LECTURES ON SHIMURA CURVES: ARITHMETIC FUCHSIAN GROUPS

PETE L. CLARK

1. What is an arithmetic Fuchsian group?

The class of Fuchsian groups that we are (by far) most interested in are the **arithmetic** groups. The easiest way to describe arithmetic Fuchsian groups is that class of groups containing all groups commensurable with $PSL_2(\mathbb{Z})$ and also all groups which uniformize compact Shimura curves. This is however, not a very well-motivated definition: structurally, what do classical modular curves and Shimura curves have in common that elevates them above the rank and file of the common Fuchsian group?

The goal of this section is to try to answer this question. We will not give complete proofs.

Definition: For a Fuchsian group (of the first kind) Γ, the **trace ring** R is the subring of $\mathbb{R}$ generated by the traces of elements of Γ. (Note that the trace of an element of $PSL_2(\mathbb{R})$ is well-defined only up to ± 1, but this is good enough for the definition to make sense.) Similarly, the **trace field** is defined to be the subfield of $\mathbb{R}$ generated by the traces of the elements of Γ. Define $B = R[\Gamma]$ to be the R-algebra generated by the matrix entries, and similarly for $k_0[\Gamma]$.

For instance, if $\Gamma = PSL_2(\mathbb{Z})$, $k_0 = \mathbb{Q}$ and $k_0[\Gamma] = M_2(\mathbb{Q})$.

Consider the following properties of a Fuchsian group Γ (of the first kind):

(WA1) k_0 is a number field and R is an order in k_0.
(WA2) There exists $\sigma \in PSL_2(\mathbb{R})$ and a number field $K \subset \mathbb{R}$ such that $\sigma^{-1}\Gamma\sigma \subset PSL_2(\mathcal{O}_K)$.
(WA3) The trace field is a number field k_0, and $k_0[\Gamma]$ is a quaternion algebra B/k_0, and $R[\Gamma]$ is an order of B.

Theorem 1. *The properties (W1), (W2) and (W3) are all equivalent.*

We shall call a Fuchsian group satisfying the equivalent conditions of the theorem **weakly arithmetic**.

Problem 4.1: Show that being weakly arithmetic is a commensurability invariant.

Problem 4.2: a) Show that the triangle groups $\Delta(a, b, c)$ are weakly arithmetic, and indeed that the trace ring R is $\mathbb{Z}[2\cos(\frac{\pi}{a}), 2\cos(\frac{\pi}{b}), 2\cos(\frac{\pi}{c})]$.

b)** Compute the quaternion algebra $k[\Gamma]$ in terms of $a,\ b,\ c$.[1]

Problem 4.3: a) Give a cardinality argument to show that most Fuchsian groups are not weakly arithmetic.
b) Give an explicit example of such a Fuchsian group (of the first kind!).

Now consider the following properties of a Fuchsian group of the first kind.

$(A1)$ Γ has infinite index in its commensurator $\mathrm{Comm}(\Gamma)$.
$(A2)$ Γ is commensurable with a group derived from a quaternion algebra.
$(A2')$ $\Gamma^2 = \{\gamma^2 \mid \gamma \in \Gamma\}$ is derived from a quaternion algebra.
$(A2'')$ The trace ring R is an order in the trace field, a totally real number field, and for every nonidentity real embedding $\iota : k_0 \hookrightarrow \mathbb{R}$, $\iota(\mathrm{Tr}(\Gamma))$ is a bounded subset of $\mathbb{R}$.
(A3) There exists $G/\mathbb{Q}$ a connected, noncommutative, almost $\mathbb{Q}$-simple algebraic group, a $\mathbb{Q}$-embedding $\iota : G \hookrightarrow GL_n$, and a homomorphism $\tau : G(\mathbb{R}) \to PSL_2(\mathbb{R})$ of real Lie groups, surjective and with compact kernel, such that the group $\tau(G(\mathbb{Q}) \cap GL_n(\mathbb{Z})) \subset PSL_2(\mathbb{R})$ is commensurable with Γ.

Theorem 2. *a) The three conditions (A1), (A2) and (A3) are equivalent, and a Fuchsian group satisfying them is said to be **arithmetic**.*
b) Arithmetic groups are weakly arithmetic.
c) The converse does not hold: e.g., there exist only finitely many triples (a, b, c) such that $\Delta(a, b, c)$ is arithmetic. The complete list is due to Takeuchi.

The equivalences$(A2) \iff (A2') \iff (A2'')$ are due to Takeuchi. It is a nice result – e.g., using it, it is easily seen which Hecke grups $\Delta(2, q, \infty)$ are arithmetic – but is probably of lesser importance in the grand scheme of things. Condition (A3) generalizes well to "arithmetic lattices" inside semisimple Lie groups, an important area of mathematics with work done by (e.g.) Borel, Serre and Margulis. The equivalence of (A2) and (A3) is one of those annoying facts for which it had been hard to find a reference, but now there is a paper by Mochizuki in which he gives the complete proof. The equivalence of (A1) and (A3) is a result of Margulis. It is very enlightening to keep in mind the (A3) characterization of arithmetic Fuchsian groups: it should be viewed as saying that for algebraic curves uniformized by arithmetic Fuchsian groups (and only those curves) there is a substantial supply of Hecke operators.

Remark: The terminology "weakly arithmetic" has been made up on the spot to distinguish between the weaker and the stronger sets of equivalent conditions. However, this notion turns out to be almost equivalent to a notion introduced in a recent (2000) paper of Schaller and Wolfart: they define a **semi-arithmetic** Fuchsian group to be one satisfying the equivalent axioms (WA1) through (WA3) with the additional assumption that the trace field is totally real. It seems too soon to say what size of a role this class of groups plays in mathematics – in a certain technical sense, it is an open question to find "interesting" semi-arithmetic groups apart from the triangle groups.

[1] E.g., write a computer program that will compute the invariants of the quaternion algebra $k_0[\Gamma]$.

2. Construction of a Fuchsian group from an order in a quaternion algebra

Let F be a totally real field, and let B/F be a quaternion algebra. If $[F : \mathbb{Q}] = g$, enumerate the real places of F by $\infty_1, \ldots, \infty_g$. In case $\mathcal{O}$ is an order in a quaternion algebra over F which is ramified at $g-1$ of the infinite places of F, we can associate to $\mathcal{O}$ an arithmetic Fuchsian group, which is cocompact unless B is the split quaternion algebra. By definition, a Shimura curve is the quotient of $\mathcal{H}$ by such a Fuchsian group (equivalently, by any cocompact arithmetic group). Notice that the situation is much cleaner when $F = \mathbb{Q}$: any order in any indefinite quaternion algebra will do.

It seems reasonable then to take a down-to-earth approach to the construction when $F = \mathbb{Q}$ and an adeles / algebraic groups approach in the general case.

2.1. **Background on quaternion algebras.** Let F be any field of characteristic different from 2. Then any quaternion algebra over F can be given in the form $\langle a,\ b\rangle$, where $a,\ b \in F^\times$, that is, as the F-algebra generated by two elements i and j satisfying the relations, $i^2 = a,\ j^2 = b,\ ij = -ji$ (and it follows that $(ij)^2 = -ab$.

Especially, as F-vector space we have

$$\langle a,\ b\rangle = F \cdot 1 \oplus F \cdot i \oplus F \cdot j \oplus F \cdot ij.$$

For an arbitrary element

$$x = x_0 \cdot 1 + x_1 \cdot i + x_2 \cdot j + x_3 \cdot k,$$

we define the reduced trace $T(x) = 2x_0$, the canonical involution $\overline{x} = T(x) - x$, and the reduced norm

$$N(x) = x \cdot \overline{x} = x_0^2 - ax_1^2 - bx_2^2 - abx_3^2 \in F.$$

Exercise 4.X: Let ρ be the left-regular representation of $B = \langle a,\ b\rangle$, i.e., the map which views $x\cdot$ as an endomorphism of the underlying 4-dimensional F-vector space of B.
a) Show that the trace of $\rho(x\cdot)$ is $2T(x)$.
b) Show that the norm of $\rho(x\cdot)$ is $N(x)^2$.
Conclude that T, $x \mapsto \overline{x}$ and N are well-defined independent of the choice of the presentation $B = \langle a,\ b\rangle$.

Proposition 3. *Fix $a \in F^\times$. A quaternion algebra B/F is isomorphic to $\langle a,\ b\rangle$ for some b if and only if there exists a subfield of B isomorphic to $F(\sqrt{a})$.*

Proof: The "only if" direction is clear, since the subfield $F[i]$ is of the required form. Conversely, suppose that there exists an embedding $F(\sqrt{a}) \hookrightarrow B$. Let σ be the nontrivial Galois automorphism of $E = F(\sqrt{a})/F$. By the Skolem-Noether theorem, σ extends to an inner automorphism on all of B: there exists $u_\sigma \in B^\times$ such that for all $e \in E$, $E^\sigma = u_\sigma^{-1} e u_\sigma$. Since $e = \sigma^2(e) = u_\sigma^{-2} e u_\sigma^2$, u_σ^2 lies in the centralizer of E, which is easily seen to be E itself. Clearly u_σ is not in E, so

$$F \subset F[u_\sigma^2] \cap E \subset F[u_\sigma] \cap E = F.$$

Thus $u_\sigma^2 \in F$, say $u_\sigma^2 = b$. Let i be one of the two elements $\sqrt{a}$ in E. By construction, $u_\sigma^{-1} i u_\sigma = -i$, so $B \cong \langle a,\ b\rangle$.

Proposition 4. *Let B/F be a quaternion algebra, and $E = F(\sqrt{a})$ a quadratic subfield of B. Then $B \otimes_F E \cong M_2(E)$.*

Proof: Define

$$I = \begin{bmatrix} \sqrt{a} & 0 \\ 0 & -\sqrt{a} \end{bmatrix}, J = \begin{bmatrix} 0 & b \\ 1 & 0 \end{bmatrix}.$$

It is easy to check that the E-algebra generated by these two matrices is on the one hand $B \otimes_F E$ and on the other hand $M_2(E)$.

2.2. **The down-to-earth approach.** Let $B/\mathbb{Q}$ be an indefinite rational quaternion algebra. Let $\mathcal{O}$ be an order in B. For example, one can write $B \cong \langle a,\ b\rangle$ with $a,\ b \in \mathbb{Z}$ and take

$$\mathcal{O} = \{x = x_0 + x_1 \cdot i + x_2 \cdot j + x_3 \cdot ij \mid x_i \in \mathbb{Z}.\}$$

Let $\mathcal{O}^1$ be the elements of $\mathcal{O}$ of reduced norm 1.

To say that B is indefinite is to say that there exists an isomorphism $\iota : B \otimes_{\mathbb{Q}} \mathbb{R} \cong M_2(\mathbb{R})$; we choose such an isomorphism, which allows us in particular to regard $\mathcal{O}$ as being embedded in $M_2(\mathbb{R})$, $\mathcal{O}^\times$ in $GL_2(\mathbb{R})$ and $\mathcal{O}^1$ in $SL_2(\mathbb{R})$.

Define $\Gamma = \Gamma(B, \mathcal{O}) = \mathcal{O}^1/\pm 1$ to be the image of $\mathcal{O}^1$ in $PSL_2(\mathbb{R})$.

Exercise 4.X: Show that the groups Γ associated to different orders $\mathcal{O}$ are commensurable.

Theorem 5. *Γ is a discrete subgroup of $PSL_2(\mathbb{R})$. It is cocompact if (and only if) B is a division algebra.*

Proof: Note that for two commensurable subgroups $\Gamma_1,\ \Gamma_2 \subset PSL_2(\mathbb{R})$, the discreteness (resp. cocompactness) of Γ_1 implies the discreteness (resp. cocompactness) of Γ_2. So we will assume that $\mathcal{O}$ is the order constructed from $B \cong \langle a,\ b\rangle$ above. By Proposition 4, we have an explicit embedding of $\mathcal{O}$ into $M_2(\mathbb{R})$, namely

$$x = x_0 + x_1 \cdot i + x_2 \cdot j + x_3 \cdot ij \mapsto \begin{bmatrix} x_0 + x_1\sqrt{a} & x_2 + x_3\sqrt{a} \\ b(x_2 - x_3\sqrt{a}) & x_0 - x_1\sqrt{a} \end{bmatrix}.$$

It is easy to see that there exists an open neighborhood U of the identity matrix such that $\mathcal{O}^1 \cap U = 1$. Choosing U such that $|g_{11} - 1|,\ |g_{22} - 1| < 1$, we have $|(x_0 + x_1\sqrt{a}) + (x_0 - x_1\sqrt{a}) - 2| < 1$, or $|2x_0 - 2| < 1$, which implies that $x_0 = 1$ and $x_1 = 0$. Then choosing g_{12} and g_{21} to be sufficiently closer to zero we get that $x_2 = x_3 = 0$.

Compactness:

2.3. **The adelic approach.** Let $G/\mathbb{Q}$ be any reductive algebraic group, and let $G' = [G, G]$ be its (semisimple) derived subgroup. Recall that the adelic points of G, $G_{\mathring{A}} = G(\mathring{A})$ is topologized as the restricted direct product of the locally compact topological groups $\{G(\mathbb{Q}_p)\}_{p \leq \infty}$ with respect to the family of compact subgroups $G(\mathbb{Z}_p)$. (For this to make sense, we choose some extension of G to a group scheme over $\mathbb{Z}$. Any two such extensions differ only in finitely many primes, so give the same topology. Perhaps it is psychologically easier to think of G as actually being a group over $\mathbb{Z}$, as will be the case in our applications.)

Theorem 6. *a)* $G(\mathbb{Q})$ *is discrete in* $G_{\overset{\circ}{A}}$.
b) $G(\mathbb{Q})\backslash G_{\overset{\circ}{A}}$ *is compact if and only if* $G_{/\mathbb{Q}}$ *is anisotropic, i.e., does not admit* $\mathbb{G}_m$ *as a subgroup.*

Now take F a totally real number field and B/F a quaternion algebra. If $[F:\mathbb{Q}] = g$, let $\infty_1, \ldots, \infty_g$ denote the real places of F. Let $G/\mathbb{Q}$ be the algebraic group whose points in a field K of characteristic zero are $G(K) = B^\times \otimes_{\mathbb{Q}} K$. Let Its derived subgroup G' is such that for all K, $G'(K)$ consists of elements of $G(K)$ of reduced norm 1. Note that G' is anisotropic over $\mathbb{Q}$ if and only if B is a division algebra.

For an adelic group, we use a subscripted 0 to denote the projection to the non-Archimedean components and a subscripted ∞ for projection onto the infinite place. Let T_0 be any compact open subgroup of G'_0, and put $T = T_0 G'_\infty$, $\Gamma_T = T \cap G'(\mathbb{Q})$. Finally, let Γ denote the projection of Γ_T onto G_∞.

Theorem 7. Γ *is a discrete subgroup of* G'_∞, *and* $\Gamma\backslash G'_\infty$ *is compact if* B *is a division algebra.*

Proof: ....

We are now ready to decode all this: label the infinite places so that B is split at the first r of them for some $0 \leq r \leq g$. Let H denote the unique division quaternion algebra over $\mathbb{R}$, and let H^1 denote the elements of reduced norm 1 in H: note that this is the unit sphere in H and is therefore compact. We have

$$G_\infty \cong GL_2(\mathbb{R})^r \times (H^\times)^{g-r}.$$
$$G'_\infty \cong SL_2(\mathbb{R})^r \times (H^1)^{g-r}.$$

We shall assume that $r > 0$, i.e., that the quaternion algebra is split at at least one real place (otherwise the aforementioned group Γ would be finite). Let us also continue to write Γ for the projection of Γ to the factor $SL_2(\mathbb{R})^r$. The key result is:

Theorem 8. $\Gamma \subset SL_2(\mathbb{R})^r$ *is discrete, and compact if* B *is a division algebra.*

Proof: This follows immediately from the preceding results and from the following general fact about groups to be found as Prop. 1.10 in Shimura's book:

Proposition 9. *Let* G_1, G_2 *be locally compact groups,* $\Gamma \subset G_1 \times G_2$ *a closed subgroup, and* Γ_1 *the projection of* Γ *to* G_1. *Suppose that* G_2 *is compact. Then:* Γ_1 *is closed in* G_1, $\Gamma\backslash(G_1 \times G_2)$ *is compact if and only if* $\Gamma_1\backslash G_1$ *is compact, and if* Γ *is discrete in* $G_1 \times G_2$, *then* Γ_1 *is discrete in* G_1.

In particular, let $\mathcal{O}$ be an order in B. We take Γ to consist of those elements in $\mathcal{O}$ with reduced norm 1. This is "the Γ" corresponding to an adelic group $T_0 = \widehat{\mathcal{O}^\times}$. Then what the proceeding theory is telling us is that Γ is a discrete subgroup of $SL_2(\mathbb{R})^r$, hence that it acts as a discrete group of transformations on r copies of the upper halfplane $\mathcal{H}^r$. Thus, if we want to maintain our current setting of discrete subgroups acting on $\mathcal{H}$, we need to enforce the condition that $r = 1$, i.e., that B is ramified at all but 1 infinite place. In other words, we get a family of compact, arithmetic Riemann surfaces $\Gamma\backslash\mathcal{H}$, the **Shimura curves**.

On the other hand, it suggests a possible broadening of perspective: taking e.g. $\mathcal{O} = M_2(\mathcal{O}_F)$ (so that B is the split quaternion algebra over F), we get a nice action of $\mathcal{O}$ on $\mathcal{H}^g$, the quotient space being a **Hilbert modular variety**. In general one gets a mixture of the two cases, or what is usually referred to as a **higher-dimensional quaternionic Shimura variety**.