The Modular Group and its Actions

A. Muhammed Uludağ
(with an appendix by Hakan Ayral)

December 28, 2013

To Prof. Rolf-Peter Holzapfel, with admiration and respect

Abstract

This is a leisurely introduction to the modular group and some of its actions. We start by defining the modular group as an encoding of the euclidean algorithm and we present some basic facts about it. We then study and relate the following actions of the modular group: on itself by conjugation, on the Farey tree, on the circle, on binary quadratic forms, on planar framed lattices, on the upper half plane & on the hyperbolic plane, and on dessins.

Keywords. Euclidean algorithm, continued fractions, modular group, Farey tree, modular graphs, surface triangulations, lozenges, dessins, çarks, binary quadratic forms, bracelets.

Contents

1 Introduction

The projective group of integral two by two matrices of determinant one is called the *modular group*[1] and denoted $\mathrm{PSL}_2(\mathbb{Z})$.

This group is without doubt one of the most significant objects in mathematics. It and its subgroups appear in many apparently independent contexts in a surprising fashion. It is connected to many fundamentally important notions in an essential manner; for example; hyperbolic geometry, elliptic

[1]For the moment we shall not distinguish between the modular group $\mathrm{PSL}_2(\mathbb{Z})$ and its extensions $\mathrm{PGL}_2(\mathbb{Z})$, $\mathrm{SL}_2(\mathbb{Z})$ and $\mathrm{GL}_2(\mathbb{Z})$.

curves, modular curves, modular forms and functions, dessins, sporadic simple groups, circle homeomorphisms, etc.

The modular group appears as the foremost object in many lists. Thus one can introduce and study it in many different contexts; for example as:

- the automorphism group of the Farey tree,
- the free product $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}$,
- the braid group on three strands modulo its center
- the mapping class group of the torus (or of the punctured torus)
- the projective group of two by two integral matrices $\mathrm{PSL}_2(\mathbb{Z})$,
- a group of isometries of the hyperbolic plane or a group of conformal transformations of the upper half plane,
- a group of homeomorphisms of the circle,
- the automorphism group of the group $\mathbb{Z} \times \mathbb{Z}$,
- the outer automorphism group of the free group $\mathbb{Z} * \mathbb{Z}$.

In view of its importance, it is very strange that there are almost no treatises dedicated exclusively to the modular group. Our aim here is to trace the origin of the modular group till its source, namely the Euclidean algorithm, the place where the categories of quantity and relation make contact, resulting in a big bang of concepts, notions and ideas. From our point of view the modular group is nothing but an encoding of the Euclidean algorithm, expressed in an operator language via continued fractions, which is then translated into the matrix language and to the language of group theory.

We will use the basic group-theoretical fact that $\mathrm{PSL}_2(\mathbb{Z})$ is isomorphic to the free product $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}$. We will then study and relate the following actions of the modular group:

- the left-action by automorphisms of the infinite trivalent plane tree and by homeomorphisms on its boundary,
- the left-action on the set of pairs of trivalent ribbon graphs with an edge (preserving the underlying ribbon graph),
- the left-conjugation action on itself,
- the right-action on the binary quadratic forms of Gauss (preserving the set of values assumed by the form),

- the left-action on the circle S^1 by Möbius transformations
- the left action on the set of planar framed lattices, area-preserving linear transformations of the Euclidean plane,
- the left-action on the upper half plane $\mathcal{H}$ by Möbius transformations,
- the left-action on the set of dessins, preserving the monodromy group.

Of course, it is impossible to claim completeness of this list: every time $\mathrm{PSL}_2(\mathbb{Z})$ manifests itself as a subgroup of some group G, it acts on G by left-multiplication and by conjugation. There are surely many significant actions not accounted for here, some perhaps being discovered as these lines are being read.

This text is intended to be a first-order approximation to a monograph on the modular group. We don't claim that the text is self-contained; many proofs have been omitted to be able to include as many basic facts about the modular group as possible. There are some exercises throughout the text with varying levels of difficulty. Whenever a new action is encountered in the text, we signal this by a ♡ symbol.

Acknowledgements. This research has been funded by the TÜBİTAK grant 110T690 and a Galatasaray University Research Grant. I am grateful to A. Papadopoulos, L. Ji and S-T Yau for inviting me to attend the Kunming meeting on Group Actions and to publish in the Handbook of Group Actions. This paper is a result of a joint book project on the modular group by M. Uludağ and İ. Portakal in turkish. I am indebted to Hakan Ayral, İrem Portakal, Ayberk Zeytin, Merve Durmuş, İsmail Sağlam, Fırat Yaşar, Çiğdem Karakoç and Tuğçe Çolak for various discussions.

2 Rise of the modular group

2.1 Anthyphairesis (Euclidean Algorithm)

Suppose that we are given two sticks and we are asked to compare their lenghts. One way to perform this task is to take a ruler, measure these magnitudes in terms of some unit, and divide the numbers we get. The resulting dimensionless number, namely the ratio r gives a precise information about the comparison of the sticks' lengths.

Stick A−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−−

Stick B −−−−−−−−−−−−−−−−−−−

However, there is something ugly about the usage of a ruler (and therefore a unit) in order to define a dimensionless number. It is possible to avoid this by applying the following comparison process, called *anthyphairesis* (reciprocal substraction):

2.1.1 Anthyphairesis, first version.

Perform the following operations: (Operation $T^\bullet$) Cut off from the stick A multiples of the stick B until the remainder C becomes shorter then B, then (Operation U) replace A with B, replace B with C and repeat the procedure.

This is the primordial algorithm of Euclid (fl. 300 BC), who shows that the process terminates provided A and B are commensurable and yields the greatest common measure at the last step[2]. If we record how many B's have been cut from A's at each loop, and if the algorithm terminates, we get a finite sequence $\lfloor n_0, n_1, \dots n_k \rfloor$ of integers in terms of which we may express the ratio r as a regular continued fraction, as follows:

Suppose that the process terminates at step k. Then we may encode the algorithm as follows:

$$(T^\bullet)^{n_k} \cdot U \cdot (T^\bullet)^{n_{k-1}} \cdot U \cdots (T^\bullet)^{n_1} \cdot U \cdot (T^\bullet)^{n_0} (r) = 0,$$

or, reversing the procedure, we may write

$$T^{n_0} \cdot U \cdot T^{n_1} \cdot U \cdots T^{n_{k-1}} \cdot U \cdot T^{n_k} (0) = r,$$

where $T^\bullet = T^{-1}$. Now set $x_0 := A$, $x_1 := B$. Then one has

$$\begin{array}{lll}
x_0 = n_0 x_1 + x_2 & (0 < x_2 < x_1) & \Longrightarrow [x_0 : x_1] = n_0 + [x_2 : x_1] \\
x_1 = n_0 x_2 + x_3 & (0 < x_3 < x_2) & \Longrightarrow [x_0 : x_1] = n_1 + [x_2 : x_1] \\
x_2 = n_2 x_3 + x_4 & (0 < x_4 < x_3) & \Longrightarrow [x_2 : x_3] = n_2 + [x_4 : x_3] \\
. & & \\
. & & \\
x_{k-2} = n_{k-2} x_{k-1} + x_k & (0 < x_{k+1} < x_k) & \Longrightarrow [x_{k-2} : x_{k-1}] = n_{k-2} + [x_k : x_{k-1}] \\
x_{k-1} = n_k x_k + 0 & (x_{k+1} = 0) & \Longrightarrow [x_{k-1} : x_k] = n_k,
\end{array}$$

and therefore

$$r = [x_0 : x_1] = n_0 + \frac{1}{[x_1 : x_2]} = n_0 + \cfrac{1}{n_1 + \cfrac{1}{[x_2 : x_3]}} = n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{[x_3 : x_4]}}} = \dots$$

[2]Beware that Euclid's claim and its proof are much more precise and focused than the account we give here for recreational purposes.

$$= n_0 + \cfrac{1}{n_1 + \cfrac{1}{\ddots + \cfrac{1}{\lfloor x_{k-1} : x_k \rfloor}}} = n_0 + \cfrac{1}{n_1 + \cfrac{1}{\ddots + \cfrac{1}{n_k}}}$$

This is called a regular (or simple or ordinary) continued fraction. We see that T acts as $r \mapsto r+1$ on ratios, whereas U acts as $r \mapsto 1/r$. In case the lengths of the sticks A and B are not commensurable[3], the process does not terminate and we obtain an infinite regular continued fraction.

Exercise. Compute the continued fraction expansion of $\frac{55}{34}$.

The numbers n_k, called the *partial quotients of* r, can be expressed as

$$n_k = n_k(r) = \left\lfloor \frac{1}{\tau^{k-1}(r)} \right\rfloor, \qquad \text{provided } \tau^{k-1}(r) \neq 0,$$

where $\tau : [0,1) \to [0,1)$ is the Gauss continued fraction map defined by

$$\tau(0) = 0 \text{ and } \tau(r) = \frac{1}{r} - \left\lfloor \frac{1}{r} \right\rfloor \qquad (r \neq 0).$$

Here $\lfloor r \rfloor$ denotes the integral part of r.

Proposition 2.1. *The monoid $\mathcal{M}$ generated by the operators T and U acting on ratios is isomorphic to the monoid of projective matrices*

$$\mathrm{PGL}_2(\mathbb{N}) := \left\{ \begin{bmatrix} p & q \\ r & s \end{bmatrix} \quad : \quad p,q,r,s \in \mathbb{N}, \quad ps - qr = \pm 1 \right\}.$$

(In fact, $\mathrm{PGL}_2(\mathbb{N})$ is exactly the same monoid as $\mathrm{GL}_2(\mathbb{N})$; we use the former one; for the sake of consistency of notation.)

Proof. The map $\mathcal{M} \to \mathrm{PGL}_2(\mathbb{N})$ sending

$$T \to \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } U \to \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

is a homomorphism of monoids. Surjectivity is shown by reducing a matrix $M \in \mathrm{PGL}_2(\mathbb{N})$ by successively multiplying with T^{-1} and U to get the identity. In fact, $\mathcal{M}$ is freely generated by T and U. $\square$

[3]Commensurability means the existence of a common unit of measure.

The words in $\mathcal{M}$ coming from a terminating anthyphairesis process never end with a U, so these form a submonoid of $\mathcal{M}$ which is isomorphic to the following submonoid of $\mathrm{PGL}_2(\mathbb{N})$:

$$\mathcal{M}^+ := \left\{ \begin{bmatrix} p & q \\ r & s \end{bmatrix} : p,q,r,s \in \mathbb{N},\, ps - qr = \pm 1,\, p \le q\,(\implies r \le s) \right\} \cup \{I\}$$

We have a bijective map $W \in \mathcal{M}^+ \to W(0) \in \mathbb{Q}_{\ge 0}$. This correspondence extends to a correspondence between the set of infinite words in $\mathcal{M}$ and $\mathbb{R}_{\ge 0}$, as follows.

Denote by $\partial\mathcal{M}$ the set of all infinite words in $\mathcal{M}$ (the set of infinite words in $\mathcal{M}^+$ yields the same set). In other words, $\partial\mathcal{M}$ is the set of non-terminating anthyphairesis processes. The monoid $\mathcal{M}$ acts on $\partial\mathcal{M}$ by concatenation from the left.

Terminating processes can be represented by non-terminating ones, i.e. the infinite word

$$T^{n_0}UT^{n_1}U\cdots T^{n_{k-1}}UT^{n_k}UT^{\infty}$$

represents the finite word

$$T^{n_0}UT^{n_1}U\cdots T^{n_{k-1}}UT^{n_k}.$$

Now let $W = T^{n_0}\prod_{i=1}^{\infty} UT^{n_i} \in \partial\mathcal{M}$ and denote

$$W_l = T^{n_0}\prod_{i=1}^{l} UT^{n_i}.$$

Then it can be shown that the sequence $W_l(0)$ always converges, and one can define a map

$$W \in \partial\mathcal{M} \to \lim_{l\to\infty} W_l(0) \in \mathbb{R}_{\ge 0}.$$

In other words the continued fractions

$$n_0 + \cfrac{1}{n_1 + \cfrac{1}{\ddots}}$$

uniquely represent all non-negative real numbers, provided $n_0 \ge 0$ and $0 < n_i \le \infty$ for $i > 0$. The action of $\mathcal{M}$ on $\partial\mathcal{M}$ becomes the action of $\mathrm{PGL}_2(\mathbb{N})$ on $\mathbb{R}_{\ge 0}$ by linear fractional transformations.

Exercise. Describe the continued fraction expansion of $\sqrt{2}$.

∞

Instead of proving the claims of convergence, uniqueness, etc, by some delicate arguments, we will adopt a more fundamental strategy and associate to W directly a Dedekind cut of $\mathbb{Q}$. This requires the extension of the above correspondence to negative numbers as well, and it may appear that replacing the monoid $\mathrm{PGL}_2(\mathbb{N})$ by the group $\mathrm{PGL}_2(\mathbb{Z})$ should do the job. However, this choice brings in the semiregular continued fractions

$$\pm n_0 \pm \cfrac{1}{n_1 \pm \cfrac{1}{\ddots}}$$

and the uniqueness of representation is lost: every real number has uncountably many semiregular continued fraction expansions. Our plan is to show that if we take instead the modular group $\mathrm{PSL}_2(\mathbb{Z})$ as our starting point, then there is a *seamless* correspondence between infinite words and all real numbers. This will bring the further benefit of providing a clear insight into the uncountable multitude of semiregular continued fraction expansions.

2.1.2 Anthyphairesis, second version.

Perform the following operations: (Operation $T^\bullet$) As in the first version, cut off from the stick A multiples of the stick B until the remainder becomes shorter than B, but this time repeat this step one more time so that you will have a stick C of "negative" length, then (Operation S) replace A with B, replace B with B-C and enter the loop.

This is a variation of Euclid's algorithm which terminates provided A and B are commensurable. If we record how many B's have been cut from A's at each loop, and if the algorithm terminates, we get a finite sequence $\lceil n_0, n_1, \dots n_k \rceil$ of integers in terms of which we may express the ratio r as a "minus-" continued fraction, (sometimes also called "backwards" continued fraction) as follows:

Suppose that the process terminates at step k. Then we may encode the algorithm as follows:

$$(T^\bullet)^{n_k} \cdot S \cdot (T^\bullet)^{n_{k-1}} \cdot S \cdots (T^\bullet)^{n_1} \cdot S \cdot (T^\bullet)^{n_0} (r) = 0,$$

or, reversing the procedure, we may write

$$T^{n_0} \cdot S \cdot T^{n_1} \cdot S \cdots T^{n_{k-1}} \cdot S \cdot T^{n_k} (0) = r, \quad (n_0 \geq 1, \quad n_1, n_2 \cdots \geq 2)$$

This gives

$$r = n_0 - \cfrac{1}{n_1 - \cfrac{1}{\cdots - \cfrac{1}{n_{k-1} - \cfrac{1}{n_k}}}}.$$

We see that T acts as $r \mapsto r+1$ on the set of ratios, whereas S acts as $r \mapsto -1/r$. As in the previous case, the process does not terminate if the lengths A and B are not commensurable and we obtain an infinite semiregular continued fraction.

Exercise. Compute the "minus-" continued fraction expansion of $\frac{7}{129}$.

Exercise. Show that

$$\lceil n_0, n_1, \dots \rceil = \lfloor n_0 + 1, \underbrace{2, 2, \dots, 2}_{n_1 - 1}, n_2 + 2, \underbrace{2, 2, \dots, 2}_{n_3 - 1}, n_4 + 2, \dots \rfloor$$

.

Note that, in the "minus"-continued fraction $r = \lceil n_0, n_1, \dots n_k \rceil$, we require that $n_0 \geq 1$ and $n_i \geq 2$ for $i > 0$. If we remove this restriction then we get the somewhat surprising result:

Proposition 2.2. *The monoid generated by T and S is in fact a group, isomorphic to the modular group*

$$\mathrm{PSL}_2(\mathbb{Z}) := \left\{ \begin{pmatrix} p & q \\ r & s \end{pmatrix} \quad : \quad p, q, r, s \in \mathbb{Z}, \quad ps - qr = 1 \right\} \Big/ \langle \pm I \rangle$$

Proof. The monoid is a group because $TS = L$ is of order 3, so that $T = LS \implies T^{-1} = SL^2 = STSTS$. Now denote the projectivization of $\begin{pmatrix} p & q \\ r & s \end{pmatrix}$ by $\begin{bmatrix} p & q \\ r & s \end{bmatrix}$ and define the map sending

$$T \to \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } S \to \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Hence $\mathrm{PSL}_2(\mathbb{Z})$ is generated by S and L. In fact the modular group is freely generated by these elements, in other words one has

$$\mathrm{PSL}_2(\mathbb{Z}) \simeq \langle S \rangle * \langle L \rangle \simeq \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}.$$

□

Loosely speaking, we may say that the modular group encodes the set of all comparison (anthyphairesis) processes "by excess", whereas the group $\mathrm{PGL}_2(\mathbb{Z})$ encodes the set of all comparison processes. Since mathematics is ultimately about comparing magnitudes, this explains why we encounter the modular group so often in mathematics. Our goal in the next section is to show that this loose remark can be made much more precise. But before this we present some useful facts about the modular group.

2.2 Miscellaneous facts about the modular group

Let us first recall that the trace of a matrix is a class function; i.e. $tr(XMX^{-1}) = tr(M)$ for every X and M. The absolute value of the trace is well defined on $\mathrm{PSL}_2(\mathbb{R})$. A non-identity element M of $\mathrm{PSL}_2(\mathbb{R})$ is called *elliptic* if $|tr(M)| < 2$, *parabolic* if $|tr(M)| = 2$ and *hyperbolic* if $|tr(M)| > 2$. One has the following standard result:

Lemma 2.1. a) *An element $M \neq I$ of $\mathrm{PSL}_2(\mathbb{Z})$ is of finite order if an only if it is elliptic. It is of order 2 if and only if $|tr(M)| = 0$, in this case it is conjugate to S. It is of order 3 if and only if $|tr(M)| = 1$, in this case it is conjugate to L or L^2.* **b)** *Any parabolic element is conjugate to a translation T^n for some n.*

An element of a group is called *primitive* if it is not a positive power of some other element of the group. There are infinitely many primitive conjugacy classes of elements of $\mathrm{PSL}_2(\mathbb{Z})$, which can be described with Lyndon words (equivalently by bracelets or by graphs called çarks), see Section 6.

Lemma 2.2. *Any pair of elliptic generators of the modular group is simultaneously conjugate to the pair $\{L, S\}$ or to the pair $\{L^{-1}, S\}$.*

2.2.1 Encoding the elements of $\mathrm{PSL}_2(\mathbb{Z})$

There are several ways of representing the elements of $\mathrm{PSL}_2(\mathbb{Z})$. It seems appropriate to collect them here.

1. **Projective matrices.** This is the definition. Group operation is the matrix product.

2. **Linear fractional transformations.** Here the group operation is the composition of transformations.

3. **Words.** Since $\mathrm{PSL}_2(\mathbb{Z}) \simeq \mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z} = \langle S, L \mid S^2 = L^3 = 1\rangle$, any element of the modular group can be represented by a word in S and L.

The group operation is concatenation of words. Any word in S and L can be written as $W = S^\alpha (LS)^{n_0}(L^2S)^{n_1}(LS)^{n_2}(L^2S)^{n_3}\dots(LS)^{n_k}S^\beta$, where $n_0, n_k \geq 0$, $n_1, \dots, n_{k-1} > 0$ and $\alpha, \beta \in \{0, 1\}$. We may represent this expression as $W = S^\alpha[n_0, n_1, \dots, n_k]S^\beta$. Alternatively, one can choose another set of generators of the group, for example, T and S. One should be careful because then these generators will not be free generators for the group in general.

4. **Paths.** Any element of the modular group can be represented by a path on the bipartite Farey tree in a unique way if we choose a base edge of the tree. The group operation is concatenation of paths. This is the content of the next Section.

5. **Conjugal.** Let $M \in \mathrm{PSL}_2(\mathbb{Z})$ and set $M_0 := M$. For $i \geq 0$ and X_i a cyclically reduced word in L and S, write $M_{i-1} = M_i X_i M_i^{-1}$ until the sequence (X_i) stabilizes (say after the nth step). Then M can be represented by the sequence $(X_1, X_2, \dots, X_n)$.

2.2.2 Functions on the modular group

We have already considered the trace function on $\mathrm{PSL}_2(\mathbb{Z})$. This is a class function. Another class function associates to M the length of the minimal word in L, S in its conjugacy class. There are some other functions of importance on the modular group. By using the word representation, we define the function

$$\lambda(S^\alpha[n_0, n_1, \dots, n_k]S^\beta) := (-1)^\alpha(\alpha + 2(n_0 + n_1 + \cdots + n_k) - \beta).$$

(This also defines a function on the bipartite Farey tree $\mathcal{F}$ constructed in the next section: it associates to the edge marked by $M \in \mathrm{PSL}_2(\mathbb{Z})$ its signed distance from the edge marked by I. In particular $\lambda(\{I\}) := 0$.) We also define the L-turn and L^2-turn counting functions by

$$\lambda^+(S^\alpha[n_0, n_1, \dots, n_k]S^\beta) := (-1)^\alpha(n_0 + n_2 + \dots)$$

$$\lambda^-(S^\alpha[n_0, n_1, \dots, n_k]S^\beta) := (-1)^\alpha(n_1 + n_3 + \dots).$$

One has thus $\lambda = 2\lambda^+ + 2\lambda^- + (-1)^\alpha(\alpha - \beta)$. Another important function measures the difference between L-turns and L^2-turns:

$$\lambda^\delta := \lambda^+ - \lambda^-$$

2.2.3 Modular group and the braid group

Artin's braid group $\mathbb{B}_3$ admits the presentation $\mathbb{B}_3 = \langle a, b \,|\, aba = bab \rangle$. The modular group is a quotient of this group. For indeed, if we put $x = aba$ and $y = ab$, then we get $a = y^{-1}x$ and $b = x^{-1}y^2$. Hence this change of generators is invertible. The relation $aba = bab$ in terms of these new generators reads as $x^2 = y^3$, i.e. $\mathbb{B}_3$ is isomorphic to the group $\langle x, y \,|\, x^2 = y^3 \rangle$. It follows that the element $x^2 = y^3 = (ab)^3$ is central and generates the kernel of the map sending (x, y) to $(S, L) \in \mathrm{PSL}_2(\mathbb{Z})$.

As a consequence of this discussion, we see that the modular group acts on $\mathbb{B}_3$ by inner automorphisms. ♡

2.2.4 Congruence subgroups of the modular group

In Section 6 we will show how subgroups of the modular group correspond to graphs (called modular graphs) on oriented surfaces. Here we introduce the classical subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ of immense importance in number theory. The modular graphs corresponding to these subgroups are not as easily described.

The *principal congruence modular group of level* N is the subgroup

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

This subgroup is normal in $\mathrm{PSL}_2(\mathbb{Z})$ since it is the kernel of the reduction homomorphism $\mathrm{PSL}_2(\mathbb{Z}) \to \mathrm{PSL}_2(\mathbb{Z}/N\mathbb{Z})$, which is also surjective. The following subgroups are called the *modular groups of Hecke type.*

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) \,:\, b \equiv c \equiv 0 \mod N \right\},$$

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

One has

$$\Gamma(N) \subset \Gamma_1(N) \subset \Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z}) \quad \text{and} \quad \Gamma(1) = \Gamma_1(1) = \Gamma_0(1) = \mathrm{SL}_2(\mathbb{Z}).$$

We also have the group

$$\Gamma_0^+(N) := \left\langle \Gamma_0(N), W_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix} \right\rangle < \mathrm{GL}_2(\mathbb{Q}).$$

The latter is not a subgroup of the modular group, but it is commensurable with it when $N = 2$. The element $W_N(z) = -1/Nz$ is called the *Fricke* *involution.*

Question. What kind of numbers give rise to comparison processes that can be expressed as sequences of words in a subgroup (in particular a congruence subgroup) of the modular group? What is the measure of this set of numbers?

2.2.5 Some quotients of the modular group

We note in passing some known facts about the quotients of $\mathrm{PSL}_2(\mathbb{Z})$.

Proposition 2.3. *The alternating group A_n is a quotient of $\mathrm{PSL}_2(\mathbb{Z})$ for $n \geq 9$.*

G. A. Miller proved in 1901, that for $n \geq 9$, the alternating group A_n is generated by an element of order 2 and an element of order 3, see [13]. Dey and Wiegold [35] gave explicit generators in 1971. It is also known that most finite simple groups are generated by two elements of order 2 and 3, so these are quotients of the modular group as well.

One-relator quotients of $\mathrm{PSL}_2(\mathbb{Z})$, in terms of its presentation as $\mathbb{Z}/2\mathbb{Z} * \mathbb{Z}/3\mathbb{Z}$ have also been an object of study. To begin with we have the triangle groups

$$T_{2,3,n} := \langle x, y, z \,|\, x^2 = y^3 = (xy)^n = 1 \rangle.$$

This group is finite when $n = 2, 3, 4, 5$, infinite solvable when $n = 6$ and it contains non-abelian free subgroups for $n > 6$. In general, a one-relator quotient of the modular group is a group with presentation

$$T_{2,3,W} := \langle x, y, z \,|\, x^2 = y^3 = W^n = 1 \rangle,$$

where $W \in \mathrm{PSL}_2(\mathbb{Z})$. These groups are called *generalized triangle groups.* Evidently, this group depends only on the conjugacy class of W (we will give an explicit description of these conjugacy classes in Section 6). It is known that $T_{2,3,W}$ is infinite for $n \geq 6$ and contains non-abelian free subgroups for $n > 6$, see [1]. There are some experimental results on the nature of the remaining groups ([3]) as well as some asymptotic results ([12]).

3 The bipartite Farey tree $\mathcal{F}$ and its boundary

3.1 Construction of $\mathcal{F}$

From $\mathrm{PSL}_2(\mathbb{Z})$ we construct the *bipartite Farey tree $\mathcal{F}$*, whose edges are identified with the elements of the modular group. The set of vertices of $\mathcal{F}$ is defined as $V(\mathcal{F}) = V_{\otimes}(\mathcal{F}) \sqcup V_{\bullet}(\mathcal{F})$, where $V_{\otimes}(\mathcal{F}) = \{\{W, WS\} \colon W \in \mathrm{PSL}_2(\mathbb{Z})\}$ is the set of degree-2 vertices and $V_{\bullet}(\mathcal{F}) = \{\{W, WL, WL^2\} \colon W \in \mathrm{PSL}_2(\mathbb{Z})\}$ is the set of degree-3 vertices. Two distinct vertices v and v' are joined by an edge if and only if the intersection $v \cap v'$ is non-empty and in this case the edge between the two vertices is the only element in the intersection. The edges incident to the vertex $\{W, WL, WL^2\} \in V_{\bullet}$ are W, WL and WL^2, and these edges inherit a natural cyclic ordering which we fix for all vertices as (W, WL, WL^2). Thus $\mathcal{F}$ is an infinite bipartite ribbon graph. It is a tree since $\mathrm{PSL}_2(\mathbb{Z})$ is freely generated by S and L.

$M \in \mathrm{PSL}_2(\mathbb{Z})$ acts on $\mathcal{F}$ from the left by ribbon graph automorphisms ♡
as follows:

$$\begin{array}{rcl} W \in E(\mathcal{F}) & \mapsto & MW \in E(\mathcal{F}) \\ \{W, WS\} \in V_{\otimes}(\mathcal{F}) & \mapsto & \{MW, MWS\} \in V_{\otimes}(\mathcal{F}) \\ \{W, WL, WL^2\} \in V_{\bullet}(\mathcal{F}) & \mapsto & \{MW, MWL, MWL^2\} \in V_{\bullet}(\mathcal{F}). \end{array}$$

The action is free on $E(\mathcal{F})$ since this is no other than the left-regular action ♡
of $\mathrm{PSL}_2(\mathbb{Z})$ on itself.

3.2 Boundary of $\mathcal{F}$

A *path* in a graph $\mathcal{G}$ is a sequence of edges $e_1, e_2, \ldots, e_k$ of $\mathcal{G}$ such that e_i and e_{i+1} are coincident at a vertex, for each $1 \leq i < k$. The objects of the *fundamental groupoid* of $\mathcal{G}$ are the edges of $\mathcal{G}$ and the morphisms are defined to be non back-tracking oriented paths. Composition in the groupoid is defined as the concatenation of paths wherever possible.

Since the edges of $\mathcal{F}$ are labeled by reduced words in the letters L and S, a path in the bipartite Farey tree $\mathcal{F}$ is a sequence of reduced words (W_i) in L and S, such that $W_i^{-1} W_{i+1} \in \{L, L^2, S\}$ for every i. Since $\mathcal{F}$ is connected and simply connected, there is a unique non-backtracking path through any pair of edges.

An *end* of $\mathcal{F}$ is an equivalence class of infinite (but not bi-infinite) non-backtracking paths in $\mathcal{F}$, where eventually coinciding paths are considered to be equivalent. In other words, an end of $\mathcal{F}$ is the equivalence class of an infinite sequence of finite reduced words (W_i) in L and S with $W_i^{-1} W_{i+1} \in \{L, S\}$ for every i, where sequences with coinciding tails are equivalent. One may also view the end (W_i) of $\mathcal{F}$ as the pair (W_1, W), where W_1 is the starting edge and W is the infinite word

$$W = \prod_{i=1}^{\infty} W_i^{-1} W_{i+1}$$

in L and S. Here infinite words with different starting edges are taken to be equivalent.

Remark 3.1. *Ends of $\mathcal{F}$ are orbits of the following free action of* $\mathrm{PSL}_2(\mathbb{Z})$ ♡ *on infinite, non-backtracking paths: If $M \in \mathrm{PSL}_2(\mathbb{Z})$, construct the sequence (M_i) with*

$$I = M_1, M_2, \ldots, M_{k-1}, M_k = M,$$

where $M_i^{-1} M_{i+1} \in \{S, L, L^2\}$ for $1 \le i \le k-1$. Then the action

$$\mathrm{PSL}_2(\mathbb{Z}) \times \{\textit{paths on } \mathcal{F}\} \longrightarrow \{\textit{paths on } \mathcal{F}\}$$

sends $(M, (W_i))$ to the homotopy class of the path

$$W_1 M^{-1} = W_1 M^{-1} M_1, W_1 M^{-1} M_2, \ldots, W_1 M^{-1} M_k = W_1, W_2, \ldots$$

In plain words, the action concatenates (or deconcatenates) a path representing M to the head of the path (W_i).

Let us denote by $\partial\mathcal{F}$ the set of ends of $\mathcal{F}$. The action of $\mathrm{PSL}_2(\mathbb{Z})$ on $\mathcal{F}$ extends to an action on the set $\partial\mathcal{F}$, the element $M \in \mathrm{PSL}_2(\mathbb{Z})$ sending the ♡ path (W_i) to the path (MW_i). (Note that this is not the $\mathrm{PSL}_2(\mathbb{Z})$-action on the set of paths of $\mathcal{F}$, introduced in the above remark.) This action is neither free nor transitive, see the next section for details.

Given an edge e of $\mathcal{F}$ and an end b of $\mathcal{F}$, there is a unique path in the class b which starts at e. Hence for any edge e, we may identify $\partial\mathcal{F}$ with the set of infinite paths that start at e. We denote this set by $\Pi_1^\infty(\mathcal{F}, e)$ and call it the *fundamental groupoid* [4] *of $\mathcal{F}$ at infinity, based at e.* We endow $\Pi_1^\infty(\mathcal{F}, e)$ with the product topology. Intuitively, closeness of two paths isk determined by the number of common edges.

Given any edge e' of $\mathcal{F}$, the spaces $\Pi_1^\infty(\mathcal{F}, e)$ and $\Pi_1^\infty(\mathcal{F}, e')$ are canonically homeomorphic. This homeomorphism is given by pre-composing with the unique path joining e to e'. The above-mentioned action of the modular group on the set $\partial\mathcal{F}$ induces an action of $\mathrm{PSL}_2(\mathbb{Z})$ by homeomorphisms of the topological space $\Pi_1^\infty(\mathcal{F}, e)$, for any choice of a base edge e.

3.3 Circle and the continued fraction map

The space $\Pi_1^\infty(\mathcal{F}, e)$ is homeomorphic to the Cantor set. We want to "contract the holes" of this Cantor set to obtain the continuum, as follows. Define a *rational end* of $\mathcal{F}$ to be an eventually left-turn or eventually right-turn path.

[4] This set is not a groupoid, but is a part of the fundamental groupoid of the properly defined completion $\bar{\mathcal{F}} = \mathcal{F} \cup \partial\mathcal{F}$

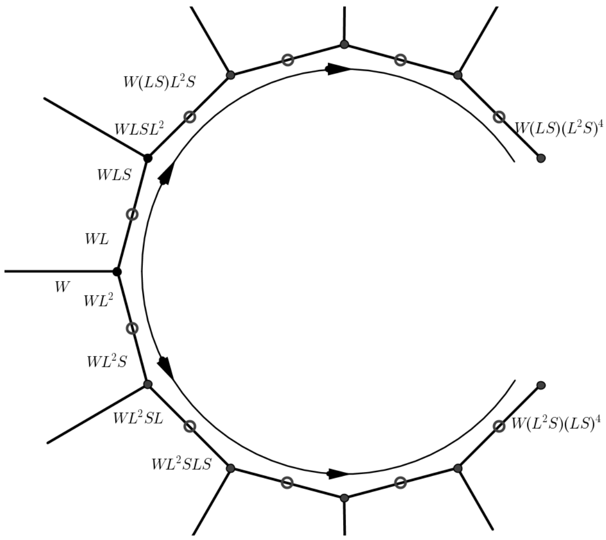


Figure 1: The left and right turn paths which bifurcate from the vertex $\{W, WL, WL^2\}$.

Now introduce the equivalence relation $\sim$ on $\partial\mathcal{F}$ as: left- and right- rational paths which bifurcate from the same vertex are equivalent, see Figure 1.

This equivalence relation can be understood as follows. $\Pi_1^\infty(\mathcal{F}, e)$ is a cyclically ordered topological space in a natural way. Pairs of rational ends are not separated by a third point whereas between any other distinct pairs of points there is always another point. This equivalence relation sets equal points which are not separated by a third point. Recall that, as sets, $\Pi_1^\infty(\mathcal{F}, e)$ and $\partial\mathcal{F}$ are in bijection. On the quotient space $\Pi_1^\infty(\mathcal{F}, e)/\sim$ there is the quotient topology induced by the topology on $\Pi_1^\infty(\mathcal{F}, e)$ such that the projection map

$$\Pi_1^\infty(\mathcal{F}, e) \longrightarrow \Pi_1^\infty(\mathcal{F}, e)/\sim$$

is continuous. We shall denote this quotient space by S^1_e. This equivalence relation is preserved under the canonical homeomorphisms $\Pi_1^\infty(\mathcal{F}, e) \longrightarrow \Pi_1^\infty(\mathcal{F}, e')$ and is also respected by the $\mathrm{PSL}_2(\mathbb{Z})$-action. Therefore we have the commutative diagram

$$\begin{array}{ccc} \Pi_1^\infty(\mathcal{F}, e) & \longrightarrow & \Pi_1^\infty(\mathcal{F}, e') \\ \downarrow & & \downarrow \\ S^1_e & \longrightarrow & S^1_{e'} \end{array}$$

where the horizontal arrows are canonical homeomorphisms and the vertical arrows are projections. Moreover, $\mathrm{PSL}_2(\mathbb{Z})$ acts by homeomorphisms on S^1_e, for any e. ♡

Now, $\mathcal{F}$ comes equipped with a distinguished edge, the one marked with the identity element I of $\mathrm{PSL}_2(\mathbb{Z})$. Hence all spaces S^1_e are canonically homeomorphic to S^1_I.

Any element of S^1_I can be represented by an infinite word in L and S. Regrouping occurrences of LS and L^2S, any such word x of S^1_I can be thus written in one of the following forms:

$$\begin{aligned} x &= (LS)^{n_0}(L^2S)^{n_1}(LS)^{n_2}(L^2S)^{n_3}(LS)^{n_4}\cdots \text{ or} \\ x &= S(LS)^{n_0}(L^2S)^{n_1}(LS)^{n_2}(L^2S)^{n_3}(LS)^{n_4}\cdots, \end{aligned}$$

where $n_0, n_1 \cdots \geq 0$. Since our paths do not have any backtracking we have $n_0 \geq 0$ and $n_i > 0$ for $i = 1, 2, \cdots$. The pairs of words

$$\begin{array}{ll} (LS)^{n_0}\cdots(LS)^{n_k+1}(L^2S)^{\infty} \text{ and } (LS)^{n_0}\cdots(LS)^{n_k}(L^2S)(LS)^{\infty}, & (k \text{ even}) \\ (LS)^{n_0}\cdots(L^2S)^{n_k+1}(LS)^{\infty} \text{ and } (LS)^{n_0}\cdots(L^2S)^{n_k}(LS)(L^2S)^{\infty}, & (k \text{ odd}) \end{array}$$

correspond to pairs of rational ends and represent the same element of S^1_I. For irrational ends this representation is unique.

Set $U(z) = 1/z$. Noting that

$$\begin{aligned} LS(z) = T = 1 + z &\implies (LS)^n(z) = n + z \text{ and} \\ L^2S(z) = 1/(1 + 1/z) = UTU(z) &\implies (L^2S)^n = UT^nU, \end{aligned}$$

we can rewrite the element $x \in S^1_I$ in the form

$$\begin{aligned} x &= T^{n_0}UT^{n_1}UT^{n_2}T^{n_3}UT^{n_4}\cdots \\ x &= ST^{n_0}UT^{n_1}UT^{n_2}T^{n_3}UT^{n_4}\cdots \end{aligned}$$

We shall employ the usual notation for the continued fractions

$$[n_0; n_1, n_2, \dots] := n_0 + \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{\ddots}}}.$$

Define the continued fraction map $\kappa : S^1_I \to \hat{\mathbf{R}}$ by

$$\kappa(x) = \begin{cases} [n_0, n_1, n_2, \dots] & \text{if } x = T^{n_0}UT^{n_1}UT^{n_2}T^{n_3}U\dots \\ -1/[n_0, n_1, n_2, \dots] & \text{if } x = ST^{n_0}UT^{n_1}UT^{n_2}T^{n_3}U\dots \end{cases}$$

Here is the long-promised seamless correspondence with the reals.

Theorem 3.2. *The continued fraction map is a homeomorphism.*

Proof. First note that this map is well defined as it respects the equivalence of pairs of rational ends: $[n_0, \dots n_k + 1, \infty]$ and $[n_0, \dots n_k, 1, \infty]$ represent the same number. Moreover, it is bijective from the set of rational ends modulo equivalence of pairs onto the set of rational numbers. With this identification of pairs of rational ends with $\mathbb{Q}$, every infinite path determines a unique Dedekind cut. To be more precise, given an infinite path $x = T^{n_0}UT^{n_1}UT^{n_2}\cdots \in S^1_I$, the set A is defined to be the set of all rational paths which are to the right of x and to the left of the path $x_{-\infty} = (SR^2)^\infty$, and the set B is defined to be the set of all paths which are to the left of x and to the right of the path $x_\infty = (RS)^\infty$. □

To each pair of rational ends, we associate the $\bullet$-type vertex of $\mathcal{F}$ from which they bifurcate and thus we obtain a canonical identification of $V_\bullet$ with $\mathbb{Q}$.

As a consequence of the above theorem, we see that the continued fraction map conjugates the $\mathrm{PSL}_2(\mathbb{Z})$-action on S^1_I to its action on $\hat{\mathbf{R}}$ by Möbius ♡ transformations and the following result becomes evident:

Theorem 3.3. *For any real $r \in \mathbb{R}$, the orbit $\mathrm{PGL}_2(\mathbb{Z}) \cdot r$ is exactly the set of numbers having a continued fraction expansion with the same tail as r. In other words, the orbits of the $\mathrm{PGL}_2(\mathbb{Z})$ consists of those numbers whose continued fraction expansions coincide at infinity.*

3.4 Periodic paths and the real CM set

Let $\gamma := (W_1, W_2, \dots, W_n)$ be a finite path in F. Then the *periodization* of γ is the path γ^ω defined as

$$W_1, W_2, \dots, W_n, W_nW_1^{-1}W_2, W_nW_1^{-1}W_3, \dots, W_nW_1^{-1}W_n,$$
$$(W_nW_1^{-1})^2W_2, \dots, (W_nW_1^{-1})^2W_n, (W_nW_1^{-1})^3W_2, \dots$$

In plain words, γ^ω is the path obtained by concatenating an infinite number of copies of a path representing $W_1^{-1}W_n$, starting at the edge W_1. If $W_1^{-1}W_n$ is elliptic, then γ^ω is an infinitely backtracking finite path. If not, γ^ω is actually infinite and represents an end of $\mathcal{F}$. We call these *periodic ends of* $\mathcal{F}$. An element M and its positive powers have the same periodization. Thus we have a map

$$\mathit{periodization}: \{\mathit{primitive\ non-elliptic\ elements}\} \longrightarrow \{\mathit{ends\ of\ } \mathcal{F}\}$$

whose image consists of periodic ends. The modular group action on $\partial\mathcal{F}$ preserves the set of periodic ends and the periodization map is $\mathrm{PSL}_2(\mathbb{Z})$-equivariant.

Given an edge e of $\mathcal{F}$ and a periodic end b of $\mathcal{F}$, there is a unique path in the class b which starts at e. In this way the set of periodic ends of $\mathcal{F}$ is identified with the set of eventually periodic paths based at e. This set is dense in $\Pi_1^\infty(\mathcal{F}, e)$ and preserved under the canonical homeomorphisms between the spaces $\Pi_1^\infty(\mathcal{F}, e)$ and $\Pi_1^\infty(\mathcal{F}, e')$. Every periodic end has a unique $\mathrm{PSL}_2(\mathbb{Z})$-translate, which is a purely periodic path based at e. Finally, the set of periodic ends descends to a well-defined subset of S_e^1. The image of this set under the continued fraction map consist of the set of eventually periodic continued fractions, i.e. the set of real quadratic irrationalities (also called the "real CM-set"). The $\mathrm{PSL}_2(\mathbb{Z})$-action preserves the real CM-set, and its $\heartsuit$
orbits consists of the eventually periodic continued fractions with the same period.

Exercise. Show that the $\mathrm{PSL}_2(\mathbb{Z})$-action on S_e^1 is not free and find its fixed points.

4 The projective general linear group over $\mathbb{Z}$

4.1 Anthyphairesis, third version

This is a combination of the previous two versions: cut off from the stick A multiples of the stick B until the remainder C is shorter than B, and repeat this once more if C is greater then B/2. Replace A with B, replace B with C and enter the loop.

If we record how many B's have been cut from A's at each loop, and if the algorithm terminates, we get a finite sequence $[|n_0, \epsilon_0, n_1, \epsilon_1, \ldots n_k|]$ of integers in terms of which we may express the ratio r as a "nearest-integer-" continued fraction

$$r = n_0 + \cfrac{\epsilon_0}{n_1 + \cfrac{\epsilon_1}{n_2 + \cfrac{\epsilon_2}{\cdots + \cfrac{\cdots}{n_{k-1} + \cfrac{\epsilon_{k-1}}{n_k}}}}} \tag{1}$$

Here the numbers $\epsilon_i = \epsilon_i(r) \in \{-1, +1\}$ are determined by the ratio r. In case the lengths of the sticks A and B are not commensurable, the process does not terminate and we have an infinite semiregular continued fraction.

Exercise. What are the numbers r with $\epsilon_i(r) = +1$ (or $\epsilon_i(r) = -1$) for all

i in their nearest continued fraction expansion? What is the measure of this set?

Exercise. Given a real number r, determine the set of reals having the same sign sequence in their nearest continued fraction expansion as r.

Exercise. Does this define a submonoid? i.e. given two nearest continued fractions $[|n_0, \epsilon_0, n_1, \epsilon_1, \ldots n_k|]$ and $[|m_0, \epsilon_0, m_1, \epsilon_1, \ldots m_k|]$, is it true that

$$[|n_0, \epsilon_0, n_1, \epsilon_1, \ldots n_k, 1, m_0, \epsilon_0, m_1, \epsilon_1, \ldots m_k|]$$

is the nearest continued fraction expansion of some number?

4.2 Anthyphairesis, your own version

If you devise a scheme to choose between the first and the second versions at the end of every loop, you will have your own Anthyphairesis and your own semiregular continued fraction expansion of the form (1). This choice amounts to choosing a sequence $(\epsilon_i) \in \{-1, +1\}^\omega$ as a function of the real number r whose semiregular continued fraction expansion is to be defined. Choosing the constant sequence $(+1)^\omega$ yields the classical version ("plus"-continued fractions), whereas the sequence $(-1)^\omega$ yields the "minus"-continued fractions. In the case of closest integer continued fractions the sequence (ϵ_i) is not constant and depends on r.

Here, the $\pm$ signs are to be chosen in a fixed manner according to your scheme. But one may also devise a scheme to choose these signs as a function of r as well. There are many such schemes studied in the literature, such as α-continued fractions Nakada, [30], their generalization known as Japanese continued fractions by Marmi, Moussa and Yoccoz [27] and an even more general one introduced recently by Masarotto [28]. Lehner introduced "the mother of all continued fractions" see [5]. Another scheme has been devised by Zagier and studied by Katok et al. [21]. These semiregular continued fractions are obtained by iterating an appropriate generalization/modification of the Gauss map and studied in the spirit of discrete dynamical systems.

Exercise. Invent your own semiregular continued fraction expansion.

4.3 Anthyphairesis, the most primitive version

Perform the following operations: If A is greater then B, write $A = B + C$; if not; write $A = B - C$. Replace A with B, B with C and enter the loop until no B remains.

Applying this to the ratio $r = [x_0 : x_1]$ gives

$x_0 = x_1 \pm x_2 \implies [x_0 : x_1] = 1 \pm [x_2 : x_1]$
$x_1 = x_2 \pm x_3 \implies [x_1 : x_2] = 1 \pm [x_3 : x_2]$
$x_2 = x_3 \pm x_4 \implies [x_2 : x_3] = 1 \pm [x_4 : x_3]$
.
.
$x_{k-2} = x_{k-1} \pm x_k \implies [x_{k-2} : x_{k-1}] = 1 \pm [x_k : x_{k-1}]$
$x_{k-1} = x_k + 0 \quad (x_{k+1} = 0) \implies [x_{k-1} : x_k] = 1,$

and therefore

$$r = 1 \pm \cfrac{1}{1 \pm \cfrac{1}{\ddots \pm \cfrac{1}{1}}}.$$

We may encode this process by a sequence of signs $[\pm, \pm, \dots]$, which can be compressed by grouping and counting consecutive repetitions of signs. This is the most primitive one among all anthyphairesis processes we encountered. We call this the *janissary*[5] semiregular continued fraction expansion. While implementing this algorithm we see sometimes same numbers coming up twice; but don't worry, the algorithm eventually wins.

Example For $r = 13/32$ we get

$$\begin{array}{lcl} 13 = 32 - 19 & \implies & 13/32 = 1 - 19/32 \\ 32 = 19 + 13 & \implies & 32/19 = 1 + 13/19 \\ 19 = 13 + 6 & \implies & 19/13 = 1 + 6/13 \\ 13 = 6 + 7 & \implies & 13/6 = 1 + 7/6 \\ 6 = 7 - 1 & \implies & 6/7 = 1 - 1/7 \\ 7 = 1 + 6 & \implies & 7/1 = 1 + 6/1 \\ 1 = 6 - 5 & \implies & 1/6 = 1 - 5/6 \\ 6 = 5 + 1 & \implies & 6/5 = 1 + 1/5 \\ 5 = 1 + 4 & \implies & 5/1 = 1 + 4/1 \\ 1 = 4 - 3 & \implies & 1/4 = 1 - 3/4 \\ 4 = 3 + 1 & \implies & 4/3 = 1 + 1/3 \\ 3 = 1 + 2 & \implies & 3/1 = 1 + 2/1 \\ 1 = 2 - 1 & \implies & 1/2 = 1 - 1/2 \\ 2 = 1 + 1 & \implies & 2/1 = 1 + 1 \\ 1 = 1 + 0 & \implies & 1 = 1 + 0 \end{array} \quad \implies \quad \tfrac{13}{32} = 1 - \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 - \cfrac{1}{1 + \cfrac{1}{1 - \cfrac{1}{\cdots}}}}}}}$$

Or, in the language of operators,

$$\frac{13}{32} = TSTUTUTUTSTUTSTUTUTSTUTUTSTUTUT(0).$$

[5] Who walks in their traditional parade by taking two steps forward and one step back.

Alternatively, any such word can be written as a word in $\tilde{T} = TU$ and $L = TS$.

Exercise. Describe the janissary expansion of a general integer.

Exercise. Describe the janissary expansion of the golden ratio $\frac{1+\sqrt{5}}{2}$.

Question. How can one pass from the janissary expansion to the simple continued fraction expansion?

Question. Describe the monoid and the group related to janissary expansions in terms of matrices.

4.4 $\mathrm{PGL}_2(\mathbb{Z})$.

To summearize the situation, we started with the classical case where every real number has a unique regular continued fraction expansion. The picture then was quite clear. But now we have a multitude of continued fraction expansions. In fact, every real number has uncountably many semiregular continued fraction expansions. Is there a way to understand these in a unified manner? Answering this question will require the construction of a "thickening" of the bipartite Farey tree $\mathcal{F}$, on which $\mathrm{PGL}_2(\mathbb{Z})$ acts. It is possible to make this construction in a purely combinatorial manner, as we did with the bipartite Farey tree $\mathcal{F}$. However, since this requires some preparation we prefer to take a shortcut and make use of the $\mathrm{PGL}_2(\mathbb{Z})$ -action on the upper half plane.

Recall that the operator formulation of semiregular expansions makes use of the operators T, U and S. Since T and S already generates the modular group, the result below should not come as a surprise:

Theorem 4.1. *The monoid generated by the operators T, U and S is isomorphic to the projective group of invertible two by two integral matrices:*

$$\mathrm{PGL}_2(\mathbb{Z}) := \left\{ M = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \mid p, q, r, s \in \mathbb{Z};\ \det(M) = \pm 1 \right\} / \pm I.$$

Proof. As before, T is invertible in the monoid since $T^{-1} = SL^2 = STSTS$ and S and T are invertible since they are involutive. Hence they generate a group. The kernel of

$$1 \longrightarrow \mathrm{PSL}_2(\mathbb{Z}) \to \mathrm{PGL}_2(\mathbb{Z}) \xrightarrow{\det} \{\pm 1\}$$

is generated by T and S, and since U is sent to -1, the group generated must be the group $\mathrm{PGL}_2(\mathbb{Z})$. $\square$

The following presentation is well known.

$$PGL_2(\mathbb{Z}) \simeq \langle V, U, K \,|\, V^2 = U^2 = K^2 = (VU)^2 = (KU)^3 = 1\rangle$$

with $V(z) = -z$, $U(z) = \dfrac{1}{z}$, $K(z) = 1 - z$, so that $S = UV$ and $L = KU$. The following amalgamated product description is also well known.

$$PGL_2(\mathbb{Z}) \simeq \Sigma_3 *_{\mathbb{Z}/2\mathbb{Z}} (\mathbb{Z}/2\mathbb{Z})^2$$

In fact, one may express V in terms of U and $T = KV$ as follows:

$$V = T^{-1}UTUT^{-1}U \implies K = TV = UTUT^{-1}U.$$

It turns out that $\mathrm{PGL}_2(\mathbb{Z})$ is generated by T and U. Moreover, K is conjugate to U. Hence if we re-write the above presentation in generators T and U, the relation K^2 becomes redundant and we have the presentation

$$PGL_2(\mathbb{Z}) \simeq \langle T, U |\ \overset{(Rel.1)}{(T^{-1}UTUT^{-1}U)^2} = \overset{(Rel.2)}{U^2} = \overset{(Rel.3)}{(T^{-2}UTU)^2} = \overset{(Rel.4)}{(UTUT^{-1})^3} = 1\rangle.$$

In fact, this presentation is still redundant as one may deduce the first relation from the rest. Indeed $(T^{-2}UTU)^2 = 1 \quad (Rel.3) \implies$

$$T^{-2}UTUT^{-2}UTU = 1 \implies T^{-1}UTUT^{-1} = TUT^{-1}UT$$

Multiplying the latter equality by $UT^{-1}UTUT^{-1}U$ from the left, we get

$$UT^{-1}UTUT^{-1}U \cdot T^{-1}UTUT^{-1} = UT^{-1}UTUT^{-1}U \cdot TUT^{-1}UT$$

$$\implies U(T^{-1}UTUT^{-1}U)^2U = UT^{-1}(UTUT^{-1})^2UT =$$
$$UT^{-1}(UTUT^{-1})^3TU = 1 \quad (Rel.4)$$

which implies the first relation. Hence the presentation

$$\mathrm{PGL}_2(\mathbb{Z}) \simeq \langle T, U |\, U^2 = (UTUT^{-2})^2 = (UTUT^{-1})^3 = 1\rangle.$$

5 Action on the upper half plane $\mathcal{H}$

A discrete cocompact subgroup Λ in $\mathbb{C}$ is called a *planar lattice*. The quotient $\mathbb{C}/\Lambda$ is the 2-torus with a complex structure (i.e. an elliptic curve). A *framed planar lattice* is a triple $(\Lambda; \lambda_1, \lambda_2)$, where Λ is a lattice and (λ_1, λ_2) is an ordered basis for Λ such that $\mathrm{Im}(\lambda_2/\lambda_1) > 0$. The framed lattice $(\Lambda; \lambda_1, \lambda_2)$ is isomorphic to the framed lattice $(\Lambda; \omega_1, \omega_2)$ if and only if $\lambda_2/\lambda_1 = \omega_2/\omega_1$.

Consequently, every framed lattice is isomorphic to a unique framed lattice of the form $(\mathbb{Z} + \mathbb{Z}\tau; 1, \tau)$, where τ lies in the complex upper half plane $\mathcal{H} := \{z \mid \operatorname{Im} z > 0\}$.

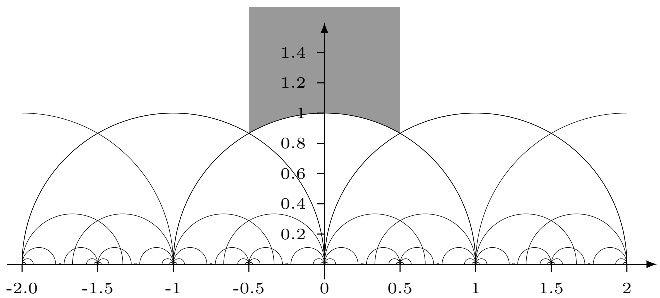
The group $\mathrm{SL}_2(\mathbb{Z})$ acts on the set of framings of a lattice by ♡

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} : (\Lambda; \lambda_1, \lambda_2) \longrightarrow (\Lambda; r\lambda_2 + s\lambda_1, p\lambda_2 + q\lambda_1),$$

thus it acts on the set of parameters $\tau = \lambda_2/\lambda_1 \in \mathcal{H}$ by

$$\begin{pmatrix} p & q \\ r & s \end{pmatrix} : \tau \longrightarrow \frac{p\tau + q}{r\tau + s}.$$

The orbits of this action are identified with the isomorphism classes of elliptic curves, see [17].



Noting that the subgroup $\langle \pm I \rangle$ of $\mathrm{SL}_2(\mathbb{Z})$ acts trivially on $\mathcal{H}$, we see that $\mathrm{PSL}_2(\mathbb{Z})$ acts from the left on $\mathcal{H}$ by linear fractional (or Möbius) transformations. ♡ This action is by conformal isomorphisms and it also respects the hyperbolic metric

$$ds^2 := \frac{dx^2 + dy^2}{y^2}.$$

This action is discontinuos on $\mathcal{H}$. It is not free since it stabilizes the points $S(i) = i$ and $L(\omega) = \omega$ where $\omega = \frac{1+\sqrt{-3}}{2}$. Hence for any $M \in \mathrm{PSL}_2(\mathbb{Z})$, the element MSM^{-1} stabilizes $M(i)$ and MLM^{-1} stabilizes $M(\omega)$. If we remove the $\mathrm{PSL}_2(\mathbb{Z})$-orbits of i and ω from $\mathcal{H}$, the action becomes free and its quotient space is a Riemann surface with a canonical hyperbolic metric. It is well known that it is the twice-punctured plane $\mathbb{C}\backslash\{0, 1\}$. If we insist on keeping the points with finite stabilizers then the quotient $\mathcal{M} := \mathrm{PSL}_2(\mathbb{Z})\backslash\mathcal{H}$ is an orbifold, called the *modular curve*. It has two orbifold points with local groups $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$. The modular curve $\mathcal{M}$ admits the structure of a canonical hyperbolic cone-manifold. It is the moduli space of elliptic curves.

Consider the circular arc in the boundary of the standard fundamental region, connecting the point i to the point ω. The $\mathrm{PSL}_2(\mathbb{Z})$-orbit of this arc is a topological realization of the bipartite Farey tree, to be denoted by $\mathcal{F}_{top}$. The modular group acts on $\mathcal{F}_{top}$ by construction, and its quotient space is a one-dimensional orbifold, which we call the *modular arc*. The modular arc embeds as a one-dimensional sub-orbifold of the modular curve, as the arc connecting its two orbifold points.

Elements of $\mathrm{PGL}_2(\mathbb{Z})$ of determinant -1 exchange the lower and upper half planes, so this group does not act on the upper half plane. Nevertheless, we may represent it as a group of self-maps of the upper half plane, by interpreting the elements of determinant -1 as antiholomorphisms, as follows:

$$M = \begin{bmatrix} p & q \\ r & s \end{bmatrix} \rightsquigarrow \begin{cases} \frac{pz+q}{rz+s} & \text{if } \det(M) = 1 \\ \\ \frac{p\bar{z}+q}{r\bar{z}+s} & \text{if } \det(M) = -1 \end{cases}$$

This representation is called the *extended modular group* and denoted by $\mathrm{PSL}^*(\mathbb{Z})$ (see [24]). Thus we have the representation:

$$\mathrm{PSL}_2^*(\mathbb{Z}) \simeq \left\langle V(z) = -\bar{z},\, U(z) = \frac{1}{\bar{z}},\, K(z) = 1 - \bar{z} \right\rangle$$

where U, V, K are reflections. For example U fixes the unit half-circle in $\mathcal{H}$. Alternatively, one has the representation

$$\mathrm{PSL}_2^*(\mathbb{Z}) \simeq \left\langle T(z) = z + 1,\, U(z) = \frac{1}{\bar{z}} \right\rangle$$

Now PSL_2^* is a group generated by reflections and as a fundamental region for $\mathrm{PSL}_2^*(\mathbb{Z})$ we can choose the half of the standard fundamental region for $\mathrm{PSL}_2(\mathbb{Z})$; namely the region

$$\Delta := \left\{ z \in \mathcal{H} \,:\, 0 \leq \Re(z) \leq \frac{1}{2}, \quad 1 \leq |z| \right\}.$$

The quotient $\mathcal{H}/{\mathrm{PSL}_2}^*(\mathbb{Z})$ is an orbifold in the sense of Kato [18]. We call this orbifold *modular tile* and denote it by $\mathcal{M}^*$. It is the quotient of the modular curve $\mathcal{M}$ by a reflection that fixes pointwise the three arcs through the elliptic points and the infinity. Thus $\mathcal{M}$ is a double covering of $\mathcal{M}^*$, corresponding to the subgroup $\mathrm{PSL}_2(\mathbb{Z}) \subset \mathrm{PSL}_2^*(\mathbb{Z})$. Under its canonical hyperbolic metric, $\mathcal{M}^*$ is a triangle with interior angles $\pi/2, \pi/6, 0$. The fundamental group with $\pi_1(\mathcal{M}^*) \simeq \mathrm{PSL}_2^*(\mathbb{Z})$ is generated by reflections of this triangle along its sides.

We name the edges and vertices of Δ as follows:

$$V : \left\{ z \in \mathcal{H} \,:\, \Re(z) = 0, \quad \Im(z) > \frac{1}{2} \right\}$$

$$U : \left\{ z \in \mathcal{H} \, : \, 0 \le \Re(z) \le \frac{1}{2}, \quad |z| = 1 \right\}$$

$$K : \left\{ z \in \mathcal{H} \, : \, \Re(z) = 0, \quad \Im(z) > \frac{\sqrt{3}}{2} \right\}$$

$$S : (0,1), \quad \{L, R\} : \left(\frac{1}{2}, \frac{\sqrt{3}}{2} \right).$$

Translates of Δ tessellate the upper half plane into triangular cells, each cell bounded by three arcs marked U, V and K and corners marked by S and $\{L, R\}$ (third corners lie at infinity). We denote this tessellation by $\mathrm{PSL}_2(\mathbb{Z})\Delta$.

We say that two cells are *neighbors* if they have a non-empty intersection. A *path* in $\mathrm{PSL}_2(\mathbb{Z})\Delta$ is a sequence of neighboring cells.

Any infinite path starting at Δ can be encoded as a (non-reduced) word in the letters S, U, V, K, L, R by registering the labels of the intersections of cells. It determines an infinite semiregular continued fraction, if we respectively substitute $-1/z$, $1/z$, $-z$, $1-z$, $1-1/z$ and $1/1-z$ for the letters S, U, V, K, L and R. Conversely, every semiregular continued fraction with partial quotients in $\mathbb{Z}$ can be expressed (not uniquely) as a word in S, U, V and K. However, such a path may come back to the neighborhood of a cell infinitely often. If this does not happen then we say that the path, or the corresponding semiregular continued fraction, is *non-wandering*.

Now the realization of the bipartite Farey tree $\mathcal{F}$ sits inside the tessellation $\mathrm{PSL}_2(\mathbb{Z})\Delta$ as the union of the orbits of the modular arc connecting two elliptic points on the boundary of Δ, and every non-wandering path determines a unique infinite path on $\mathcal{F}$, starting on the edge marked I. As we have seen, every path can be encoded as a word in the letters S and L. This infinite path is not back-tracking free, meaning that the encoding infinite word may not be reduced. Thanks to the non-wandering property, these back-tracks can be eliminated to yield a unique, well-defined infinite path on $\mathcal{F}$ without back-tracking. We say that two paths in the tessellation are *homotopic* if they give rise to the same path on $\mathcal{F}$ in this way. Since the latter represent the real numbers we obtain the following tautological result:

Theorem 5.1. *Two non-wandering semiregular continued fractions converge to the same real number if and only if they are represented by homotopic paths of the tessellation. Wandering semiregular continued fractions do not converge.*

Thus semiregular continued fraction expansion schemes can be seen as ways of choosing a representative from the homotopy class of paths representing real numbers.

(In order to get a somewhat clearer picture of the paths in the tessellation, fix some $\epsilon \geq 0$ and consider the $\mathrm{PSL}_2^*(\mathbb{Z})$-orbit of the cell

$$I_\epsilon := \left\{ z \in \mathcal{H} \,:\, 0 \leq \Re z \leq \frac{1}{2}, \quad 1 \leq |z| \leq 1 + \epsilon \right\} \subseteq \Delta.$$

We denote the orbit by $\mathcal{F}_\epsilon$. Evidently, $\mathcal{F}_0$ is the bipartite Farey tree $\mathcal{F}$ (or rather its realization in $\mathcal{H}$) and $\mathcal{F}_\infty$ is the tessellation $\mathrm{PSL}_2(\mathbb{Z})\Delta$ of the upper half plane. One has $I_\infty = \Delta$. For a small $\epsilon > 0$, the set $\mathcal{F}_\epsilon$ is a thin strip around $\mathcal{F}$ and admits by construction a faithful $\mathrm{PSL}_2^*(\mathbb{Z})$-action. Orbits of the above strip, which we call *cells*, induce a tessellation of $\mathcal{F}_\epsilon$. Then a path in the tessellation $\mathrm{PSL}_2(\mathbb{Z})\Delta$ can be viewed as a sequence of neighboring cells in $\mathcal{F}_\epsilon$.)

6 Modular graphs and dessins

Let Γ be any subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. Then Γ acts on $\mathcal{F}$ from the left and to Γ we associate a quotient graph $\Gamma\backslash\mathcal{F}$ whose edges are the Γ-orbits of edges of $\mathcal{F}$; equivalently, they are the right-cosets of Γ. Vertices of $\Gamma\backslash\mathcal{F}$ are the Γ-orbits of vertices of $\mathcal{F}$. To be more precise, one has

$$\begin{aligned} E(\Gamma\backslash\mathcal{F}) &= \{\Gamma\cdot W : W \in \mathrm{PSL}_2(\mathbb{Z})\} \\ V(\Gamma\backslash\mathcal{F}) &= V_\otimes(\mathcal{F}\backslash\Gamma) \cup V_\bullet(\mathcal{F}\backslash\Gamma), \end{aligned}$$

where

$$\begin{aligned} V_\otimes(\Gamma\backslash\mathcal{F}) &= \{\Gamma\cdot\{W, WS\} : W \in \mathrm{PSL}_2(\mathbb{Z})\} \\ V_\bullet(\Gamma\backslash\mathcal{F}) &= \{\Gamma\cdot\{W, WL, WL^2\} : W \in \mathrm{PSL}_2(\mathbb{Z})\}. \end{aligned}$$

The edge connecting two distinct vertices v and v' of $\Gamma\backslash\mathcal{F}$ is the set

$$v \cap_\Gamma v' := \bigcup_{\nu \in v, \nu' \in v'} \nu \cap \nu'$$

which is a right-coset of Γ if non-empty. There are no other edges. This incidence relation induced from $\mathcal{F}$ gives a well-defined incidence relation and we obtain a bipartite graph with a special edge, i.e. the one labeled with the coset Γ.

Definition 6.1. *Let* Γ *be* any *subgroup of the modular group. The graph* $\Gamma\backslash\mathcal{F}$ *is called a* ***modular graph****.*

The edges incident to the vertex $\Gamma\{W, WL, WL^2\}$ are ΓW, ΓWL, ΓWL^2, and these edges inherit a natural cyclic ordering $(\Gamma W, \Gamma WL, \Gamma WL^2)$ from the vertex. Hence $\Gamma\backslash\mathcal{F}$ is a ribbon graph.

Since the set of edges of $\Gamma\backslash\mathcal{F}$ is identified with the set of right-cosets of Γ, the graph $\Gamma\backslash\mathcal{F}$ has $[\mathrm{PSL}_2(\mathbb{Z}) : \Gamma]$ many edges. For instance, for $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, the quotient graph $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathcal{F}$ is the *modular arc* with one edge that looks as in Figure 2. Every modular graphs is a covering (in the appropriate sense) of the modular arc. $\mathcal{F}$ is its universal cover.

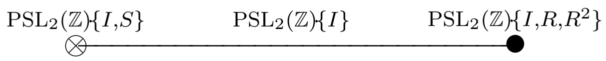


Figure 2: The modular arc.

In general $\Gamma\backslash\mathcal{F}$ is a bipartite ribbon graph possibly with one pending vertex for each conjugacy class of an elliptic element of Γ. Conversely, any connected bipartite ribbon graph G, with $V(G) = V_{\otimes}(G) \sqcup V_{\bullet}(G)$, such that every $\otimes$-vertex is of degree 1 or 2 and every $\bullet$-vertex is of degree 1 or 3, is modular since the universal covering of G is isomorphic to $\mathcal{F}$. With some effort one can define the fundamental group of $\Gamma\backslash\mathcal{F}$ (as in [25]) so that there is a canonical isomorphism $\pi_1(\Gamma\backslash\mathcal{F}, \Gamma) \simeq \Gamma < \mathrm{PSL}_2(\mathbb{Z})$, with the canonical choice of Γ as a base edge.

Theorem 6.2. *Subgroups* Γ *of the modular group (or equivalently the fundamental groups* $\pi_1(\Gamma\backslash\mathcal{F})$*) are free products of copies of* $\mathbb{Z}$, $\mathbb{Z}/2\mathbb{Z}$ *and* $\mathbb{Z}/3Z$*.*

For a proof, see [24]. Note that two distinct isomorphic subgroups Γ_1, Γ_2 of the modular group may give rise to non-isomorphic ribbon graphs $\Gamma_1\backslash\mathcal{F}$ and $\Gamma_2\backslash\mathcal{F}$. We shall see shortly that $\mathbb{Z}$-subgroups generated by a hyperbolic element give examples of this phenomena. In other words, the fundamental group does not characterize the graph.

In topology, there is a well-known correspondence between subgroups of the fundamental group of a space and the coverings of that space. The following two results are orbifold (or "stacky") analogues of this correspondence for coverings of the modular curve, stated in terms of graphs. For more details on fundamental groups and covering theory of graphs see [25].

Proposition 6.1. *If* Γ_1 *and* Γ_2 *are conjugate subgroups of* $\mathrm{PSL}_2(\mathbb{Z})$*, then the graphs* $\Gamma_1\backslash\mathcal{F}$ *and* $\Gamma_2\backslash\mathcal{F}$ *are isomorphic as ribbon graphs. Hence there is a 1-1 correspondence between modular graphs and conjugacy classes of subgroups of the modular group.*

Proof. Let $\Gamma_2 = M\Gamma_1 M^{-1}$. The desired isomorphism is then the map

$$\begin{aligned}\varphi : E(\Gamma_1\backslash\mathcal{F}) &\to E(\Gamma_2\backslash\mathcal{F})\\ \Gamma_1 W &\mapsto \Gamma_2 MW.\end{aligned}$$

Note that one has $\varphi(\Gamma_1\cdot\{I\}) = \Gamma_2 M$. Suppose now that $\varphi : E(\Gamma_1\backslash\mathcal{F}) \to E(\Gamma_2\backslash\mathcal{F})$ is a ribbon graph isomorphism and let $\varphi(\Gamma_1 I) = \Gamma_2 M$. This induce an isomorphism of fundamental groups

$$\varphi_* : \pi_1(\Gamma_1\backslash\mathcal{F}, \Gamma_1) \simeq \pi_1(\Gamma_2\backslash\mathcal{F}, \Gamma_2 M).$$

Since φ is a ribbon graph isomorphism, these two groups are also isomorphic as subgroups of the modular group. The former group is canonically isomorphic to Γ_1 a whereas the latter group is canonically isomorphic to

$$M^{-1}\pi_1(\Gamma_2\backslash\mathcal{F}, \Gamma_2)M \simeq M^{-1}\Gamma_2 M.$$

Therefore modular graphs parametrize conjugacy classes of subgroups of the modular group, whereas the edges of a modular graph parametrize subgroups in the conjugacy class represented by the modular graph. In conclusion we get:

Theorem 6.3. *There is a 1-1 correspondence between modular graphs with a base edge (G, e) (modulo ribbon graph isomorphisms of pairs (G, e)) and subgroups of the modular group.*

As a simple consequence we see that there exist uncountably many subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ as there exist uncountably many modular graphs.

Theorem 6.4. *There is a 1-1 correspondence between modular graphs with two base edges (G, e, e') (modulo ribbon graph isomorphisms of pairs (G, e, e')) and cosets of subgroups of the modular group.*

From the perspective of group actions, we see that the conjugation action of $\mathrm{PSL}_2(\mathbb{Z})$ on its subgroups translates into an action on the pairs (G, e). This action fixes the graph G and moves the edge e, where S moves e to its $\otimes$-negihbor and R rotates it around its $\bullet$-neighbor. If we consider this action simultaneously on the set of edges of G, we obtain the *monodromy action* of $\mathrm{PSL}_2(\mathbb{Z})$ corresponding to G. It is determined by two permutations, one of order 2 (corresponding to S) and one of order 3 (corresponding to R). Since G is connected, this action is transitive on $E(G)$. ♡

6.1 Triangulations and lozenges

Besides the fundamental group, another basic invariant of $\Gamma\backslash\mathcal{F}$ is its genus, which is defined to be the genus of the Riemann surface $\mathcal{H}/\Gamma$.

Every modular graph $\Gamma\backslash\mathcal{F}$ has a canonical piecewise analytical realization $\Gamma\backslash\mathcal{F}_{top}$ on the Riemann surface $\Gamma\backslash\mathcal{H}$ with edges being geodesic arcs. Equivalently, these edges are lifts of the modular arc by $\Gamma\backslash\mathcal{H} \longrightarrow \mathrm{PSL}_2(\mathbb{Z})\backslash\mathcal{H}$. If instead we lift the geodesic arc connecting the $\otimes$- elliptic point to the cusp to the surface $\Gamma\backslash\mathcal{H}$, then we obtain another graph on the surface, which is called an *ideal triangulation*. Lifting the remaining geodesic arc gives rise to yet another type of graph, called a *lozenge tiling*. So there is a triality of these graphs.

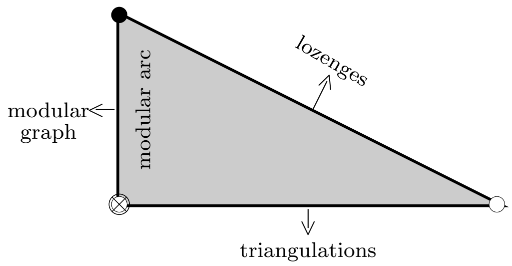


Figure 3: The modular curve. It consists of two triangles, the second is on the back of the page, glued to this one.

Conversely, given a triangulation of an oriented surface, we construct the "trial" modular graph by placing $\bullet$-vertices on the centers of the triangles, $\otimes$-vertices on the midpoints of the arcs of the triangulation, and by connecting all $\bullet$-vertices to neighboring $\otimes$-vertices. Hence combinatorial triangulations (including degenerate ones) of oriented topological 2-manifolds parametrize conjugacy classes of subgroups of the modular group. A choice of an oriented arc of a triangulation determines a coherent choice of an edge trial modular graph, i.e. a subgroup in the conjugacy class determined by the triangulation. In other words triangulations with an oriented arc parametrize subgroups of the modular group. The conjugation action of $\mathrm{PSL}_2(\mathbb{Z})$ on its subgroups translates into its action on triangulations which moves the oriented arc. ♡

Exercise. Determine the lozenge tilings of surfaces corresponding to subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ and describe the action of $\mathrm{PSL}_2(\mathbb{Z})$ on them.

6.2 Dessins

The Riemann surface underlying the modular curve $\mathrm{PSL}_2(\mathbb{Z})\backslash\mathcal{H}$ is the Riemann sphere with one puncture, and can be identified with $\mathbb{C}$. Up to an affine transformation, one may assume that the orbifold points $\otimes$ and $\bullet$ are respectively the points 0 and 1. For an arbitrary subgroup Γ of $\mathrm{PSL}_2(\mathbb{Z})$, the quotients $\Gamma\backslash\mathcal{H}$ will be a covering of $\mathbb{C}$ branched at $0, 1$ and ∞. The branching behavior cannot be arbitrary, i.e. if we set

$$\beta_\Gamma : \Gamma\backslash\mathcal{H} \longrightarrow \mathrm{PSL}_2(\mathbb{Z})\backslash\mathcal{H}$$

then locally around the origin, β_Γ will behave like z or z^2 and around the point 1, it will behave like z or z^3. There is no restriction on branching at infinity however.

In general, a finite covering $\beta : R \to \mathbb{P}^1(\mathbb{C})$ of the Riemann sphere which is branched at most at 0, 1 and ∞ is called a *Belyi morphism*. The celebrated theorem of Belyi states that a Riemann surface R can be defined over a number field if and only if there exists a Belyi morphism on R. A pair (R, β) of a Riemann surface with a Belyi morphism defined on it, is accordingly called a *Belyi pair*. Two Belyi pairs (R_1, β_1) and (R_2, β_2) are considered equivalent, if there exists an isomorphism $f : R_1 \to R_2$ with $\beta_2 \circ f = \beta_1$.

To the Belyi pair (R, β) we associate a bipartite ribbon graph $G = \beta^{-1}(\mathcal{I})$ on R, which is defined to be the lift of the arc $\mathcal{I}$ connecting 0 and 1 in $\mathbb{P}^1(\mathbb{C})$. The ribbon structure is induced from the orientation of R. This graph has the property that $R - G$ consists of a disjoint union of simply connected discs. In general, a pair (R, G) of an oriented 2-manifold with a finite graph $G \subset R$ satisfying this property is called a *dessin*.

We have the following proposition borrowed from [11]:

Proposition 6.2. *A Belyi pair (R, β) is up to equivalence uniquely determined by:*

- *a dessin up to equivalence;*
- *a bipartite connected ribbon graph up to equivalence;*
- *a finite index subgroup of the rank-2 free group $\mathbb{F}_2$ up to conjugation;*
- *a monodromy map $\alpha : \mathbb{F}_2 \to \Sigma_d$, i.e. a transitive action of $\mathbb{F}_2$ on the set $1, ..., d$ up to conjugation in the symmetric group Σ_d.*

The outer automorphism group of $\mathbb{F}_2$ being $GL(2, \mathbb{Z})$, we see that there is an action of the modular group on dessins, see the forthcoming thesis [32] ♡ for some details. This action preserves the monodromy group.

Since there are no restrictions on their vertex degrees whatsoever, one may conclude that dessins are generalizations of finite modular graphs. However, the situation is more intricate than this. First note that any dessin G on R

determines a triangulation of R, by connecting the midpoints of the discs in $R - G$ to the vertices of G. Alternatively, this is the triangulation obtained by lifting to R the triangulation of $\mathbb{P}^1(\mathbb{C})$ which consists of two triangles with vertices at 0, 1 and ∞. As we saw in the preceding subsection, the graph "trial" to this triangulation is modular and determines a subgroup of the modular group. The point is that the quotient of the upper half plane by the congruence subgroup of level two is precisely the thrice-punctured sphere:

$$\Gamma(2)\backslash\mathcal{H} \simeq \mathbb{P}^1(\mathbb{C})\backslash\{0, 1, \infty\}.$$

The automorphism group of the surface $\mathbb{P}^1(\mathbb{C})\backslash\{0, 1, \infty\}$ is the group $\Sigma_3 \simeq \mathrm{PSL}_2(\mathbb{Z})/\Gamma(2)$, and thus we have a branched Galois covering

$$\mathbb{P}^1(\mathbb{C})\backslash\{0, 1, \infty\} \longrightarrow \mathrm{PSL}_2(\mathbb{Z})\backslash\mathcal{H}$$

of degree 6. Yet another way of saying this is

$$\pi_1(\mathbb{P}^1(\mathbb{C})\backslash\{0, 1, \infty\}) \simeq \Gamma(2) \simeq \mathbb{F}_2.$$

Hence, in a sense, every dessin is already determined by some modular graph.

Note that the free group $\mathbb{F}_2$ appears also as the derived subgroup of the modular group, its quotient surface being a once-punctured torus. Subgroups of $\mathbb{F}_2$ determines coverings of this torus, branched at one point. One may consider the quadrangulated surfaces obtained by lifting the standard quadrangulation of the torus (with only one quadrangle) to these coverings. Being the group of outer automorphisms of $\mathbb{F}_2$, the group $GL(2, \mathbb{Z})$ acts on these quadrangulated coverings as well. ♡

6.3 Categories of coverings of the modular curve

The category (directed poset) of coverings of the modular arc with a base edge is equivalent to the category of subgroups of the modular group with inclusions as morphisms. We denote the former system with $\mathbf{Cov}^*(\otimes\!\!-\!\!\bullet)$ and the latter with $\mathbf{Sub}(\mathrm{PSL}_2(\mathbb{Z}))$. Since intersection of subgroups is a subgroup, both systems are directed. The category $\mathbf{Cov}^*(\otimes\!\!-\!\!\bullet)$ is the category of modular graphs with a base edge, where morphisms are coverings of modular graphs with a base edge. Forgetting base edges yields the covering category of the modular arc, denoted as $\mathbf{Cov}(\otimes\!\!-\!\!\bullet)$. The three categories have uncountably many objects as there exists uncountably many modular graphs.

The category $\mathbf{Cov}^*(\otimes\!\!-\!\!\bullet)$ has the full sub-category $\mathbf{FCov}^*(\otimes\!\!-\!\!\bullet)$, which consists of base-edged coverings of finite degree, equivalent to the category of finite-index subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ under inclusion. We denote the latter category by $\mathbf{FSub}(\mathrm{PSL}_2(\mathbb{Z}))$. Since the intersection of subgroups of finite index is a subgroup of finite index, both systems are directed. The category $\mathbf{FCov}^*(\otimes\!\!-\!\!\bullet)$ is countable, since its objects are finite modular graphs with a base edge. Forgetting base edges yields the category finite coverings of the modular arc, denoted as $\mathbf{FCov}(\otimes\!\!-\!\!\bullet)$.

The inverse limit of the system of ambient pointed topological surfaces $\Gamma\backslash\mathcal{H}$, where Γ is an object of $\mathbf{FSub}(\mathrm{PSL}_2(\mathbb{Z}))$, was studied under the name *punctured (non-compact) solenoid* by Penner. By virtue of Belyi's theorem, this limit can be viewed as a kind of universal arithmetic curve. The limit of the realizations of the category of finite modular graphs $\mathbf{FCov}^*(\otimes\!\!-\!\!\bullet)$ is a compact subspace of the punctured solenoid, which we call the *ribbon solenoid*. In other words, this is the limit

$$\varprojlim \Gamma\backslash\mathcal{F}_{top} \subset \varprojlim \Gamma\backslash\mathcal{H}.$$

Let $\widehat{\mathrm{PSL}_2(\mathbb{Z})}$ be the profinite completion of the modular group. Then $\mathrm{PSL}_2(\mathbb{Z})$ acts continuously on $\widehat{\mathrm{PSL}_2(\mathbb{Z})}$ by multiplication. Hence the $\mathrm{PSL}_2(\mathbb{Z})$-action ♡
$(z,t) \in \mathcal{H} \times \widehat{\mathrm{PSL}_2(\mathbb{Z})} \to (Mz, tM^{-1})$ is also continuous. Therefore any sub- ♡
group Γ of the modular group acts continuously on $\mathcal{H} \times \widehat{\mathrm{PSL}_2(\mathbb{Z})}$, and the quotient space is homeomorphic to the punctured solenoid if Γ is of finite index, see [33].

6.4 Çarks

This section is borrowed from [31]. Let M be a hyperbolic element of the modular group. Then it generates a subgroup $\langle M\rangle \simeq \mathbb{Z}$. A *çark* is by definition a modular graph of the form $\c{C}_M := \langle M\rangle\backslash\mathcal{F}$ where

$$\pi_1(\langle M\rangle\backslash\mathcal{F}) = \langle M\rangle \simeq \mathbb{Z},$$

so the çark $\langle W\rangle\backslash\mathcal{F}$ is a graph with only one circuit, which we call the *spine* of the çark. Every çark has a canonical realization as a graph $\langle M\rangle\backslash\mathcal{F}_{top}$ embedded in the surface $\langle M\rangle\backslash\mathcal{H}$, which is an annulus since M is hyperbolic. In fact $\langle M\rangle\backslash\mathcal{H}$ is the annular uniformization of the modular curve $\mathcal{M}$ corresponding to $M \in \pi_1(\mathcal{M})$. Again by hyperbolicity of M, this graph will have infinite "Farey branches" attached to the spine in the direction of both boundary components of the annulus.

If M is parabolic, then $\langle W\rangle\backslash\mathcal{F}$ has Farey branches attached to the spine in only one direction, and its topological realization $\langle M\rangle\backslash\mathcal{F}_{top}$ sits on a punctured disc. If M is elliptic, $\langle W\rangle\backslash\mathcal{F}$ is a tree with a pending edge which abuts

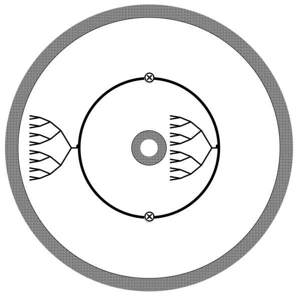


Figure 4: The çark $\mathcal{F}/\langle SR^2SR\rangle$.

at a vertex of type $\otimes$ when M is of order 2 and of type $\bullet$ when M is of order 3. Its topological realization $\langle M\rangle\backslash\mathcal{F}_{top}$ sits on a disc with an orbifold point.

By Proposition 6.1 the graphs $Ç_M$ and $Ç_{XMX^{-1}}$ are isomorphic for every element X of the modular group and from Theorem 6.3 we deduce the following result, see [7]:

Corollary 6.1. *There are bijective correspondences between the sets*

$$\left\{ \text{ çarks } \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{conjugacy classes of } \mathbb{Z}\textit{-subgroups in } \mathrm{PSL}_2(\mathbb{Z}) \\ \textit{generated by a single hyperbolic element} \end{array} \right\}.$$

$$\left\{ \begin{array}{c} \textit{çarks} \\ \textit{with a base edge} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{subgroups } \langle M\rangle \textit{ of } \mathrm{PSL}_2(\mathbb{Z}) \\ \textit{generated by a single hyperbolic element} \end{array} \right\}$$

The subgroup $\langle M\rangle$ is also generated by M^{-1} and the çark cannot distinguish beyween these two elements. If we specify a direction of its spine then we can distinguish one among the generators $\{M, M^{-1}\}$ of this subgroup. Such çarks are said to be *directed.*

Corollary 6.2. *There are bijective correspondences between the sets*

$$\left\{ \text{ directed çarks } \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{conjugacy classes of} \\ \textit{hyperbolic elements of } \mathrm{PSL}_2(\mathbb{Z}) \end{array} \right\}$$

$$\left\{ \begin{array}{c} \textit{directed çarks} \\ \textit{with a base edge} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{hyperbolic elements} \\ \textit{of } \mathrm{PSL}_2(\mathbb{Z}) \end{array} \right\}.$$

6.5 Braceletes and necklaces

A çark can be encoded as follows: First remove all $\otimes$-vertices of the çark. Next, turn once around the spine. Upon meeting a $\bullet$-vertex on which a branch is attached by R, cut that branch and tag that $\bullet$-vertex with a "0". In a similar fashion, upon meeting a $\bullet$-vertex on which a branch attached by R^2, cut that branch and tag that $\bullet$-vertex with a "1". We obtain a finite graph called a *(binary) bracelet* which is by definition an equivalence class of binary strings under cyclic permutations (i.e. rotations) and reversals. Conversely, by using the convention $0 \leftrightarrow R$ and $1 \leftrightarrow R^2$ we can reconstruct the çark from its bracelet. Bracelets are the orbits of the dihedral group action on the set of words in $0, 1$, where the action is by cyclic permutations and reversals of words.

An orbit of the set of words in 0,1 under cyclic permutations is called a *(binary) necklace*. These encode directed çarks. Choosing the ordering $0 < 1$ and imposing the lexicographic ordering of the words, one may choose a minimal representative in each orbit. The minimal representative of a primitive (aperiodic) word is called a *Lyndon word*. They appear in numerous contexts. In our case they are

$$0, 1, 01, 001, 011, 0001, 0011, 0111, 00001, 00011, 00101, 00111, 01011, 01111 \ldots$$

7 The action on binary quadratic forms

A homogeneous function of degree two in two variables is called a *binary quadratic form*. A such form $f(x, y) = Ax^2 + Bxy + Cy^2$ is said to be *indefinite* if its discriminant $\Delta(f) = B^2 - 4AC$ is positive. If A, B, C are integers then f is called *integral* and if $\gcd(A, B, C) = 1$ then it is called *primitive*.

The group $\mathrm{PSL}_2(\mathbb{Z})$ acts on the set of integral binary quadratic forms by ♡

$$\begin{aligned} \mathit{Forms} \times \mathrm{PSL}_2(\mathbb{Z}) &\to \mathit{Forms} \\ (f, U) &\mapsto U \cdot f := f(U(x, y)) \end{aligned}$$

leaving the discriminant invariant. Two forms are said to be *equivalent* if they belong to the same $\mathrm{PSL}_2(\mathbb{Z})$-orbit. The $\mathrm{PSL}_2(\mathbb{Z})$-orbit of f is denoted $[f]$. The stabilizer of f is called its *automorphism group*, denoted by $\mathrm{Aut}(f)$. If $\Delta(f) < 0$ one has $\mathrm{Aut}(f) \simeq \mathbb{Z}$, and the two generators of this group are called the *fundamental automorphisms* of f.

To any given hyperbolic element, say $M = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \mathrm{PSL}_2(\mathbb{Z})$ we associate the following primitive and indefinite binary quadratic form:

$$f_M = \frac{\operatorname{sgn}(p + s)}{\gcd(q, s - p, r)} \big(r, s - p, -q\big) \tag{2}$$

Elaborating on the maps $\langle M\rangle\backslash\mathcal{F} \longleftrightarrow M \longrightarrow f_M$ the following result is shown [34], [31]:

Proposition 7.1. *There are natural bijections between*

$$\left\{\begin{matrix}\textit{primitive}\\ \textit{directed çarks}\end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix}\textit{classes of primitive, indefinite}\\ \textit{binary quadratic forms}\end{matrix}\right\}$$

$$\left\{\begin{matrix}\textit{primitive directed çarks}\\ \textit{with a base edge}\end{matrix}\right\}. \longleftrightarrow \left\{\begin{matrix}\textit{primitive indefinite}\\ \textit{binary quadratic forms}\end{matrix}\right\}$$

8 Et cetera

There are actions of $\mathrm{PSL}_2(\mathbb{Z})$ by conjugation and by multiplication on various groups containing the modular group: Thompson's group, $\mathrm{SL}_2(\mathbb{R})$, $\mathrm{SL}_n(\mathbb{Z})$, $\mathrm{Sp}_n(\mathbb{Z})$, etc, which we had no time to consider here. We conclude with the following faithful representation which stems from the action of the group $\mathrm{S}L(2,\mathbb{R})$ on its lie algebra $\mathrm{sl}(2,\mathbb{R})$: ♡ ♡

$$\begin{bmatrix} a & b\\ c & d\end{bmatrix} \longrightarrow \begin{bmatrix} a^2 & 2ac & c^2\\ ab & ad+bc & cd\\ b^2 & 2bd & d^2\end{bmatrix}$$

9 Appendix. Gosper's Algorithm for Continued Fraction Arithmetic

Ralph William (Bill) Gosper is known for a few algorithms, including the hashlife algorithm to accelerate the cellular automata computations of Game of Life type, and another one to find closed form identities for hypergeometric identities. In [10], he described an algorithm to perform arithmetic operations directly on continued fraction representations of numbers. This was quite surprising, because it was believed that in Khinchin's words "even the problem of finding the continued fraction for a sum from the continued fractions representing the addends is ... unworkable in computational practice" [22]. To this day, this algorithm is not quite well-known so we include it in the paper.

9.1 Univariate Case

Before looking at arithmetic operations with two operands, let's investigate the univariate case, where we want to compute $\frac{a+bx}{a'+b'x}$ with $a, a', b, b' \in \mathbb{Z}$ being some integer constants and x the number in continued fraction form on which we want to perform arithmetic operations, like calculating the multiplicative or additive inverse of x or a linear function of it without resorting to convert to decimal form. For this case it's convenient to think of these constants as a 2×2 matrix like $\begin{bmatrix} a & b \\ a' & b' \end{bmatrix}$. It follows that mentioned arithmetic operations can be represented as specific matrices with $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ corresponding to additive inverse which is $-x$, and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ corresponding to multiplicative inverse $\frac{1}{x}$ to name a few.

Gosper's algorithm is a stateful algorithm, such that as the computation progress the values in the matrix change as explained below and act as the state of the algorithm like some auxiliary variables for intermediate values.

At each moment Gosper's algorithm proceeds with one of two operations, which are "pulling" the following c.f. term from the variable (x in the representation above) or "emitting" the next c.f. term of the computation. Which of these two operations to perform is decided by comparing the integral part of the ratios $\frac{a}{a'}$ and $\frac{b}{b'}$ for equality. If $\lfloor \frac{a}{a'} \rfloor = \lfloor \frac{b}{b'} \rfloor$ then an "emit" operation is performed, if not then one more c.f. term is "pulled" from variable till the equality holds.

The rationale for this decision procedure is simple; for $x = 0$ the rational function $\frac{a+bx}{a'+b'x}$ takes the value $\frac{a}{a'}$ and for $x = \pm\infty, \lim_{x \to \pm\infty} \frac{a+bx}{a'+b'x} = \frac{b}{b'}$, therefore when the integral part of these two ratios agree the continued fraction representation of the rational function is guaranteed to follow by this integral term, irrespective of the value of the variable.

Now that we have a decision procedure to decide which operation to perform at any point, first we define the "emit" operation where we receive the next term of the rational function and update the constants accordingly. In the case $\lfloor \frac{a}{a'} \rfloor = \lfloor \frac{b}{b'} \rfloor = n, n \in \mathbb{Z}$ will be the emitted next term of computed c.f. and the matrix values representing the state will change as $\frac{a+bx}{a'+b'x} = n + \frac{1}{(\frac{\alpha+\beta x}{\alpha'+\beta' x})}$. To determine α, α', β and β' (which are the new values of a, a', b, b' respectively) we further write

$$n + \frac{1}{(\frac{\alpha+\beta x}{\alpha'+\beta' x})} = n + \frac{\alpha' + \beta' x}{\alpha + \beta x} = \frac{n(\alpha + \beta x) + \alpha' + \beta' x}{\alpha + \beta x}$$

$$= \frac{(n\alpha + \alpha') + (n\beta + \beta')x}{\alpha + \beta x};$$

thus $\frac{a+bx}{a'+b'x} = \frac{(n\alpha+\alpha')+(n\beta+\beta')x}{\alpha+\beta x}$ and $\alpha = a'$, $\beta = b'$, $(n\alpha+\alpha') = a \Rightarrow \alpha' = a - na'$, $(n\beta + \beta') = b \Rightarrow \beta' = b - nb'$.

Therefore we can write the change in matrix representation for the "emit" operation as $\begin{bmatrix} a & b \\ a' & b' \end{bmatrix} \to \begin{bmatrix} a' & b' \\ a - na' & b - nb' \end{bmatrix}$.

When the integral part of the ratios does'nt agree we perform the second operation, which "pulls" the next c.f. term from the variable and change the matrix with this extra information. Let n be the next c.f. term of x so we can write $x = n + \frac{1}{x'}$ where x' is the remaining c.f. tail of x; then we can rewrite $\frac{a+bx}{a'+b'x}$ as

$$\frac{a + b(n + \frac{1}{x'})}{a' + b'(n + \frac{1}{x'})} = \frac{a + bn + \frac{b}{x'}}{a' + b'n + \frac{b'}{x'}} = \frac{ax' + bnx' + b}{a'x' + b'nx' + b'} = \frac{b + (a + bn)x'}{b' + (a' + b'n)x'}$$

providing us with the operation's transformation, on matrix representation which can also be written as $\begin{bmatrix} a & b \\ a' & b' \end{bmatrix} \to \begin{bmatrix} b & a + bn \\ b' & a' + b'n \end{bmatrix}$. The decision procedure to select which operation to perform, ensures that the minimum number of terms are used from the variable; on the other hand "pulling" terms can be applied anytime, independently of whether the integral part of the ratios agree or not, as what this operation provides is to transform the rational function to a better rational approximation by changing the terms of function with larger ones based on the information obtained from the term read from variable.

In order to perform operations on continued fractions with finite length, the algorithm can be applied with one extra term of ∞ added to the end of variable's c.f. expansion. In this way when this infinity term is read the matrix transforms to $\begin{bmatrix} a & b \\ a' & b' \end{bmatrix} \to \begin{bmatrix} b & b \\ b' & b' \end{bmatrix}$ which represents $\frac{b}{b'}$ as the last term of the algorithm's output which can be converted to c.f. representation on its own.

9.2 Bivariate Case

The algorithm for the univariate case can be extended to permit arithmetic operations over two operands represented in c.f. form. As in the univariate case, we setup a rational function but this time with variables x and y for the two operands. Let $\frac{a+bx+cy+dxy}{a'+b'x+c'y+d'xy}$ represent the arithmetic operation we want to carry and constants $a, a', b, b', c, c', d, d' \in \mathbb{Z}$. As in the previous case, we can identify this rational function with a matrix of form $\begin{bmatrix} a & b & c & d \\ a' & b' & c' & d' \end{bmatrix}$.

Just like in the univariate case, common arithmetic operations can be represented with specific matrices, like addition being represented as $\left[\begin{smallmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{smallmatrix}\right]$ which corresponds to $\frac{x+y}{1}$, subtraction represented as $\left[\begin{smallmatrix} 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{smallmatrix}\right]$ or $\left[\begin{smallmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{smallmatrix}\right]$

which corresponds to $\frac{x-y}{1}$ and $\frac{-y+x}{1}$ respectively, multiplication as $\left[\begin{smallmatrix}0 & 0 & 0 & 1\\ 1 & 0 & 0 & 0\end{smallmatrix}\right]$ which corresponds to $\frac{xy}{1}$, and division as $\left[\begin{smallmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\end{smallmatrix}\right]$ or $\left[\begin{smallmatrix}0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\end{smallmatrix}\right]$ which corresponds to $\frac{x}{y}$ and $\frac{y}{x}$ respectively. Obviously more complicated linear functions or polynomials can easily be represented with a suitable choice of starting constants.

For the bivariate case deciding whether to "emit" a new term for the result or "pull" another term from one of the variables is similar to the univariate case, such that we consider all the extremal values $\langle x, y\rangle$ pair can attain, which consist of $\langle x = 0, y = 0\rangle, \langle x = \pm\infty, y = 0\rangle, \langle x = 0, y = \pm\infty\rangle$ and $\langle x = \pm\infty, y = \pm\infty\rangle$. As before, if the integral value of the rational function agrees on all four cases, then this value can be emitted as the next c.f. term of the result and the matrix can be updated to reflect the decrease due to the term being emitted. For $\langle x = 0, y = 0\rangle$ the value of the rational function becomes $\frac{a}{a'}$ and for the remaining cases the rational function takes the values of $\frac{b}{b'}, \frac{c}{c'}$ and $\frac{d}{d'}$ respectively; therefore the decision procedure for choosing which operation to perform reduces to check whether the equality $\lfloor\frac{a}{a'}\rfloor = \lfloor\frac{b}{b'}\rfloor = \lfloor\frac{c}{c'}\rfloor = \lfloor\frac{d}{d'}\rfloor$ holds, in which case we can "emit" a new term.

Let's name the new values of $a, a', b, b', c, c', d, d'$ after the extraction of an integral value n from the rational function as $p, p', q, q', r, r', s, s'$ respectively; then we can write

$$\frac{a + bx + cy + dxy}{a' + b'x + c'y + d'xy} = n + \frac{1}{\left(\frac{p+qx+ry+sxy}{p'+q'x+r'y+s'xy}\right)}$$

$$= \frac{n(p + qx + ry + sxy) + p' + q'x + r'y + s'xy}{p + qx + ry + sxy}$$

$$= \frac{(np + p') + (nq + q')x + (nr + r')y + (ns + s')xy}{p + qx + ry + sxy}.$$

This equation allow us to express the new matrix values in terms of old ones, such that $p = a', q = b', r = c', s = d'$ and furthermore $a = np + p' \Rightarrow p' = a - np = a - na'$ and $q' = b - nb', r' = c - nc', s = d - nd'$ in a similar fashion. With new matrix values expressed in terms of old ones, we can write the transformation on a matrix resulting from emitting a c.f. term n as

$$\begin{bmatrix} a & b & c & d\\ a' & b' & c' & d'\end{bmatrix} \to \begin{bmatrix} a' & b' & c' & d'\\ a - na' & b - nb' & c - nc' & d - nd'\end{bmatrix}$$

When the integral part of the four ratios do not agree, new terms must be read from the variables to transform the function into a better rational approximation of what remains from the variables.

References

[1] G. Baumslag, J.W. Morgan, P.B. Shalen, *Generalized triangle groups*, Math. Proc. Camb. Phil. Soc. (1987) **102**, 25.

[2] J. Buchmann and U. Vollmer, Binary quadratic forms, Algorithms and Computation in Mathematics, vol. 20, Springer, Berlin, 2007.

[3] C. Marston, G. Havas, and M. F. Newman, *On one-relator quotients of the modular group.*, Groups St Andrews 2009 in Bath 1 (2011): 183.

[4] J. H. Conway, (with the assistance of Francis Y. C. Fung), The sensual (quadratic) form, Carus Mathematical Monographs, vol. 26, Mathematical Association of America, Washington, DC, 1997.

[5] K. Dajani, C. Kraaikamp, *The mother of all continued fractions*, Colloq. Math. 84/85 (2000), part. 1, 109123.

[6] K. Dajani, C. Kraaikamp, *Ergodic Theory of numbers*, Carus Mathematical Monoraph, 29. Mathematical Association of America, ISBN: 0-88385-034-6.

[7] M. Durmuş, *Farey Graph and Binary Quadratic Forms*, Master's thesis, İstanbul Technical University, Turkey, 2012.

[8] J.-P. Duval, *Generation d'une section des classes de conjugaison et arbre des mots de Lyndon de longueur borné,*, Theoret. Comput. Sci. 60 (1988) 255-283.

[9] C. F. Gauss, Disquisitiones arithmeticae, Translated into English by Arthur A. Clarke, S. J, Yale University Press, New Haven, Conn., 1966.

[10] R. W. Gosper, *Continued fraction arithmetic,* HAKMEM Item 101B, MIT Artificial Intelligence Memo, (1972), 239.

[11] F. Herrlich and G. Schmithusen, *Dessins d'enfants and origami curves* in: Handbook of Teichmller Theory Edited by Athanase Papadopoulos (2009).

[12] I. Kapovich, P.E. Schupp, *Random quotients of the modular group are rigid and essentially incompressible*, Journal fr die reine und angewandte Mathematik (2009).

[13] G.A. Miller, *On the groups generated by two operators*, Bull. Amer. Math. Soc. 7 (1901), no. 10, 424426

[14] *OEIS (The on-line encyclopedia of integer sequences)*, published electronically at http://oeis.org/.

[15] Y.T. Ulutaş, Y., I.N. Cangül, *One Relator Quotients of the Modular Group*, Bull. Inst. Math. Academia Sinicia 32 (2004): 291-296.

[16] E. Girondo, G. Gabino, Introduction to compact Riemann surfaces and dessins d'enfants, Vol. 79. Cambridge University Press, 2011.

[17] R. Hain, *Lectures on Moduli Spaces of Elliptic Curves*, arXiv: 0812.1803v2 [math.AG]

[18] M. Kato, *On uniformizations of surfaces.* Homotopy Theory and Related Topics (Kyoto 1984) (Advanced Studies in Pure Mathematics, 9). North-Holland, Amsterdam, 1987, pp. 149172.

[19] S. Katok and I. Ugarcovici, *Symbolic dynamics for the modular surface and beyond* , Bull. Amer. Math. Soc. (N.S.) **44** (2007), no. 1, 87–132.

[20] Katok, Svetlana, and Ilie Ugarcovici. *Structure of attractors for (a, b)-continued fraction transformations.* arXiv preprint arXiv:1004.4200 (2010).

[21] S. Katok, *Coding Of Closed Geodesics After Gauss And Morse*, Geometriae Dedicata 63.2 (1996): 123-145

[22] A. Ya. Khincin, *Continued Fractions*, University of Chicago Press, 1964.

[23] F. Klein, *Über die transformation elfter ordnung der elliptischen functionen*, Mathematische Annalen **15** (1879), no. 3-4, 533–555.

[24] R. S. Kulkarni, *An arithmetic-geometric method in the study of the subgroups of the modular group*, Amer. J. Math. **113** (1991), no. 6, 1053–1133.

[25] J.H. Kwak, R. Nedela, *Graphs and their Coverings*, http://www.savbb.sk/ nedela/graphcov.pdf

[26] S. Lando and A. Zvonkin, *Graphs on surfaces and their applications*, Encyclopaedia of Mathematical Sciences, Springer, 2004.

[27] S. Marmi, P. Moussa, J-C. Yoccoz, *The Brjuno functions and their regularity properties*, Comm. Math. Phys. 186. (1997), no. 2, 265293.

[28] *Valentina Masarotto Metric and arithmetic properties of a new class of continued fraction expansions*, Master Thesis, Universita di Padova

[29] P. A. MacMahon, *Applications of a Theory of Permutations in Circular Procession to the Theory of Numbers*, Proc. London Math. Soc. **S1-23**, no. 1, 305.

[30] H. Nakada, *Metrical theory for a class of continued fractions transformations*, Tokyo J. math 4 (1981), 399-426.

[31] A. M. Uludağ, A. Zeytin, M. Durmuş, *Binary quadratic forms as dessins*, (preprint)

[32] F. Yaşar, *Grothendieck's dessin theory*, on-going Master thesisi Koç University

[33] Penner, Robert C. *Decorated Teichmüller Theory*, Vol. 1. European Mathematical Society, 2012.

[34] P. Sarnak, *Reciprocal geodesics*, Analytic number theory, Clay Math. Proc., vol. 7, Amer. Math. Soc., Providence, RI, 2007, pp. 217237

[35] J. Wiegold, *The Schur multiplier: an elementary approach*, Groups St Andrews 1981, London Math. Soc. Lecture Note Ser. 71 (Cambridge University Press, 1982), 137 154.