Introductory Lectures on $SL(2, Z)$ and modular forms.

W.J. Harvey, King's College London

November 2008 .

Week 1.
1) We begin with a definition. The *modular group* is the subgroup $SL(2,\mathbb{Z})$ of the matrix group $SL(2,\mathbb{R})$ consisting of matrices with integer entries and determinant 1.

There is an important action of $SL(2,\mathbb{R})$ on the upper half plane $\mathcal{U} = \{z = x + iy \mid y > 0\}$, as *fractional linear* (Mobius) transformations:

$$\tau \mapsto \frac{a\tau + b}{c\tau + d}. \tag{1}$$

It is readily verified that the kernel of this action (the subgroup which acts trivially, fixing all points) is the centre $C_2 \cong \langle \pm I_2 \rangle$. The quotient group $PSL(2,\mathbb{R})$ acts faithfully on $\mathcal{U}$ and also on the boundary $\widehat{\mathbb{R}} = \mathbb{R} \cup \infty \cong S^1$; it is shown in elementary accounts of complex analysis or hyperbolic geometry (see for example [**?**], [**?**]) that the group action on $\widehat{\mathbb{R}}$ is triply transitive.

2) The hyperbolic metric ds_h^2 is given by the formula

$$ds_h^2 = y^{-2}(dx^2 + dy^2) \tag{2}$$

on $\mathcal{U}$: this is a Riemannian metric on the upper half plane. Because the action defined in equation (1) is *transitive* on the points of $\mathcal{U}$ and also on the set of unit tangent directions at each point (two simple exercises for the reader), the space is a symmetric space in the sense of differential geometry.

Invariance of the metric under the (differential of the) action by $SL(2,\mathbb{R})$ is verified by a simple calculation (Exercise 1). Thus the real Mobius transformations form the (direct) *isometry group* of $\mathcal{U}$: it is not hard to show that the action is transitive and so no larger direct isometry group is possible .

Note that the space $\mathcal{U}$ is in fact Poincaré's model of *hyperbolic plane geometry*. Geodesics between any two points are defined by circular arcs (or line segments) orthogonal to the boundary real line. From this, we are able to use geometric ideas such as polygonal shape, convexity, length and area to illuminate the group activity on $\mathcal{U}$.

3) The significance of $\mathcal{U}$: lattices in the complex plane and their classification by 'marked shape'.

The upper half plane serves as parameter space for a range of interesting objects: Gauss used it to classify binary quadratic forms, and several classical authors established a link with complex tori and elliptic functions. We look at this second aspect here.

Complex structures on a torus. The first item we consider is the shapes of complex tori, surfaces of genus 1 with a complex analytic structure.

The classic Weierstrass theory of elliptic (i.e. doubly periodic meromorphic) functions, which will be summarised later on, depends on this standard model for a torus coming from a choice of generating set for the lattice of periods, isomorphic to the (free abelian rank 2) fundamental group of the torus,

$$\pi_1(X) = \mathbb{Z} + \mathbb{Z}.$$

Underlying this, there is the important concept of a *homotopy-marking* for the surface, which underpins the theory of deformations, the rigorous study of varying shapes of torus. Intuitively, this boils down to considering the effect of changing the shape of a fundamental parallelogram tile for this lattice group of plane translations; a stricter method delivers a very precise description, a genuine space of shapes, the first *space of moduli,* the precursor of a widespread pattern of description for types of algebraic variety of specified type, a tool with great influence in algebraic geometry and elsewhere.

From the *uniformisation theorem* or, alternatively, the Riemann-Roch theorem, one sees directly that a marked complex structure on the torus is tantamount to this choice of two nonzero complex numbers $\{\lambda_1,\ \lambda_2\}$ which are linearly independent over the reals, representing the monodromy of a chosen non-trivial holomorphic 1-form around the two generating loops; this determines a lattice subgroup $\Lambda = \langle m\lambda_1 + n\lambda_2 \mid m, n \in \mathbb{Z}$ of the additive group $\mathbb{C}$ such that the complex torus is isomorphic to the quotient space $\mathbb{C}/\Lambda$. A more topological way to specify a marking begins from a choice of base point and then two simple based loops whose homotopy classes generate $\pi_1(X, x_0)$. This determines (either by lifting paths to the universal covering plane or by integration) two Euclidean line segments, joining 0 to the complex numbers λ_1 and λ_2 respectively, which may be regarded as a *geometric marking* of the torus. The standard picture of the torus $X = X(\lambda_1, \lambda_2)$ is then obtained by identification of opposite sides of the parallelogram with corners at the points $0, \lambda_1, \lambda_2$ and $\lambda_1 + \lambda_2$.

When are two marked complex torus structures equivalent*?* This means that the tori are to be conformally homeomorphic, by a mapping which is produced by a conformal (hence complex linear) map in the covering plane between the two given markings. This happens if and only if there is a nonzero complex scaling factor and perhaps a switch of ordering of the numbers, after which we may assume that the marking pair is given by $\lambda_1 = 1,\ \lambda_2 = \tau$ with $\operatorname{Im} \tau > 0$. Thus, $\tau \in \mathcal{U}$.

When do two such pairs determine the same quotient Riemann surface? The corresponding normalised lattices $\Lambda_\tau = \langle m + n\tau \mid m, n \in \mathbb{Z}$ and $\Lambda_{\tau'}$ have to coincide , which means that there are integers $a,\ b,\ c,\ d$ such that

$$1 = c\tau + d \tag{3}$$
$$\tau' = a\tau + b, \tag{4}$$

and satisfying $ad - bc = 1$, so that the process can be done in reverse. Thus the two notions of equivalence taken together produce the action [**?**] of the modular group on the space T_1 of marked complex tori, given by the action of the homogeneous modular group $\Gamma(1) = PSL_2(\mathbb{Z})$ by fractional linear automorphisms:

$$\tau \mapsto \frac{a\tau + b}{c\tau + d}$$

of the hyperbolic plane $\mathcal{U} = \{Im(\tau) > 0\}$, as in .

In this way, we encounter the classic prototype for a discrete group action, as first considered by Klein and by Poincaré, the modular group $\Gamma_1 \cong PSL(2, \mathbb{Z})$ operating on the upper half plane.

4) Classification of group elements into types or conjugacy classes, and the corresponding mappings of $\mathcal{U}$.

For this, you can either use the classification of real (invertible) matrices by eigenvalues or go by the closely related fixed point properties, which drive the present geometrical approach. Recall first that a complex Möbius transformation distinct from the identity has either one or two fixed points in the Riemann sphere $\mathbb{CP}^1$; the real coefficients force restrictions in the geometric types which occur.

Definition. A real Möbius map is called *elliptic* if it has one fixed point inside $\mathcal{U}$. It is *parabolic* if it has one boundary fixed point, *hyperbolic* if it fixes two boundary points.

Typical examples of parabolic transformation are real translations $T(z) = z + b$, with $b \neq 0$: each preserves as a set every horizontal line. This is the family of *horocycles* at ∞ . A conjugate map has the same property with respect to the family of circles tangent to the boundary circle at the fixed point.

Hyperbolic transformations are conjugate to real dilations $U_\lambda(z) = \lambda z$, with $\lambda > 1$: Each one fixes a pair of points, and preserves the hyperbolic geodesic joining them, acting on this as a hyperbolic translation from the repelling fixed point (in the examples it is 0) towards the attracting fixed point (at ∞ for U_λ). An interesting example of a hyperbolic transformation in $SL(2, \mathbb{Z})$ is the so-called Arno'ld's cat mapping) given by the matrix

$$T = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

This generates a semigroup of positive iterates, $T \circ T = T^2, T \circ T^2 = T^3, \ldots$, which distort the unit square (thought of as a tile within the real plane) by stretching in one eigendirection and shrinking in the other to give a sequence of parallelograms: when the result is projected onto the quotient torus, one obtains a dissection of the tile first of all into three triangular pieces, but then generating a more and more fragmented pattern as n grows. However, when one takes a photographic image, which is produced by an array of $m \times m$ black or white pixellated dots, this quotient map is really just a permutation of the dots, and some large enough power of it gives the identity mapping, so that the photograph reappears, a highly paradoxical effect when viewed among the surrounding chaotic patterns. Look it up in (for instance) *Anton & Rorres*'s Elementary Linear Algebra with applications (John Wiley & Sons, 2000).