LECTURES ON SHIMURA CURVES 2: GENERAL THEORY OF FUCHSIAN GROUPS

PETE L. CLARK

1. Prologue: GL_N, PGL_N, SL_N, PSL_N

Let R be any commutative ring. Then by $GL_N(R)$ we mean the group of all $N \times N$ matrices M with entries in R, and which are invertible: $\det(M) \in R^\times$. The determininant map gives a homomorphism of groups which is easily seen to be surjective: defining the kernel to be $SL_N(R)$, we get an extension of groups

The center Z of $GL_N(R)$ consists of the scalar matrices $R^\times \cdot I_N$, so as a group isomorphic to $R^\times$. By definition, $PGL_N(R) = GL_N(R)/Z$. The determinant map factors through to give a surjective homomorphism $PGL_N(R) \to R^\times/\det(Z) = R^\times/R^{\times 2}$, whose kernel we call $PSL_N(R)$. We summarize the situation with the following diagram:

$$1 \to \mu_N(R) \to Z \cong R^\times \to R^{\times N} \to 1.$$
$$1 \to SL_N(R) \to GL_N(R) \to R^\times \to 1.$$
$$1 \to PSL_N(R) \to PGL_N(R) \to R^\times/R^{\times N} \to 1,$$

where $\mu_N(R) = R^\times[N]$ are the elements r of R such that $r^N = 1$.

Now matter what R is, $PSL_N(R)$ is an interesting group: when $R = \mathbb{F}_q$ is a finite field of cardinality at least 4, $PSL_N(\mathbb{F}_q)$ is a finite simple group. When $R = \mathbb{R}$ or $\mathbb{C}$, we get a simple Lie group. (Better yet: take $R = \mathbb{Z},\ \mathbb{Z}[\sqrt{2}],\ \mathbb{Z}[\sqrt{-1}], \ldots$)

Remarks: When $K = \mathbb{C}$, then (since every complex number is an Nth power), we have $PSL_N(\mathbb{C}) = PGL_N(\mathbb{C})$. When $K = \mathbb{R}$ and N is odd, we have $SL_N(\mathbb{R}) = PSL_N(\mathbb{R}) = PGL_N(\mathbb{R})$. When N is even, $PSL_N(\mathbb{R}) = SL_N(\mathbb{R})/\{\pm 1\}$ has index 2 in $PGL_N(\mathbb{R})$.

We now specialize to the case $R = K$ is a field. Then $PGL_N(K)$ acts by automorphisms on $\mathbb{P}^{N-1}_{/K}$. Namely, viewing this projective space as the space of elements $[x_1, \ldots, x_N] \in \mathbb{A}^N_{/K}$ not all of whose coordinates are zero modulo scalars, there is an evident action of $GL_N(K)$ which descends to $PGL_N(K)$ (since scalar matrices do not change the equivalence class). A basic result is that $PGL_N(K)$ is in fact the full group of automorphisms of $\mathbb{P}^{N-1}/K$ as an algebraic variety.

When $N = 2$ we are asking for the automorphism group of the rational function field $K(z)$, and we can give an alternate description of this as the group of functions $\frac{az+b}{cz+d}$, where $ad - bc \neq 0$ (the group law is composition of functions), which is evidently isomorphic to $PGL_2(K)$.

So in particular $PSL_2(\mathbb{C}) = PGL_2(\mathbb{C})$ acts on $\mathbb{CP}^1$ ("the Riemann sphere") by linear fractional transformations, and the analogous statement for $\mathbb{R}$ would seem to be that $PGL_(\mathbb{R})$ acts on $\mathbb{RP}^1$, which we may view as the equator of the sphere. However, for any complex matrix $g = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL_2(\mathbb{C})$, we have

$$\det(g) \cdot \Im(z) = \Im(g(z)) \cdot |cz + d|^2,$$

so that (as we saw before), it happens to be the case that elements of $PGL_2^+(\mathbb{R}) = PSL_2(\mathbb{R})$ preserve $\mathcal{H}$.

In other words, the theory of subgroups of $PSL_2(\mathbb{R})$ acting on $\mathcal{H}$ is a construction of mixed algebraic and analytic character. It seems useful to keep this in mind, especially to appreciate later analogies with the p-adic situation.

2. Introduction: Ode to the upper halfplane

A Fuchsian group is a discrete group Γ of holomorphic transformations of $\mathcal{H}$. Thus $\Gamma\backslash\mathcal{H}$ is a Riemann surface. If Γ is "of the first kind," then $\Gamma\backslash\mathcal{H}$ can be given the structure of a complex algebraic curve (which may or may not be complete).

Before we get into the study of Fuchsian groups, we should reflect upon the rôle of the upper half plane, an object which – for once! – admits a definition which is perhaps too elementary to reveal its true importance.

I. Recall that $\mathcal{H} = GL_2(\mathbb{R})/\mathbb{C}^\times = SL_2(\mathbb{R})/O(2) = PSL_2(\mathbb{R})/SO(2)$. This is much more structural definition: $SL_2(\mathbb{R})$ is a real Lie group, and $SO(2) \cong S^1$ is a maximal compact subgroup. In fact it is the unique maximal compact subgroup up to conjugacy: it is the point stabilizer of $i \in \mathcal{H}$, and its conjugates are the stabilizers of other points. Thus $\mathcal{H}$ can be viewed the parameter space of maximal compact subgroups of $SL_2(\mathbb{R})$.

In fact this is a very general phenomenon: if G is any connected real Lie group, it admits a unique up to conjugacy self-normalizing maximal compact subgroup K. The quotient space G/K, which parameterizes the conjugacy class of K, is *homeomorphic* to $\mathbb{R}^n$ for some n.

Example: Take $G = SL_n(\mathbb{R})$. Then a maximal compact subgroup is $SO(n)$ – the set of matrices A such that $A^tA = I_n$ of determinant 1. The quotient space is of dimension $n^2 - 1 - (\frac{n(n-1)}{2}) = \frac{n^2+n-2}{2}$. Note that this quantity is even when $n \equiv 1,\ 2 \pmod 4$ but is odd when $n \equiv 0,\ 3 \pmod 4$.

Recall that the Lie group $PSL_2(\mathbb{R})$ admits a bi-invariant Riemannian metric $\mathfrak{g}$. There is a natural identification of $PSL_2(\mathbb{R})$ with the unit tangent bundle to $\mathcal{H}$, and in this way $\mathfrak{g}$ defines a Riemannian metric on $\mathcal{H}$. Explicitly, we can take

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

Under this metric, $\mathcal{H}$ becomes a complete Riemannian surface of constant negative curvature, a **hyperbolic plane** $\mathbb{H}^2$.

Theorem 1. *$PSL_2(\mathbb{R}) = \mathrm{Isom}^+(\mathcal{H}, ds)$ is the group of orientation-preserving isometries of $\mathcal{H}$.*

On the other hand, $PSL_2(\mathbb{R})$ acts on $\mathcal{H}$ by holomorphic automorphisms, and we also have

Theorem 2. *$PSL_2(\mathbb{R}) = \mathrm{Aut}_{\mathbb{C}}(\mathcal{H})$ is the group of automorphisms of $\mathcal{H}$ as a complex manifold.*

There is a general philosophy (going back to Klein) that one should regard two geometric structures as equivalent if their automorphism groups are the same. In our day we recognize this as a principle of descent. In any case, it is a remarkable property of $\mathcal{H}$ that its holomorphic and geometric automorphism groups are equivalent.

Moreover, the quotient of a Lie group by a maximal compact subgroup, $\mathcal{H}$ is simply connected (and indeed contractible). Recall that this almost determines its complex structure uniquely, by the following important result.

Theorem 3. *(Uniformization Theorem) Let Y be a connected Riemann surface.*
a) Its universal cover $\tilde{Y}$ is naturally a Riemann surface.
b) We have $Y = \Pi\backslash\tilde{Y}$, where Π is a discrete group of fixed-point free holomorphic automorphisms of $\tilde{Y}$, isomorphic to the fundamental group $\pi_1(Y)$.
c) There are, up to isomorphism, only three possibilities for $\tilde{Y}$: $\mathbb{CP}^1$, $\mathbb{C}$ and $\mathcal{H}$.

Remark: The first two parts of the theorem are rather straightforward (given the basics of covering space theory); the content resides in the the third part.

It seems enlightening to compare $\mathcal{H}$ to $\mathbb{CP}^1$ and $\mathbb{C}$.

$\mathbb{CP}^1$ does not uniformize any Riemann surface other than itself.

Exercise 2.1: Prove it. (I can think of at least four proofs: (i) using the fact that if $Y \rightarrow X$ is a finite unramified covering map of degree n, then $\chi(X) = \chi(Y)/n$. (ii) Using the Riemann-Hurwitz formula. (iii) Using the Lefschetz trace formula. (iv) Using Luroth's theorem on subfields of rational function fields.)

The holomorphic automorphism group of $\mathbb{CP}^1$ is $PGL_2(\mathbb{C}) = PSL_2(\mathbb{C})$, which acts on $\mathbb{CP}^1$ by linear fractional transformations with complex coefficients. On the other hand, $\mathbb{CP}^1$ admits a natural Riemannian metric ds (the one it inherits from viewing it as the unit 2-sphere in $\mathbb{R}^3$) with constant positive curvature and $\mathrm{Isom}^+(\mathbb{CP}^1, ds) = SO(3)$. There is a natural inclusion

$$SO(3) \hookrightarrow PSL_2(\mathbb{C})$$

but the latter group (a noncompact 6 dimensional real manifold) is very much larger than the former (a compact 3 dimensional real manifold). Thus not every holomorphic map preserves the geometric structure. (On the other hand, $PSL_2(\mathbb{C})$ is the group of orientation preserving isometries of $\mathbb{H}^3$, hyperbolic three space.)

On the other hand, $\mathbb{CP}^1$ is far from contractible, so is not such a promising candidate for uniformization.

$\mathbb{C}$ on the other hand is contractible. It also has the structure of a Lie group, so that the orbit of 0 under any discrete subgroup of holomorphic automorphisms of $\mathbb{C}$ is actually a discrete subgroup H of $\mathbb{C}$, so $H \cong \mathbb{Z}^d$ with $0 \leq d \leq 2$. We have seen the case $d = 2$ before: H is a lattice, and we can uniformize all elliptic curves in this way. If H has rank 1, then $\mathbb{C}/H$ is isomorphic, as a complex Lie group, to $\mathbb{C}^\times$, or if you like, to a genus zero curve with two points removed. If H has rank zero, then $\mathbb{C}/H = \mathbb{C} = \mathbb{A}^1$, a genus zero curve with one point removed.

The group of holomorphic automorphisms of $\mathbb{C}$ is $\mathbb{C} \rtimes \mathbb{C}^\times$, the semidirect product of the translations by the complex linear maps. This is a nonabelian, but solvable, subgroup of the full Mobius group $PSL_2(\mathbb{C})$. On the other hand $\mathbb{C}$ has the Euclidean metric (zero curvature!), and its group of orientation preserving isometries is $\mathbb{C} \rtimes SO(2)$. Observe that this is still smaller than the full group of holomorphic automorphisms: homotheties do not preserve the metric.

Exercise 2.2: Notice that in all three cases we at least had an inclusion $\mathrm{Isom}^+(\tilde{Y}, ds) \hookrightarrow Aut_{\mathbb{C}}(\tilde{Y})$. We could have shown this directly; how?

It follows (process of elimination!) that every compact Riemann surface of genus at least 2 arises as the quotient of $\mathcal{H}$ by a discrete subgroup Γ which admits a presentation

$$\Pi(g, 0) = \langle \alpha_1, \ldots, \alpha_g, \beta_1, \ldots, \beta_g \mid \prod_{i=1}^{g} [\alpha_i, \beta_i] = 1 \rangle.$$

Now suppose that Y is a Riemann surface obtained from a compact Riemann surface of genus $g \geq 0$ by removing $n \geq 0$ points. (Equivalently, $Y = C(\mathbb{C})$ for a possibly affine algebraic curve C.)

Exercise 2.3:
a) Suppose $n > 0$. Show (or recall) that the fundamental group of Y is isomorphic to

$$\Pi(g, n) = \langle \alpha_1, \ldots, \alpha_g, \beta_1, \ldots, \beta_g, \gamma_1, \ldots, \gamma_n \mid \prod_{i=1}^{g} [\alpha_i, \beta_i] \cdot \prod_{j=1}^{n} \gamma_n = 1 \rangle,$$

which is itself isomorphic to a free group on $2g + n - 1$ generators.
b) Deduce that the universal cover of Y is $\mathbb{P}^1\mathbb{C}$, $\mathbb{A}^1\mathcal{C}$ or $\mathcal{H}$ according to whether the Euler characteristic $\chi(Y)$ is positive, zero, or negative.

The notion of getting a Riemann surface from a Fuchsian group Γ by the construction $\Gamma \backslash \mathcal{H}$ is, however, even more general than the above construction indicates. For instance, we saw at the beginning of the course that $PSL_2(\mathbb{Z}) \backslash \mathcal{H} \cong \mathbb{A}^1$; here the quotient map $J : \mathcal{H} \to \mathbb{A}^1$ is clearly not a uniformization map in the above sense. As we embark upon a a general study of Fuchsian groups and then try to wend our way back to moduli of elliptic curves (and other things), this is a good example to keep in mind: exactly what is preventing J from being a uniformization map, and what is the modular interpretation of this?

Problem 2.1: Can one, in fact, obtain every complex algebraic curve as $\Gamma \backslash \mathcal{H}$ for a suitable Fuchsian group Γ?

5groups (the triangle groups) whose quotient is $\mathbb{P}^1$. If you

3. Foundations of Fuchsian Groups

3.1. **I. Topology.** Definition: Let G be any group acting on a topological space X.[1] The action is said to be **freely discontinuous** if each point $x \in X$ has a neighborhood U such that the translates $\{gU\}_{g\in G}$ are pairwise disjoint.

Under such a circumstance, the map $q : X \to G\backslash X$ is a (Galois) covering map. In particular, q is a local homeomorphism, so that given a sheaf of functions $\mathcal{F}$ on X, the pushed forward sheaf $q_*\mathcal{F}$ on $G\backslash X$ has stalk at $q(P)$ isomorphic to the stalk of $\mathcal{F}$ at $P \in X$. (This is a fancy way of saying that any additional local structure which X may have is inherited by $G\backslash X$.)

This is a very nice state of affairs but is obviously too specialized for applications: essentially, we must allow ramified coverings. The following definition identifies a reasonable group action.

Definition: An action G on a space X is **discontinuous at x** $\in X$ if there exists a neighborhood U of x such that the set $\{g \in G \mid gU \cap U \neq \emptyset\}$ is finite. An action is **discontinuous** if it is discontinuous at every point of X.

There are several other natural definitions of properly discontinuous group actions in the literature, and it is natural to wonder whether your favorite definition is equivalent to the given one. We record some of the more useful equivalences.

Proposition 4. *Let G be a locally compact group acting on a locally compact metrizable space X. TFAE:*
a) G acts discontinuously.
b) For all x in X, the orbit Gx is discrete, and the stabilizer G_x is finite.
c) Given any compact subset $K \subset X$ and any $x \in X$, the set of $G(x,K) = \{g \in G \mid gx \in K\}$ is finite.

Proof: $c) \implies b)$: If for some $x \in X$, Gx were not discrete, then there would exist a sequence of distinct elements $g_n \in G$ and $y \in X$ such that $g_n x \to y$. But then for every neighborhood U of Y, we'd have $G(x,U)$ is infinite. Since by assumption y has some neighborhood with compact closure, this contradicts condition c). Moreover, since $\{x\}$ is compact, G_x is finite.
$b) \implies c)$: If for some x and K, $G(x,K)$ were infinite, then (using the assumption of finiteness of the stabilizers), there exists a sequence of distinct elements g_n of g such that $g_n x$ has an accumulation point, contradicting the discreteness of the orbit.
$b) \implies a)$: For any point $x \in X$, there exists a ball $B = B_\epsilon(x)$ such that $gx \in B \implies gx = x$. Now let $B_1 = B_{\epsilon/2}(x)$. Then $gB_1 \cap B_1 \neq \emptyset \implies gx = x$, and by assumption this is true for only finitely many g.

[1] It shall go without (further) comment that all such actions are continuous.

a) $\implies$ *b)*: Suppose that for some x, Gx has a limit point y. Then any neighborhood of y will meet infinitely many of its images under elements of g, a contradiction. Finally, discontinuity at x clearly implies the finiteness of the stabilizer G_x.

Remark: The group Homeo(X) of self-homeomorphisms of X has a natural (compact-open) topology. It is quite clear that a group $G \subset$ Homeo(X) which acts discontinuously on X is discrete: otherwise, there would exist a sequence of distinct elements g_n of G converging to the identity, so that for every point x in X, x is an accumulation point of the orbit Gx.

One should ask: if $G \subset$ Homeo(X) is a discrete group, must it act discontinuously on X? The answer in general is negative.

Exercise 2.4: Let $G \subset PSL_2(\mathbb{R})$ be subgroup with at least one infinite orbit Gz. Show that G does not act properly discontinuously on all of $\mathcal{C}\mathbb{P}^1$.

The following definition refines these considerations.

Let H be a subgroup of $PSL_2(\mathbb{R})$. A point $z \in \mathcal{C}\mathbb{P}^1$ is said to be a **limit point** of H if there exists $w \in \mathcal{C}\mathbb{P}^1$ and a sequence h_n of distinct elements of H such that $h_n w \to z$. The **limit set** $\Lambda(H) \subset \mathcal{C}\mathbb{P}^1$ is the set of all limit points.

Exercise 2.5: Let $H \subset_2 (\mathbb{R})$ and $z \in \mathbb{C}\mathbb{P}^1$. Show that $z \in \Lambda(H) \iff H$ is *not* discontinuous at z.

In light of this discussion, the following result is therefore a bit surprising.

Theorem 5. *Let Γ be a subgroup of $PSL_2(\mathbb{R})$. The following are equivalent:*
a) Γ is discrete.
b) Γ acts discontinuously on $\mathcal{H}$.
c) $\Lambda(\Gamma)$ is a closed subset of $\mathbb{R}\mathbb{P}^1$.
A subgroup satisfying these equivalent conditions is called ***Fuchsian****.*

Proof: From the above discussion, it is clear that each of the other conditions implies the discreteness of Γ (because, if Γ were not discrete, its limit set would be all of $\mathbb{C}\mathbb{P}^1$). We shall see in the next section that if Γ is discrete, its point stabilizers are finite cyclic groups. Thus all that needs to be shown is that if Γ is discrete then for all $z \in \mathcal{H}$, Gz does not accumulate in $\mathcal{H}$ (although it may well have accumulation points on the boundary). Also *b)* $\iff$ *c)* follows from the preceding exercise.

a) $\implies$ *b)*: By Proposition 4, it is enough to show that for every compact subset $K \subset \mathcal{H}$ and every $z \in \mathcal{H}$, $\Gamma(z, K)$ is finite. Since $\Gamma(z, K) = \Gamma \cap PSL_2(\mathbb{R})(z, K)$ and Γ is discrete, it is enough to show that ${}_2(\mathbb{R})(z, K)$ is compact. Indeed, we can clearly show the same statement with $SL_2(\mathbb{R})$ instead.

In other words, we would like to show that the set E of matrices $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $ad - bc = 1$ and $\frac{az+b}{cz+d} \in K$ is a closed, bounded subset of $\mathbb{R}^4$. Consider the continuous map $\beta : SL_2(\mathbb{R}) \to \mathcal{H}$ given by evaluation at z. Then $E = \beta^{-1}(K)$, so evidently E is closed. To see boundedness, we may assume that WLOG K is the subset $\{z = x + iy \mid m \leq y \leq M\}$ for some $m < M$, since these are cofinal in the compact subsets of $\mathcal{H}$. Then

$$|\frac{az+b}{cz+d}| = \frac{\Im(z)}{|cz+d|^2} \geq m,$$

so that we get the inequalities

$$|cz+d| \leq \sqrt{\frac{\Im(z)}{m}},$$

$$|az+b| \leq M\sqrt{\frac{\Im(z)}{m}}.$$

The boundedness of the vector $[a, b, c, d]$ follows easily, and we're done.

Remark: The moral here is that there is no general principle that guarantees discreteness of a group of transformations implies the discontinuity of its action. Indeed, given a discrete subgroup $\Gamma \subset PSL_2(\mathbb{C}) = PGL_2(\mathbb{C})$, one has no advance information about the structure of the limit set $\Lambda(\Gamma)$. Instead one defines $\Omega(\Gamma) = \mathbb{CP}^1 \setminus \Gamma$, so that Ω is an open subset of the Riemann sphere on which Γ acts properly discontinuously. The only catch here is that Ω could be empty: indeed this occurs for $\Gamma = PSL_2(\mathbb{Z}[i])$. A discrete subgroup $\Gamma \subset PSL_2(\mathbb{C})$ for which $\Lambda(\Gamma)$ is proper in $\mathbb{CP}^1$ is called a **Kleinian group**.

Example: Let us illustrate the above theory for cyclic subgroups $\Gamma = \langle\gamma\rangle$ of $PSL_2(\mathbb{R})$.

Case 1: γ is elliptic. Then Γ is Fuchsian if and only if it is finite. If so, then clearly $\Lambda(\Gamma) = \emptyset$.

Case 2: γ is parabolic. Then Γ is Fuchsian, and $\Lambda(\Gamma)$ consists of a single element on $\mathbb{RP}^1$; conversely, all points on $\mathbb{RP}^1$ arise this way.

Case 3: γ is hyperbolic. Then Γ is Fuchsian, and $\Lambda(\Gamma)$ is a 2 element subset of $\mathbb{RP}^1$; all such subsets arise.

Exercise 2.6: Work out the details of this example.

Exercise 2.7: What is the limit set of $PSL_2(\mathbb{Z})$? (Hint: recall that the limit set is closed!)

Definition: A Fuchsian group Γ is **of the first kind** if $\Lambda(\Gamma) = \mathbb{RP}^1$.

We will only have commerce here with Fuchsian groups "of the first kind." On the other hand, Fuchsian groups of the second kind are topologically interesting: it turns out that if $\Lambda(\Gamma)$ is a proper subset of $\mathbb{RP}^1$ whose cardinality is at least 3, then it is a perfect nowhere dense subset, i.e., homeomorphic to the Cantor set (or, if you like, to $\mathbb{Z}_p$).

3.2. **II. Geometry.** Let us review some aspects of the equality $\mathcal{H} = \mathbb{H}^2$, i.e., of the upper halfplane endowed with the Riemannian metric

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

As mentioned above, we have $\mathrm{Isom}^+(mathbbH^2) = PSL_2(\mathbb{R})$.

Proposition 6. *The geodesics on $\mathbb{H}^2$ are the semicircles orthogonal to the real axis and the vertical lines (which we may view as the case of infinite radius).*

Sketch proof: We shall not review the official definition of a geodesic in the context of Riemannian geometry (they are solutions to a certain partial differential equation). Recall that given an element of the unit tangent bundle of a Riemannian manifold M (i.e., a point P on the manifold together with a unit vector v in the tangent space of that point), there is an $\epsilon > 0$ and a smooth function $C : [0, \epsilon) \to M$ such that $C(0) = P$, $C'(0) = v$: less formally, we get a geodesic by starting at any point and pushing off in any direction. The characteristic property of a geodesic curve C is that given any point P on C, then there exists a neighborhood U of P such that if $Q \in C \cap U$, then the arc of C connecting P to Q has minimal length among all paths from P to Q.

The image of a geodesic under an isometry of the manifold is (clearly) another geodesic, and it is easy to see that the images of vertical lines under $PSL_2(\mathbb{R})$ yield all semicircles orthogonal to the real axis, and also that these are "enough" curves: i.e., each element of the unit tangent bundle lies on exactly one of them. Finally, it is easy to see that vertical line segments are geodesics: since ds does not differ from the Euclidean metric under change of x coordinate, the length of any path from $x + ia$ to $x + ib$ is at least as long as the length of its projection onto this line segment. Thus the minimal length is attained by the vertical line segment, namely

$$\int_0^1 \frac{dy/dt}{dy} dt = \log(b/a).$$

Remark: Note that each "full" geodesic has infinite length. Recall that this means that if we travel around the hyperbolic plane at bounded speed, we never reach the boundary in finite time. One calls such a Riemannian manifold **complete**.

The following result is presented just for culture; we will not need it.

Corollary 7. *The distance between two points $z,\ w \in \mathcal{H}$ is given by*

$$d(z, w) = \log \frac{|z - w| + |z - w|}{|z - \overline{w}| - |z - w|} = \log[w, z^*, z, w^*],$$

where

$$[z_1, z_2, z_3, z_4] = \frac{(z_1 - z_2)(z_3 - z_4)}{(z_2 - z_3)(z_1 - z_4)}$$

is the cross ratio, and the geodesic joining z to w intersects $\partial\mathcal{H}$ at z^* and w^*.

Proof: See [**?**, Theorem 1.2.5-6].

The Riemannian metric induces a volume form (or here, an area element), in this case $d\mu = \frac{dxdy}{y^2}$, which again must be invariant under $PSL_2(\mathbb{R})$ since the metric is.

Definition: A **hyperbolic polygon** is a closed subset of $\overline{\mathcal{H}}$ bounded by geodesic segments. The intersection of such segments is called a **vertex**. Note that this definition takes place in the *extended* hyperbolic plane: it is permissible for some (or all) of the vertices to be on the boundary. (It is useful to draw sketches in the unit disk model of the hyperbolic plane.)

Note that the fundamental equality $\mathrm{Isom}+(\mathbb{H}^2) = \mathrm{Aut}_{\mathbb{C}}(\mathcal{H})$ implies that there is no notion of "similar figures" in hyperbolic geometry: any transformation which preserves angles also preserves lengths! This may sound like a shortcoming but is actually the source of a magnificent richness: the angle sum in a hyperbolic n-gon is not predetermined, as in Euclidean geometry: rather, by pulling the vertices closer or farther apart there is enough room to make polygons with arbitrarily small angle sum. We illustrate with the following special case:

Theorem 8. *(Gauss-Bonnet) For any $0 \leq \alpha,\ \beta, \gamma$ be three numbers such that $\alpha + \beta + \gamma < \pi$. Then there exists a unique (up to isometry) hyperbolic triangle Δ with these angles, whose area is $\mu(\Delta) = \pi - \alpha - \beta - \gamma$.*

Proof: [**?**, Theorem 1.4.2].

Remark: It is indeed possible to take $\alpha = \beta = \gamma = 0$; for example, take the region outside the unit circle and with $|x| \leq \frac{1}{2}$. This region has hyperbolic area π. (On the other hand, the interior of the unit semicircle has infinite hyperbolic area.)

3.3. **III. Linear Algebra.** Let $\gamma = \pm \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be an element of $PSL_2(\mathbb{R})$. Put $T(\gamma) = |a + d|$. We say γ is
elliptic if $T(\gamma) < 2$,
parabolic if $T(\gamma) = 2$,
hyperbolic if $T(\gamma) > 2$.

We will call a subgroup $H \subset PSL_2(\mathbb{R})$ **of hyperbolic type** if all its elements are hyperbolic.

Note that the characteristic polynomial of a representative of γ in $SL_2(\mathbb{R})$ is
$$P(t) = t^2 \pm T(\gamma)t + 1,$$
with discriminant $T(\gamma)^2 - 4$. Thus a hyperbolic element has distinct real eigenvalues, a parabolic element has a repeated real eigenvalue, and an elliptic element has a conjugate pair of complex eigenvalues.

Exercise 2.8: Let $1 \neq \gamma \in SL_2(\mathbb{R})$.
We may view γ as an automorphism of $\mathcal{H}$ and of $\partial\mathcal{H} = \mathbb{RP}^1$.
a) If γ is elliptic, show that it has a unique fixed point in $\mathcal{H}$ (and a conjugate fixed point in $\mathcal{H}^-$).
b) If γ is parabolic, show that it has a unique fixed point in $\partial\mathcal{H}$.
c) If γ is hyperbolic, show that it has two distinct fixed points in $\partial\mathcal{H}$.

An element of $\mathcal{H}$ which is the fixed point of an elliptic element of Γ is called an **elliptic point**, and an element of $\partial\mathcal{H}$ which is the fixed point of a hyperbolic element of Γ is called a **hyperbolic point**. An element $z \in \partial\mathcal{H}$ which is the fixed point of a parabolic element is called a **cusp**.

Observe that every hyperbolic point or cusp for Γ is an element of the limit set $\Lambda(\Gamma)$.

One says that two elliptic points (resp. cusps) which are in the same Γ-orbit are **equivalent**. An equivalence class of elliptic points is called an **elliptic cycle**; the order of an elliptic cycle is the order of the stabilizer of any of its representative elliptic points.

Given a hyperbolic element γ, there is a unique half-circle in $\mathcal{H}$ with boundary points two fixed points, called the **axis** of γ.

Exercise 2.9: a) Show that a hyperbolic element is conjugate to $\begin{bmatrix} \lambda & 0 \\ 0 & 1/\lambda \end{bmatrix}$ with a uniquely determined λ.
b) Show that any parabolic element is conjugate to $\pm \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.
c) Show that an elliptic element is conjugate to $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ for a unique $\theta \in [0, 2\pi)$.

Exercise 2.10: a) Show that $PSL_2(\mathbb{R})$ acts doubly transitively on $\partial\mathcal{H}$. (That is, given any two ordered pairs of distinct points (P_1, P_2), (Q_1, Q_2) of $\partial\mathcal{H}$, there exists $\gamma \in PSL_2(\mathbb{R})$ such that $\gamma P_i = Q_i$ for $i = 1,\ 2$.
b) Does $PSL_2(\mathbb{R})$ act doubly transitively on $\mathcal{H}$? (Hint: Is there a geometric invariant that must be preserved?)

Proposition 9. *The set of elliptic points for a fixed Fuchsian group* Γ *does not accumulate in* $\mathcal{H}$.

Proof: Suppose $\{z_n\}_{n=1}^{\infty}$ is a sequence of distinct elliptic points converging to some $w \in \mathcal{H}$. Because of XXX, there exists a neighborhood U of w such that

$$\gamma U \cap U \neq \emptyset \implies \gamma(w) = w.$$

For sufficiently large n, $z_n \in U$ and $z_n \neq w$, and there exists $\gamma \in \Gamma$ such that $\gamma(z_n) = z_n$. Then $\gamma U \cap U \neq \emptyset$, so that $\gamma w = w$. But this means that γ has two distinct fixed points in $\mathcal{H}$, a contradiction.

3.4. **IV. Group theory.**

Proposition 10. *Two nonidentity elements of* $PSL_2(\mathbb{R})$ *commute if and only if they have the same fixed point set.*

Proof: If $\alpha\beta = \beta\alpha$, then α preserves the fixed point set of β. If we suppose that α is parabolic, it is conjugate to $z \mapsto z \pm 1$, so its centralizer consists of parabolic elements with cusp ∞, i.e., $z \mapsto z + \lambda$ for $\lambda \in \mathbb{R}$. If α is elliptic, then again all elements of $PSL_2(\mathbb{R})$ which preserve the fixed point set must also fix the unique element of the fixed point set lying in $\mathcal{H}$, so that β lies in the abelian group $\alpha SO(2)\alpha^{-1}$. Finally, if $\alpha(z) = \lambda z$ is hyperbolic, the fact that the centralizer is $\mathbb{G}_m(\mathbb{R})$ follows from a direct calculation.

Note that what we actually showed was that the centralizer in $PSL_2(\mathbb{R})$ of a hyperbolic / parabolic / elliptic element consists of all hyperbolic / parabolic / elliptic elements with the same fixed point set.

Corollary 11. *Let Γ be a Fuchsian group all of whose nonidentity elements have the same fixed point set. Then Γ is cyclic.*

Proof: Using the previous result, this reduces to the fact that discrete subgroups of $SO(2)$, $\mathbb{R}^\times$ and $\mathbb{R}$ are cyclic.

In particular every abelian Fuchsian group is cyclic.

Theorem 12. *Let Γ be a nonabelian Fuchsian group. Then it normalizer $N(\Gamma)$ in $PSL_2(\mathbb{R})$ is a Fuchsian group.*

Proof: Suppose that $N(\Gamma)$ is not discrete. Then there exists a sequence of distinct elements $T_i \in N(\Gamma)$ such that $T_i \to 1$. For $S \in \Gamma \setminus \{1\}$, $T_i S T_i^{-1} \to S$, and since Γ is discrete we have $T_i S T_i^{-1} = S$ for all sufficiently large i. Now choosing two arbitary elements S, S' we get that they have the same fixed point set as T_i for all sufficiently large i, and hence that they commute with each other, contradiction.

3.5. **V. Riemann surfaces.**

Proposition 13. *Let Γ be a Fuchsian group. Then $Y(\Gamma) := \Gamma\backslash Y$ has, in a canonical way, the structure of a Riemann surface.*

Proof: Let $\mathcal{E} \subset \mathcal{H}$ be the locus of fixed points of elliptic elements of Γ. By the above material, we know that $\mathcal{E}$ is a discrete (possibly empty) subset of $\mathcal{H}$. On the complement $Y \setminus \mathcal{E}$, Γ acts freely discontinuously, so passage to the quotient gives an unramified (normal) covering with group Γ. In particular the map is a local homeomorphism, so it is clear how to endow the quotient with a complex structure: formally speaking, near any point P on the quotient, a function element φ is decreed to be analytic if $\circ q$ is locally analytic near any chosen preimage $q^{-1}(P)$.

It remains to describe the complex structure on the quotient locally at the fixed point $P = q(\tilde{P})$ of an elliptic element $\gamma \in \Gamma$. Let λ be a holomorphic isomorphism from $\mathcal{H}$ to the unit disk such that $H(\gamma) = 0$. After this coordinate change γ is given locally as multiplication by $\zeta_n = e^{2\pi i a/n}$, where $n = \#\langle\gamma\rangle$. We can thus define a function element on the quotient to be locally analytic at P if its preimage as a function element near $0 \in D$ is analytic and has valuation at 0 divisible by n. This gives an analytic structure on the quotient space.

4. Fundamental regions

For a full rank lattice $\Lambda \subset \mathbb{C}$, one best visualizes the quotient space $\mathbb{C}/\Lambda$ by considering the side identifications on a fundamental parallelogram for Λ. This went without explicit mention in our discussion of uniformized elliptic curves (here, after all, the topological picture is the same in every case). The analogous geometric construction for Fuchsian groups Γ acting on $\mathcal{H}$ is much richer, and plays a unifying role in the theory.

Definition: A **fundamental set** $S \subset \mathcal{H}$ for Γ is a set of representatives for the Γ orbits of $\mathcal{H}$: in other words, $\mathcal{H} = \coprod_{g\in\Gamma} gS$. It is obvious that fundamental sets exist for group actions on any set (provided that you believe in the axiom of choice).

Comparison with the case of lattices in $\mathbb{C}$ suggests however that this is not the definition we really want: a parallelogram is not a fundamental set until we choose to delete one of each pair of opposite sides. This choice is arbitrary and, in fact, inconvenient: in order to see what we are glueing, it would be better to have both edges. On the other hand we should try to impose some additional nice properties for a fundamental region: at the moment, a fundamental region need not be measurable.

It is, of course, not *a priori* clear which properties can be required of a fundamental region for an arbitrary Fuchsian group (e.g. connectedness, yes; compactness, no), nor necessarily which are desirable. But in the interest of streamlining the presentation, we jump to the following definition.

Definition: A subset $R \subset \mathcal{H}$ is called a **fundamental region** for Γ if it satisfies the following conditions:
FR1) R is equal to the closure of its interior, R°.
FR2) $\bigcup_{g \in \Gamma} gR = \mathcal{H}$, and for $1 \neq g \in \Gamma$, $R^\circ \cap gR^\circ = \emptyset$.
FR3) R° is hyperbolically convex (so is in particular connected).
FR4) ∂R has measure zero, and is a countable union of **edges** C_i, which are either closed geodesic arcs or closed intervals of the real line (**free edges**). If $C_i \neq C_j$, then $C_i \cap C_j$ is empty or consists of a single point.
FR5) The induced tesselation $\{gR \mid g \in \Gamma\}$ is **locally finite**: any compact set meets only finitely many translates of R.

Example: Let $\Gamma = \langle \gamma \rangle$ be a cyclic subgroup generated by the hyperbolic transformation $\gamma(z) = \lambda z$, for $\lambda \neq 1$. We may assume that $\lambda > 1$ (otherwise replace γ by γ^{-1}). Let

$$R := \{z \in \mathcal{H} \mid 1 \leq |z| \leq \lambda\}.$$

That R is a fundamental domain for Γ is self-evident. The quotient Riemann surface is a topological cylinder (in particular, it is not an affine or complete algebraic curve)[2] It is not hard to see that R has infinite hyperbolic area.

Exercise 2.11: Show that, in fact, any fundamental region with a free side has infinite hyperbolic area.

Exercise 2.12: Carry out the corresponding discussion for a cyclic group generated by a parabolic element.

Example: Let $\Gamma = PSL_2(\mathbb{Z})$. We claim that

$$R = \{z \in \mathcal{H} \mid \Re(z) \in [-\frac{1}{2}, \frac{1}{2}], |z| \geq 1\}$$

[2]However, it inherits the structure of a Riemannian surface of constant negative curvature, so is – I suppose – what geometers would call a "pseudosphere."

is a fundamental domain for Γ. Fix $z \in \mathcal{H}$ and $\gamma = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(\mathbb{Z})$. Then $\Im(\gamma(z)) = \Im(z)/|cz+d|^2$. Since $+\mathbb{Z}$ is a lattice in $\mathbb{C}$, there exists a minimum $|cz+d|$ for $(c,d) \neq (0,0)$, which in turn implies that for fixed z, $\max_{\gamma \in \Gamma} \Im(\gamma(z))$ exists. Choose γ so as to maximize $y = \Im(\gamma(z))$, where $w = \gamma(z) = x+iy$. Let $S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then

$$\Im(S\gamma(z)) = \Im(-1/w) = \frac{y}{|w|^2} \leq y,$$

and we conclude that $|w| \geq 1$. Now put $T(z) = z+1$, so that $\Im(T^h(\sigma(z)) = \Im(\sigma(z))$ for all integers h, so that $|T^h(\sigma(z))| \geq 1$. Certainly we can choose h so as to put the real part of $\sigma(z)$ in the interval $[\frac{-1}{2}, \frac{1}{2}]$, and we have shown that every element of $\mathcal{H}$ is equivalent to an element of R.

It remains to be seen that distinct elements of R° are not Γ-equivalent, so assume for a contradiction that $z' = \sigma(z)$. WLOG, $\Im(z) \leq \Im(z') = \Im(z)/|cz+d|^2$. Thus

$$|c|\Im(z) \leq |cz+d| \leq 1.$$

If $c = 0$, then $a = d = \pm 1$ so $z' = z \pm b$, impossible. Therefore $c \neq 0$. Looking at R, we clearly have $\Im(z) > \sqrt{3}2$, so the equation implies $c = 1$, so $|z \pm d| \leq 1$. But if $z \in R^\circ$ and $|d| \geq 1$, we have $|z+d| > 1$. Hence $d = 0$, so $|z| \leq 1$, and z is not in R°. This completes the proof.

Note that the right hand side of the fundamental region is a hyperbolic triangle with angles $(\pi/2, \pi/3, 0)$. Thus by Gauss-Bonnet it has hyperbolic area $\pi/6$, hence the given fundamental domain for $PSL_2(\mathbb{Z})$ has area $\pi/3$. (This could, of course, be shown by a direct computation.)

Exercise 2.13: a) Deduce from the proof that $PSL_2(\mathbb{Z})$ is generated by S and T.
b) Thus $PSL_2(\mathbb{Z})$ is also generated by S and $W = ST$.
c) Use the tesselation of $\mathcal{H}$ by translates of R to show that $PSL_2(\mathbb{Z})$ admits the presentation

$$\langle S, W \mid S^2 = W^3 = 1 \rangle;$$

in combinatorial group theory one would say that $PSL_2(\mathbb{Z})$ is the **free product** of $\mathbb{Z}/2\mathbb{Z}$ and $\mathbb{Z}/3\mathbb{Z}$. (This begins to illustrate the supremacy of geometric methods in "abstract" group theory. It is not at all straightforward to give a direct algebraic proof of this, and harder still to imagine how one would guess the result in the first place.)

Exercise 2.14: Use the above presentation for $PSL_2(\mathbb{Z})$ to solve the **congruence subgroup problem**; in other words, show that there exist finite index subgroups $\Gamma \subset PSL_2(\mathbb{Z})$ which do not contain $\Gamma(N)$ for any N. (Suggestions: (i) Γ' can be taken to be normal. (ii) What are the possible composition factors of the groups $PSL_2(\mathbb{Z}/N\mathbb{Z})$? (iii) Look up the list of finite simple groups which can be generated by an element of order 2 and an element of order 3 (it's a long list).)

Note that the fundamental region for a Fuchsian group – or even the tesselation that it determines – is far from unique. For instance, we could translate the fundamental region R for $PSL_2(\mathbb{Z})$ to the right by any real number in $(0,1)$ and get a different tesselation. (In this case the tesselations are conjugate under the action of $PSL_2(\mathbb{R})$, but this need not occur in general.) It is thus important to identify properties which are independent of the choice of the fundamental region.

Proposition 14. *Any two fundamental regions for the same Fuchsian group have the same hyperbolic area (which may be infinite).*

Proof: Let R_1 and R_2 be two fundamental regions. By definition, the areas are equal to the areas of their interiors, so

$$R_1 \supset R_1 \cap (\bigcup_{g\in\Gamma} g(R_2^\circ)) = \bigcup_{g\in\Gamma} R_1 \cap g(R_2^\circ).$$

The sets on the right hand side are pairwise disjoint, so

$$\mu(R_1) \geq \sum_{g\in\Gamma} \mu(R_1 \cap gR_2^\circ) = \sum_{g\in\Gamma} \mu(gR_1 \cap R_2^\circ) \geq \mu(\bigcup_g gR_1 \cap R_2^\circ) = \mu(R_2^\circ) = \mu(R_2).$$

By symmetry, we conclude $\mu(R_1) = \mu(R_2)$.

In view of this proposition, it makes sense to define the **covolume** of a Fuchsian group Γ as $\mu(R)$ for any fundamental region R for Γ. We will write $\mathrm{vol}(\Gamma)$. A Fuchsian group is said to be **v-finite** if $\mathrm{vol}(\Gamma) < \infty$.

Proposition 15. *Let* $\Gamma' \subset \Gamma$ *be a subgroup of a Fuchsian group. Then*

$$v(\Gamma') = [\Gamma : \Gamma']\,\mathrm{vol}(\Gamma).$$

Exericse 2.X: Prove it.

We now turn to the construction of a fundamental region satisfying all the desired properties. We first choose any $p \in \mathcal{H}$ which is not an elliptic element for Γ. Then define

$$D_P(\Gamma) = \{z \in \mathcal{H} \mid \forall \gamma \in \Gamma,\ d(z,p) \leq d(z, \gamma(p))\}$$

where d denotes the hyperbolic distance. In other words, we just take those points z all of whose Γ translates are at least as far from p as z itself.

Theorem 16. $D_p(\Gamma)$ *is a fundamental region for* Γ*.*

Sketch proof: We need the following alternate construction of the Dirichlet region. Namely, Then $D_p(\Gamma) = \cap_{g\in\Gamma\setminus 1} H_p(g)$, where

$$H_p(g) = \{z \in \mathcal{H} \mid d(z,p) \leq d(z, g(p))\}.$$

Reflecting a bit on the analogies with Euclidean geometry, it is plausible (and also provable!) that $H_p(g)$ can be described as follows: for $1 \neq g\Gamma$, let $L_p(g)$ be the unique geodesic connecting p and $g(p)$(our assumption on p guarantees that these are distinct points). Let $T_p(g)$ be the perpendicular bisector of $L_p(g)$. then $T_p(g)$ divides $\mathcal{H}$ into two halfplanes[3], and the one containing p is $H_p(g)$.

At least given a belief that hyperbolic geometry works out nicely, much of the proof follows from this: as $D_p(\Gamma)$ an intersection of halfplanes, it is evidently hyperbolically convex (in particular, connected), and its boundary consists of geodesic

[3]Quarter planes?

segments and free edges. It is less clear that $D_p(\Gamma)$ is locally finite; for this, see [**?**, Theorem 3.5.1].

One of the beautiful features of the Dirichlet region is that it displays the cusps and the elliptic points: namely, each equivalence class of cusp or ellipic points shows up exactly once as a vertex of any Dirichlet region. The basic idea here is that since cusps and elliptic points have nontrivial stabilizers, they could not lie in the interior of any fundamental region. (Moreover, a cusp is a limit point of a Γ-orbit in $\mathcal{H}$ so must be represented in the closure of a fundamental set.)

There is one annoying little problem, however: an elliptic point of order n will correspond to a vertex of the fundamental domain with angle $2\pi/n$, so when $n = 2$ we get an angle of π, and it is impossible to see such vertices. This occurs in the example of the modular group: there is a "vertex" at $z = i$ despite the fact that this does not appear to be an orbifold point of the fundamental region. Luckily, the theory of elliptic points of order 2 is resolved by the theory of side pairings.

Side pairings: The sides of a Dirichlet domain arise as the nonempty sets $g(R) \cap R$. This picks out a subset Γ^* of elements of Γ and a surjective map Φ from Γ^* to the set of sides S of R. This map is in fact a bijection. Moreover, Γ^* is closed under inversion: if $gR \cap R \neq \emptyset$, then so is its image under g^{-1}. Thus there is a natural *pairing* amongst the sides of R, induced by the action of Γ^*. Note that a side gets paired with itself if and only if the corresponding element g has order 2. This allows us to resolve the problem of the previous paragraph: such a side must contain an order two elliptic element, and we can subdivide the side into two subsegments and further analyze the side pairing action.

Definition: We say a Fuchsian group is **geometrically finite** if it admits a Dirichlet region with finitely many sides (in the above refined sense).

Theorem 17. Γ *is generated by the side pairing elements* Γ^*. *In particular, if* Γ *is geometrically finite, it is finitely generated. Indeed, if there exists a Dirichlet domain with* s *sides, then* Γ *can be generated by* s *elements. If* Γ *has no elliptic points of even order, then it can be generated by* $\frac{s}{2}$ *elements.*

5. Basic theorems on Fuchsian groups

We know list some of the fundamental theorems on Fuchsian groups. For the sake of brevity, we omit the proofs (in many cases, they follow by careful consideration of what has already been shown).

Theorem 18. *(Poincaré) Every Fuchsian group of the first kind admits a finite index normal subgroup without elliptic elements.*

Remark: In contrast to the other results of this section, this *does not* follow more or less directly from what we have already done: some nontrivial geometry and finite group theory are required. See Stillwell's book on geometry of surfaces for a wonderful discussion.

Theorem 19. *For a Fuchsian group* Γ, *the following are equivalent:*
a) There exists a finite orbit Γz.

b) The limit set $\Lambda(\Gamma)$ *has at most two elements.*
c) Γ *is either abelian or conjugate to a group generated by* $g(z) = \lambda z,\ h(z) = \frac{-1}{z}$*.*
Such a Fuchsian group is said to be ***elementary****.*

Remark: The terminology "elementary" is standard but misleading: the elementary groups are the Fuchsian groups which are too small to be interesting or useful. In many instances, they are exceptions to the general theory.

Exercise 2.15: Which Riemann surfaces arise as $\Gamma\backslash\mathcal{H}$ for an elementary Fuchsian group?

Before presenting the next result, let us discuss some topological aspects of the classification of Riemann surfaces. Namely, any connected orientable surface can be given the structure of a Riemann surface (sometimes in many different ways). Let us divide the underlying topological surface S into four categories:

I: S is compact, of genus $g \geq 0$. Equivalently, S is the set of complex points of a projective complex algebraic curve. We shall say S is of type $(g, 0, 0)$.

II: S is obtained from a compact Riemann surface of genus g with $n > 0$ **punctures** removed. Equivalently, S is the set of complex points of an affine complex algebraic curve. We say S is of type $(g, n, 0)$.
III. S is obtained from a surface of the form II above by removing r pairwise disjoint closed disks. We say S is of type (g, n, r). For instance, the unit disk is of type $(0, 0, 1)$, and the punctured unit disk is of type $(0, 1, 1)$. Note that for $r > 0$, these are not associated to algebraic curves.

Each of these Riemann surfaces are **topologically finite**: their homology groups $H^i(S, \mathbb{Q})$ are finite-dimensional.

IV. There are also surfaces which are **topologically infinite**, obtained (loosely speaking) by allowing g, n and/or r to take the value ∞.

Theorem 20. *For a nonelementary Fuchsian group* Γ*, the following are equivalent:*
a) Γ *is geometrically finite.*
b) Γ *is finitely generated.*
c) $\Gamma\backslash\mathcal{H}$ *is topologically finite.*
If these conditions are satisfied, Γ *has only finitely many elliptic cycles, say of orders* $m_1, \ldots, m_d$ *(*$d = 0$ *is possible). If* $\Gamma\backslash\mathcal{H}$ *is of type* (g, n, r)*, one says that the* ***signature*** *of* Γ *is* $(g; m_1, \ldots, m_d; n; r)$*. Then* Γ *admits a presentation with generators* $\alpha_1, \ldots, \alpha_g, \beta_1, \ldots, \beta_g, \gamma_1, \ldots, \gamma_d, \rho_1, \ldots, \rho_n, \iota_1, \ldots, \iota_r$ *and relations*

$$\gamma_i^{m_i} = 1,\ \prod_{i=1}^{g} [\alpha_i, \beta_i] \cdot \gamma_1 \cdots \gamma_d \cdot \rho_1 \cdots \rho_d \cdot \iota_1 \cdots \iota_r = 1.$$

For algebraic purposes, we are not interested in the Riemann surfaces which have boundary curves. The following result explains which Fuchsian groups have $\Gamma\backslash\mathcal{H}$ of type I or II.

Theorem 21. *For a Fuchsian group* Γ*, the following are equivalent:*
a) Γ *is v-finite.*

b) A Dirichlet region for Γ is geometrically finite and has no free sides.
c) $\Lambda(\Gamma) = \partial\mathcal{H}$.
d) $\Gamma\backslash\mathcal{H}$ is of type $(g, n, 0)$.
A Fuchsian group satisfying these equivalent conditions is said to be ***of the first kind****.*

Since we will only have truck with Fuchsian groups of the first kind, we now abbreviate the type $(g, n, 0)$ to (g, n).

Theorem 22. *Let Γ be a Fuchsian group of the first kind, with elliptic cycles of orders $m_1, \ldots, m_d$, and such that $\Gamma\backslash\mathcal{H}$ has type (g, n). Then*

$$\mathrm{vol}(\Gamma) = 2\pi\left((2g-2) + \sum_{i=1}^{d}(1 - \frac{1}{m_i})\right). \tag{1}$$

Remark: This is essentially a theorem of hyperbolic geometry.

Note that, among other things, Equation 1 gives a necessary condition for the existence of a Fuchsian group of a given signature, namely that the parenthesized quantity be positive. It turns out that this condition is also sufficient:

Theorem 23. *(Poincaré-Maskit) If $(2g-2) + n + r + \sum_{i=1}^{d}(1 - \frac{1}{m_i}) > 0$, there exists a nonelementary Fuchsian group with signature $(g; m_1, \ldots, m_d; n; r)$.*

Exercise 2.17: Let Γ be a Fuchsian group.
a) Show that $\mathrm{vol}(\Gamma) \geq \frac{\pi}{21}$, with equality attained if and only if Γ has signature $(0; 2, 3, 7; 0)$. (We shall later see that there exists a *unique* conjugacy class of Fuchsian groups with this signature.)
b) If Γ has parabolic elements, show that $\mathrm{vol}(\Gamma) \geq \frac{\pi}{3}$, with equality attained if and only if Γ has signature $(0; 2, 3; 1)$ (again, it will turn out that such a Γ is necessarily conjugate to the modular group).

Problem 2.2: Let Γ be a Fuchsian group of signature $(0; 2, 3, 7; 0)$.
a) Show that Γ has a presentation of the form

$$\langle x,\ y,\ z \mid x^2 = y^3 = z^7 = xyz = 1\rangle.$$

(Note that the perspective of abstract generators and relations reveals essentially nothing about the structure of this group, not even that it is infinite. You might try to figure out what group results if the 7 is replaced by a 5.)
b) By Theorem 18, there exists Γ' a finite-index normal subgroup of Γ of hyperbolic type. Let $G = \Gamma/\Gamma'$. Show that at least one composition factor of G is a simple group.
c) Consider the compact Riemann surface $X' := \Gamma'\backslash\mathcal{H}$. If g is its genus, show that $\#G = 84(g-1)$.
d)* Show that there does not exist such a Γ' with $g \leq 2$, and that there exists a unique Γ' with $g = 3$.
e)** For which values of g does there exist such a Γ'?[4]

[4]Don't actually try working on this before you consult the literature: this is an open problem upon which a tremendous amount of work has been done. For instance, it is known that one *can* take G to be the Monster.

Note that it is implicit in our discussion that for a Fuchsian group Γ of the first kind, the Riemann surface $\Gamma\backslash\mathcal{H}$ can be compactified by adding finitely many points, one for each equivalence class of cusps. This description will be suitable for our purposes (because, in fact, we will be most interested in the case when there are no cusps!). A somewhat finer analysis would be necessary in order to develop, e.g., the theory of Fourier expansions of modular forms. See [**?**, Chapter 1] for an excellent account of these matters.

Corollary 24. *For a Fuchsian group of the first kind, the following are equivalent:*
a) $\Gamma\backslash\mathcal{H}$ *is of type* $(g, 0, 0)$*, i.e., is compact.*
b) Γ *is cocompact, i.e.,* $\Gamma\backslash PSL_2(\mathbb{R})$ *is compact.*
c) Γ *has no parabolic elements.*

Note that if Γ has parabolic elements, so does every finite index subgroup. On the other hand, the following (highly nontrivial) theorem says that elliptic elements can be eliminated.
In particular, cocompact Fuchsian groups admit finite index normal subgroups of hyperbolic type.