ be a an order-preserving[12] permutation of the index set I, and suppose that we are given a family of *isomorphisms*

$$f_i : X_\sigma(i) \to X_i$$

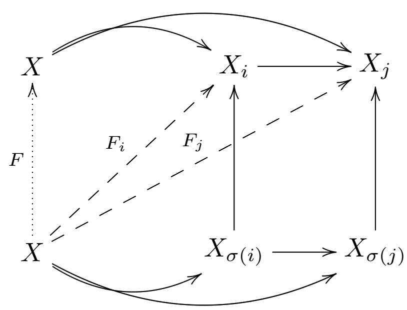
compatible in the sense that for $i > j$ there is a commutative diagram



Then define a family of maps $F_i : X \to X_i$ by

$$F_i = f_i \circ p_{\sigma(i)}$$

This gives a commutative diagram



with uniquely induced map $F : X \to X$. This idea applies to the p-power congruence subgroups with the natural isomorphisms

$$\Gamma\backslash\mathfrak{H} \to g\Gamma g^{-1}\backslash\mathfrak{H} \qquad \text{by} \qquad \Gamma \cdot z \to g\Gamma g^{-1} \cdot gz$$

Thus, $SL_2(\mathbb{Z}[\frac{1}{p}])$ acts on the projective limit. ///

Thus, so far, we have natural actions of $SL_2(\mathbb{Z}_p)$ and of $SL_2(\mathbb{Z}[\frac{1}{p}])$ on the projective limit $\lim_n \Gamma(p^n)\backslash\mathfrak{H}$, which is also expressible as the limit over p-power congruence subgroups. We certainly would like to combine the two actions.

[3.0.6] Claim: We have a natural action of $SL_2(\mathbb{Q}_p)$ on $\lim_n \Gamma(p^n)\backslash\mathfrak{H}$.

Proof: [... iou ...] ///

There are many reasons to prefer $GL_2(\mathbb{Q}_p)$ to $SL_2(\mathbb{Q}_p)$, as topological groups. Luckily, we can easily *redo* the above with $GL(2)$ rather than $SL_2()$, with $GL_2(\mathbb{Q})$ acting on the union $\mathfrak{H}\cup\overline{\mathfrak{H}}$ of upper and lower half-planes. Let

$$\widetilde{\Gamma}(N) = \{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL_2(\mathbb{Z}) : \begin{bmatrix} a & b \\ c & d \end{bmatrix} = 1 \bmod N, \text{ (entry-wise)}\}$$

be the principal congruence subgroups in $GL_2(\mathbb{Z})$, rather than $SL_2(\mathbb{Z})$. Then the same discussion as with $SL_2()$ gives

$$\lim_n \widetilde{\Gamma}(1)/\widetilde{\Gamma}(p^n) \approx \lim_n GL_2(\mathbb{Z}/p^n) \approx GL_2(\mathbb{Z}_p)$$

and natural actions of $GL_2(\mathbb{Z}_p)$ and $GL_2(\mathbb{Z}[\frac{1}{p}])$ on

$$\lim_n \widetilde{\Gamma}(p^n)\backslash(\mathfrak{H} \cup \overline{\mathfrak{H}})$$

[12] This sense of *order-preserving* is what one should expect, namely, for $i > j$ we have $\sigma(i) > \sigma(j)$.

which we assemble to give an action of $GL_2(\mathbb{Q}_p)$.

[3.0.7] Remark: The topological group $GL_2(\mathbb{Q}_p)$ is *not* compact, and the fact that a family of modular curves admits $GL_2(\mathbb{Q}_p)$ as part of its group of automorphisms has enormous potential. We are not yet equipped exploit the appearance of this group of symmetries, but will prepare to do so.

---

4. *Enabling the action of* $SL_2(\mathbb{R})$

We have found automorphism groups $SL_2(\mathbb{Q}_p)$ or $GL_2(\mathbb{Q}_p)$ of families of modular curves, but the original and most immediate group action, that of $SL_2(\mathbb{R})$, has been *disabled*.

That is, the upper half-plane $\mathfrak{H}$ is a homogeneous space for $SL_2(\mathbb{R})$, being

$$\mathfrak{H} \approx SL_2(\mathbb{R})/SO(2) \quad (SO(2) \text{ the isotropy group of } i \in \mathfrak{H})$$

but the group $SL_2(\mathbb{R})$ does *not* act on any individual quotient

$$\Gamma\backslash\mathfrak{H} \;\approx\; \Gamma\backslash SL_2(\mathbb{R})/SO(2) \qquad\qquad (\Gamma \text{ a congruence subgroup})$$

since the Γ gets in the way: $SL_2(\mathbb{R})$ normalizes *no* such Γ. Similarly, $SL_2(\mathbb{R})$ cannot act on any projective limit of such quotients, because, still, conjugation by $SL_2(\mathbb{R})$ does not stabilize any good *collection* of subgroups Γ. Thus, the $SL_2(\mathbb{R})$-homogeneity of quotients $\Gamma\backslash\mathfrak{H}$ is difficult to see or use in this form. For the same reasons, the left action of $GL_2(\mathbb{R})$ on

$$GL_2(\mathbb{R})/(SO(2) \times \{\text{scalar matrices}\}) \;\approx\; \mathfrak{H} \cup \overline{\mathfrak{H}}$$

gets unfortunately submerged in quotients

$$\widetilde{\Gamma}\backslash GL_2(\mathbb{R})/(SO(2) \times \{\text{scalars}\}) \;\approx\; \widetilde{\Gamma}\backslash(\mathfrak{H} \cup \overline{\mathfrak{H}})$$

and in limits.

We *could* have $SL_2(\mathbb{R})$ or $GL_2(\mathbb{R})$ acting on the *right* if the $SO(2)$ weren't there. Indeed, this turns out to be a powerful argument to give up the otherwise appealing complex structure on $\mathfrak{H}$, and consider

$$\begin{array}{ccc} \Gamma\backslash SL_2(\mathbb{R}) & \text{instead of} & \Gamma\backslash\mathfrak{H} \\ \lim_\Gamma \Gamma\backslash SL_2(\mathbb{R}) & \text{instead of} & \lim_\Gamma \Gamma\backslash\mathfrak{H} \\ \widetilde{\Gamma}\backslash GL_2(\mathbb{R}) & \text{instead of} & \widetilde{\Gamma}\backslash(\mathfrak{H} \cup \overline{\mathfrak{H}}) \\ \lim_{\widetilde{\Gamma}} \widetilde{\Gamma}\backslash GL_2(\mathbb{R}) & \text{instead of} & \lim_{\widetilde{\Gamma}} \widetilde{\Gamma}\backslash(\mathfrak{H} \cup \overline{\mathfrak{H}}) \end{array}$$

where Γ denotes congruence subgroups of $SL_2(\mathbb{Z})$, and $\widetilde{\Gamma}$ denotes congruence subgroups of $GL_2(\mathbb{Z})$. Then

[4.0.1] Claim: We have natural isomorphisms

$$\lim_n \Gamma(p^n)\backslash SL_2(\mathbb{R}) \;\approx\; SL_2(\mathbb{Z}[\frac{1}{p}]) \setminus (SL_2(\mathbb{R}) \times SL_2(\mathbb{Q}_p))$$

$$\lim_n \widetilde{\Gamma}(p^n)\backslash GL_2(\mathbb{R}) \;\approx\; GL_2(\mathbb{Z}[\frac{1}{p}]) \setminus (GL_2(\mathbb{R}) \times GL_2(\mathbb{Q}_p))$$

where the subgroups $SL_2(\mathbb{Z}[\frac{1}{p}])$ and $GL_2(\mathbb{Z}[\frac{1}{p}])$ are diagonally imbedded, and are *discrete*.

Proof: [... iou ...] ///
