Journal of the Inst. of Math. Jussieu (2005) **4**(3), 443–508 © Cambridge University Press 443
doi:10.1017/S1474748005000101 Printed in the United Kingdom

ON ARITHMETIC CURVES IN THE MODULI SPACES OF CURVES

PIERRE LOCHAK

Institut de Mathématiques de Jussieu, 175 rue du Chevaleret, F-75013 Paris, France (lochak@math.jussieu.fr)

(Received 6 February 2003; accepted 7 July 2003)

Abstract We study several types of curves and higher-dimensional objects inside the moduli spaces of curves, insisting on their arithmetic properties in the perspective of Grothendieck–Teichmüller theory. On the way we explicitly identify those curves which were originally associated by W. Veech with certain rational polygonal billiards.

Keywords: moduli spaces of curves; Grothendieck–Teichmüller theory; rational polygonal billiards

AMS 2000 *Mathematics subject classification:* Primary 14D22; 32G15; 14H30

1. Introduction

In this paper we argue, using concrete examples, that Thurston's viewpoint on the topology of surfaces and its subsequent ramifications are quite relevant when trying to explore the landscape delineated in Grothendieck's *Esquisse d'un programme*. A salient feature of Thurston's work and legacy in this area is the recognition of the importance of Teichmüller space, an essentially analytic object, in these *a priori* topological matters; one can take as a starting point the fact that the mapping class group, defined as the group of isotopy classes of diffeomorphisms of a surface, coincides with the isometry group of the corresponding Teichmüller space endowed with the Teichmüller metric. This classical result of Royden enables one to bridge the gap between topology and analysis. Here we try and go one step further, moving into the arithmetic nature of certain objects (old and new). Actually the intermediate step, namely from analytic to algebraic geometry, will turn out to be a source of beautiful and difficult problems.

We will be mostly interested here in finding and studying curves which are immersed in the moduli spaces of curves, are geodesic for the associated Teichmüller metric and can be defined by algebraic equations with coefficients in number fields. Some higher-dimensional loci will also naturally come in. In particular, we isolate two classes of curves of rather different natures. The first kind we call *origamis* because they can be defined in a purely topological and in fact combinatorial way by assembling squares of paper into a closed topological surface. A particular form of origamis was defined by W. Thurston in his famous announcement on his work on diffeomorphisms of surfaces [**68**] and they were taken up by W. Veech in [**69**]. We show here that they carry amazing arithmetic information, which we are certainly not able to decode yet, by showing that they are defined by equations with coefficients sitting in certain number fields. From now on we will call an object 'arithmetic' if it can be defined over a number field. In some sense that property makes origamis into a natural higher-dimensional analogue of the so-called 'dessins d'enfants'. In order to prevent possible misunderstandings, we note that curves sitting inside moduli spaces of curves generate an incredibly rich structure, with three types of arithmetics at work which one might call modular (Riemann), period (Jacobi) and hyperbolic (Fuchs) arithmetics. For elliptic curves they correspond respectively to the j and τ functions and to the Fricke–Klein real parameters for the Teichmüller space of the curve punctured at the origin (in order to make it hyperbolic; in general this Fuchsian arithmetic is reflected in the traces of the elements of the uniformizing Fuchsian group, which encode the Poincaré length spectrum of the underlying Riemann surface). We emphasize that we are dealing here primarily with the first type of arithmetic, the connection with the second type being notoriously mysterious.

In terms of the Thurston–Bers classification of diffeomorphisms (see [**68**] and [**17**]), which in large part extends the classification for the torus, that is for the group $SL_2(\mathbb{Z})$, origamis lie mostly on the parabolic and hyperbolic sides. The second class of objects we consider lies more on the elliptic side and we call these *eigencurves* or more generally *eigenloci* in the higher-dimensional case. Saying that they are of elliptic nature is the same as saying that they are connected with the torsion elements in the mapping class groups, or else with the automorphisms of curves. Here the situation is even more mysterious than in the case of origamis, because one starts by constructing loci in Teichmüller space and there is at present no criterion available ensuring algebraization, that is telling when the projection to moduli space is algebraic. This is a fascinating and to all appearances very deep problem which was first posed by Veech in a particular case connected with the study of rational polygonal billiards. So the objects we consider here do have, at least in a special case, a concrete incarnation in terms of dynamical systems. From a modular viewpoint, eigencurves, of which only preciously few have been detected at present, arise quite naturally, corresponding to discs in Teichmüller space which are globally but not pointwise fixed under the action of torsion elements of the mapping class group. With this viewpoint in mind we will show that the famous curves detected by Veech can in fact be written down in a completely explicit and elementary fashion, which among other things makes it plain that they too are arithmetic (defined over number fields). We hope that leaving any other motivation aside, this result can be of interest to people working in the area.

As mentioned at the very beginning the motivation behind this paper actually comes from Grothendieck–Teichmüller theory and the study of the action of the arithmetic Galois groups on the algebraic fundamental groups of certain geometric objects. We believe that the topologico-analytic framework provided by the theory developed by Thurston, H. Masur, Veech and many others can and perhaps should play an important role in trying to implement certain suggestions contained in Grothendieck's *Esquisse* (reproduced in [**63**], with an English translation), and beyond. We remark that Grothendieck alludes to the work of Thurston which was being developed at the same time that Grothendieck himself was journeying in his *Longue marche à travers la théorie de Galois* and of which he knew only through rather distant hearsay. He terms it 'chirurgie hyperbolique' (hyperbolic surgery), when it may be fair to say that he himself was after an at present still very conjectural 'chirurgie arithmétique'.

For the above reasons we have written this text having in mind a reader coming from an algebraic geometric background, with perhaps little knowledge of some fairly 'standard' analytic and dynamical systems features. We believe, however, that some results might be of interest to people coming from the latter areas, who should have no trouble skipping the sketchy reminders we have included. We have also taken advantage of the existence of efficient electronic reference databases: recalling a theme and the name of an author nowadays enables one to easily retrieve the relevant paper(s). Finally, we have written a conclusion in the form of an introduction in which we enlarge the landscape a bit, trying to convey some of the intuitions which might subtend further exploration of this truly fascinating mathematical stage. Needless to say that last section is highly not self-contained and quite partial in both senses of the word, hardly dipping into deeper waters.

2. Teichmüller discs, flat structures and geodesic curves

We have assembled in this section most notions and results of topological or analytic nature we will make use of. For background material on Teichmüller and moduli spaces of curves from an analytic viewpoint, the reader can consult any of the several existing textbooks, e.g. [**37**] (or see Chapter 1 of [**7**] for an introduction which meets our needs). We deal with Riemann surfaces of finite type and cusps (or 'punctures' or 'marked points') play an important role; they are often omitted in classical texts on Teichmüller theory (with [**14**] being a notable exception) so that they have to be restored at the appropriate places.

The standard textbooks on quadratic and Jenkins–Strebel differentials are [**65**] and [**19**]. The Teichmüller geodesic discs were first studied in detail in [**50**] where the reader can also find nice manageable versions of Teichmüller's and Strebel's main theorems. Flat structures were emphasized in this context by Veech in [**69**], where they are called F-structures. An account from a quasi-conformal viewpoint can be found in [**13**] and a nice review with proofs appears in [**52**]. For a dense survey of results with references, see the first section of [**25**]; one can also consult [**24**], [**31**], [**39**] and [**49**] as well as references therein.

Our main goal in this section is to build up the picture described in Proposition 2.10, which necessitates, however, the introduction of a fair amount of material.

2.1. Teichmüller space and its cotangent bundle

Let S be a differentiable surface of finite hyperbolic type (g, n); that is $S = S_{g,n}$ is obtained by removing n ($n \geqslant 0$) points from a closed surface of genus g ($g \geqslant 0$) and it is assumed to have strictly negative Euler characteristic, i.e. $2g - 2 + n > 0$. Let $\mathcal{T}(S)$ be the associated *Teichmüller space*. A point $t = (X, f) \in \mathcal{T}$ is determined by a Riemann surface X and a *marking*, that is a diffeomorphism $f : S \to X$. Two pairs (X, f) and

(X', f') are equivalent if $f' \circ f^{-1}$ is homotopic to an isomorphism (that is a biholomorphic map) from X to X'; the elements of $\mathcal{T}$ are equivalence classes of such pairs. The marking f may thus be regarded as defined only up to homotopy or equivalently up to isotopy. We will often write $\mathcal{T}(S) = \mathcal{T}_{g,n}$ as $\mathcal{T}(S)$ depends effectively only on the type of the model surface S, although the isomorphism $\mathcal{T}(S) \simeq \mathcal{T}(S')$ for two models S and S' is not canonical. We will abbreviate $\mathcal{T}_{g,n}$ to $\mathcal{T}$ when the type (g, n) is unambiguous. Distinguished points should be considered either as marked (marked points) or deleted (punctures or cusps). We will sometimes rely on the reader for the choice when it is made clear by the context, but for simplicity we will consistently use the following notation, thinking of punctures: if X is of type (g, n), we let $\hat{X}$ denote the associated projective curve, obtained by completing X (or erasing the marked points thinking in those terms). So $X = \hat{X} \setminus \{P_1, \ldots, P_n\}$, where the P_i are the punctures. Let $P_X = P_1 + \cdots + P_n$ denote the corresponding divisor.

The cotangent space to $\mathcal{T} = \mathcal{T}_{g,n}$ is described by the classical Kodaira–Spencer deformation theory: viewing $\mathcal{T}$ as the space of complex structures on S modulo diffeomorphisms isotopic to the identity let $\omega = \omega_{\hat{X}}$ denote the sheaf of holomorphic differentials on $\hat{X}$, ω^* its dual, namely the sheaf of holomorphic vector fields; deformation theory identifies the tangent space $T_t\mathcal{T}$ at $t = (X, f)$ as the H^1 of the holomorphic vector fields vanishing on P_X (see, for example, [**37**, § 7.2.4] for the case $n = 0$). So $T_t\mathcal{T} \simeq H^1(\hat{X}, \omega^*(-P_X))$; dualizing one finds $T_t^*\mathcal{T} \simeq H^0(\hat{X}, \omega^{\otimes 2}(P_X))$. In other words, the cotangent space to $\mathcal{T}$ at the point t can be identified with the vector space of quadratic differentials on $\hat{X}$ which are holomorphic on X and have at most simple poles. We denote this vector space by $\mathcal{Q}(X)$ or just $\mathcal{Q}$ if X is unambiguously defined. The Riemann–Roch formula enables one to compute or cross-check that $\dim_{\mathbb{C}}(\mathcal{Q}(X)) = \dim_{\mathbb{C}}(\mathcal{T}_{g,n}) = 3g - 3 + n$.

2.2. Local representation, foliations, Strebel differentials

Given X as above, any element $q \in \mathcal{Q} = \mathcal{Q}(X)$ can be written locally on X as $q = \phi(z)\,\mathrm{d}z^2$, where z is a local coordinate and $\mathrm{d}z^2$ is short for $(\mathrm{d}z)^{\otimes 2}$. Looking near $z = 0$ and after performing a locally biholomorphic change of variables, one can assume that $\phi(z) = z^k$ for some integer $k \geqslant -1$, since we are considering differentials with at most simple poles. Next $z^k\,\mathrm{d}z^2 = c(\mathrm{d}(z^{k/2+1}))^2$ (c an unimportant constant). So if k is even, and in particular if $k = 0$ (corresponding to a regular point of q), q is locally the square of a 1-form, namely $q = c\,\mathrm{d}\zeta^2$ with $\zeta = z^{k/2+1}$. If k is odd, however, in particular near a simple pole ($k = -1$), one needs to pass to a twofold cover in order to write q (locally) as the square of a 1-form. It is easy to see that this construction can in fact be globalized. Namely, given q there exists a twofold cover $\tilde{X}$ of X, say $\pi : \tilde{X} \to X$ and a 1-form $\tilde{\omega}$ such that $\tilde{q} = \pi^* q = \tilde{\omega}^2$. Moreover, *because q has at most simple poles*, the form $\tilde{\omega}$ can be extended (as well as $\tilde{q}$) to a *holomorphic* form on the completion of $\tilde{X}$. This reduces to a local analysis and boils down to the fact that $z^{-1}(\mathrm{d}z)^2 = 2(\mathrm{d}\zeta)^2$ with $z = \zeta^2$.

A quadratic differential $q \in \mathcal{Q}$ defines a *metric* with line element $\mathrm{d}s_q^2 = |\phi(z)|\,|\mathrm{d}z|^2$ and a *measure* μ_q with $\mathrm{d}\mu_q = |\phi|\,\mathrm{d}z \wedge \mathrm{d}\bar{z}$; one easily checks that these expressions are independent of the local coordinate. Because complex transformations are conformal, i.e. preserve angles, it makes sense to integrate a quadratic differential in a given direction θ. More precisely, this is well defined *modulo* π (not 2π; because $\mathrm{d}(\mathrm{e}^{\mathrm{i}\theta}z)^2 = \mathrm{e}^{2\mathrm{i}\theta}\,\mathrm{d}z^2$). So for any direction θ, i.e. any $\theta \in \mathbb{R}/\pi\mathbb{Z}$, we have a *foliation* $\mathcal{F}_\theta(q)$ whose leaves are obtained by integrating q in the direction θ, that is for $\mathrm{Arg}(q) = \theta \pmod{\pi}$. In particular, the horizontal (respectively, vertical) foliation corresponds to $\theta = 0$ (respectively, $\theta = \pi/2$). We also note that $\mathrm{Arg}(\mathrm{e}^{\mathrm{i}\theta}q) = \theta + \mathrm{Arg}(q)$, so that $\mathcal{F}_{\theta'}(\mathrm{e}^{\mathrm{i}\theta}q) = \mathcal{F}_{\theta'-\theta}(q)$. In particular, $\mathcal{F}_\theta(q)$ is the horizontal foliation of the rotated form $\mathrm{e}^{-\mathrm{i}\theta}q$. We have detailed this rather obvious action of the rotations because we will need to consider all directions simultaneously in what follows.

Looking again at the local expression $q = z^k\,\mathrm{d}z^2$, one finds that q is integrable, that is X has finite μ_q measure, if and only if q has at most simple poles. We thus record the following classical and important proposition.

Proposition 2.1. *If $t = (X, f) \in \mathcal{T}_{g,n}$ is a point in Teichmüller space, $T_t^*\mathcal{T}_{g,n} \simeq \mathcal{Q}(X)$, where $\mathcal{Q}(X) = H^0(\hat{X}, \omega^{\otimes 2}(P_X))$ is the $(3g-3+n)$-dimensional $\mathbb{C}$-vector space of integrable quadratic differentials on X.*

The space $\mathcal{Q}(X)$ is endowed with the norm defined by $\|q\| = \mu_q(X)$. We also see from the local expression that near a point $P \in X$ which is a zero of order k of q, the horizontal foliation $\mathcal{F}_0(q)$ consists of $k+2$ half-lines abutting at P and partitioning a neighbourhood into $k+2$ sectors. This extends verbatim to $k = 0$ (regular point) and $k = -1$ (simple pole). And this is of course also true for the foliations $\mathcal{F}_\theta(q)$, which are obtained by rotating the horizontal foliation.

We note that a quadratic differential q can be written as the square of a 1-form, that is $q = \omega^{\otimes 2}$ for some 1-form ω, if and only if the foliations $\mathcal{F}_\theta(q)$ are orientable—all of them or just one of them, which is equivalent. We then simply say that q is a square, or is *orientable*. For this to happen, it is of course necessary that q be *locally* orientable, that is only even values of k may occur. In particular, a differential with at least one simple pole or one simple zero is not orientable.

We now introduce the so-called *Jenkins–Strebel differentials* (we will henceforth omit the first name for the sake of brevity only), again paying attention to the poles and to the fact that we will need to deal with all directions at once. Given $q \in \mathcal{Q}(X)$ and a direction $\theta \in \mathbb{R}/\pi\mathbb{Z}$, the *critical graph* $C_\theta(q)$ is, by definition, the union of the critical leaves of the foliation $\mathcal{F}_\theta(q)$, that is the trajectories starting from or ending at the zeros and simple poles of q. Recall that there is only one trajectory starting from a simple pole of q, so that $C_\theta(q)$ may have endpoints ('loose ends'). Then we have the following definition.

Definition 2.2. The integrable quadratic differential $q \in \mathcal{Q}(X)$ is Strebel in the direction θ if the critical graph $C_\theta(q)$ is compact. We then call θ a *Strebel direction* for q.

We say simply that q is Strebel if it is Strebel for $\theta = 0$, that is if the horizontal critical graph is compact. Of course q is Strebel in the direction θ if and only if $\mathrm{e}^{-\mathrm{i}\theta}q$ is Strebel. If q is Strebel in the direction θ, the complement $X \setminus C_\theta(q)$ is a disjoint union of topological cylinders, which from a conformal viewpoint are entirely characterized by their respective moduli. We will see in § 2.6 that there is a fascinating interplay between the Strebel property and the description of the points at infinity of geodesic curves in the moduli spaces of curves.

2.3. Examples

Here we give explicit expressions for bases of $\mathcal{Q}(X)$ in the case where X is either a sphere with marked points (type $(0,n)$) or a hyperelliptic curve. Simple as they may appear these examples are good to bear in mind, given that there are very few other computable cases.

Let first X be of type $(0,n)$, that is $X = \mathbb{P}^1\mathbb{C} \setminus \{z_1, \dots, z_n\}$. Cover as usual $\mathbb{P}^1\mathbb{C}$ with two charts with local coordinates z and w, such that z (respectively, w) omits the points at infinity (respectively, the origin) and $wz = 1$. Assume first that none of the z_i is at infinity. Then any $q \in \mathcal{Q}(X)$ can be written as

$$q(z) = \left(\sum_{1 \leqslant i \leqslant n} \frac{a_i}{z - z_i} \right) \mathrm{d}z^2, \tag{2.1}$$

with n complex parameters a_i. However, these are connected by three relations which express that q must be regular at infinity, since this is assumed not to be a marked point. More precisely, for small w we have

$$\frac{1}{z - z_i} = \sum_{k \geqslant 0} z_i^k w^{k+1}$$

(if one of the z_i is 0, use the convention that $0^0 = 1$) and $\mathrm{d}z^2 = w^{-4}\,\mathrm{d}w^2$. This yields the expansion of q near infinity:

$$q(w) = \sum_{k \geqslant 0} \left(\sum_i a_i z_i^k \right) w^{k-3}\,\mathrm{d}w^2.$$

So the three relations read

$$\sum_{1 \leqslant i \leqslant n} a_i z_i^k = 0, \quad k = 0, 1, 2, \tag{2.2}$$

and we find that $\mathcal{Q}(X)$ is indeed of dimension $n - 3$, noting that the relations are linear in the parameters a_i. If one of the z_i is at infinity, say $z_n = \infty$, the above has to be modified only slightly: in (2.1) the sum runs over the values $i = 1, \dots, n-1$, whereas in (2.2) the value $k = 2$ is omitted because q may have a simple pole at infinity. In other words, one gets $n - 1$ parameters a_i and two relations.

Another basis of $\mathcal{Q}(X)$ shows itself to be useful. We assume here for simplicity that none of the z_i lies at the origin (if not, just translate) and in fact, as a first step, that $z_i \neq 0, \infty$ (all i). Note that everything we do is actually $PGL_2(\mathbb{C})$ invariant and these assumptions are simply for the sake of writing down closed formulae. We set $P(z) = \prod_i (z - z_i)$ and define

$$q_j = z^j \frac{\mathrm{d}z^2}{P(z)}, \quad j = 0, 1, \dots, n-4, \tag{2.3}$$

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

so that q_j is regular at infinity. If $z_n = \infty$ we omit it in the definition of P, which is thus of degree $n-1$ and still get the same expression for the q_j with the same values of j; the last elements q_{n-4} then has a simple pole at $z_n = \infty$. Of course these two bases are related by decomposing rational functions on X into their simple elements.

Let now X be given as the hyperelliptic curve with equation in the affine plane (x, y):

$$y^2 = \prod_{i=1}^{n}(x - x_i) = P(x). \tag{2.4}$$

We assume for simplicity that n is odd, that is the completion of X has only one point at infinity. The case of even n requires minor changes only. We write $n = 2g+1$ where $g \geqslant 1$ is the genus of X and let X_0 denote the quotient of X by the hyperelliptic involution ι. So X_0 is of type $(0, n+1)$, because ∞ is a ramified point for the involution. The marked points on the sphere X_0 are the x_i and ∞.

Now $\mathcal{Q}(X)$ can be spanned by the union of two types of differentials:

$$\mathcal{Q}(X) = \langle \{q_j; j = 0, \ldots, n-3\} \cup \{\tilde{q}_k; k = 0, \ldots, g-3\}\rangle. \tag{2.5}$$

The q_j are lifted from the ones with the q_j on X_0 (see (2.3)) and we give them the same name (rather than $\iota^* q_j$) for simplicity. So $q_j = x^j\,\mathrm{d}x^2/y^2$ and the index runs to $n-3$ because there are effectively $n+1$ marked points. The $\tilde{q}_k$ in turn do not descend to X_0 and are given by $\tilde{q}_k = x^k\,\mathrm{d}x^2/y$. That the index k runs to $k = g-3$ comes again from an easy analysis at infinity (where there is a uniformizing parameter t with $x = t^{-2}$, $y \sim t^{-(2g+1)}$). One checks that the dimension of $\mathcal{Q}$ is $d = (n-2) + (g-2) = 3g-3$ as it should be.

We finally remark that the q_j are not orientable in the genus 0 situation because they have at least one simple pole. Their holomorphic lifts q_j are indeed orientable exactly for j even, that is if the only zero, which lies at the origin, has even order. As for the $\tilde{q}_k$, they are not orientable.

2.4. Geodesic discs and flat structures

Let again $t = (X, f) \in \mathcal{T} = \mathcal{T}_{g,n}$ be a point in Teichmüller space, $q \in \mathcal{Q} = \mathcal{Q}(X)$ an integrable quadratic differential on X. For ease of notation we will sometimes identify the tangent and cotangent spaces of Teichmüller and moduli spaces.

Definition 2.3. The *Teichmüller disc* (or just 'disc') $D(t, q) \subset \mathcal{T}$ with centre t in the direction q is the complex geodesic ray for the Teichmüller distance emanating from t along q.

In real rather than complex terms, $D = D(t, q)$ is thus the union of the real geodesic rays originating from t along the tangent vectors $\mathrm{e}^{\mathrm{i}\theta} q$ as θ varies in $\mathbb{R}/\pi\mathbb{Z}$. Of course $D(t, \lambda q) = D(t, q)$ for any $\lambda \in \mathbb{C}^*$; if one assumes that q is normalized by $\|q\| = 1$, there remains the choice of a phase $\lambda \in S^1 \simeq \mathbb{R}/2\pi\mathbb{Z}$.

A concrete description of such a disc is essentially implied by Teichmüller main theorem, as made precise and revisited by Ahlfors, Bers, etc. (see, for example, [**37**] and [**50**]).

Fortunately for the non-analyst these deep analytic results actually lead to a clear geometric and algebraic picture. Let $\mathcal{U}$ be the open unit disc equipped with the Poincaré metric; we also use the upper half-plane model $\mathcal{H}$. The respective automorphism groups are $PSU(1,1)$ and $PSL_2(\mathbb{R})$ and using the classical Cayley isomorphism we pass freely from one model to the other. The first important piece of information is that the disc $D = D(t,q)$ is given as an *isometric* embedding $\mathcal{U} \hookrightarrow \mathcal{T}$ of $\mathcal{U}$ (or $\mathcal{H}$) into $\mathcal{T}$. This is a consequence of the extremality properties of such mappings (see, for example, [**37**] or [**19**]). In other words, we have the following proposition.

Proposition 2.4. *The Poincaré and Teichmüller metrics coincide on a Teichmüller geodesic disc.*

In order to make the correspondence more explicit, one can introduce the so-called *flat structures*. Let S be a genus g closed differentiable surface with a set Σ of r marked points. A flat structure on (S,Σ) is determined by an atlas (V,ϕ): the $V = (V_i)_{i\in I}$ define an open covering of $S\setminus\Sigma$; the charts $\phi_i : V_i \to \mathbb{R}^2$ are such that on the intersection $V_i \cap V_j$ the transition map $\phi_{ij} = \phi_j \circ \phi_i^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ is of the form $\phi_{ij}(z_i) = \pm z_j + c_{ij}$. Here we use complex notation, z_i (respectively, z_j), being defined by the chart ϕ_i (respectively, ϕ_j). The c_{ij} are constants. A flat structure is an equivalent class of such objects for obvious equivalence relations; for instance one can enlarge the atlas by adding redundant charts, etc. The point is that the transition maps are restricted to be translations, possibly composed with reflections. We denote such a structure by u or more precisely by (S,Σ,u).

Using complex notation was not an abuse because a flat structure defines in particular a *complex structure* on $S\setminus\Sigma$: indeed the transition maps are holomorphic. Moreover, one can endow $S\setminus\Sigma$ with a metric obtained by locally pulling back the ordinary Euclidean metric from the plane $\mathbb{R}^2$: indeed the transition maps are isometries of that metric. The surface S then acquires a *Euclidean structure with cone-type singularities*; the singularities occur at the points of Σ (see below for more). In particular, the Gaussian curvature is zero except at the singular points, hence the name 'flat structures'. Finally, again because of the form of the transition maps, *lines* or directions $\theta \in \mathbb{R}/\pi\mathbb{Z}$ are well defined on (S,Σ,u). These three properties are the main ingredients, together with the correspondence between flat structures and quadratic differentials. Indeed, we have the following proposition.

Proposition 2.5. *Flat structures are in one-to-one correspondence with quadratic differentials.*

Let us describe this simple but fundamental correspondence. First, starting from (S,Σ,ϕ), the expression $\mathrm{d}z^2$ is well defined away from the singularities, precisely because of the form of the transition maps: $z \to \pm z + c$. So one can glue these local expressions ($\mathrm{d}z_i^2 = \mathrm{d}z_j^2$) into a quadratic differential $q(u)$ on $S\setminus\Sigma$. Then at a cone point $P\in\Sigma$, the total cone angle has to be a multiple of π as otherwise the notion of direction would not be defined. This of course also results from the form of the transition maps. Denoting this angle by $2c\pi$ for some half-integer c, one finds that the quadratic differential $q(u)$ has order $k = 2(c-1)$ at P, this k being as at the beginning of § 2.2. This is a consequence of the local behaviour at P; it includes the cases $k = -1$ ($c = \frac{1}{2}$, a simple pole) and $k = 0$ ($c = 1$, an apparent singularity).

One can compute the genus $g(S)$ of S using the Gauss–Bonnet or Euler formula. Each $P \in \Sigma$ contributes a quantity $\kappa_P = 2\pi - 2c_P\pi = -k_P\pi$ to the curvature, where $2c_P\pi$ is the cone angle at P and k_P is the order of $q(u)$. Only simple poles produce positive curvature. By Gauss–Bonnet the local contributions add up to $2\pi\chi(S) = 2\pi(2 - 2g(S))$. Hence the useful formula:

$$2g(S) - 2 = \Sigma_{P\in\Sigma}(c_P - 1) = \tfrac{1}{2}\Sigma_{P\in\Sigma}k_P. \tag{2.6}$$

Of course the equality of the first and last terms simply says that the divisor of q has degree $4g(S) - 4$, as it should since this is twice the degree of the canonical bundle.

In the opposite direction, start from X of type (g, n) with a quadratic differential q having at most simple poles at the n marked points. Let $\Sigma_q \subset \hat{X}$ be the union of the divisor of q together with the marked points, i.e. the divisor P_X. Generically with respect to q the set Σ_q is just the divisor of q but we prefer to include the marked points by convention, at the expense of possible apparent singularities. We let $X_q = \hat{X} \setminus \Sigma_q$. If m is the number of zeros of q, not counting multiplicities, the surface X_q is thus of type (g, r) with $r = m + n$.

On X_q the differential q can be locally written as $q = \mathrm{d}z^2$ and the local coordinate z is defined exactly up to a sign and an additional constant. So these local coordinates glue into a flat structure $u(q)$ on X_q. So given q and X we find a flat structure $(\hat{X}, \Sigma_q, u(q))$. The constructions sketched above are clearly inverse of each other.

One more word about orientability. Define a flat structure to be *orientable* if and only if there exists an atlas such that all the transition maps are translations, i.e. one can get rid of the signs $\pm$. It is clear from the discussion above that q and $u(q)$ (respectively, u and $q(u)$) are simultaneously orientable or non-orientable.

We may now return to the concrete description of the Teichmüller discs. The main point is to define an action of $PSL_2(\mathbb{R})$ on flat structures as follows. Given (S, Σ, u) a flat structure determined via an atlas as above and $A \in PSL_2(\mathbb{R})$, the flat structure $(S, \Sigma, u' = Au)$ is defined by postcomposing charts with A. Warning: this operation *does* modify the complex structure and this is actually the whole point. We will nonetheless use the customary complex notation: the symbol Az actually denotes the 2-vector obtained by applying A to $z = (x, y)$. Moreover, Az is defined only up to sign since A is. These ambiguities cancel in what follows. So if the z_i (respectively, z'_i) are local coordinates for u (respectively, u') with charts $\phi_i : V_i \to \mathbb{R}^2$ (respectively, ϕ'_i and the same V_i) one has $\phi'_i = A \circ \phi_i$, that is $z'_i = Az_i$. Then over $V_i \cap V_j$, where $z_j = \pm z_i + c_{ij}$, this gives $z'_j = \pm z'_i + Ac_{ij}$.

When is it that $u' = Au$ defines the same flat structure as u? Clearly, this happens if and only if A is complex analytic, that is if and only if A is a rotation: $A = (\pm)\mathrm{e}^{\mathrm{i}\theta}$ in complex notation. Then the local change of coordinates is just $z' = \mathrm{e}^{\mathrm{i}\theta}z$. Now start from a (co)tangent vector in the Teichmüller space $\mathcal{T}_{g,n}$ of type (g, n), so from a triplet (X, f, q) where X is of type (g, n), f is a marking from a model surface and q is an integrable quadratic differential on X. We get a flat structure $u(q)$ on X_q as above and we can consider its orbit under the action of $PSL_2(\mathbb{R})$. This orbit naturally lives in the (co)tangent space $T^{(*)}\mathcal{T}_{g,r}$ and we may project it down to $\mathcal{T}_{g,r}$. By the remark above the action actually factors and we get an action of $\mathcal{H} = PSL_2(\mathbb{R})/PSO_2(\mathbb{R})$. Teichmüller's fundamental theorem implies the following proposition.

Proposition 2.6. *The orbit $\mathcal{H} \cdot u(q)$ coincides with the Teichmüller disc $D(t, q)$.*

Together with Propositions 2.4 and 2.5, this provides the desired description of Teichmüller discs. In particular, $D(t, q)$ can be viewed as the set of complex structures which underlie the flat structures $Au(q)$, as A runs through $PSL_2(\mathbb{R})$, the stabilizer of X being $PSO_2(\mathbb{R}) \simeq \mathbb{R}/\pi\mathbb{Z}$. Note that if $A = e^{i\theta}$, we get $z' = e^{i\theta} z$, hence $(dz')^2 = e^{2i\theta}\, dz^2$, that is $q' = e^{2i\theta} q$. So $\theta \in \mathbb{R}/\pi\mathbb{Z}$ fixes the complex structure and acts on q via $q \to e^{2i\theta} q$. There is also an action of the rotation group on the disc $D(t, q)$ (just as on the unit disc $\mathcal{U}$ or on $\mathcal{H}$) defined simply by $q \to e^{i\theta} q$ for the rotation of angle θ.

We add one more word in order to make the above procedure quite precise. In effect the flat structure $u(q)$ has been constructed on X_q, which is of type (g, r) with $r = m+n$, m being the number of zeros of q. We already have a marking f for the surface X of type (g, n); it can be transformed into a marking of X_q, by modifying the marking f so as to fix the additional marked points. This operation is not canonical but this ambiguity will play no role here. By doing so we have picked a point in $\mathcal{T}_{g,r}$; consider then the orbit of $u(q)$ in that space, then map it to $\mathcal{T}_{g,n}$ using the forgetful map $\mathcal{T}_{g,r} \to \mathcal{T}_{g,n}$ which erases the last m points. This map is *not* an isometry but it is indeed an isometry between these two discs, because the $PSL_2(\mathbb{R})$ action fixes the zeros of q. So the flat structure $u(q)$ does describe a geodesic disc in $\mathcal{T}_{g,n}$. We will return to this in some more detail in § 2.7 below.

2.5. From Teichmüller to moduli space

We now want to descend the situation from Teichmüller to moduli space. Let us first fix notation; $S = S_{g,n}$ is a model surface of type (g, n), $\Gamma(S) = \pi_0(\mathrm{Diff}^+(S))$ is the associated mapping class group or *Teichmüller modular group* (modular group for short), namely the group of isotopy classes of diffeomorphisms of S fixing the marked points or punctures pointwise; that is we consider here the *pure* modular group (see § 2.7 for variants). We also write $\Gamma(S) = \Gamma_{g,n}$ or just Γ, although this is a story with too many traditional Γ and we will try to avoid confusion. This group acts on $\mathcal{T}(S)$ as follows: if $t = (X, f) \in \mathcal{T}(S)$ and $g \in \Gamma(S)$, then $g \cdot t = (X, f \circ g^{-1})$. The action is proper and discontinuous and the *moduli space* $\mathcal{M}(S)$ is the quotient $\Gamma(S)\backslash\mathcal{T}(S)$, which in this section is viewed as a complex orbifold. We write $\mathcal{M}(S) = \mathcal{M}_{g,n} = \mathcal{M}$ and denote by $\pi : \mathcal{T} \to \mathcal{M}$ the canonical projection.

The modular groups also act naturally on flat structures. Namely, let (S, Σ, u) be a flat structure, where now $S = S_g$ is compact of genus g and Σ is a set of r distinct points. Let $\Gamma(S, \Sigma)$ be the group of isotopy classes preserving the elements of Σ pointwise: $\Gamma(S, \Sigma) \simeq \Gamma_{g,r}$. Suppose the flat structure is given as in § 2.4 by an atlas (V, ϕ); then if $g \in \Gamma(S, \Sigma)$, the flat structure $g \cdot u$ on (S, Σ) is defined by the atlas $(g(V), \phi \circ g^{-1})$. That is if $\phi_i : V_i \to \mathbb{R}^2$ is a chart for u, $\phi_i \circ g^{-1} : g(V_i) \to \mathbb{R}^2$ is a chart for $g \cdot u$. We isolate the following obvious but important proposition.

Proposition 2.7. *The* (*left*) *action of the modular group on flat structures commutes with the* (*right*) *action of* $PSL_2(\mathbb{R})$.

Consider now a Teichmüller disc $D(t,q) \subset \mathcal{T}$ with $X = \pi(t) \in \mathcal{M}$. Here we consider that $\mathcal{T} = \mathcal{T}_{g,r}$, adding the m zeros of q to the n original marked points; again see § 2.7 about erasing the m additional points. The modular group $\Gamma_{g,r}$ acts on $\mathcal{T}_{g,r}$ and we can look at the stabilizer $S(D) = \mathrm{Stab}_\Gamma(D) \subset \Gamma_{g,r}$, where stabilizer means *global* stabilizer, not pointwise. There may actually be a non-trivial normal subgroup $H \subset S(D)$ which fixes D *pointwise*, and H is *finite* because it is contained in the finite automorphism group of the hyperbolic Riemann surface X (for more on this, see § 4). We let $\Gamma(D) = \mathrm{Stab}_\Gamma(D)/H$. So we have an exact sequence:

$$1 \to H \to S(D) \to \Gamma(D) \to 1, \tag{2.7}$$

which in general is not split. As for $H \subset \Gamma_{g,r}$, it is the group of the common automorphisms to all the Riemann surfaces underlying the elements of D, i.e. all points of $\pi(D)$; here H should be regarded only up to conjugacy in Γ and $\pi(D) \subset \mathcal{M}(H)$, anticipating notation of § 4. That is, $\pi(D)$ is contained in the special locus defined by the conjugacy class of H as a finite subgroup of $\Gamma_{g,r}$.

The next step consists of observing that there is a canonical embedding $\mu : \Gamma(D) \hookrightarrow PSL_2(\mathbb{R})$, namely the *monodromy representation* or map, which in the present terms is obtained as follows. Given $g \in \Gamma(D)$, we can let it act on the flat structure $u(q)$, as explained above. The description of D in terms of an action of $PSL_2(\mathbb{R})$ given in § 2.4 shows that D is globally fixed under the action of g if and only if there exists an element $A = A(g) \in PSL_2(\mathbb{R})$ such that $g \cdot u(q) = A(g)u$. Locally on a chart (V, ϕ), this means that $\phi \circ g^{-1} = A \circ \phi$; in other words, g acts in an affine way on the Riemann surface X considered as a Euclidean surface with singularities. More exactly there must be a diffeomorphism representing the isotopy class g which effects this and which we still denote by g. Because of the form of the transition maps for a flat structure, A is clearly independent of the chart. So one can define μ by $\mu(g) = A(g) \in PSL_2(\mathbb{R})$ and this map is into by the very definition of $\Gamma(D)$.

We also note (see [**69**, Proposition 2.7]) that $\mu(\Gamma(D))$ is a *discrete* subgroup of $PSL_2(\mathbb{R})$. In the sequel we will often identify $\Gamma(D)$ and $\mu(\Gamma(D))$ calling it the *affine group* of the disc D; up to an extension by a finite group it coincides with the stabilizer of D under the action of the Teichmüller modular group. We summarize the above for clarity in the following proposition.

Proposition 2.8. *The following are naturally associated with a Teichmüller disc* $D = D(t,q) \subset \mathcal{T} = \mathcal{T}_{g,r}$.

(i) *Two subgroups* $S(D)$ *and* H *of the Teichmüller modular group* $\Gamma = \Gamma_{g,r}$, *the group* H *being finite.*

(ii) *An injective map* $\mu : \Gamma(D) \hookrightarrow PSL_2(\mathbb{R})$ *of their quotient* $\Gamma(D) = H \backslash S(D)$ (*called the affine group of* D); *the image of* μ *is discrete and not cocompact.*

This natural interplay between surface groups sitting inside the Teichmüller modular groups and discrete subgroups of $PSL_2(\mathbb{R})$ provided by the monodromy representation looks to us to be relevant and indeed fascinating from the viewpoint of Grothendieck–Teichmüller theory.

Generically with respect to Lebesgue measure on the unit cotangent bundle of Teichmüller space the stabilizer $S(D)$ is in fact the trivial group, so are *a fortiori* H and $\Gamma(D)$. We will be interested in the cases, which are highly exceptional from the measure-theoretic viewpoint and have been highlighted by Veech, where $\Gamma(D)$ is as 'as large as possible', namely is a lattice, meaning that it has finite covolume in $PSL_2(\mathbb{R})$. The canonical projection $\pi(D)$ to moduli space is given as $C = C(X, q) = \Gamma(D)\backslash D \simeq \Gamma(D)\backslash\mathcal{H}$ and the point is that of course *C is a curve if and only if $\Gamma(D)$ is a lattice*. Note that $\Gamma(D)$ may contain elliptic elements, so that C should be regarded as an *orbifold*. Actually it is better to keep track of H by viewing C as the quotient $S(D)\backslash\mathcal{H}$, where the action of H is, however, not effective. This generalizes the case of the moduli of elliptic curves: one has $\Gamma_{1,1} \simeq SL_2(\mathbb{Z})$, rather than $PSL_2(\mathbb{Z})$, with the non-split exact sequence (2.7), where $H = \mathbb{Z}/2\mathbb{Z}$ has the elliptic involution as its only non-trivial element. It also generalizes to special loci (see § 4 below). In fact it is always best to think in terms of orbifolds (and their algebraic counterparts, namely stacks) which keep track of the automorphisms of the various objects, which here carry very significant information.

Before turning to the study of these curves or orbifolds C, we note that this situation of a disc with a large stabilizer represents a small and rather exceptional part of the flat structures, corresponding to a countable dense set of the cotangent bundle of Teichmüller space [**70**]. There is at present a vast literature on these matters, falling naturally into two categories: the first one pertains more to ergodic theory, exploring the generic translation surface and studying in particular the geodesic flow on the cotangent bundle of Teichmüller space (for recent progress and references, see [**16**]); the second type of study explores precisely the situation we are interested in, of discs projecting to curves in moduli space. The problem of detecting these algebraic, and *a fortiori* arithmetic leaves inside the Teichmüller geodesic foliation looks very interesting and difficult. It also seems to require new ideas and techniques, going beyond the ones existing at present, even under the geometric and general form developed in [**6**] (see also [**9**] for a summary). An important point is that we are not working with algebraic foliations and certainly not algebraically integrable ones; we are trying to detect certain algebraic leaves among highly non-algebraic and in fact uniquely ergodic ones.

2.6. Geodesic curves and Strebel directions

In this section, we use the constructions introduced above in order to give a modular description of results which are in essence due to Veech with hints in Thurston's famous delayed announcement (see [**68**, p. 430]). We have cast them in a form which should be more to the liking of algebraic geometers and especially practitioners of Grothendieck–Teichmüller theory than the usual phrasing in topological and dynamical systems parlance. Arithmetic properties will be added later on.

Let us first briefly repeat the set-up for convenience. We start from a Riemann surface X of type (g, n) and $q \in \mathcal{Q} = \mathcal{Q}(X)$ an integrable quadratic form. Let X_q be the associated curve of type (g, r) adding in as marked points the zero set of q. Let $u(q)$ denote the corresponding flat structure. We get a point on $\mathcal{M}_{g,r}$; we choose a marking $f : S = S_{g,r} \to X_q$ and lift it to a point $t = (X_q, f) \in \mathcal{T}_{g,r}$. The arbitrariness of f does not play a serious role in what follows. We thus construct a disc $D = D(t, q) \subset \mathcal{T}_{g,r}$, with associated affine group $\Gamma(D) \subset \Gamma_{g,r}$; there is also a natural injective map $\mu : \Gamma(D) \hookrightarrow PSL_2(\mathbb{R})$, whose image is discrete. Let again $\pi(D) = C = C(X, q) = \Gamma(D) \backslash \mathcal{H}$ be the image of D in $\mathcal{M}_{g,r}$ via the canonical projection. For the time being we do not use the natural orbifold structure of C.

Let $a \in \Gamma(D)$ be such that $\mu(a) = A \in PSL_2(\mathbb{R})$ is a *non-trivial idempotent*, i.e. the eigenvalues of A are ± 1 and $A \neq \pm 1$ (everything is up to sign). Then A acting on $\mathcal{H}$ has exactly one fixed point at infinity or, to put it differently, there is a non-zero $v \in \mathbb{R}^2$, unique up to scalar multiplication, such that $Av = \pm v$. Let $\theta \in \mathbb{R}/\pi\mathbb{Z}$ be the direction of v, where $\mathbb{R}/\pi\mathbb{Z}$ is viewed as the space of lines in $\mathbb{R}^2$ passing through the origin. Then we have the following proposition, whose second part will be elucidated further after the statement.

Proposition 2.9. *Under the above assumptions, θ is a Strebel direction for q. Moreover, there is a power of a which is a product of powers of Dehn twists along the cylinders determined by the critical graph $C_\theta(q)$.*

For the proof, which is actually not difficult, we refer to [**69**, Proposition 2.4]; this proposition appears already in [**68**, p. 430] and in a paper by I. Kra (see reference in [**69**]). It is one of the key facts of the theory, and it does *not* depend on the global structure of $\Gamma(D)$, in particular whether its image via μ is a lattice or not.

Recall from § 2.2 that a *critical trajectory* in the direction θ is a leaf of the foliation $\mathcal{F}_\theta(q)$ which contains a zero or a pole of q in its closure. A *connection* is a critical trajectory which joins two such points. The first part of the proposition thus asserts that *any critical trajectory in the direction θ is a connection* and in fact any leaf of $\mathcal{F}_\theta(q)$ is either closed or a connection. The element a preserves the foliation, permuting connections. Since there are finitely many of them, there is a power of a which fixes the connections and this is the one which is considered in the second part of the lemma.

Let us detail this second part, introducing notation which will be used in the sequel. The complement $X_q \setminus C_\theta(q)$ is a disjoint union of cylinders $c_1, \dots, c_t$. Pull back the c_i to the surface S via the marking f so that the preimage $s_i = f^{-1}(c_i)$ is a topological cylinder on S. Its fundamental group is generated by an isotopy class α_i of simple loops, defining the *core* of s_i. The collection $\underset{\sim}{\alpha} = (\alpha_1, \dots, \alpha_t)$ is called a *multicurve* on S. Generally speaking, a multicurve is a collection of non-intersecting simple loops which are mutually non-isotopic and are not isotopic to the trivial loop or to loops surrounding the punctures. They are always considered up to isotopy; so non-intersecting means that there are representatives which are non-intersecting, not mutually isotopic means they are not equal as isotopy classes, etc. Below we usually skip these distinctions and sometimes also identify the situation on X_q and on S; this effectively amounts to considering the situation up to conjugacy in $\Gamma_{g,r}$, which is the defining difference between Teichmüller and moduli spaces, namely the forgetting of the marking.

Returning to the multicurve $\underset{\sim}{\alpha} \subset S$, we let $\tau_i \in \Gamma_{g,r}$ denote the Dehn twist along α_i, diffeomorphisms being also regarded up to isotopy. Since the α_i are disjoint, the τ_i commute. The second assertion of the proposition says that in $\Gamma_{g,r} = \Gamma(S)$ one has

$$a^{\ell} = \prod_{i=1}^{i=t} \tau_i^{\ell_i}, \tag{2.8}$$

for some positive numbers ℓ and ℓ_i $(i = 1, \dots, t)$.

Because of the importance of that proposition, we sketch the appealing and simple image it provides in terms of flat structures. Let $u(q)$ be the flat structure on X_q, which can now be viewed as a Euclidean surface with conic singularities. Directions are well defined on X_q and for any $\theta \in \mathbb{R}/\pi\mathbb{Z}$ the leaves of the foliation $\mathcal{F}_\theta(q)$ are mapped into straight lines. Following Thurston (see [**24**] for details) one defines a *developing map* $\mathrm{Dev} : \tilde{X}_q \to \mathbb{R}^2$, where $\tilde{X}_q$ is the universal cover of X_q equipped with the pulled-back flat structure. The map 'Dev', which is defined up to a translation, is obtained by continuation along straight lines starting from point in $\tilde{X}_q$. It can be completed by adding in the singularities $\Sigma_q \subset X_q$ and their preimages $\tilde{\Sigma}_q \subset \tilde{X}_q$. Critical trajectories now appear as open half-lines or half-segments originating at points of Σ_q; they can also be lifted to $\tilde{X}_q$ and developed via the map 'Dev' on the plane $\mathbb{R}^2$. Connections are open segments joining two points of Σ_q or $\tilde{\Sigma}_q$; they can of course return to the same point on X_q, not so on $\tilde{X}_q$ because they are homotopically non-trivial on X_q. On the plane $\mathbb{R}^2$, the first half of the proposition says that starting from *any* point in $\mathrm{Dev}(\tilde{\Sigma}_q)$ in the direction θ, one eventually encounters another such point. The set $\mathrm{Str}(q) \subset S^1 = \mathbb{R}/\pi\mathbb{Z}$ of the Strebel directions of q coincides with the set defined by the directions of the eigenvectors of the non-trivial unipotent elements in $\Gamma(D)$. This is a *subset* of the directions of the connections of q, which in turn is a certain subset of the directions of the lines connecting the array of points $\mathrm{Dev}(\tilde{\Sigma}(q)) \subset \mathbb{R}^2$.

Let us now see how this and further information can be viewed in terms of curves in the moduli space, introducing the further assumption that $\Gamma(D)$ $(D = D(X, q))$ is a lattice. For simplicity we write $\mathcal{M} = \mathcal{M}_{g,r}$ and we let $\bar{\mathcal{M}}$ denote the stable (Deligne–Mumford) completion of $\mathcal{M}$, constructed here if one wants in an analytic way, as in the work of L. Bers, by adding Riemann surfaces with nodes to $\mathcal{M}$. Then we have the following proposition.

Proposition 2.10. *Assume that C is a curve, i.e. assume that $\mu(\Gamma(D)) \subset PSL_2(\mathbb{R})$ is a lattice. Then the following hold.*

(i) *The curve $C = \pi(D)$ is* not *complete; let $\bar{C}$ denote its completion. The set of cusps $\bar{C} \setminus C \subset \bar{\mathcal{M}} \setminus \mathcal{M}$ consists of a finite number $k \geqslant 1$ of points on the divisor at infinity $\mathcal{D} = \bar{\mathcal{M}} \setminus \mathcal{M}$.*

(ii) *The only singularities of C, if there are any, are transverse self-intersections.*

(iii) *For any direction $\theta \in \mathbb{R}/\pi\mathbb{Z} \simeq S^1$, either θ is a Strebel direction or the foliation $\mathcal{F}_\theta(q)$ on X is uniquely ergodic. The set $\mathrm{Str}(q)$ of Strebel directions coincides with the set of directions of the connections of q. It is countably infinite and dense on the circle $S^1 \simeq \mathbb{R}/\pi\mathbb{Z}$.*

(iv) *$\Gamma(D)$ acts naturally on $\mathrm{Str}(q)$ and there are k orbits under this action, corresponding to the cusps of C.*

(v) *Let $\theta \in \mathrm{Str}(q)$ be a Strebel direction and $\underset{\sim}{\alpha}(\theta)$ be the multicurve on X defined by the cores of the cylinders in the decomposition induced by the critical graph $C_\theta(q)$. Then θ determines a cusp $P \in \bar{C} \setminus C$ and P lies on the stratum of the divisor at infinity $\bar{\mathcal{M}} \setminus \mathcal{M}$, which is labelled by $\underset{\sim}{\alpha}(\theta)$.*

(vi) *The cyclic parabolic subgroup of $\mu(\Gamma(D)) \subset PSL_2(\mathbb{R})$ determined by P is generated by a unipotent element A. Setting $a = \mu^{-1}(A) \in \Gamma_{g,r}$, there is an integer ℓ such that a^ℓ is a product of powers of the commuting Dehn twists along the elements of the multicurve $\underset{\sim}{\alpha} = \underset{\sim}{\alpha}(\theta)$.*

(vii) *The moduli of the cylinders determined by $C_\theta(q)$ are mutually commensurable.*

Remark. The first part of (iii) is the so-called 'Veech dichotomy' (cf. [**69**, Proposition 2.11]). Comparing with Proposition 2.9 above and its rephrasing before the statement of the present proposition, we see that *under the assumption that $\Gamma(D)$ is a lattice*, the set $\mathrm{Str}(q)$ of Strebel directions now *coincides* with the set of directions of the connections for q: as soon as one critical leaf of the foliation $\mathcal{F}_\theta(q)$ has compact closure, the whole critical graph $C_\theta(q)$ enjoys the same property.

With the amount of information introduced above, a large part of the assertions of the proposition are easy. Specifically, part (i) has been mentioned already; that is $\Gamma(D)$ is discrete but *not* cocompact. It is also a consequence of (iii) since there always exists at least one connection (think, for example, in terms of flat structures). As for (ii), the only singularities of C, as an immersed curve, come from self-intersections, and these are transverse because C is geodesic. We have already mentioned the first assertion in (iii) as being classical. Again the main point with respect to to the lemma consists of showing that *if* $\Gamma(D)$ is a lattice, any connection is associated with a parabolic element of $\Gamma(D)$. Recall that a foliation is uniquely ergodic if it has exactly one transverse invariant measure. See, for instance, [**52**] for a historical survey of this material.

The group $\Gamma(D)$ acts on the set of lines through the origin, as a subgroup of $PSL_2(\mathbb{R})$. It also acts via Möbius transformations on the points at infinity of $\mathcal{H}$. The last part of (iii) and assertion (iv) are classical in the second setting because $\Gamma(D)$ has finite covolume (is Fuchsian of the first kind). Comparing the two actions, one gets these assertions for the first action too.

Assertion (v) should be clear from the discussion following Proposition 2.9. Let us denote by $\mathcal{A}_{g,r}$ the set of multicurves on $S = S_{g,r}$. It can be made into a finite-dimensional, but not locally finite, CW-complex. Its role was first emphasized by W. Harvey and it has played, together with several variants, a prominent role in the topological study of mapping class groups, Teichmüller spaces, etc. (see, for instance, [**51**]; see also § 3.2 below). There is a natural action of the modular group $\Gamma_{g,r}$ on $\mathcal{A}_{g,r}$ where for $h \in \Gamma_{g,r}$ and $\underset{\sim}{\alpha} \in \mathcal{A}_{g,r}$, the element $h \cdot \underset{\sim}{\alpha} \in \mathcal{A}_{g,r}$ is obtained by considering h as a diffeomorphism up to isotopy and letting it act on the components of $\underset{\sim}{\alpha}$. We now recall that $\bar{\mathcal{M}}_{g,r}$ has a natural stratification, which may be called the *stable stratification*, whose strata are labelled by the multicurves modulo action of the modular group, that is by $\Gamma_{g,r} \backslash \mathcal{A}_{g,r}$. The stratum labelled by the orbit of a multicurve $\underset{\sim}{\alpha}$ corresponds to the Riemann surfaces pinched along the loops of $\underset{\sim}{\alpha}$, which are replaced with nodes. The open stratum $\mathcal{M}_{g,r}$ of smooth curves corresponds to $\underset{\sim}{\alpha} = \emptyset$. Now (v) results from the fact that going to infinity in a Strebel direction, θ consists of pinching the geodesics uniquely associated with the multicurve $\underset{\sim}{\alpha}(\theta)$ (see, for example, [**50**] for details). Again in the moduli space $\mathcal{M}_{g,r}$ the direction $\theta \in \mathrm{Str}(q)$ is actually defined only up to the action of $\Gamma(D)$ on $\mathrm{Str}(q)$ and the multicurve $\underset{\sim}{\alpha}(\theta)$ up to action of the modular group, both being well defined only in Teichmüller space.

Assertion (vi) is a restatement of the second part of Proposition 2.9, which we have included for completeness. Finally, assertion (vii) is a consequence of (vi) (see [**69**, § 9]), more precisely of the existence of the parabolic element a^ℓ. Indeed the powers $\tau_i^{\ell_i}$ (cf. (2.8) above) of the twists on the cylinders have to be induced by one and the same affine map $A^\ell \in PSL_2(\mathbb{R})$ and this implies that the ratios of the moduli are rational (see also [**72**, § 4.2]). This last assertion (vii) gives a powerful *necessary* condition for $\Gamma(D)$ to be a lattice; indeed for *any* Strebel direction, the moduli of the cylinders of the associated decomposition must have commensurable ratios. This is the 'Veech criterion', which seems to be the only general tool at present for determining whether or not $\Gamma(D)$ is a lattice. It rules out many possible candidates but making sure that the ones which passed the test do give rise to lattices is quite difficult and has been done only in a few cases which more or less all follow the same pattern (see § 5 below).

Proposition 2.10 illustrates the interplay between the behaviour of a quadratic differential q on X and the associated curve $C(X, q)$ in the moduli space of X (with added marked points at the zeros of q), when the latter curve indeed exists, i.e. when the geodesic disc $D(t, q)$ (for any t in the fibre $\pi^{-1}(X) \subset \mathcal{T}$) projects to a curve. This is only part of the story. In particular, it focuses essentially on what happens at infinity so that it explores only part of the *parabolic* aspects of the situation. We will have occasions to present glimpses of the elliptic and hyperbolic aspects later.

2.7. On permuting, adding and erasing marked points

Up to now we have used pure (or coloured) Teichmüller modular groups (or mapping class groups) only, and have also added marked points (punctures) to the curves (surfaces). We now say a few words about these operations, which are sometimes far from being innocuous, or even well understood.

Let us first say a word about allowing permutations of marked points. This is for instance unavoidable in genus 0, because then permutations of the marked points are the only possible automorphisms. It is also useful to bear in mind in general because many constructions have natural amplifications in that direction. Allowing permutations means that the marked points are considered only setwise. Namely, if X is of type (g, n), the data now consist in the complete curve $\hat{X}$ together with the divisor P_X. This gives rise to the moduli space we denote by $\mathcal{M}_{g,[n]}$, the notation being intended to suggest that points are considered as a set. The space $\mathcal{M}_{g,[n]}$ is again considered here as a complex orbifold and there is a natural orbifold Galois cover $\mathcal{M}_{g,n} \to \mathcal{M}_{g,[n]}$ with covering group $\mathcal{S}_n$, the permutation group on n objects. We denote by $\Gamma_{g,[n]}$ the orbifold fundamental group of $\mathcal{M}_{g,[n]}$: $\Gamma_{g,[n]} = \pi_1^{\mathrm{orb}}(\mathcal{M}_{g,[n]})$. The above cover then gives rise to the short exact sequence

$$1 \to \Gamma_{g,n} \to \Gamma_{g,[n]} \to \mathcal{S}_n \to 1. \tag{2.9}$$

We note that the notion of colour disappears at the level of Teichmüller spaces, or rather the points can be considered as coloured *ipso facto*. In other words, there is just one Teichmüller space $\mathcal{T}_{g,n}$: $\mathcal{T}_{g,[n]} \simeq \mathcal{T}_{g,n}$. This isomorphism is not unique; it is realized by choosing a labelling of the marked points at a given point $t = (X, f) \in \mathcal{T}_{g,n}$, so that the choices are parametrized again by $\mathcal{S}_n$, and then follow it via the marking, which is possible since everything is defined up to isotopy. We note finally that there are of course partially coloured versions, obtained as the preimage of subgroups of $\mathcal{S}_n$ in (2.9). In particular, one can consider groups of the form $\mathcal{S}_1 \times \mathcal{S}_2 \times \cdots \times \mathcal{S}_p$, where $\mathcal{S}_i$ is the group of permutation on n_i objects and $n_1 + \cdots + n_p = n$. It amounts to allowing only permutations among points of a certain type. For instance, if k is again the order of q at a marked point, one allows only permutations of points with the same order. This completes our short reminder about permutations.

Adding and erasing marked points is another natural, more mysterious operation. One useful remark to start with: adding marked points can be used in order to rigidify the situation, as an alternative to using level structures. Precisely, by using the Riemann–Hurwitz formula, it is easy to see that *a non-trivial automorphism of a smooth projective curve of genus g over a field of characteristic* 0 *fixes at most* $2g + 2$ *points*. This translates into the fact that $\mathcal{M}_{g,n}$ *is a smooth variety for* $n > 2g + 2$. In other words, it is an orbifold with trivial orbifold structure because objects have no automorphisms. In terms of algebraic geometry $\mathcal{M}_{g,n}$ can be viewed as a smooth $\mathbb{Q}$-scheme (a stack with trivial stack structure) for $n > 2g+2$. This can be extended to $\mathbb{Z}$, but with a bigger lower bound on n.

We now come to the question of the connection between marked points and geodesics, which is quite relevant for our purpose. There is a natural fibration $p : \mathcal{T}_{g,n+1} \to \mathcal{T}_{g,n}$ defined by 'erasing the last point'. That is, given a marked curve of type $(g, n + 1)$, one gets a marked curve of type (g, n) by erasing the last point and keeping the same diffeomorphism as marking. This 'forgetful functor' induces the fibration p. The main point here is that p is *not* geodesic for the respective Teichmüller metrics on $\mathcal{T}_{g,n+1}$ and $\mathcal{T}_{g,n}$. This is easy to explain (and prove): Teichmüller's theorem describes the geodesics as locally affine maps and there is no reason why such a map should respect an added marked point. Concretely speaking, think of an elliptic curve E_τ, i.e. for our purpose a parallelogram defined by 1 and $\tau \in \mathcal{H}$. Deforming E_τ into $E_{\tau'}$ in a geodesic way consists of applying a *linear* change of coordinates $L(\tau, \tau')$ mapping the parallelogram $(1, \tau)$ into $(1, \tau')$. Now add marked points P_τ and $P_{\tau'}$, thus working effectively in $\mathcal{T}_{1,2}$: the origin of the curves is marked by assumption and working with parallelograms provides marking diffeomorphisms for the respective tori. Since $P_{\tau'}$ is arbitrary, there is clearly no reason why L should respect the marking, i.e. in general $L(\tau,\tau')(P_\tau) \neq P'_\tau$.

This strongly suggests that *the forgetful map p decreases distances*; that this is actually the case is an immediate consequence of the fact that the Teichmüller and the Kobayashi distances coincide (see, for example, [**37**, § 6.4] for the closed case).

The map p is clearly equivariant, so descends to a map $p : \mathcal{M}_{g,n+1} \to \mathcal{M}_{g,n}$, using the same name for simplicity. Note that it does *not* descend to $\mathcal{M}_{g,[n+1]}$ because there is no notion of 'last point' down there. In fact it can clearly be defined on the quotient of $\mathcal{M}_{g,n+1}$ by the stabilizer of the last object in the permutation group $\mathcal{S}_{n+1}$, which is actually sometimes useful. Now recall that $\mathcal{M}_{g,n+1}$ is nothing but the *universal curve* sitting over $\mathcal{M}_{g,n}$. In other words, being careful to consider $\mathcal{M}_{g,n}$ as a complex orbifold (or a $\mathbb{Q}$-stack for future purposes), i.e. as a *fine* moduli space, there is a tautological fibration $p : \mathcal{C}_{g,n} \to \mathcal{M}_{g,n}$ where the fibre at a point consists of the curve it represents. And there is indeed a canonical isomorphism $\mathcal{C}_{g,n} \simeq \mathcal{M}_{g,n+1}$ which can be visualized by considering that for $X \in \mathcal{M}_{g,n}$ the $(n+1)$st marked point in $p^{-1}(X)$ is running over the curve X, where the n marked points are to be avoided. All this can be constructed in an algebraic way and is defined over $\mathbb{Q}$ (indeed over $\mathbb{Z}$).

We now go back to geodesic discs and our usual setting, that is $t = (X, f) \in \mathcal{T}_{g,n}$ and $q \in \mathcal{Q}(X)$. These data define a geodesic disc $D = D(t, q) \subset \mathcal{T}_{g,r}$, adding the zeros of q as marked points (see the end of § 2.4). Adding these points is useful and indeed indispensable when discussing the Strebel properties, because one then needs to consider the critical graph of q. But we also have the following useful assertion: *the image in $\mathcal{T}_{g,n}$ of the disc D under the point-erasing map $p : \mathcal{T}_{g,r} \to \mathcal{T}_{g,n}$ is geodesic*. That is the disc D does indeed map to a geodesic disc when erasing the zeros of q. In fact, putting it the other way round, the image $p(D)$ is the geodesic disc defined in the first place by the cotangent vector $(t, q) \in T^*\mathcal{T}_{g,n}$ and it lifts to $D \subset \mathcal{T}_{g,r}$ because one respects the additional marked points by definition: at any point (t', q') of D, considered in the cotangent space $T^*\mathcal{T}_{g,r}$, the m additional marked points are the zeros of q'. Another way to put it is to say that $p(D)$ is geodesic, and so is D, because the map p is distance decreasing and the distance on D does not exceed that on $p(D)$, so they must coincide. From now on we will usually notationally confuse $p(D) \subset \mathcal{T}_{g,n}$ and $D \subset \mathcal{T}_{g,r}$.

3. Arithmetic origamis

The class of objects we consider in this section is derived from a dynamical construction due to Thurston (in [**68**, § 6]) which was subsequently generalized by Veech (in [**69**, § 9]). Here we add the fact that these objects also carry arithmetic information, emphasizing that we mean 'modular arithmetic', in the terminology of the introduction. They were already known to have analytic and 'hyperbolic arithmetic' content (see, for example, [**68**], [**69**], [**24**] and [**39**]).

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

3.1. Origamis: definition, arithmeticity and first consequences

We give a constructive and combinatorial definition of origamis, which just as for 'dessins d'enfants' can easily be formalized using the language of elementary topology. Here is the cut-and-paste version.

Definition 3.1. An *origami* $\mathcal{O}$ is an assemblage of identical squares of paper satisfying the following two conditions:

(i) the squares are glued along their edges, shaping a connected surface without boundary; and

(ii) at any vertex there abut an even number of squares.

An origami is thus nothing but a closed (i.e. compact without boundary) topological surface divided into squares. Here 'square' is to be taken topologically; it can for instance be defined as a 2-cell bounded by a topological circle with four distinct marked points on it. We simply want to insist that the definition is indeed topological or combinatorial. The only singularities are of conical type and occur at the vertices where the number of abutting squares is not equal to four. The main point of this section is that we have the following proposition.

Proposition 3.2. *An origami $\mathcal{O}$ uniquely determines the following objects.*

(i) *A type (g,n); a family of geodesic discs D in the Teichmüller space $\mathcal{T}_{g,n}$ which are equivalent modulo the action of the Teichmüller modular group; finally a well-defined geodesic disc in $\mathcal{T}_g(=\mathcal{T}_{g,0})$.*

(ii) *A smooth arithmetic curve $B = B(\mathcal{O})$.*

(iii) *An arithmetic immersion $\phi : B \to \mathcal{M}_{g,n}$ such that $C = C(\mathcal{O}) = \phi(B) = \pi(D)$ is the canonical projection of $D \subset \mathcal{T}_{g,n}$ to $\mathcal{M}_{g,n}$.*

(iv) *A distinguished point $c \in C = \phi(B) \subset \mathcal{M}_{g,n}$ corresponding to an arithmetic curve of type (g,n).*

So the curve B and the immersion ϕ are arithmetic, and this is the novelty here. We will mention some consequences below, but they should be seen as just a beginning. We use the letter B partly because the alphabet is restricted and partly for 'base', since we get a stable arithmetic fibration $\phi^*(\mathcal{C}_{g,n}) \to B$ with $\mathcal{C}_{g,n}$ the universal curve sitting over $\mathcal{M}_{g,n}$.

Let us here denote by S the topological surface shaped by the origami; g is the genus of S. It also determines the integer $r \geqslant 0$, namely the number of singularities, or else the number of vertices such that the number of abutting squares is not four. These numbers and the total angle at each singular point are connected by the Gauss–Bonnet formula in the form of (2.6) in § 2.4; we take up the notation which is used there, noting that by condition (ii) in Definition 3.1 the angle at a singularity is a multiple of π. The number n ($\leqslant r$) is then the number of singularities at which the angle is equal to π, that is the number of vertices at which exactly two squares abut. One can then work either with (g,r) or with (g,n) (see the end of § 2.7). We assume for simplicity that (g,n), and so *a fortiori* (g,r), is hyperbolic $(2g-2+n>0)$; minor changes would be requested in the few exceptional and not especially interesting cases where this is not the case.

Next the construction determines a flat structure u on S by *declaring* each square to be indeed conformally the standard unit square, then lifting the quadratic differential $q_0 = \mathrm{d}z^2$ to a differential q on S and setting $u = u(q)$. Proceeding as in § 2.4 we get a geodesic disc D, which we may view in $\mathcal{T}_{g,r}$ or $\mathcal{T}_{g,n}$. This proves the first part of (i) of the proposition, which expresses that we ignore the marking here, so work modulo action of the modular group. But we actually have a more precise information than in § 2.4, where there was no canonical marking: an origami actually comes equipped with a partial marking, more precisely enough information to determine a point in the Teichmüller space $\mathcal{T}_g$. Hence the second part of item (i). This relatively fine and not so crucial point should become clearer after we have rephrased the construction in terms of multicurves in § 3.2 below. We thereafter freely write $\mathcal{T}$ and $\mathcal{M}$ without further ado, bearing in mind the discussion in § 2.7.

Next we prove that $D \subset \mathcal{T}$, or rather any disc in the family which we now fix once and for all, descends to a curve $C \subset \mathcal{M}$, to which end we have to show that the corresponding affine group $\Gamma(D)$, or more precisely $\mu(\Gamma(D))$, in the notation of § 2.5, is a lattice in $PSL_2(\mathbb{R})$. This comes directly from [**68**] and is explained in detail in [**24**], where more information on the geometry of this construction can be found. It goes roughly as follows: let $\check{T}$ be the once punctured topological torus and V the set of vertices (singular of not) of the origami $\mathcal{O}$. The punctured surface $\check{S} = S \setminus V$ is a finite unramified cover of $\check{T}$. This covering can be considered as a covering of surfaces with flat structures, since that on S has been constructed precisely by lifting the standard structure on T. Given such a covering it is relatively easy to show that the affine groups of the top and bottom surfaces are commensurable, here in fact commensur*ate*. The crux of the matter is the interpretation of $\Gamma(D)$ in terms of affine diffeomorphisms as in § 2.5. We refer to [**24**] for a full proof, specifically to Theorem 4.9 therein (see also their Remark 2.2). As a result, we find that $\Gamma(D)$ is not only a lattice but indeed is also commensurable with $PSL_2(\mathbb{Z})$, the latter group being the affine group of the punctured torus. We will denote $\Gamma(D)$ by $\Gamma(\mathcal{O})$ as this group effectively depends on $\mathcal{O}$ only.

One can actually say much more, as exemplified by Theorem 5.5 in [**24**] which in part summarizes known results. In particular, *origamis are in one-to-one correspondence with the geodesic discs in Teichmüller space whose affine groups are arithmetic*, that is commensurable with $PSL_2(\mathbb{Z})$; this use of 'arithmetic' has to do of course with hyperbolic, not modular arithmetic.

Next, and most relevant for us, we want to add the (modular) arithmetic information. We just saw that the projection $C = \pi(D) \subset \mathcal{M}$ is a curve; it has at most transverse intersections (ordinary double points) because D is geodesic. We let $B = \tilde{C}$ denote its normalization and $\phi : B \to \mathcal{M}$ be the corresponding immersion: $\phi(B) = C$. Now $B = \tilde{C}(\mathcal{O})$ can be written as the quotient $\Gamma(\mathcal{O})\backslash\mathcal{H}$, with $\Gamma(\mathcal{O})$ commensurable with $PSL_2(\mathbb{Z})$ (forgetting about the interesting orbifold structure encoded in the group H). The action is

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

proper and discontinuous, not necessarily free. This shows that B is defined over a number field, proving (ii). Here we are using the following theorem, which is a hyperbolic and slightly non-standard version of arithmeticity criteria which in essence are due to Weil and nowadays are part of Grothendieck's descent theory. The following theorem obtains.

Theorem 3.3. *Let X be a smooth complex curve which can be written as a quotient $X \simeq G\backslash\mathcal{H}$, where $G \subset PSL_2(\mathbb{R})$ is commensurable with a Fuchsian triangular group. Then X can be defined over a number field.*

We refer especially to [**3**] for this kind of statements using triangle groups. It may be useful to add that the celebrated Belyi theorem, which caused real astonishment to Grothendieck when it appeared in 1978 (see, for example, [**59**] for its usual statement and proof) has to do with the converse of the above statement. We will not need it but still state for clarity the corresponding hyperbolic version.

Theorem 3.4 (hyperbolic unramified version of Belyi's theorem). *A smooth complex curve X can be defined over a number field if and only if there exists a finite set $Z \subset X$ such that the affine curve $\check{X} = X \setminus Z$ is uniformized by a Fuchsian group $G \subset PSL_2(\mathbb{R})$ with G commensurable to $PSL_2(\mathbb{Z})$.*

Note that we used Theorem 3.3 in order to confirm that B is defined over a number field, because in Theorem 3.4 the group G does act freely on $\mathcal{H}$. Theorem 3.3 is like a ramified version of the direct part of Theorem 3.4. It may also be useful to bear in mind that the fundamental groups of C and $B = \tilde{C}$ are easily related: $\pi_1(C)$ is an extension of $\pi_1(B)$ by $\mathbb{Z}^t$, where t is the number of ordinary double points of C. This is still valid over $\mathbb{C}$ or $\bar{\mathbb{Q}}$ from the algebraic viewpoint, replacing the groups by their respective profinite completions: see the discussion in [**22**, § I.11]. Here one should rather use orbifold (respectively, stack) fundamental groups, so that $\pi_1(B) \simeq \Gamma(\mathcal{O})$; for the algebraic theory see [**57**] and [**45**], which, however, can be cut short under the present, relatively simple circumstances.

Next we show that the immersion ϕ is arithmetic, to which end we use specialization which concretely tells us that it is enough to show that the graph of ϕ has infinitely many $\bar{\mathbb{Q}}$-points. But since ϕ is the modular map, this amounts to showing that the image curve $C = C(\mathcal{O}) \subset \mathcal{M}$ itself has infinitely many $\bar{\mathbb{Q}}$-points. It is plain that this is indeed the case. In order to confirm this point, return to the concrete description of D at the end of § 2.4. Running along D corresponds to deforming the initial unit square into a rectangle, which we can assume has sides 1 and $\tau \in \mathcal{H}$; the latter provides a parameter for D. Given τ, one finds a point $P_\tau \in C$ which can be described explicitly: namely let again $\check{S}/\check{T}$ be the topological covering determined by $\mathcal{O}$; give T the structure of the elliptic curve E_τ with modular invariant $j = j(\tau)$; $\check{E}_\tau$ is E_τ with the origin deleted. Then $P_\tau \in C$ parametrizes the curve X_τ such that $X_\tau/\check{E}_\tau$ is the unramified covering with underlying topological covering $\check{S}/\check{T}$. Now X_τ is arithmetic, i.e. defined over $\bar{\mathbb{Q}}$, if and only if E_τ is; this happens exactly when $j(\tau) \in \bar{\mathbb{Q}}$, so indeed infinitely often, for instance when τ is imaginary quadratic. We remark that the phrasing above is not quite correct, as one should pay attention to the differences between coarse and fine moduli, and between residue fields and residue gerbes; but for our concerns here these differences are irrelevant (for a more correct formulation see the discussion after Corollary 3.5 below). This completes the proof of (iii).

Finally, (iv) has been added only to point out that the curve C has a convenient distinguished $\bar{\mathbb{Q}}$-point. Namely, one can take $\tau = \mathrm{i} = \sqrt{-1}$, which is indeed defined over $\bar{\mathbb{Q}}$ and corresponds to considering actual *squares* of paper in Definition 3.1. We actually used that point above when defining the flat structure $u(\mathcal{O})$ associated with $\mathcal{O}$.

It is plain that Proposition 3.2 should be considered as the beginning rather than the end of the game. Below and in the next subsection we will list a few consequences and remarks. We refer to [**53**] for further investigation. We will use the language of arithmetic Galois actions and Grothendieck–Teichmüller theory rather freely, with very short reminders; the interested reader should be able to trace the necessary references from these and from § 6 below.

Let $\mathcal{O}$ be an origami, $\check{S}/\check{T}$ the corresponding unramified topological covering of the punctured unit square. Let $C = C(\mathcal{O}) \subset \mathcal{M}$ be as above and let $K = K(\mathcal{O})$ be the field of definition of C. Let $\tau \in \mathcal{H}$ such that $j(\tau) \in K$. We get as above a point $P_\tau \in C$ corresponding to an arithmetic étale covering $X_\tau/\check{E}_\tau$. Under these assumptions we get the following corollary.

Corollary 3.5. *The field of moduli of the covering $X_\tau/\check{E}_\tau$ is contained in K.*

In fact let $G_K = \mathrm{Gal}(\bar{K}/K)$ denote as usual the absolute Galois group of K. For $\sigma \in G_K$, we find that the transform ${}^\sigma P_\tau$ is a point of C and moreover ${}^\sigma E_\tau = E_\tau$. So ${}^\sigma P_\tau$ corresponds to a covering ${}^\sigma X_\tau/\check{E}_\tau$ with the same topological type $\mathcal{O}$ as $X_\tau/\check{E}_\tau$. Hence the two coverings are isomorphic, proving the corollary.

We defined K as the field of definition of C; let us make this point a little more precise. The moduli space $\mathcal{M}$ is here considered as a Deligne–Mumford stack over $\mathbb{Q}$, i.e. we consider the generic fibre of the stack constructed in [**11**]. Formally (see, for example, [**40**, Chapter 11]), one then looks at the generic point of C and the corresponding gerbe; the latter has a coarse moduli space, necessarily the spectrum of a field, and this defines K. It may be a good exercise to unravel this definition in down-to-earth terms. In particular, if C is not contained in a special locus (see § 4 below) that is if the group H of § 2.5 is trivial, the generic point of C is schematic and $K(\mathcal{O})$ is just its residue field. We remark that the finite group $H = H(\mathcal{O})$ of generic automorphisms can be read off the topological origami $\mathcal{O}$: it corresponds to the combinatorial automorphisms of $\mathcal{O}$ preserving orientation as well as the horizontal and vertical directions.

Corollary 3.5 should be considered only as a sample statement, illustrating the fact that the very existence of K enables one to make statements which are uniform with respect to the isomorphism type of the base elliptic curve. In other words, the arithmeticity of the curve $C(\mathcal{O})$ organizes covers of elliptic curves with given topological type in a fashion which deserves more detailed investigation. Now what about the action of that quotient of $G_{\mathbb{Q}}$ which does not fix C? A first answer is given by the following statement, which is a slightly weakened form of Theorem 5.4 in [**53**].

Proposition 3.6 (M. Möller). *The absolute Galois group $G_{\mathbb{Q}}$ acts on the set of origamis and the action is faithful in the following sense: for any non-trivial $\sigma \in G_{\mathbb{Q}}$*

there exists an origami $\mathcal{O}$ with associated immersed curve $C = C(\mathcal{O})$ such that the action of σ maps C to an origami curve distinct from C.

More explicitly let $\mathcal{O}$ be an origami, with associated data B, ϕ and C as in Proposition 3.2; the first part of the proposition asserts in particular that ${}^{\sigma}C$ is an origami curve, that is there exists an origami, denoted ${}^{\sigma}\mathcal{O}$ such that ${}^{\sigma}C = C({}^{\sigma}\mathcal{O})$. This may look rather surprising from the present viewpoint as for instance C is a geodesic curve and there is *a priori* just no connection between the action of the arithmetic Galois group and the highly transcendental fact of being geodesic for the Teichmüller metric. At any rate that $G_{\mathbb{Q}}$ permutes origamis is certainly much less 'obvious' than the fact that it permutes dessins d'enfant, since the latter simply amounts to the fact that $G_{\mathbb{Q}}$ permutes the finite unramified coverings of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$. The second statement of the proposition says that for any $\sigma \in G_{\mathbb{Q}}$, $\sigma \neq 1$, there exists an origami $\mathcal{O}$ such that $C({}^{\sigma}\mathcal{O}) \neq C(\mathcal{O})$, which is a stronger assertion than ${}^{\sigma}\mathcal{O} \neq \mathcal{O}$.

Corollary 3.5 and Proposition 3.6 can also be thought of in terms of Hurwitz curves, as an origami appears as such a curve together with a natural immersion and some extra data coming from the canonical Strebel differential and directions. From this viewpoint it is rather their geodesic character which may be unexpected, that is they carry a lot of analytic information tightly woven with arithmetic information.

We note a preliminary connection with dessins (I am indebted to L. Zapponi for these and more specific remarks, to be developed elsewhere). Let $C = C(\mathcal{O})$ as above, $\hat{C}$ its completion and $P \in \hat{C}$ a cusp of C. It corresponds to a stable curve which is a covering of the degenerate elliptic curve with invariant $j = \infty$. This is nothing but a, not necessarily connected, dessin d'enfant identifying this degenerate elliptic curve with the thrice-punctured projective line. Moreover, the Galois action on the curves $C(\mathcal{O})$ induces an action on the cusps which is compatible with the Galois action on the corresponding dessins. In fact it seems that the Galois action on origamis contains that on dessins, looking at infinity in the moduli spaces of curves. We hope these matters will be further investigated.

It may be useful to put such assertions as Corollary 3.5 and Proposition 3.6 in perspective, underlining the contrast with anabelian geometry, which asserts that the outer Galois action is highly sensitive to, in fact recognizes the isomorphism class of, the curve (for a starting point on anabelian geometry, we refer to the pioneering letter of Grothendieck reproduced in [**63**] and to the introductory paper by F. Pop in the same volume). Let us spell out such a statement under the present circumstances. Let E be an arithmetic elliptic curve and let us assume for the purpose of simplicity and illustration that $j(E) \in \mathbb{Q}$. Let again $\check{E} = E \setminus \{0\}$. Distinguishing more carefully between the various fundamental groups, we have $\pi_1^{\mathrm{top}}(\check{T}) \simeq F_2$, where π_1^{top} is the topological fundamental group and $F_2 = \mathbb{Z} * \mathbb{Z}$ is the free group on two generators (we do not need to worry about basepoints here). We then have

$$\pi_1^{\mathrm{geom}}(\check{E}) = \pi_1^{\mathrm{alg}}(\check{E} \otimes \bar{\mathbb{Q}}) \simeq \hat{F}_2,$$

where π_1^{alg} is the algebraic fundamental group and π_1^{geom} is the geometric fundamental group, i.e. obtained after extending scalars to a separably closed field. Finally, $\hat{F}_2$

is the profinite completion of F_2. This being said we get an outer Galois representation $\rho_E : G_{\mathbb{Q}} \to \mathrm{Out}(\pi_1^{\mathrm{geom}}(\check{E})) \simeq \mathrm{Out}(\hat{F}_2)$. Anabelian geometry then asserts the following theorem.

Theorem 3.7. *With the above assumptions and notation, if ρ_E and $\rho_{E'}$ are isomorphic as outer representations, the curves E and E' are isomorphic.*

In other words, the representation ρ_E recognizes the isomorphism type $j(E) \in \mathbb{Q}$ of the curve E. The above theorem is a deep result, originally due to A. Tamagawa under the above assumptions, that is for hyperbolic affine curves over number fields. It is now also (amply) covered by later results of S. Mochizuki. We refer to [**56**] for a nice historical account with references. Here we also used the fact that $\hat{F}_2$ is centrefree, an expected but actually far from trivial statement (see for instance the original paper by G. Belyi containing his famous result for a proof): it implies that the outer Galois action determines $\pi_1^{\mathrm{alg}}(E)$ as an extension of $G_{\mathbb{Q}}$ by $\pi_1^{\mathrm{geom}}(E)$.

In apparent contrast with Theorem 3.7, Proposition 3.6 asserts for example that if one consider coverings of $\check{E}$ and $\check{E}'$ which are isomorphic as topological coverings, any element σ of the Galois group will map them to topologically isomorphic coverings: given $\mathcal{O}$, in order to compute ${}^{\sigma}\mathcal{O}$, one can use any point of $C(\mathcal{O})$ such that the modular invariant of the underlying elliptic curve is fixed under the action of σ.

Let us close this subsection with a few simple but useful remarks. First one can restate Definition 3.1 as follows: in order to define an origami, take a chequered notepad, cut out a connected but not necessarily simply connected region bounded by horizontal and vertical sides and perform identifications on the boundary. Condition (i) in Definition 3.1 says that the result should be a closed surface; condition (ii) amounts to requiring that one pairs horizontal (respectively, vertical) sides with horizontal (respectively, vertical) sides. One can also give a purely combinatorial version of this construction, using permutation groups, much as for dessins. That is one describes via labelling of the squares and permutations the unramified topological coverings of the punctured torus, just as for dessins one describes the coverings of the thrice-punctured sphere.

Very little is known about origamis in terms of explicit computations concerning their geometric structure, and nothing at all about the attached (modular) arithmetic, Galois action, etc. For instance, a simple class of examples can be defined as follows: an L-shaped origami of type (m, n) (m and n are integers greater than 1 and the pair is unordered) is given by a row of m squares and a column of n squares with one square in common at the bottom left corner. The identifications are the obvious ones, that is, except for the single common square the m and n strips are made into tori. One thus gets a closed surface of genus 2. Let $\Gamma(m, n) \subset PSL_2(\mathbb{R})$ be the affine group of $L(m, n)$, the L-shaped origami of type (m, n). Here we are in an orientable situation, so we could be more precise and work with $SL_2(\mathbb{R})$. All we know *a priori* is that $\Gamma(m, n)$ is commensurable with $PSL_2(\mathbb{Z})$. It is not difficult to see that $\Gamma(2, 2) \subset PSL_2(\mathbb{Z})$ is the index 3 subgroup generated by $P\Gamma(2)$ (the free subgroup of matrices congruent to the identity modulo 2) and the involution $z \to -1/z$ (letting $PSL_2(\mathbb{R})$ act on $\mathcal{H}$ via Möbius transformations as usual) which corresponds to the symmetry with respect to the diagonal occurring when $m = n$. It seems, however, that the group $\Gamma(m, n)$ is not known for general values of (m, n). There are obviously interesting open questions in that spirit such as: what are the groups which actually occur as the affine groups of origamis? Does one get a cofinal family in $PSL_2(\mathbb{Z})$? Not to mention the hoard of arithmetic questions which the above suggests.

Finally, we note that one can work over integers and reduce the result modulo almost all primes (i.e. all but finitely many), getting curves in positive characteristics. This is done using a well-known general procedure for 'clearing denominators' which we describe briefly and not in full generality. The same considerations apply to the objects we will meet in later sections and we will not repeat them. Let $\mathcal{X}$ be a smooth integral projective scheme over $\mathbb{Z}$; one can think of the stable compactification of $\mathcal{M}_{\mathbb{Z}}$, the moduli space of curves, as constructed in [**11**]. Here we ignore the difference between schemes and stacks (manifolds and orbifolds), as it can be easily dealt with for our purpose, especially in the case of the moduli stacks of curves. Let $X = \mathcal{X} \otimes \mathbb{Q}$ be the generic fibre. Let $Y \subset X$ be a closed (integral) subscheme defined over some number field K (e.g. the curve $C(\mathcal{O})$ attached to an origami). Then one can 'spread out' Y over $\mathbb{Z}$, that is there exists a flat model $\mathcal{Y} \subset \mathcal{X}$ defined over the ring of integers $\mathcal{O}_K$ of K, obtained by taking the schematic closure of Y in $\mathcal{X}$, which is automatically flat over $\mathbb{Z}$. Practically speaking, one can write down equations for Y with coefficients in K, eliminate denominators by multiplying out by integers and then make sure that the coefficients of each of the resulting equations (with coefficients in $\mathcal{O}_K$) are globally coprime, that is they do not belong to any common prime ideal of $\mathcal{O}_K$. The last condition is equivalent to flatness. One may now reduce $\mathcal{Y}$ modulo primes $\mathfrak{p} \in \operatorname{Spec}(\mathcal{O}_K)$, which also amounts to considering the intersections $Y_p = \mathcal{Y} \cap \mathcal{X}_p$ where $X_p = \mathcal{X} \otimes \mathbb{Z}/p\mathbb{Z}$ is the fibre of $\mathcal{X}$ at p ($p \in \mathbb{Z}$ a prime number) and $\mathfrak{p} \in \mathcal{O}_K$ lies over p. The reduction is 'good' for almost all primes.

More precisely, for the curve C attached to an origami we avoid the following phenomena:

(i) the normalization $B = \tilde{C}$ has bad reduction;

(ii) the possible marked points coalesce;

(iii) the possible transverse intersections degenerate—since they are transverse in characteristic 0, this amounts to avoiding the primes dividing the discriminants of the finitely many quadratic forms describing these intersections.

All in all we get for almost all primes a curve $C_p = C_p(\mathcal{O}) \subset \mathcal{M}_{\mathbb{Z}} \otimes \mathbb{F}_p$ with distinct marked points and transverse self-intersections, which is defined over a finite extension of the field $\mathbb{F}_p$. Among the many questions which come to mind is: is the geodesic character of the original characteristic 0 object reflected in its reductions?

3.2. Origamis and multicurves

Here we reformulate the construction in terms of multicurves, which is actually the original language in which Thurston presented his construction. We especially wish to point out the tantalizing similarity of the situation with the one exploited in [**30**] and [**55**], again as a possible bridge between topology and arithmetic. We will briefly explain the connection with Grothendieck–Teichmüller theory at the end of this subsection but these admittedly cryptic indications should possibly be skipped at first reading and taken up after getting acquainted with § 6 and part of the related material.

Start from an origami as in Definition 3.1. It defines the surface S and cuts it into squares. The set V of vertices has cardinality $s \geqslant 1$, the set $\Sigma \subset V$ of singularities (vertices with a number of abutting squares not equal to four) has cardinality $r \leqslant s$ and the set of marked points (vertices with two abutting squares) has cardinality $n \leqslant r$. Because the number of squares abutting at any vertex is even, one can distinguish between horizontal and vertical sides of the squares (of course one can permute the names 'horizontal' and 'vertical' in what follows). Let $\underline{\gamma}_{\mathrm{h}}$ (respectively, $\underline{\gamma}_{\mathrm{v}}$) be the union of the horizontal (respectively, vertical) sides. These are considered as piecewise smooth curves on S and are defined only up to isotopy. One can then take the respective duals $\underset{\sim}{\alpha}_{\mathrm{h}}$ and $\underset{\sim}{\alpha}_{\mathrm{v}}$ of $\underline{\gamma}_{\mathrm{v}}$ and $\underline{\gamma}_{\mathrm{h}}$ (beware of the inversion of horizontal and vertical here): take one point in each square and join them by traversing exactly once the vertical (respectively, horizontal) sides of each square, so as to get $\underset{\sim}{\alpha}_{\mathrm{h}}$ (respectively, $\underset{\sim}{\alpha}_{\mathrm{v}}$). Note that the construction still makes sense for the n squares corresponding to the marked points, where the horizontal (respectively, vertical) sides are identified. It is in fact easy to check that $\underset{\sim}{\alpha}_{\mathrm{h}}$ and $\underset{\sim}{\alpha}_{\mathrm{v}}$ are indeed *multicurves* on the surface S with s special points, that is we get two elements of $\mathcal{A}_{g,s}$, the set of multicurves on a surface of type g, s (see § 2.6).

Let us spell out part of the dictionary, which should be clear by now. The origami was constructed by lifting the quadratic differential $q_0 = \mathrm{d}z^2$ on the unit square to a differential q. The horizontal and vertical directions are Strebel directions for q and after possibly changing q into $-q$ one has that γ_{h} (respectively, γ_{v}) is the critical graph in the horizontal (respectively, vertical) direction (up to a homotopy fixing V). The differential q has $s-r$ removable singularities, $r-n$ zeros and n simple poles. Now $\underset{\sim}{\alpha}_{\mathrm{h}}$ and $\underset{\sim}{\alpha}_{\mathrm{v}}$ are (homotopic to) unions of horizontal (respectively, vertical) non-critical trajectories. They are the $\underset{\sim}{\alpha}(\theta)$ which occur in Proposition 2.10 for $\theta = 0$ and $\theta = \frac{1}{2}\pi$, respectively. Again we can erase points as in § 2.7, going in particular from $\mathcal{M}_{g,s}$ to $\mathcal{M}_{g,r}$, then to $\mathcal{M}_{g,n}$ and $\mathcal{M}_g$; each time one has to identify the curves in $\underset{\sim}{\alpha}_{\mathrm{h}}$ and $\underset{\sim}{\alpha}_{\mathrm{v}}$ which become homotopic after erasure. We will not elaborate further on this point here.

One can go the other way, starting from two multicurves $\underset{\sim}{\alpha}$ and $\underset{\sim}{\alpha}'$ which *fill up* S, that is are such that the complement of the union $\underset{\sim}{\alpha} \cup \underset{\sim}{\alpha}'$ is a disjoint union of cells. One then takes the dual to recover the graphs $\underline{\gamma}$ and $\underline{\gamma}'$. Because the complement of $\underset{\sim}{\alpha} \cup \underset{\sim}{\alpha}'$ is homotopically trivial, one gets the last part of (i) in Proposition 3.2. Note that there may be (topological) automorphisms of the structure defined by $(S, \underset{\sim}{\alpha}, \underset{\sim}{\alpha}')$, corresponding to ramification of the projection $\mathcal{T}_g \to \mathcal{M}_g$, but that does not prevent the corresponding geodesic disc D to be well defined in $\mathcal{T}_g$.

The basic point here is that the two multicurves correspond to two independent Strebel directions (i.e. not in the same orbit for the action of the modular group) and are enough to recover the whole structure. We remark that in general, it is *not* enough to find two such directions in order to make sure that a disc in Teichmüller space descends to a curve in moduli space (see Corollary 4 of [**25**] in the language of dynamical systems and § 5 below).

Let us rephrase the above more formally for possible future use. Let $\mathcal{M}$ be the moduli space of curves. We do not make the type explicit; it is (g, m) for $m = s$, r, n or 0 and we do not discuss the passage (point erasing) from one type to another. Let $\bar{\mathcal{M}}$ be the stable completion, $\mathcal{A}$ the set of multicurves, Γ the modular group. The elements of the quotient $\Gamma\backslash\mathcal{A}$ parametrize the strata of the stable stratification of $\bar{\mathcal{M}}$ (see § 2.6). An origami $\mathcal{O}$, hence the arithmetic immersion $\phi : B \to \mathcal{M}$ it defines, can equally be determined as an element of $\Gamma\backslash(\mathcal{A} \times \mathcal{A})$, where the action of Γ is diagonal and the only condition is that the two multicurves fill up the surface, that is cut it into cells. We remark in order to prevent possible confusion that pairs of strata of $\bar{\mathcal{M}}$ are parametrized by the product $\Gamma\backslash\mathcal{A} \times \Gamma\backslash\mathcal{A}$.

It seems quite striking that such combinatorial information near infinity on $\mathcal{M}$ uniquely determines immersed curves where both the curve and the immersion are defined over number fields. Let us compare this situation with what already exists in the same vein. First the local situation near the points of maximal degeneracy (dimension zero strata of the stable stratification) is well understood in characteristic 0. It amounts to studying Mumford curves in equicharacteristic 0 (without assuming that the residue field is algebraically closed), in other words one has to study the deformation of graphs of projective lines over, say, $\mathbb{Q}[\![t]\!]$. This has been done in [**35**] and other papers by the same authors, with recent additions and improvements by T. Ichikawa. One starts from combinatorial data, describing a maximal multicurve ('pants decomposition'), or equivalently a graph of projective lines with three distinguished points on each of them.

The global viewpoint (still at infinity though!) is implicit in [**12**], in genus 0. Briefly, MacLane coherence relations for quasi-tensor categories describe the structure at infinity of the genus 0 moduli spaces. This is explained in detail in [**7**]. Much more topology is needed in order to go to higher genera: complexes of curves describe the structure at infinity of the moduli spaces of curves of any finite type. A concise exposition can be found in [**27**] (see also, for example, [**18**]). In [**30**], one uses a two-dimensional complex introduced (but not studied) in the appendix of [**29**]. The fact that one can find such a finite-dimensional connected, simply connected, finitely presented complex governing the automorphisms of the profinite Teichmüller tower (see § 6.1 below) is in essence the content of the phrase: there exists a lego at infinity, etc. Again in genus 0, this is embodied more simply in the MacLane coherence relations and their consequences, like the Yang–Baxter equations.

The Grothendieck–Teichmüller lego *at infinity* for all finite types (g, n) is now essentially understood. Now what about moving inward? The above discussion of origamis suggests the following problem: consider *bimulticurves*, that is pairs of multicurves which fill up the given topological surface. In particular, one of the multicurves can be maximal, defining a maximally degenerate point. Then build these bimulticurves into a complex, by introducing the usual moves for complexes of curves, namely associativity moves (alias Whitney or fusion moves) and simple moves (see any paper on complexes of curves, starting with [**29**]; see in particular [**30**] in this context), or perhaps different types of moves. The question is: can one build a finite, perhaps just two-dimensional but not locally finite complex which is connected, simply connected and finitely presented and has bimulticurves as vertices? This and related questions would represent a deeper incursion into the arithmetic of the problem, given that contrary to what happened at infinity not everything is defined over $\mathbb{Q}$.

4. Special loci and algebraic eigenloci

Origamis were constructed starting from parabolic elements of the modular groups, namely products of commuting Dehn twists corresponding to a pair of multicurves (see §§ 2.6 and 3.2). Another salient feature is that there is no algebraization problem, that is any disc constructed in that way in Teichmüller space descends to a curve in moduli space. We now come to another type of curves immersed in the moduli spaces of curves, and in fact more generally to loci of higher dimensions in these spaces, which do not seem to have been investigated yet. They are elliptic by nature, inasmuch as they are built starting from torsion elements of the modular groups, and there is a very serious and interesting algebraization problem attached to the construction.

The sequence of events in this and the next section is organized roughly as follows. Fix a hyperbolic type (g, n) and the attending objects $\mathcal{T}$, $\mathcal{M}$, Γ. The group Γ acts isometrically on $\mathcal{T}$; the fixed-point set descends to $\mathcal{M}$ where it breaks into totally geodesic *and* algebraic varieties which we call special loci for short. We review a few useful facts about special loci in § 4.1, including some results having to do with their relevance to Galois and Grothendieck–Teichmüller actions. Then we move to eigenloci which we define first in Teichmüller space as loci which are globally preserved but not necessarily pointwise fixed by finite subgroups of Γ. Most of these loci are not algebraic and one comes across the same algebraization problem that was evidenced by Veech (see the end of § 2.5 above). However, we show in § 4.2 that it is easy in genus 0 and in the hyperelliptic case to write down many algebraic and in fact rational eigenloci explicitly. We come in § 5 to geodesic eigenloci. We first explain the connection with representation theory (hence the name 'eigenlocus'), computing the dimensions of such loci in general by means of the classical Chevalley–Weil formula. Finally, we go back to the one-dimensional case and to the geodesic eigencurves which appear in the study or rational polygonal billiards. These are geodesic by definition and Veech was the first to prove algebraicity in some very special cases. In § 5.3 we show how to identify the curves he studied with the explicit arithmetic eigencurves computed in § 4.2, using the rigidity of the monodromy representation for hyperbolic stable fibrations.

4.1. A few facts about special loci

We fix a finite hyperbolic type (g, n), so the Teichmüller space $\mathcal{T} = \mathcal{T}_{g,n}$ the moduli spaces $\mathcal{M}_{g,n}$ and $\mathcal{M}_{g,[n]}$ (see § 2.7) and the modular groups $\Gamma_{g,n}$ and $\Gamma_{g,[n]}$. Permuting the points *is* a non-trivial issue here; it is in fact the only issue in the genus 0 case, which is precious because it contains all the information about braid groups and is essentially the only one where things can be made explicit. We recall that permutations do not exist at the level of Teichmüller space: $\mathcal{T}_{g,n} = \mathcal{T}_{g,[n]}$.

Let us start in an analytic context and point out some important geometric facts which it actually took quite some time to elucidate (see in particular [**47**] and [**20**] and references therein, especially by the last-named author); we will then briefly discuss some more algebraic and arithmetic issues. Accordingly, until further notice, the objects are regarded analytically; in particular all groups are discrete by default.

Let $H \subset \Gamma_{g,[n]} = \Gamma$ be a non-trivial finite subgroup. The group H acts on $\mathcal{T}$ and we let $\mathcal{T}(H)$ denote the fixed-point set. Then $\mathcal{T}(H)$ is biholomorphically equivalent to the Teichmüller space $\mathcal{T}_{\gamma,\nu}$, where (γ, ν) is the type of the quotient of any marked Riemann surface in $\mathcal{T}$ under the action of H. More precisely, consider $t = (X, f) \in \mathcal{T}$; by definition there is a subgroup of $\mathrm{Aut}(X)$ which is isomorphic to H and such that the action on X is topologically that of H. We write $Y = X/H$ for the quotient; Y is equipped with an orbifold structure because the action of H may not be free and this gives a type (γ, ν) where ν is the sum of the number of marked points and orbifold points. So the first important piece of information is as follows.

Fact 4.1. $\mathcal{T}(H) \simeq \mathcal{T}_{\gamma,\nu}$. *In words: deforming a* marked *Riemann surface with symmetries is the same as unconditionally deforming the quotient marked surface.*

The above isomorphism is not so difficult to prove *if* one knows that $\mathcal{T}(H)$ is not empty. That this is the case follows from the positive answer to the Nielsen realization problem, given in all generality by S. Kerckhoff and whose beautiful proof also features an offshoot of Thurston's program or vision. In other words, for any finite subgroup H of the modular group, there exists at least one Riemann surface which has a subgroup of its automorphism group topologically acting like H.

The next step consists of projecting $\mathcal{T}(H)$ down to the moduli space $\mathcal{M}$, getting $\mathcal{M}(H) = \pi(\mathcal{T}(H))$. Here one takes $\mathcal{M}$ to be either the pure moduli space $\mathcal{M}_{g,n}$ or the space $\mathcal{M}_{g,[n]}$, or any space in between those two, obtained by allowing certain permutations only. Essentially by definition the *special loci* $\mathcal{M}(H) \subset \mathcal{M}$ for varying H constitute the branch locus of the infinite Galois covering $\pi : \mathcal{T} \to \mathcal{M}$. Using level structures, one can build a *finite* Galois covering so that this still holds true for that covering: in other words, the discrete modular groups are virtually torsion free. Now comes the following fact.

Fact 4.2. $\mathcal{M}(H)$ *is an irreducible quasi-projective variety.*

In contrast with the eigenloci to be introduced below, we see that $\mathcal{T}(H)$ always descends to an algebraic variety, whose dimension is easily computable; namely

$$\dim(\mathcal{M}(H)) = \dim(\mathcal{T}(H)) = 3\gamma - 3 + \nu,$$

where γ and ν have a clear purely topological meaning. Let us note that $\mathcal{M}(H)$ actually depends only on the conjugacy class of H in the suitable modular group Γ (i.e. such that $\mathcal{M} = \Gamma\backslash\mathcal{T}$). The special loci $\mathcal{M}(H)$ are totally geodesic for the Teichmüller metric, being the projections of the fixed-point sets of the action of Γ on $\mathcal{T}$, which is isometric. We note two interesting special cases. First, $\dim(\mathcal{M}(H)) = 0$, that is $\mathcal{M}(H)$ is a point if and only if the quotient $Y = X/H$ is a sphere with three orbifold or deleted points. These are the so-called 'curves with many automorphisms', which play an important role in structuring the moduli spaces and bring us back to the topics surrounding Belyi's theorem (see Theorems 3.3 and 3.4 above). Second, and more generally, assuming that the quotient Y has genus 0, $\mathcal{M}(H)$ is essentially isomorphic to an irreducible component of one of the so-called Hurwitz spaces. One can thus recover much of the geometry of the Hurwitz spaces by looking inside the moduli spaces of curves, something which is quite satisfying, in particular in terms of a purported 'universality' of Grothendieck–Teichmüller theory.

Now let H_0 be an *abstract* finite group and consider $\mathcal{M}[H_0]$, defined to be the union of the $\mathcal{M}(H)$ such that $H \subset \Gamma$ is isomorphic to H_0. Then we have the following fact.

Fact 4.3. *$\mathcal{M}[H_0]$ is a quasi-projective variety defined over $\mathbb{Q}$; the $\mathcal{M}(H)$ are defined over number fields.*

Here and above in Fact 4.2, we were again not quite correct, phrasing things in terms of coarse rather than fine moduli. One should really see $\mathcal{M}$ as a $\mathbb{Q}$-stack; the closed substack $\mathcal{M}[H_0]$ is defined over $\mathbb{Q}$ by its very definition. Its finitely many components, namely the $\mathcal{M}(H)$, are thus defined over number fields. We refer to the discussion following Corollary 3.5 above for some more detail and to [**45**] for a broader and more algebraic perspective.

The loci $\mathcal{M}(H)$ are in general singular, not necessarily normal, but the fact that they are totally geodesic implies that their singularities are not too complicated, in particular all (self-)intersections are proper. Let $\widetilde{\mathcal{M}(H)}$ denote the normalization of $\mathcal{M}(H)$; the algebraic and analytic meanings of 'normalization' coincide by standard comparison results. The normal variety $\widetilde{\mathcal{M}(H)}$ has a nice description in terms of covers, very similar to what we have seen for discs. Indeed, let $\Gamma(H) \subset \Gamma$ denote the normalizer of H inside Γ, which geometrically can be described as the *global* stabilizer of $\mathcal{T}(H)$ under the action of Γ. We then have the following fact [**20**, **47**].

Fact 4.4. *The normalization $\widetilde{\mathcal{M}(H)}$ of $\mathcal{M}(H)$ is isomorphic to the quotient $\Gamma(H)\backslash\mathcal{T}(H)$.*

The group $\Gamma(H)$ does not act effectively on $\mathcal{T}(H)$, as by definition the elements of H fix $\mathcal{T}(H)$ *pointwise* and the normalization $\widetilde{\mathcal{M}(H)}$ should be considered as an orbifold, including the non-effective action of H: this gives $\pi_1^{\mathrm{orb}}(\widetilde{\mathcal{M}(H)}) \simeq \Gamma(H)$ and displays the geometric meaning of the tautological exact sequence

$$1 \to H \to \Gamma(H) \to \Delta(H) \to 1. \tag{4.1}$$

This sequence is very reminiscent of (2.7) in § 2.5 but we changed notation a bit, perhaps not so wisely. Here Γ is the full stabilizer and Δ is the quotient which acts effectively. It is actually best to think in terms of orbifolds and stacks and the groups which act trivially, representing generic inertia, play a prominent role (see [**45**] for the general algebraic picture). In § 2.5 the group H was not the main focus and in fact it was generically trivial. Here it moved to the front stage; the two sequences coincide for one-dimensional special loci. Note the elliptic case: $g = n = 1$, $H = \mathbb{Z}/2\mathbb{Z}$ generated by the elliptic involution, $\widetilde{\mathcal{M}(H)} = \mathcal{M}(H) = \mathcal{M}$, $\Gamma(H) = \pi_1^{\mathrm{orb}}(\mathcal{M}) = SL_2(\mathbb{Z})$ and $\Delta(H) = PSL_2(\mathbb{Z})$.

We can now state a partial analogue of Fact 4.1 at the level of moduli spaces. By Fact 4.1 we have $\Delta(H) = \Gamma(H)/H \subset \Gamma_{\gamma,\nu}$, the modular group of the quotient surface.

Moreover, it is not difficult to show that $\Delta(H)$ has finite index in $\Gamma_{\gamma,\nu}$ (see [**47**], [**20**] as well as [**62**] for explicit examples in genus 0). As a result we get the following fact.

Fact 4.5. *The normalization* $\widetilde{\mathcal{M}(H)}$ *of* $\mathcal{M}(H)$ *is isomorphic to a finite orbifold unramified cover of the moduli space* $\mathcal{M}_{\gamma,\nu}$ *of the quotient surface.*

The special loci, or loci of curves with non-trivial automorphisms, arise naturally and have attracted the attention of geometers since long ago and there is a vast literature on the subject. In particular, the strata of maximal dimensions in $\mathcal{M}_g$, corresponding to cyclic groups of prime order have been analysed in [**10**]. They can be enumerated for any given g with their respective dimensions but much remains to be understood in this domain. There also exist quite a few case studies: in particular, we have a complete chart of the special loci of $\mathcal{M}_3$, which is already quite intricate. Note that in genus 0, it is easy to describe all the possible automorphisms (see also § 4.2 below). One gets in principle a full picture of the arrangement of the special loci in $\mathcal{M}_{0,[n]}$, which is interesting in the perspective of Grothendieck–Teichmüller theory (see [**61**, **62**]).

On the other hand, special loci had not been investigated in terms of arithmetics until very recently, perhaps in part because to this end, it is really desirable to take a stacky perspective. One main geometric point is that the special loci determine a stratification of $\mathcal{M}$, namely a decomposition into finitely many algebraic locally closed sets such that the Zariski closure of a stratum is a union of strata. This is made precise by H. Popp [**58**] in the framework of complex algebraic geometry. This kind of stratification can actually be defined and studied for any separated locally Noetherian Deligne–Mumford stack (see [**40**, Theorem 11.5] and [**45**]), which serves to highlight the generality and naturality of the phenomena. We would like to underline, partly again in the perspective of Galois and Grothendieck–Teichmüller theories, the thought-provoking parallel between the stable stratification and the stratification defined by the special loci, the latter being almost never of codimension 1. At the group-theoretic level, this is connected with the parallel and contrast between Dehn twists, generating free procyclic groups and finite-order elements of the modular groups (see [**41**] and [**21**] for a first approach). At the geometrical level, let us go a little further into the meaning of the exact sequence (4.1). Let $H \subset \Gamma$ be again a finite subgroup of the modular group Γ, defined up to conjugacy. If X is a Riemann surface corresponding to a point of $\mathcal{M}(H)$, we get a possibly ramified Galois covering $\pi : X \to Y = X/H$, which is entirely determined by the sheaf $\pi_*(\mathcal{O}_X)$ on Y. Going to the universal curve $\mathcal{C}$ over $\mathcal{M}$ we find that the special locus $\mathcal{M}(H)$ carries a natural sheaf, which records the action of the automorphism group at any point.

Problem: study the class of the extension defined by (4.1) in terms of that sheaf on $\mathcal{M}(H)$, in particular for H cyclic of prime order. This is investigated further in [**45**]. The problem of whether or not sequence (4.1) splits was especially pointed out in [**62**], where it is shown by explicit methods that it does split for cyclic subgroups in genus 0. Note that already in the elliptic case (see after Fact 4.4), the extension $SL_2(\mathbb{Z}) \to PSL_2(\mathbb{Z})$ is central non-split. Finally, we point out that for H cyclic, this all has to do with higher spin curves, which were studied in a series of recent papers by T. J. Jarvis.

The corresponding picture 'at infinity' is provided by the stable stratification. In particular, take $H \simeq \mathbb{Z}$ generated by a Dehn twist h; then we have an exact sequence as in (4.1), where $\Gamma(H)$ is now the centralizer of h in Γ and $\Delta(h)$ is the fundamental group of the irreducible component $\mathcal{M}'$ of the divisor at infinity defined by h. The divisor $\mathcal{M}' \subset \mathcal{M}$ carries a natural invertible sheaf, namely its conormal line bundle, and the extension class defined by (4.1) is given by the Chern class of that line bundle. Can we get a similar description for special loci and their attached sheaves? The connecting concept between the two situations, namely Dehn twists (parabolic elements) on the one hand and automorphisms of curves (elliptic elements) on the other, is that of *inertia*. In the algebraic situation Dehn twists generate the procyclic inertia groups attached to the rank one valuations corresponding to the components of the divisor at infinity. Moreover, they are copies of $\hat{\mathbb{Z}}(1)$ as Galois modules and one can even describe the Galois action: it is cyclotomic up to conjugacy. Although this is in essence folklore, it takes quite some work to put it on a firm basis. The starting point is [**23**] which uses formal schemes, so that one has to modify the strategy when dealing with stacks (parts of the results of [**23**] have been rediscovered again and again in the 1980s, under the general heading 'branch cycle argument' but in a less general framework). What happens with automorphisms and special loci also has to do with inertia. Only this is inertia for stacks, which is connected with the automorphism groups of the objects. One gets for instance the following proposition, as an application of more general results.

Proposition 4.6 (see [45]). *Let $\mathcal{M} = \mathcal{M}_{g,[n]}$ be viewed as a $\mathbb{Q}$-stack, and $\Gamma = \Gamma_{g,[n]}$ be the corresponding profinite modular group, the geometric stack fundamental group of $\mathcal{M}$. Let h be a finite-order element in the natural discretization of Γ, that is a mapping class of finite oder. Consider the natural outer action of $G_{\mathbb{Q}}$ on Γ; then there is an open subgroup $G_K \subset G_{\mathbb{Q}}$ such that for any $\sigma \in G_K$, $\sigma(H) \sim H$, where $H = \langle h \rangle$ is the finite cyclic group generated by h and '$\sim$' denotes conjugacy in Γ. Moreover, the number field K can be taken to be the field of definition of the irreducible closed substack $\mathcal{M}(H)$. Finally, this determines an actual action of G_K on H (via some character).*

This result features a close analogue of what happens for inertia attached to a divisor with strict normal crossings. It is, however, less precise than the main result of [**23**] which says something about the full decomposition group, the analogue of which we did not even define.

Next and last: what about the Grothendieck–Teichmüller group? More precisely, the existing group $\mathbb{\Gamma}$ is defined using only the geometry at infinity (see § 6 below and references therein). In particular, it acts like the Galois group on Dehn twists, alias inertia at infinity. What about its action on the torsion elements of the modular groups, alias stack inertia, for which the action of the Galois group is given as in the theorem above? Here we do have a result, but only in the case of genus 0, namely the following theorem.

Theorem 4.7 (see [61]). *Let $\mathcal{M} = \mathcal{M}_{0,[n]}$ be the moduli space of unordered pointed sphere ($n \geqslant 3$), viewed as a $\mathbb{Q}$-stack, and $\Gamma = \Gamma_{0,[n]}$ its geometric fundamental group. Let h be an element of finite order in Γ and $F \in \mathbb{\Gamma}$ with $\lambda(F) = 1$. Then $F(h)$ is conjugate to h.*

Here λ is the Grothendieck–Teichmüller analogue of the cyclotomic character: if $\sigma \in G_{\mathbb{Q}} \subset \mathbb{\Gamma}$, $\lambda(\sigma) = \chi(\sigma)$. In genus 0, all special loci are easily seen to be defined over $\mathbb{Q}^{\mathrm{ab}}$, the

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

maximal abelian extension of Q, so that with the notation of Proposition 4.6, $K \subset \mathbb{Q}^{\mathrm{ab}}$. The upshot is of course that Theorem 4.7 almost extends Proposition 4.6 from $G_{\mathbb{Q}}$ to $\mathbb{\Gamma}$ in genus 0. 'Almost', because it uses $\mathbb{Q}^{\mathrm{ab}}$ instead of the finite extension $K \subset \mathbb{Q}^{\mathrm{ab}}$ on which $\mathcal{M}(H)$ is defined; there exists at present only a primitive embryo of Galois correspondence in Grothendieck–Teichmüller theory, so that open subgroups cannot often be defined in terms of the available coordinates, even those corresponding to abelian number fields.

In any case Theorem 4.7 does represent a first step towards displaying how the information at infinity propagates inward. Its proof uses typical tools from Grothendieck–Teichmüller theory, dealing directly with profinite groups and their automorphisms (see [**44**] for related techniques). It would be extremely interesting to extend it to higher genera, which seems hard for several reasons which we skip here. To summarize, Proposition 4.6 gives an infinite set of conditions which the Galois group fulfils by nature, whereas Theorem 4.7 says that some of them are indeed fulfilled by $\mathbb{\Gamma}$, although the latter group was defined by conditions living at infinity. So we get more and more conditions to be fulfilled by the ideal Grothendieck–Teichmüller group: some come from embedding curves, e.g. origamis, others from special loci and the attending stratification, still others will arise from the algebraic eigenloci defined below. Moreover, as explained above in Fact 4.5 the special loci, or rather their normalizations, are themselves coverings of moduli spaces which again parallels what happens for the stable stratification. This naturally gives rise to another set of conditions defined and partly explored in [**61**]. The geometric and arithmetic plot keeps thickening.

4.2. Algebraic eigenloci: definition and explicit examples

As in § 4.1 above we select a type (g, n) and a (possibly trivial) subgroup of the permutation group $\mathcal{S}_n$, giving rise to a triplet $(\mathcal{T}, \mathcal{M}, \Gamma)$, with canonical projection $\pi : \mathcal{T} \to \mathcal{M}$, so that $\mathcal{M} \simeq \Gamma \backslash \mathcal{T}$. The group Γ acts on $\mathcal{T}$ via $\gamma \cdot (X, f) = (X, f \circ \gamma^{-1})$ for $\gamma \in \Gamma$ and $t = (X, f) \in \mathcal{T}$ (cf. § 2.5). Let $H \subset \Gamma$ be a non-trivial finite group, and let $\mathcal{T}(H)$ and $\mathcal{M}(H)$ be as in § 4.1 above. Note that $\gamma \cdot \mathcal{T}(H) = \mathcal{T}(\gamma H \gamma^{-1})$ for any $\gamma \in \Gamma$; in particular, as mentioned already, $\mathcal{M}(H)$ depends only on the conjugacy class of H and the same will be true of the loci we are about to define. We now make the following general definition.

Definition 4.8. An (algebraic) *eigenlocus* $\mathcal{E} \subset \mathcal{M}$ associated with the conjugacy class of a finite subgroup $H \subset \Gamma$ of the modular group is an algebraic variety in $\mathcal{M}$ such that one (or equivalently, any) lift of $\mathcal{E}$ to $\mathcal{T}$ is stable under the action of the corresponding conjugate of H.

Let us insist that special loci, which are the fixed-point sets of the action of the modular group Γ are simultaneously algebraic, arithmetic and totally geodesic. This is not the case anymore here. In particular, there are many geodesic (in a sense to be made precise in the next section) eigenloci which are not algebraic, *a fortiori* not arithmetic. Note also that being algebraic is the same as being analytic or also just locally closed in the ordinary (analytic) topology.

The main idea is that we are looking for algebraic and indeed arithmetic loci, that is closed $\bar{\mathbb{Q}}$-substacks of $\mathcal{M}$ which generalize the special loci of the last subsection in that they correspond to subvarieties in Teichmüller space which are stable but not necessarily pointwise fixed under the action of finite subgroups of Γ: special loci appear as rather special cases of eigenloci. This generalization looks quite natural from a geometric viewpoint as well as from the standpoint of Galois and Grothendieck–Teichmüller theory: *any such eigenlocus will contain* H *in its geometric stack fundamental group* and one may hope to learn something about the Galois action; any such piece of information in turn becomes a condition which the Grothendieck–Teichmüller group has to satisfy if it is to coincide with the Galois group. Moreover, we are on the elliptic side and so aiming at an elliptic lego, more subtle than the existing parabolic lego (or lego at infinity). Concrete glimpses of these questions will appear below and an example of how to use such information can be found in [**46**].

The name 'eigenlocus' will become clearer after we discuss the geodesicity condition in § 5 below, together with the, at first encounter, surprising connection with rational polygonal billiards. A main point here consists again in trying to put together the algebraicity and the geodesicity conditions. Origamis feature a kind of, understandable, miracle: they provide geodesic discs in Teichmüller space which project to algebraic curves in moduli space which in turn are recognized to be arithmetic. Not so here. We do not know how abundant and easy to find *algebraic* eigenloci are in general, that is for general type (g,n) and we will say nothing more of the general algebraic case. Instead we will presently show how to write down in an explicit and elementary fashion many such loci in genus 0 and inside hyperelliptic loci. It will evolve in § 5 that *some* of these loci are also geodesic.

Let us first review the situation with special loci in genus 0. We take a geometric route referring to [**47**] for a more algebraic path, and to [**62**] for very explicit examples in small dimensions. Background material can be found in [**5**]. The elementary character of the analysis below should not lead the reader to undervalue the richness and complexity of the situation. Just as (almost) all modular groups are, $\Gamma_{0,n}$ is virtually torsion free *and* generated by its torsion elements (see [**41**] and [**21**] for detailed results and references). By analysing special loci we are also, among other things, studying a set of elliptic generators for all the braid groups and their profinite completions. Dehn twists, that is inertia at infinity, do not tell the full story; this motto is already present in the *Esquisse*.

So let us take up the genus 0 objects, namely $\mathcal{T}_{0,n}$, $\mathcal{M}_{0,n}$ and $\mathcal{M}_{0,[n]}$, as well as the intermediate objects in the ramified (but unramified as orbifold covering) $\mathcal{S}_n$-Galois covering $p:\mathcal{M}_{0,n}\to\mathcal{M}_{0,[n]}$. Since the covering $\mathcal{T}_{0,n}\to\mathcal{M}_{0,n}$ is unramified ($\mathcal{M}_{0,n}$ is a scheme—labelled n-pointed spheres have no automorphisms), special loci can be read off from the ramification of p, hence from subgroups of $\mathcal{S}_n$. Let us pause in order to fix notation. Write $\underline{x}=(x_i)_i$ $(i=1,\dots,n)$ for a point in $\mathcal{M}_{0,n}$. Then any $s\in S_n$ acts on $\mathcal{M}_{0,n}$ by permuting the entries; explicitly it maps $\underline{x}$ to $s\underline{x}$ such that $(s\underline{x})_i=x_{s^{-1}(i)}$. Returning to the classification of automorphisms in genus 0, the only non-trivial fact we will use is that all conformal structures on the sphere are equivalent, which is the elliptic part of the uniformization theorem. In other words, $\mathcal{T}_0$ is a point. Now any automorphism of an n-pointed sphere must preserve the underlying conformal structure, hence can be realized as an element of $PGL_2(\mathbb{C})$, and being of finite order, it is a Euclidean rotation.

By the platonic classification, it is thus cyclic, dihedral or isomorphic to the symmetry group of a platonic solid.

Given the topological n-pointed sphere $S = S_{0,[n]}$, let us say that a subgroup $H \subset \Gamma_{0,[n]}$ is *maximal* if the corresponding locus is a point, that is if S/H is a sphere with three marked points. An element $h \in \Gamma_{0,[n]}$ is maximal if it generates a maximal cyclic group $H = \langle h \rangle$. It is easy to find all maximal elements: such an h is a rotation with two ramification points (the endpoints of the axis) and ramification values, so that the quotient surface S/H has at least two marked points. These points may or may not have been marked before and the rest of the n points have to form a single orbit. We record these simple but useful finds in the following proposition.

Proposition 4.9. *Any finite subgroup $H \subset \Gamma_{0,[n]}$ is either cyclic, or dihedral or isomorphic to the symmetry group of a platonic solid. A maximal element $h \in \Gamma_{0,[n]}$ is a rotation of order $r = n$, $n-1$ or $n-2$.*

Of course one can mark as many points as one wishes, provided one marks the whole orbit under the action of the given group. Take for instance an icosahedron, divide the edges in three equal parts and mark all the dividing points but not the vertices. This produces a special locus in $\mathcal{M}_{0,[60]}$ via an injection $\mathcal{A}_5 \hookrightarrow \mathcal{S}_{60}$. One gets copies of the groups mentioned in the proposition inside $\mathcal{S}_n$.

Henceforth we will confine attention to cyclic and dihedral groups for simplicity and because maximal such subgroups will naturally arise in the next section. It would be helpful though to have a detailed picture of the full situation. After normalization, a cyclic group is generated by R acting on $\mathbb{P}^1$ via $R(z) = \zeta^{-1}z$ where ζ is an rth root of unity (the reason for writing ζ^{-1} instead of ζ will appear below). A dihedral group is generated by such an R and the involution I such that either $I(z) = -z$, $I(z) = 1/z$ or $I(z) = -1/z$, according to whether it fixes or permutes the ramification points of the rotation and to what it does on the tangent spaces at these points.

Even in this genus 0 situation we are only able to produce *rational* eigenloci; the prototypical example is as follows. Consider $\mathcal{M}_{0,n}$ with elements parametrized as $(0, x_1, \ldots, x_r, \infty) = (0, \underline{x}, \infty)$, $n = r + 2$. Having pinpointed 0 and ∞ leaves only a scaling as automorphism group, that is this representation is up to multiplication of the x_i by a common constant. Let $c = c_r \in \mathcal{S}_n$ be the r-cycle $(r, \ldots, 1)$ ($c(i) = i - 1$ modulo r) which we here use to permute the points x_i: $(c\underline{x})_i = x_{i+1}$ modulo r. Finally, let $\zeta \in \boldsymbol{\mu}_r^*$ be a primitive rth root of unity. The cycle c lifts to a maximal element γ of order $r = n-2$ in $\Gamma_{0,[n]}$ which we do not need to write down explicitly (see, however, after Proposition 4.11 below for more detail). The point $P_c = (0, 1, \zeta, \ldots, \zeta^{r-1}, \infty)$ in $\mathcal{M}_{0,n}$, i.e. the rth roots of unity together with 0 and ∞, projects to the special locus corresponding to c in $\mathcal{M}_{0,[n]}$. The action of $c \in \mathcal{S}_n$ on the point $P_c \in \mathcal{M}_{0,n}$ is equivalent to that of the automorphism R with $R(z) = \zeta^{-1}z$ as above. The special locus is reduced to the point P_c, or equivalently the quotient of the marked sphere represented by P_c by the group generated by R is a sphere with three distinguished points, namely 0, ∞ and the rth roots of unity which are just one orbit. Now consider the rational curve in $\mathcal{M}_{0,n}$

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

parametrized as

$$(0, \varphi(t), \zeta\varphi(\lambda t), \zeta^2\varphi(\lambda^2 t), \ldots, \zeta^k\varphi(\lambda^k t), \ldots, \zeta^{r-1}\varphi(\lambda^{r-1}t), \infty), \tag{4.2}$$

where the notation is as follows. The parameter t runs over the projective line, λ is an rth root of unity and φ is an algebraic function which, using a scaling and automorphisms of the t-line we may and do normalize by requiring that $\varphi(\infty) = \infty$, $\varphi(0) = \varphi'(0) = 1$; in other words, $\varphi(t) = 1 + t + O(t^2)$ near $t = 0$.

We have thus defined a rational algebraic curve $\mathcal{E}_c = \mathcal{E}_c(\lambda, \varphi) \subset \mathcal{M}_{0,n}$, with the following properties. For $t = 0$, the point $\mathcal{E}_c(t = 0)$, or rather its projection to $\mathcal{M}_{0,[n]}$ is a point with automorphism $c \in \mathcal{S}_n$; more generally we get the transformation law $(c\underline{x})_i(t) = x_{i+1}(t) = \zeta x_i(\lambda t)$. In other words, after dividing out by ζ as we may, we find that $\mathcal{E}_c$ is stable under the action of c, which induces on the t-line the rotation Λ of order dividing r defined by $t' = \Lambda(t) = \lambda t$, with fixed points $t = 0, \infty$. That was to be expected; indeed since $\mathcal{E}_c$ is rational and stable under the action of c, the latter has to induce a finite-order element of $PGL_2(\mathbb{C})$, so again a rotation of the t-line which moreover must have 0 as a fixed point. The locus $\mathcal{E}_c$ is a projective line with a certain number of points removed corresponding to the collisions $x_i = x_j$, $i \neq j$, and $x_i = 0, \infty$. These depend on the algebraic function φ and especially its degree. Now the image of $p(\mathcal{E}_c)$ in $\mathcal{M}_{0,[n]}$ is the quotient of $\mathcal{E}_c$ by Λ, assuming that γ generates the stabilizer of $\mathcal{E}_c$ in $\Gamma_{0,[n]}$ (we will see below a case where that stabilizer is actually larger). When is it that $p(\mathcal{E}_c)$ is a projective line with exactly three distinguished (i.e. marked or removed) points? Partial answer: it does happen when φ is affine. We do not know for sure but strongly suspect that it is the only case and it would actually be interesting to prove such a statement, for reasons that will appear more clearly in the sequel.

For the time being let us detail what happens for φ affine, namely $\varphi(t) = 1 + t$, with the normalization we have adopted. Setting $\xi = \lambda\zeta$, the expression (4.2) of $\mathcal{E}_c$ now reads

$$(0, 1 + t, \zeta + \xi t, \zeta^2 + \xi^2 t, \ldots, \zeta^k + \xi^k t, \ldots, \zeta^{r-1} + \xi^{r-1} t, \infty). \tag{4.3}$$

We may take for ζ any primitive rth root of unity (primitivity is for convenience only); for ξ we may take any rth root of unity, except for $\xi = \zeta$ which is excluded; in that case, corresponding to $\lambda = 1$, the expression (4.3) does not parametrize a line but only the special point P_c for any value of t. Checking dimensions, we find that for fixed ζ, that is for fixed special locus, there are $r - 1 = n - 3$ one-dimensional eigenloci (call them eigencurves) of the form (4.3); it is no coincidence that this is also the dimension of the space $\mathcal{M}_{0,n}$, as will appear more clearly in the next section. Here we just remark the following: erase the point 0 from (4.3), that is apply a point-erasing map $e : \mathcal{M}_{0,n} \to \mathcal{M}_{0,n-1}$. The dimension drops by 1 and indeed the number of associated affine eigencurves drops by 1 too: the value $\xi = 1$ ($\lambda = \zeta^{-1}$) is not permissible anymore, because erasing the point 0 restores translational symmetry. We will see below that it may also happen for certain values of λ that the eigencurve (4.3) acquires a self-intersection upon taking its image by the map e. We drop this example for the time being but will shortly return to it in order to explicitly compute its monodromy.

The rational curve (4.2) can serve as a prototype for constructing higher-dimensional rational eigenloci attached to any finite-order element of the modular group. We give a rough sketch of this easy generalization. Changing notation slightly, let now $n = rs$, where r will still be the order of the automorphism R ($R(z) = \zeta^{-1}z$) corresponding to the r-cycle $c \in \mathcal{S}_n$, and its lift $\gamma \in \Gamma_{0,[n]}$. One can of course change n to $n+1$ or $n+2$ by adding in the points 0 and/or ∞. Let $t = (t_i)_i \in \mathbb{A}^s$, $\lambda = (\lambda_i)_i \in (\boldsymbol{\mu}_r)^s$ and $\varphi = (\varphi_i)_i$ an s-tuple of algebraic functions. Consider the rational locus $\mathcal{E}_c \subset \mathcal{M}_{0,n}$ parametrized by $\mathbb{A}^s$ minus a certain number of hyperplanes (collisions) as

$$(\varphi_1(t_1), \zeta\varphi_1(\lambda_1 t_1), \dots, \zeta^{r-1}\varphi_1(\lambda_1^{r-1}t_1), \dots, \varphi_2(t_2), \zeta\varphi_2(\lambda_2 t_2), \dots, \\ \zeta^{r-1}\varphi_2(\lambda_2^{r-1}t_2), \dots, \varphi_s(t_s), \zeta\varphi_s(\lambda_s t_s), \dots, \zeta^{r-1}\varphi_s(\lambda_s^{r-1}t_s)). \quad (4.4)$$

This is a fairly obvious analogue of the eigencurve (4.2) and yet it gives a first feeling of the perfectly explicit and very rich structure that these rational eigenloci organize in the genus 0 moduli spaces. We are not going to dwell on the subject but will rather make a few remarks and ask a few questions which may or may not convince the reader that this geometric structure deserves further study.

(1) The representation (4.4) gives many rational eigenloci for cyclic groups. Does it describe them 'all'? The answer is, basically, 'yes'. Indeed, sticking to curves and maximal elements for simplicity, we have the following easy lemma.

Lemma 4.10. *As λ and φ vary, formula* (4.2) *describes all the rational eigencurves associated with the maximal r-cycle c.*

For the proof of the lemma, denoting such a curve by $\mathcal{E}_c$ as usual, one first notices that the cycle c, represented by the rotation R, induces on $\mathcal{E}_c$ an algebraic automorphism of order r, that is a rotation; taking the fixed points to be 0 and ∞, it can be written as $\Lambda(t) = \lambda t$ for a certain rth root of unity λ. Now if $x(t) = (x_i(t))$, $i = 1, \dots, r$, parametrizes a point of $\mathcal{E}_c$, applying the cycle c we find that $x_{i+1} = \zeta x_i(\lambda t)$ with the proper normalization and this implies that there exists an algebraic function φ such that $\mathcal{E}_c$ is of the form (4.2).

One then expects that taking proper care of the $PGL_2(\mathbb{C})$ action the lemma can essentially be extended to the higher-dimensional case embodied in (4.4).

(2) What about writing down rational eigenloci for all the possible automorphism groups, as described in Proposition 4.9? And do there exist algebraic eigenloci in genus 0 that are not rational? This question is prompted by the fact that the *special* loci in genus 0 are essentially copies of genus 0 moduli spaces (see [**62**] on this point). The much more abundant eigenloci of the form (4.4) look like genus 0 configuration spaces, especially for affine φ (see below). Do they 'essentially' exhaust the possibilities?

(3) Can one find all (rational) eigencurves which project in $\mathcal{M}_{0,[n]}$ to the projective line $\mathbb{P}^1$ with three distinguished points, i.e. a hyperbolic orbifold which is uniformized by a Fuchsian triangle group? The relevance of this question will hopefully appear more clearly below. As mentioned already, it is tempting to guess that formula (4.3) gives the 'general' solution.

(4) Much more difficult and obscure: can one build a genus 0 cyclotomic 'lego' using in particular the geometric structure in the moduli spaces $\mathcal{M}_{0,n}$ determined by the arithmetic eigenloci? The meaning of this question, which is in line with Grothendieck's *Esquisse*, should appear more clearly after reading § 6 below. We note that the affine eigenloci of type (4.3) are indeed defined over $\mathbb{Q}^{ab}$ and that they can be helpful in terms of Galois and Grothendieck–Teichmüller theory (see [**46**] for a first approach), especially those which fulfil the condition of question (3) above.

We now return to the computation of the monodromy in the simplest case, embodied by formula (4.3). In view of further use we modify the problem and notation in an inessential way, namely we erase the point 0 and write n instead of r for the order of the cycle. So we now have a curve $\mathcal{E}_c \subset \mathcal{M}_{0,n+1}$ with parametrization

$$(1+t, \zeta+\xi t, \zeta^2+\xi^2 t, \ldots, \zeta^k+\xi^k t, \ldots, \zeta^{n-1}+\xi^{n-1}t, \infty), \tag{4.5}$$

where ζ is a fixed primitive nth root of unity and ξ is an nth root of unity with $\xi \neq 1, \zeta$; so there are $n-2$ choices, which is also the dimension of $\mathcal{M}_{0,n+1}$. We write $\mathcal{E}_c(\xi)$ for the curve defined by (4.5).

We will narrow down our scope still further. The fact is that the curves $\mathcal{E}_c(\xi)$ are not all on an equal footing, for at least two reasons. First let $\mathcal{S}_n$ be the stabilizer of the last object in $\mathcal{S}_{0,n+1}$; we let it act on $\mathcal{M}_{0,n+1}$ and look at the images of the various $\mathcal{E}_c(\xi)$ in the quotient $\mathcal{M}_{0,n+1}/\mathcal{S}_n$. Then it may and does happen that some of these curves coincide in that quotient. This is for instance the case of $\mathcal{E}_c(\zeta^2)$ and $\mathcal{E}_c(\zeta^3)$ for $n=5$. A second discriminating factor between the various $\mathcal{E}_c(\xi)$ comes from the size of their respective stabilizers. Let $C = C_n$ denote the cyclic group of order n generated by the cycle c. By construction C stabilizes $\mathcal{E}_c(\xi) \subset \mathcal{M}_{0,n+1}$ for any ξ and the generator $c \in C$ induces the rotation Λ with $\Lambda(t) = \lambda t$, $\lambda = \zeta^{-1}\xi$. But it may and does happen that the stabilizer is bigger, namely dihedral. Let I be the involution given by $I(t) = 1/t$ and $D = D_{2n} = \langle \Lambda, I \rangle$ the group generated by Λ and I. For certain choices of n, ζ, ξ, the involution I and thus the group D stabilizes the curve $\mathcal{E}_c(\xi)$. This is the case in particular for any n and $\xi = \zeta^{-1}$, a case to which we will henceforth restrict attention. Before focusing on that case, however, we would like to point out that the two phenomena noticed above are probably worthy of attention. Understanding the geometry of the $\mathcal{E}_c(\xi)$ may be interesting by itself and, anticipating § 5.3 below, it could well be directly connected with billiards in isosceles triangles with rational angles, more specifically the fact that isosceles billiards with apex angle $k\pi/n$ give rise to 'Veech curves' if and only if $k = 1$ (see [**39**]). Another way to put it is the specificity of the quadratic differential $\omega(1, n)$ among the $\omega(k, n)$; here we use the notation of [**69**, Introduction], to which we refer the reader.

We are now ready to compute the monodromy of certain algebraic eigencurves. Recalling the case at hand we consider the curve (4.5) with $\xi = \zeta^{-1}$ and we further assume that n is odd. The latter is a completely inessential restriction, designed only to clarify notation. The case n even could be dealt with analogously but in view of applications we would have to erase the point ∞ and dealing with the two cases simultaneously makes

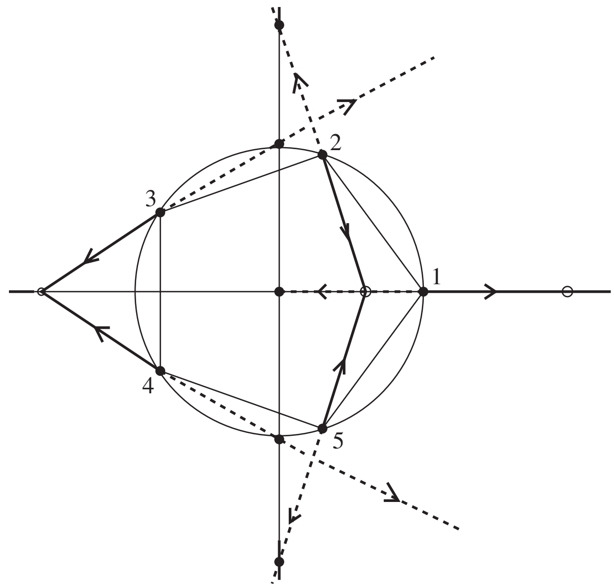

Figure 1.

life notationally quite difficult (cf. [**69**]). We can now give a precise statement; the last piece of notation appearing in it will be explained below.

Proposition 4.11. *Let $n = 2g + 1$ be an odd number, $n > 3$. The expression* (4.5) *with $\xi = \zeta^{-1}$ defines an eigencurve $\mathcal{E}_c = \mathcal{E}_c(\zeta^{-1}) \subset \mathcal{M}_{0,n+1}$, associated with* (*the cyclic group generated by*) *$\gamma = \gamma_n \in \Gamma_{0,[n+1]}$ of order n. The image of $\mathcal{E}_c$ in $\mathcal{M}_{0,[n+1]}$ is an orbifold of type $\mathbb{P}^1_{(2,n,\infty)}$, with orbifold fundamental group a Fuchsian triangular group $\Delta(2, n, \infty)$. The image of this group via the monodromy map is generated by γ and the braid $\sigma_{2,n}\sigma_{3,n-1}\cdots\sigma_{g+1,g+2}$, viewed as an element of $\Gamma_{0,[n+1]}$.*

We already know that $\mathcal{E}_c$ is an eigencurve for γ which is a maximal automorphism, i.e. the corresponding special loci is reduced to a point. We also know that γ (or rather its image $c \in \mathcal{S}_n \subset \mathcal{S}_{n+1}$) induces the rotation Λ on $\mathcal{E}_c$. As mentioned already $\mathcal{E}_c$ is also stable under the involution I. In fact changing t to $1/t$ in (4.5) and multiplying out by t, we find that I acts like the involution $i \mapsto n + 2 - i$, $2 \leqslant i \leqslant n$, that is the permutation $(2n)(3n-1)\cdots(g+1g+2) \in \mathcal{S}_n$. We also note that if $t \in \mathbb{R}$, this amounts to complex conjugation. In other words, $\mathcal{E}_c$ sits in the real locus of the quotient $\mathcal{M}_{0,n+1}/I$ (and *a fortiori* of $\mathcal{M}_{0,[n+1]}$). Speaking of coverings, we have the successive quotients $\mathcal{M}_{0,n+1} \to \mathcal{M}_{0,n+1}/C$, $\mathcal{M}_{0,n+1}/C \to \mathcal{M}_{0,n+1}/D$ as well as $\mathcal{M}_{0,n+1}/D \to \mathcal{M}_{0,[n+1]}$, the latter being unramified when restricted to the image of $\mathcal{E}_c$. It is thus enough to study $\mathcal{E}_c/D$.

In the statement of the proposition we used the notation $\mathbb{P}^1_{(p,q,r)}$ to denote the one-dimensional orbifold defined as $\mathbb{P}^1$ with three points of ramification of order p, q, r, respectively ($2 \leqslant p, q, r \leqslant \infty$); it will always be of hyperbolic type ($1/p + 1/q + 1/r < 1$). Its orbifold fundamental group is a triangle group $\Delta(p, q, r)$ generated by three elements

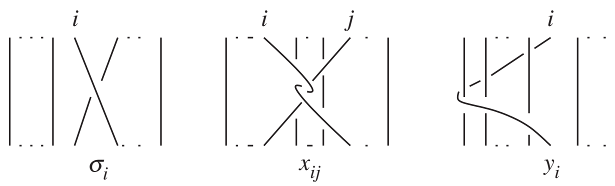

Figure 2.

x, y, z with relations $x^p = y^q = z^r = xyz = 1$. All this can easily be translated into the more algebraic language of stacks if and when it needs to be, that is when dealing with the arithmetic Galois action, as in [**46**].

Returning to the quotient $\mathcal{E}_c/D$, a little contemplation of formula (4.5) will confirm that it is indeed of type $\mathbb{P}^1_{(2,n,\infty)}$. In the parametrization we use, the three distinguished points (which could be put anywhere since $\mathbb{P}^1_{(p,q,r)}$ is rigid) lie at $t = 0$ (ramification of order n), $t = 1$ (puncture, that is infinite ramification order) and $t = -1$ (order 2). Here we actually mean the D-orbits of these values, that is, respectively, $t = \{0, \infty\}, \mu_n, -\mu_n$.

There remains to make explicit and prove the last assertion of the proposition, which requires some concrete background and notation on braid groups. However, a cinematographical understanding of the essence of the matter may be gained by staring at Figure 1. We have drawn the case $n = 5$; the generalization to any odd (or also even for that matter) n is straightforward. It will be instructive to compare this picture with Figure 3 appearing in § 5 below.

We have sketched the motion of the points $x_i(t)$ as parametrized in (4.5). For $t = 0$ we get $x_i(0) = \zeta^i$, $i = 1, \dots, n$, and the points sit at the vertices of the regular n-gon. When t runs along the segment $(0, 1)$ the points move with speed 1 along the thicker lines until at $t = 1$ a stable collision occurs, the x_i meeting in the circles with the x_{n-i} $(i = 2, \dots, g+1)$. On the other hand, as t runs along the segment $(0, -1)$ the points $x_i(t)$ follow the dashed lines and find themselves at $t = -1$ in the bullets on the imaginary axis; the point $t = -1$ is ramified for I and the configuration at $t = -1$ is stable under I, i.e. complex conjugation since t is real.

The computation of the monodromy necessitates introducing standard notation for the braid groups. We recall a bare minimum, referring essentially to Figure 2. Background material can be found, for example, in [**5**], including the detailed correspondence between braids and Dehn twists, etc.

First recall that the plane or Artin braid group on n strands B_n is generated by the σ_i, $i = 1, \dots, n-1$, with simple relations; namely, σ_i and σ_j commute for $|j - i| > 1$ and $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$. The braid group of the sphere on n strands H_n has just one added relation, namely $y_n = 0$; it is easy to see pictorially that this relation does hold for braids on the sphere. The y_i commute and the centre of B_n is free cyclic generated by $z_n = y_2y_3 \cdots y_n$ (the 'barber shop braid'). Its image in H_n is of order 2 and generates the centre of that group. Finally, the full modular group $\Gamma_{0,[n]}$ is the quotient of H_n by its centre. It is thus generated by the σ_i with the relations above and $y_n = z_n = 1$. Now the *pure* braid group of the plane on n strands is generated by the $x_{i,j}$, $1 \leqslant i < j \leqslant n$; see for instance the references quoted above for the relations these elements satisfy. The

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

pure braid group of the sphere is obtained again by adding the relation $y_n = 1$ and the pure mapping class group $\Gamma_{0,n}$ by imposing $z_n = 1$ on top of it.

We also need to recall the meaning of the $\sigma_{i,j}$ appearing in the statement of the proposition. They are introduced and used in [**43**] (see the appendix of that paper for detail; or see [**46**]). In fact σ_{ij} is defined just as σ_i, only intertwining strands i and j rather than i and $i+1$, as on the first picture of Figure 2; so $\sigma_i = \sigma_{i,i+1}$. It is actually more natural to imagine the strings of a braid hanging from a ring and attached at equally spaced points. The strings i and j are then intertwined along a chord, when viewed from above. The $\sigma_{i,j}$ can be used to study the groups as above, yielding more symmetric relations because all pairs of indices are treated on an equal footing. We also note that $x_{ij} = \sigma_{ij}^2$.

Let us return to the statement of Proposition 4.11. We have to choose a basepoint for the monodromy and we naturally pick the point on $\mathcal{M}_{0,[n+1]}$ corresponding to the curve with automorphism γ of order n, that is the point with parameter $t = 0$. This is an orbifold point and we should actually use tangential basepoints here, but the indetermination is only a global conjugacy by a power of γ, which is of no importance here. This point also corresponds to the definition of the σ_{ij} (among which the σ_i) recalled above. So we may first take γ of order n, namely $\gamma = \sigma_{n-1}\sigma_{n-2}\cdots\sigma_1$, reading from right to left. Second, and going back to Figure 1, we see that at $t = 1$ the points with labels k and $n-k$ collide. Consider the loop in $\mathcal{M}_{0,[n+1]}$ which on $\mathcal{E}_c$ is parametrized as follows: first let t vary between 0 and $1-\varepsilon$ on the real axis ($\varepsilon > 0$ small; one could use tangential basepoints on $\mathcal{E}_c$ to the same effect); then perform a half-loop counterclockwise around $t = 1$, reaching the point $t = (1-\varepsilon)^{-1}$; finally, let t run back to 0 along the real axis again. Here we recall that the group D contains the involution I which leaves $t = 1$ fixed. The above described loop in moduli space corresponds to the braid $\sigma_{2,n}\sigma_{3,n-1}\cdots\sigma_{g+1,g+2}$ or rather its image in the modular group $\Gamma_{0,[n+1]}$. It can be seen as usual from at least three closely related viewpoints:

(i) the points k and $n-k$ travel around each other and exchange places;

(ii) one effects a product of Dehn twists around loops on the marked sphere containing k and $n-k$ and no other point (these twists obviously commute)—this viewpoint will be especially relevant in proving Proposition 5.8;

(iii) the point in moduli space goes around the stratum of the divisor at infinity corresponding to this stable multicollision $k = n-k$ ($k = 1, \ldots, g+1$).

This braid has infinite order; note that its square is pure, equal to $x_{2,n}x_{3,n-1}\cdots x_{g+1,g+2}$ and together with γ they generate the monodromy group because they are images via the monodromy map of generators of $\pi_1(\mathcal{E}_c)$. This completes the proof of Proposition 4.11. One then confirms that the monodromy map is injective so that the monodromy group is itself isomorphic to $\pi_1(\mathcal{E}_c) \simeq \Delta(2, n, \infty)$. To this end it is enough to write down explicit generators x, y, z in $\Gamma_{0,[n+1]}$ which satisfy the relations of a triangular group of type $(2, n, \infty)$ as recalled above; details are given in [**46**] (which contains more than is needed here).

There remains to say but a few words on the corresponding hyperelliptic case, for future use. Let $g \geqslant 2$ and the moduli space or rather orbifold $\mathcal{M}_g$; let $\mathcal{H}_g \subset \mathcal{M}_g$ be the hyperelliptic locus. There is a natural forgetful morphism $\mathcal{H}_g \to \mathcal{M}_{0,[2g+2]}$ mapping a hyperelliptic curve to its quotient by the hyperelliptic involution and marking the ramification points. It is a homeomorphism, giving rise to the following non-split exact sequence of orbifold fundamental groups:

$$1 \to \mathbb{Z}/2\mathbb{Z} \to \pi_1^{\mathrm{orb}}(\mathcal{H}_g) \to \Gamma_{0,[2g+2]} \to 1, \tag{4.6}$$

where the kernel is of course generated by the hyperelliptic involution. Note that the above is equally valid if one considers the $\mathbb{Q}$-stack $\mathcal{M}_g$ and uses the geometric stack fundamental group, which amounts to replacing the intervening groups by their profinite completions.

Now the above can be repeated with the rational curve in $\mathcal{H}_g$ whose affine equation for the value t of the parameter is given as

$$y^2 = \prod_{k=1}^{2g+1} (x - x_k(t)), \tag{4.7}$$

with the $x_k(t)$ as in (4.5): $x_k(t) = \zeta^k + \zeta^{-k}t$. The statement of Proposition 4.11 then holds verbatim in $\mathcal{H}_g$, adding in only the hyperelliptic involution, which acts ineffectively on $\mathcal{H}_g$. Note that although (4.6) is not split, we can pull back the triangular monodromy group to $\pi_1^{\mathrm{orb}}(\mathcal{H}_g)$.

5. Geodesic eigenloci and rational polygonal billiards

In this section we take up the subject of eigenloci from a different viewpoint, assuming geodesicity in a certain sense but not algebraicity. As a special case we recover the geodesic discs and curves classically associated with polygonal billiards with rational angles. We then identify some of these curves with those eigencurves appearing in § 4.2 above, putting together their geodesic and algebraic properties.

5.1. Geodesic eigenloci: definition and first properties

We fix as usual a hyperbolic type (g, n) and the attending objects $\mathcal{T}$, $\mathcal{M}$, Γ, etc. Let $H \subset \Gamma$ be a finite group and $\mathcal{T}(H) \subset \mathcal{T}$ the fixed-point set of H (cf. § 4.1). Let now $t = (X, f) \in \mathcal{T}(H)$ and $h \in H$. Then h induces an action on the cotangent space $T_t^*\mathcal{T}$ at t, which is isomorphic to the space $\mathcal{Q}(X)$ of integrable quadratic differentials on X. We thus get a linear map $h^* : \mathcal{Q}(X) \to \mathcal{Q}(X)$; it has finite order, so can be diagonalized with eigenvalues which are roots of unity. Note that the Teichmüller metric is not Riemannian, that is it does not induce a scalar product on the cotangent space. But the Weil–Petersson metric *is* Riemannian and h^* *is* an isometry with respect to the associated scalar product (see [**37**, Theorem 7.14]). This is a close analogue of the fact that the Hecke operators induce unitary operators with respect to the Petersson scalar product (see, for example, [**64**, § 3.4]) and what we do here can be partly thought of in automorphic terms.

Example. Think back to the setting in § 4.2: we have an element $\gamma \in \Gamma_{0,[n+1]}$ of order n, projecting to the n-cycle $c \in \mathcal{S}_n \subset \mathcal{S}_{n+1}$; the special locus consists of the single point P_c and the attached automorphism is just the rotation R. Now look back at the basis of $\mathcal{Q}(P_c)$, namely the differentials q_j of (2.3) in § 2.3. Clearly, since R is multiplication by ζ^{-1}, each q_j is an eigenvalue for the action of R, with eigenvalue ζ^{-j-2}. This is still valid in the hyperelliptic situation, as described by (2.5) of § 2.3, where we are only analysing that part of the cotangent space which is pulled back from genus 0 (i.e. the q_j again). In particular, for $j = 0$ we find that the differential $q_0 = (\mathrm{d}x/y)^2$ is an eigenvector for the action of R with eigenvalue ζ^{-2}. In turn the algebraic eigenlocus appearing in Proposition 4.11 corresponds to $\xi = \zeta^{-1}$, that is $\lambda = \zeta^{-2}$. In other words, the action of the automorphism induces a rotation Λ on the eigenlocus with $\Lambda(t) = \zeta^{-2}t$. Keeping this in store for the moment we note that this is no coincidence, as will appear more clearly in § 5.3 below.

Returning to the general picture, we now add two more data: first we pick $T \subset \mathcal{T}(H)$ an irreducible algebraic subvariety; if one wants to talk about arithmetic, which is part of the motivation here, the image $Z = \pi(T) \subset \mathcal{M}$ should be defined over a number field, but this is of course not necessary as far as the geometric construction is concerned. Two examples immediately come to mind, namely $T = \{t\}$ consists of a single point, and $T = \mathcal{T}(H)$ itself, in which case $\pi(T) = \mathcal{M}(H)$ is indeed defined over a number field. Lastly we choose an irreducible representation $\rho \in \hat{H}$. For any $t \in T$ let $E_\rho(t) \subset T_t^*\mathcal{T}$ be the ρ-isotypic component, i.e. the subspace on which H acts through ρ. Then we have following lemma.

Lemma 5.1. *The subspaces $E_\rho(t)$ for $t \in T$ glue into an algebraic involutive bundle E_ρ over T.*

The fact that the fibres $E_\rho(t)$ glue into an algebraic bundle results from known results (cf. [**4**]). The fact that E_ρ is involutive is a local differential geometric fact which is a consequence of the local triviality of the cotangent bundle $T^*\mathcal{T}$. After a possible t-dependent conjugation one can locally consider $\rho = \rho_t$ as a constant representation. We will return to this local situation below (see (5.1)).

From now on, in order to avoid trivial exceptions, we assume that $E_\rho(t)$ is not reduced to the origin, that is we choose a ρ which actually occurs in the decomposition of $T_t^*\mathcal{T}$ for some $t \in T$, and hence for all $t \in T$ since we have assumed that T is connected. We may now give the following definition.

Definition 5.2. The (geodesic) *eigenlocus* associated with the finite group $H \subset \Gamma$, its irreducible representation $\rho \in \hat{H}$ and with *centre* $T \subset \mathcal{T}(H)$ is the manifold $\mathcal{T}(H, \rho, T) \subset \mathcal{T}$ spanned by the Teichmüller geodesics emanating from the points $t \in T$ along the vectors belonging to the fibre $E_\rho(t)$ of E_ρ at t.

So, thinking in terms of complex geodesics, $\mathcal{T}(H, \rho, T)$ is foliated by geodesic discs $D(t, q)$ for $q \in E_\rho(t)$. One can restrict of course to the unit bundle, that is consider normalized integrable quadratic differentials. Beware of the fact that these loci are not totally geodesic in general, except in the one-dimensional case. That they are globally stable under the action of the fixed finite group H is now an easy consequence of the definition, but not part of it, as with algebraic eigenloci (cf. Definition 4.8).

Proposition 5.3. *For any H, ρ, T, the eigenlocus $\mathcal{T}(H,\rho,T)$ is globally preserved under the action of H.*

Indeed any $h \in H$ induces as above a linear action on $T^*\mathcal{T}$ and this action preserves the bundle E_ρ. Moreover, h being an isometry, its maps a disc $D(t,q)$ ($t \in T$, $q \in T_t^*\mathcal{T}$) to another such disc, so it preserves the locus $\mathcal{T}(H,\rho,T)$.

These loci again generalize the *special* loci $\mathcal{T}(H)$ (here in Teichmüller space). Indeed we note that for *any* $T \subset \mathcal{T}(H)$, one has $\mathcal{T}(H,1,T) = \mathcal{T}(H)$. In particular, it is enough to pick one point $T = \{t\}$ corresponding to a curve having an automorphism group containing H in order to recover the whole of $\mathcal{T}(H)$; here we use the fact that $\mathcal{T}(H)$ is totally geodesic.

As a next piece of information we note that the elements of the modular group permute the eigenloci, just as they do with special loci. More precisely, we record the following easy statement.

Lemma 5.4. *The modular group Γ permutes the geodesic eigenloci; given (H,ρ,T) as above and $\gamma \in \Gamma$, we have $\gamma \cdot \mathcal{T}(H,\rho,T) = \mathcal{T}(\gamma H\gamma^{-1},\rho,\gamma \cdot T)$.*

Indeed γ permutes the geodesic discs and the linear map $\gamma^* : T_t^*\mathcal{T} \to \mathcal{T}_{\gamma\cdot t}^*\mathcal{T}$ preserves the conjugacy class of the representation.

The local situation can be analysed as follows. We are interested in points $t \in T \subset \mathcal{T}(H)$ which by definition are ramified for the canonical projection $\pi : \mathcal{T} \to \mathcal{M}$, being fixed points of some $h \in \Gamma$. In fact the corresponding point $X = \pi(t) \in \mathcal{M}$ is usually (that is except for a few low-dimensional exceptions) a singular point of the coarse moduli space. On the other hand, because $\mathcal{T}$ itself is smooth, one can in principle compute the eigenlocus $\mathcal{T}(H,\rho,T)$ by locally linearizing the action of H. To this end we use a classical averaging formula, dating back to H. Cartan and probably before. It actually works for the smooth action of any finite group H on a regular variety V. Given a local coordinate system (z) on V near a point P $(z(P) = 0)$, the new system (z') is defined via the following averaging formula:

$$z' = \frac{1}{|H|}\sum_{h\in H}(\mathrm{D}h(0))^{-1}h(z), \tag{5.1}$$

where $\mathrm{D}h$ is the differential of h. The system (z') is regular because the change of variable $z \to z'$ is tangent to the identity. Moreover, it is easily checked that in the coordinate system (z'), the action is linear, that is any $h \in H$ now acts via its differential $\mathrm{D}h(0)$ at P. Applying this to our situation, we see that given a coordinate system near a point $t \in T$, the eigenlocus $\mathcal{T}(H,\rho,T)$ locally 'coincides' in the new system with the bundle E_ρ. The above can be applied not only to $\mathcal{T}$ but also to any cover of $\mathcal{M}$ such that at least one image of t on it is a regular point. In particular, one can use finite Galois covers of $\mathcal{M}$, given for example by level structures. Moreover, the same linearization formula holds over any algebraically closed field, provided the characteristic is prime to $|H|$.

We know next to nothing about geodesic eigenloci when defined in this generality, although they appear as natural objects to look at. Given such a locus $\mathcal{T}(H,\rho,T)$ we let $\mathcal{M}(H,\rho,T) = \pi(\mathcal{T}(H,\rho,T))$ denote its projection in the moduli space $\mathcal{M}$. One of the main questions is again the *algebraization problem*: for which sets of data (H,ρ,T) is $\mathcal{M}(H,\rho,T)$ algebraic? Note that it is actually enough to show that $\mathcal{T}(H,\rho,T)$ projects to an analytic subvariety in $\mathcal{M}$, which in turn is the same as asking whether or not $\mathcal{M}(H,\rho,T)$ is locally closed for the complex topology. The fact that this always holds true for *special* loci ($\rho = 1$) boils down to the fact that Γ acts properly discontinuously on $\mathcal{T}$ (see, for example, [**20**]).

The algebraization problem inquires about which geodesic eigenloci are also algebraic eigenloci in the sense of the previous section. One may then further inquire whether or not $\mathcal{M}(H,\rho,T)$ is arithmetic. Regarding this question we note that this is easy to decide, as it depends only on the centre $Z = \pi(T)$.

Proposition 5.5. *If $\mathcal{M}(H,\rho,T)$ is algebraic, then it is arithmetic if and only if its centre $Z = \pi(T)$ is.*

Indeed this can be decided by looking at any finite cover of $\mathcal{M}$. So replace $\mathcal{M}$ with a regular cover and look at the linearization formula (5.1). It shows that the local equations of the isotypic components are indeed defined over number fields if those of the centre are. Here we use the fact that the action of Γ is itself arithmetic, i.e. defined over $\bar{\mathbb{Q}}$: as a result, in the notation of (5.1) one can pick a coordinate system (z) which is compatible with the $\bar{\mathbb{Q}}$-structure and the equations for $h(z)$ will then have coefficients in $\bar{\mathbb{Q}}$.

The special locus $\mathcal{M}(H)$ in moduli space can be considered as an orbifold, with an action of H on its preimage (any preimage) in Teichmüller space which is not effective. For non-trivial ρ however, and assuming that $\mathcal{M}(H,\rho,T)$ is algebraic, the latter defines an orbifold equipped with an effective action of H in Teichmüller space or on any sufficiently large covering. The eigenloci which descend to varieties in the moduli spaces thus hopefully provide a large collection of orbifolds which contain the finite subgroups of the Teichmüller modular groups in their orbifold fundamental groups. After putting it in a more algebraic fashion, this point may be of real interest in the perspective of Grothendieck–Teichmüller theory.

Before moving to more modest but also more concrete matters, we note that there might exist algebraic eigenloci with, so to speak, empty centres. That is one can ask the following question: does there exist $H \subset \Gamma$ a finite group and $\mathcal{E} \subset \mathcal{M}$ an algebraic variety such that H stabilizes (an irreducible component of) $\pi^{-1}(\mathcal{E}) \subset \mathcal{T}$ and acts freely (equivalently, fixed-point free) on it?

5.2. Computing the dimensions of the geodesic eigenloci: the Chevalley–Weil formula

The first *a priori* quite modest piece of information in the general case may well be the computation of the dimensions of the geodesic eigenloci. This turns out to be not so modest a task but fortunately appears as a special case of a beautiful problem first considered and solved by A. Weil and C. Chevalley; we will use the modern version given in [**15**] to which we refer for detail and complements (see also [**4**]). We include the computation for completeness and also because adding marked points here is a little less classical.

Let H, $\rho \in \hat{H}$, $T \subset \mathcal{T}$, and $\mathcal{T}(H,\rho,T)$ be as above. If ρ is the trivial representation, $\mathcal{T}(H,\rho,T)$, or rather its projection $\mathcal{M}(H,\rho,Z) \subset \mathcal{M}$ ($Z = \pi(T)$), is a special locus, the dimension of which is well known (see Fact 4.1 above). We assume henceforth that ρ is not trivial. Let $t \in T$, that is $t = (X, f)$ is a marked Riemann surface, say of hyperbolic type (g, n). We let m_ρ be the multiplicity of the representation ρ in the decomposition of $\mathcal{Q}(X)$ under the action of H; we assume that $m_\rho \neq 0$, that is ρ does occur in the decomposition. Then we obviously get

$$\dim(\mathcal{M}(H,\rho,Z)) = \dim(\mathcal{T}(H,\rho,T)) = \dim(T) + m_\rho \deg(\rho), \tag{5.2}$$

so that the point is to compute the multiplicity m_ρ ($\deg(\rho)$ denotes the degree or dimension of ρ). The computation below is actually valid over any algebraically closed field k, provided the action of H is reductive. In particular, if $\mathcal{M}(H,\rho,Z)$ is arithmetic (in particular algebraic), one can reduce modulo almost all p and everything works verbatim over $\bar{\mathbb{F}}_p$ provided:

(a) X has good reduction,

(b) the marked points do not coalesce, and

(c) the cardinal of H is prime to p.

In short, nothing terribly new happens at these primes. Below we denote the algebraically closed ground field by k out of habit; calling it $\mathbb{C}$ makes no difference.

So let X be a smooth projective curve of genus g with $n \geqslant 0$ marked points defined over the algebraically closed field k. Let $H \subset \mathrm{Aut}_k(X)$ be a group of automorphisms of X, whose cardinal is prime to the characteristic of k. Let $Y = X/H$ denote the quotient of genus $g' = g(Y)$ and $p : X \to Y$ the canonical projection, defining a possibly ramified Galois covering of degree $|H|$. We let ω denote the sheaf of differentials on X and $\mathcal{Q}$ the sheaf of quadratic differentials with at most simple poles at the marked points; $\mathcal{Q}(X)$ is the space of global sections of $\mathcal{Q}$. The n marked points are permuted under the action of H and for simplicity we will assume that they lie on unramified fibres of p; the formula we get is actually valid in the general case but we will not check it here. We can write $\mathcal{Q} = \omega^2 \otimes p^*\mathcal{P}$, where $\omega^2 (= \omega^{\otimes 2})$ is the sheaf of regular quadratic differentials on X and $\mathcal{P}$ is the line bundle on Y associated with the images of the marked points, so that $\deg(\mathcal{P}) = n/|H| = n'$, which is the number of the images of the marked points on Y; we have assumed that this set is disjoint from the branch locus of p, which will also play a role below.

For any coherent sheaf $\mathcal{F}$ on X with H-action, one defines the equivariant Lefschetz trace of $\mathcal{F}$ as

$$L_H(\mathcal{F}) = H^0(X,\mathcal{F}) - H^1(X,\mathcal{F}).$$

Here the cohomology groups are viewed as H-modules and so is $L_H(\mathcal{F})$; of course we give only the particular case adapted to our needs, that is when X is a curve. This is the computable entity and the point is to decompose $L_H(\mathcal{F})$ into irreducibles. Now we wish to apply this to the case $\mathcal{F} = \mathcal{Q}$ and by Serre's duality $H^1(X, \mathcal{Q})$ vanishes, so that $L_H(\mathcal{Q}) = H^0(X, \mathcal{Q}) = \mathcal{Q}(X)$, viewed again as an H-module. Moreover, by [**15**, Proposition 4.2] (which is a simple direct computation), for any $\mathcal{F}$ and any line bundle $\mathcal{P}$ on Y, one has

$$L_H(\mathcal{F} \otimes p^*\mathcal{P}) = L_H(\mathcal{F}) + \deg(\mathcal{P}) \cdot \mathrm{rk}(\mathcal{F}) \cdot k[H],$$

where $k[H]$ is the group ring, i.e. the regular representation of H. Applying this to our case, we get

$$\mathcal{Q}(X) = L_H(\omega^2) + n'k[H]. \tag{5.3}$$

Note that for low genera $g = 0, 1$, $H^1(X, \omega^2)$ does not vanish, so we really need to keep the Lefschetz trace in this equality. We are thus reduced to decomposing $L_H(\omega^2)$; more generally one can actually replace ω^2 with any power ω^m, $m \geqslant 1$, and the solution is analogous. We will record the result for $m = 2$, referring to [**15**] for detail (see also, for example, [**4**]).

We need a little more notation in order to state the result. Let $P_1, \ldots, P_s \in X$ be the ramification points of p and $Q_1, \ldots, Q_t \in Y$ their images. To any $P = P_i$ is attached a ramification index which depends only on its image $Q = p(P)$, so that we denote it either e_P or e_Q. The inertia subgroup $H_P \subset H$ at P is cyclic of order e_P: $H_P \simeq \mathbb{Z}/e_P\mathbb{Z}$ with generator, say, σ_P. The dual group $\hat{H}_P$ also has order e_P, generated by the character χ_P acting on k via multiplication by ζ_P, a primitive e_Pth root of unity. The completed local ring of the curve X at P is a ring of formal power series in one variable: $\hat{\mathcal{O}}_{X,P} \simeq k[\![u]\!]$, and we can choose the uniformizing parameter u and the root of unity ζ_P in such a way that the action of the inertia group is described by $\sigma_P(u) = \zeta_P u$. Let now ρ be an irreducible representation of H, i.e. an H-module (more precisely an H–k-module). Consider the tensor product $\rho_P = \hat{\mathcal{O}}_{X,P} \otimes_k \rho \simeq k[\![u]\!] \otimes_k \rho$ which as an H_P-module describes the local representation at P. Decomposing ρ_P, we let $n_{d,P,\rho}$ $(1 \leqslant d \leqslant e_P)$ denote the multiplicity of the character χ_P^d in ρ_P. This number is actually constant along the fibres, so depends only on $Q = p(P)$; we write $n_{d,P,\rho}$ or equivalently $n_{d,Q,\rho}$.

As an example, consider the case of H cyclic of order r; the case H commutative can be worked out in much the same way. Let $\sigma \in H$ be a generator, and χ a generator of $\hat{H}$, with $\chi(\sigma) = \zeta$, a primitive rth root of unity. We can write $\rho = \chi^\ell$ for some integer u, $1 \leqslant u \leqslant r$. Let $r_P = r/e_P$ denote the number of points in the fibre at P; so $r_P = r_Q$ again depends only on $Q \in Y$. Let $\sigma_P = \sigma^{r_P}$ generate H_P and $\zeta_P = \zeta^{r_P}$. The local representation at P is given by χ_P such that $\chi_P(\sigma_P) = \sigma_P^{k_P}$ for some integer k_P $(1 \leqslant k_P \leqslant e_P)$. On the other hand, the restriction of the character χ to H_P is determined by $\chi(\sigma_P) = \chi(\sigma^{r_P}) = \zeta^{r_P} = \zeta_P$. We thus find that, when restricted to H_P, $\chi = \chi_P^{j_P}$, where j_P is an inverse of k_P in $\mathbb{Z}/e_P\mathbb{Z}$, i.e. $j_P k_P = 1 \bmod (e_P)$. In turn ρ restricted to H_P is equal to $\chi_P^{\ell j_P}$ and the representation ρ_P as defined above is equal to $\chi_P^{1+\ell j_P}$. So in this case $n_{d,P,\rho}$ is 1 or 0 according to whether $1 + \ell j_P = 1 \bmod (e_P)$ or not.

Returning to the case of an arbitrary finite group H, one finds (cf. [**15**]) that the multiplicity of $\rho \in \hat{H}$ in the trace $L_H(\omega^2)$ is equal to

$$(3g'-3)\dim(\rho) + \sum_{Q\in Y}\sum_{d=0}^{e_Q-1} n_{d,Q,\rho}\biggl(1 - \frac{1}{e_Q} + \biggl\langle\frac{1-d}{e_Q}\biggr\rangle\biggr),$$

where $\langle x\rangle$ denotes the fractional part of x ($0 \leqslant \langle x\rangle < 1$). Putting this together with (5.3) and remembering that ρ occurs in the regular representation $k[H]$ a number of times equal to its dimension, we finally get the following proposition.

Proposition 5.6. *With the above notation the H-module $\mathcal{Q}(X)$ of integrable quadratic differentials decomposes as*

$$\mathcal{Q}(X) = \sum_{\rho\in\hat{H}} m_\rho \rho,$$

with the multiplicity m_ρ given by

$$m_\rho = (3g'-3+n')\deg(\rho) + \sum_{Q\in Y}\sum_{d=0}^{e_Q-1} n_{d,Q,\rho}\biggl(1 - \frac{1}{e_Q} + \biggl\langle\frac{m-1-d}{e_Q}\biggr\rangle\biggr). \qquad (5.4)$$

Together with formula (5.2) this completes the computation of the dimensions of the geodesic eigenloci.

5.3. Geodesic eigencurves and rational polygonal billiards

The name ‘eigenlocus’ was chosen in reference to the abelian case: if H is abelian, the representation ρ is one dimensional and $E_\rho(t)$ is a simultaneous eigenspace of the operators h^* ($h \in H$) acting on $T_t^*\mathcal{T} \simeq \mathcal{Q}(X)$ ($t = (X, f)$). Specializing again, just pick one finite-order diffeomorphism $h \in \Gamma$, giving rise to a cyclic group $H = \langle h\rangle$ with a preferred generator. The character ρ is then determined by $\rho(h) = \lambda$, which is a root of unity of order dividing the order of h. We then write E_λ instead of E_ρ and $E_\lambda(t)$ is thus nothing but the λ-eigenspace of h^* acting on $T_t^*\mathcal{T} \simeq \mathcal{Q}(X)$. If one further assumes that the eigenvalue λ is simple, one is lead to the important one-dimensional case. One-dimensional eigenloci, we call *eigendiscs* in Teichmüller space; an *eigencurve* is the projection of an eigendisc to moduli space, provided this projection is locally closed, in which case it is an algebraic curve. From now on we will restrict ourselves to the case of cyclic groups for simplicity.

So we consider a point $t = (X, f) \in \mathcal{T}$, a finite-order diffeomorphism h, the linearized action of h^* on $\mathcal{Q}(X)$ and an eigenvalue λ of h^*. Letting $r > 1$ denote the order of h, λ is an rth root of unity and we assume that $\lambda \neq 1$. We do *not* assume that λ is simple, thus slightly abusing the general Definition 5.2 in order to connect more obviously with §§ 2–4. We simply pick $q_\lambda \in E_\lambda \subset \mathcal{Q}(X)$, that is a λ-eigenvector of h^*. We are then clearly back to the setting of § 2. The data $(t, q_\lambda) \in T^*\mathcal{T}$ define a geodesic disc $D_\lambda = D_\lambda(t)$ which is an eigendisc. In case the projection $\pi(D_\lambda) \subset \mathcal{M}$ of the eigendisc D_λ is a curve, we denote it by C_λ. There are only preciously few known examples, which originally arose from the study of rational polygonal billiards and until very recently were all essentially patterned after the original example of Veech in [**69**], to which we will return shortly. In general the curve C_λ is determined by the pair (X, q_λ) or else by X and the flat structure $u_\lambda = u(q_\lambda)$ associated with q_λ as in § 2.4. Such curves X with these flat structures are often called *Veech curves* in a dynamical systems context.

The curve C_λ can be written as $C_\lambda = \Gamma(D_\lambda)\backslash D_\lambda$, with the exact sequence (2.7) of § 2.5, which we here rewrite as

$$1 \to H_\lambda \to S(D_\lambda) \to \Gamma(D_\lambda) \to 1. \tag{5.5}$$

Beware of the unfortunate clash of notation with respect to § 2.5. The finite kernel H_λ records the fact that C_λ may be included in a special locus: in fact $C_\lambda \subset \mathcal{M}(H_\lambda)$; generically $H_\lambda = \{1\}$ and $\mathcal{M}(H_\lambda) = \mathcal{M}(\{1\}) = \mathcal{M}$. Since $\lambda \neq 1$, $\Gamma(D_\lambda) \subset \Gamma$ contains $H = \langle h \rangle$ and $H \cap H_\lambda = \{1\}$. Now h acts on D_λ isometrically for the Teichmüller metric, which coincides with the Poincaré metric when restricted to D_λ because it is a geodesic disc. Since h has finite order, it implies that h acts as a rotation of order r (the order of λ) on the disc D_λ with centre t: in this cyclic one-dimensional case we get a copy of the Poincaré disc with the action of a finite-order rotation. The group $\Gamma(D_\lambda)$, which can also be regarded as a subgroup of $PSL_2(\mathbb{R})$, has torsion by definition and is usually *not* commensurable with $PSL_2(\mathbb{Z})$, contrary to what happens with the origamis of § 4. In that sense origamis and eigencurves are very different, although they do not describe strictly disjoint sets of geodesic curves in the moduli spaces. But they share the property of being defined over number fields, assuming in the case of eigencurves that $X = \pi(t)$ is itself defined over a number field (cf. Proposition 5.5).

As a final topic we now turn to billiards and the identification of algebraic and geodesic eigenloci (in fact eigencurves) in a very special case. We need to recall the first inputs on rational polygonal billiards. Generally speaking a 'billiard' consists of a plane bounded domain with piecewise smooth boundary inside which one studies the motion of a point (a dimensionless ball) subject only to the usual laws of reflection on the boundary. Billiards in general have been a central object of study in dynamical systems for more than a century now, ever since Poincaré and Hadamard in particular singled them out as model systems exhibiting some general phenomena in a particularly pure and striking way. For a general survey (of necessity very much incomplete) and references on billiards, we refer to [**67**]. Billiards in polygons embody the parabolic (in the sense of dynamical systems) side of the question. Because conics have their whims, we will actually use this parabolic aspect in order to study the elliptic aspects of the modular groups by means of hyperbolic curves.

Specializing again we now focus on billiards in rational polygons. For our immediate purposes it is enough to define a *rational polygon* as a convex polygon P whose angles are commensurable with π. Following [**67**, § 3.4], to which we refer for details (among many other places), we let n be the number of sides and $\pi m_i/n_i$ be the values of the angles (with m_i and n_i coprime). Using a geometric construction which has been rediscovered again and again since the thirties, one can build a compact Riemann surface by pasting together $2N$ copies of P, where N is the least common multiple of the n_i. Roughly speaking one uses reflections in the sides of P for the gluing. Put slightly differently, let D_{2N} be the dihedral group generated by the reflections in lines through the origin of the plane meeting at angle π/N. Upon reflection on the boundary, a trajectory inside P changes direction according to the action of an element of D_{2N} and this action of D_{2N} on the directions provides a rule for pasting the $2N$ copies of P. Viewing P as a domain of the complex plane, one gets a Riemann surface $X = X(P)$, whose genus $g = g(P)$ can be computed using Euler's formula. To wit one finds

$$g(P) = 1 + \tfrac{1}{2}N \sum_{i=1}^{n} \frac{m_i - 1}{n_i}. \tag{5.6}$$

We assume from now on that $g(P) \geqslant 2$, that is we exclude the four so-called integrable cases, corresponding to linear flows on the torus. Let $C_N \subset D_{2N}$ be the cyclic subgroup of index 2 consisting of the orientation-preserving transformations: it is generated by the rotation with angle $2\pi/N$. The elements of C_N translate into (complex) automorphisms of X, whereas the other elements of D_{2N}, corresponding to reflections, provide real analytic, antiholomorphic transformations. The quotient surface X/C_N has genus 0. In fact it is obtained by sewing two copies of P with opposite orientations. It follows that X/C_N is a sphere with n unordered marked points, forming a real divisor.

One then notices that the plane polygon P comes equipped with the natural differential form $\omega_0 = \mathrm{d}z$ and thus the quadratic form $q_0 = \omega_0^2 = \mathrm{d}z^2$. Considering again $2N$ copies of P with ω_0 and q_0 on each copy, these glue into ω and q on X. We get a pair (X, q) and a geodesic disc $D = D(P) \subset \mathcal{T}_g$, after fixing a marking $f : S_g \to X$; this choice plays no role in the sequel. Moreover, the automorphism group C_N is generated by a rotation R with $R(z) = \rho z$, ρ an Nth root of unity (which it is best not to fix at this point). It corresponds to a rotation Λ of the disc $D \subset \mathcal{T}_g$ given by $\Lambda(t) = \lambda t$, where the centre $t = 0$ is the point $(X, f) \in \mathcal{T}_g$. Recalling how q is constructed from q_0 one finds that $\lambda = \rho^2$.

We summarize the above in the following statement, in which we slightly modify the notation in order to tune it to that of the last subsections.

Proposition 5.7. *A plane rational n-gon P with angles $\pi m_i/n_i$ $(i = 1, \dots, n)$ defines a compact Riemann surface $X = X(P)$ of genus $g = g(P)$ as in* (5.6) *above. It has a cyclic group of automorphisms H of order $N = \mathrm{l.c.m.}(n_i)$ and the quotient surface X/H is a sphere marked with n unlabelled points which span a real divisor.*

The Riemann surface X comes equipped with a holomorphic quadratic differential q. After choosing a marking f, the point $(X, f, q) \in T^*\mathcal{T}_g$ defines a geodesic disc $D = D(P) \subset \mathcal{T}_g$ which is an eigendisc for H; the latter group acts on D via rotations of order N and centre $(X, f) \in \mathcal{T}_g$.

We note that there are obvious points of similarity but also of difference with Proposition 3.2 for origamis. The data are Euclidean, not merely topological or combinatorial and there is no guarantee that D descends to an algebraic (much less arithmetic) curve in $\mathcal{M}_g$. It may very well be that the stabilizer of D in Γ_g is trivial. Although Proposition 5.7 contains nothing new technically speaking, looking at the situation from this viewpoint may be helpful, as illustrated below. In short the above statement underlines that rational polygons do provide an ample source of geodesic eigendiscs, which, however, are of a particular kind, especially in that the quotient surface has genus 0. In other words, the centre of the disc can be seen as a point of a Hurwitz space of the simplest kind, classifying Galois covers of spheres with n marked points which are cyclic of order N. The latter space is a finite cover of $\mathcal{M}_{0,[n]}$. It might be interesting to let these structures, or equivalently (see, for example, [**38**]) spin curves, come into play.

Proposition 5.7 has nothing to do in principle with dynamical systems in general and billiards in particular. That this type of geometry has dynamical relevance was realized in the early 1980s by several people (among whom were S. Kerchkoff, Masur, J. Smillie and Veech) in the wake of Thurston's program. There exists at present a large body of literature on the subject; Veech was the first to examine (in [**69**]) the algebraization problem in that setting, that is in the present terms inquire whether the disc $D(P)$ in Proposition 5.7 projects to an algebraic curve in moduli space. He also demonstrated the dynamical relevance of that property. Among the many papers on the subject, we refer to [**69**], [**71**], [**28**], [**31**], [**72**], [**39**], [**25**] and [**52**] for recent and less recent results, various viewpoints, condensed surveys and further references. Until very recently, very few algebraizable geodesic discs were known outside of those coming from origamis, and the examples were essentially patterned after the main result of [**69**], to which we will shortly return. The situation may be changing with the recent papers [**49**] and [**8**].

The examples of algebraizable geodesic discs known to date are not explicit in the sense that the curves are shown to exist but not explicitly exhibited. The purpose of our next and last proposition is precisely to explicitly identify the eigencurves whose existence was proved in [**69**]. We thus turn to the case where the polygon P is a triangle and at first very specifically to the case of an isosceles triangle T_n with angles $(\pi/n, \pi/n, (n-2)\pi/n)$ and we assume for simplicity that $n > 3$ is odd, $n = 2g+1$; in particular T_n is an obtuse triangle. It is not difficult to confirm (see [**69**, § 8]) that the curve $X_n = X(T_n)$ is the hyperelliptic curve of genus g with affine equation $y^2 = x^n - 1$. The gluing actually goes as follows: consider the regular n-gon P_n with vertices at the nth roots of unity; let $e_n = \exp(2i\pi/n)$ and S_n the triangle with vertices $(0, 1, e_n)$. Now reflect S_n in its short side $(1, e_n)$, getting a reflected copy S_n^* of S_n; let $R_n = S_n \cup S_n^*$ be the resulting quadrilateral. Cutting R_n along its long diagonal one gets two copies of T_n. Now the curve X_n can be seen as two copies of P_n which are exchanged by the hyperelliptic involution, acting on the complex plane via $(z \mapsto -z)$. The construction above shows that it can also be obtained by gluing $2n$ copies of T_n, in accordance with Proposition 5.7 since here $N = n$. The curve X_n has an obvious automorphism of order n $(= N)$, the quotient being a sphere with three marked points. The form $\mathrm{d}z$ (respectively, the quadratic form $\mathrm{d}z^2$) on the complex plane (or on T_n) lifts to the form $\omega = \mathrm{d}x/y$ on X_n (respectively, the quadratic form $q = (\mathrm{d}x/y)^2$). Veech proved (see [**69**, Theorem 1.1]) that the geodesic disc $D_n = D(X_n, q)$ defined by (X_n, q) (plus an unimportant arbitrary marking) is algebraizable; in fact he explicitly computed its stabilizer in $\varGamma_g$, which turns out to be a triangle group of type $\Delta(2, n, \infty)$. This gives rise to an algebraic curve $C_n \subset \mathcal{M}_g$; in fact C_n is contained in the hyperelliptic locus $\mathcal{H}_g$ because the quadratic form $q = (\mathrm{d}x/y)^2 = \mathrm{d}x^2/y^2 = \mathrm{d}x^2/(x^n - 1)$ is the pull back of a form in genus 0 by the hyperelliptic involution. One can thus also consider

C_n as a curve in $\mathcal{M}_{0,[2g+2]} = \mathcal{M}_{0,[n+1]}$, which we still denote by C_n (in order to avoid confusion, our X_n is Veech's C_n in his Theorem 1.1).

Proposition 5.8. *The curve C_n viewed in $\mathcal{M}_{0,[n+1]}$ coincides with the algebraic eigencurve described in Proposition 4.11; as a curve in $\mathcal{M}_g$ it is described by formula* (4.7) *of* § 4.

In order to prove this statement we use a powerful and deep property, namely rigidity of the monodromy, which states that a non-constant map from a base curve B into a stably compactified moduli space of curves of hyperbolic type (equivalently a non-trivial stable fibration of B with smooth hyperbolic generic fibre) is entirely determined by its monodromy representation (see, for example, [**36**] and § 6.2 below). We recall that the very existence of such a map implies that the base B is itself of hyperbolic type. Here we know by Veech's main result that C_n is algebraic and the curve $\mathcal{E}_c$ in Proposition 4.11 is algebraic by definition. They both have a monodromy group which is triangular of type $\Delta(2, n, \infty)$ and we have to show that these groups coincide up to conjugacy. Both curves contain the point in $\mathcal{M}_{0,[n+1]}$ with automorphism γ of order n (the rotation of angle $2\pi/n$). We take this point as basepoint for the fundamental groups just as in Proposition 4.11. Both groups contain the element γ and the monodromy group of $\mathcal{E}_c$ contains the braid β appearing in the statement of Proposition 4.11. There remains only to show that $\beta \in \pi_1(C_n)$; actually it is enough to prove that $\pi_1(C_n)$ contains an element which is a conjugate of $\beta^{\pm 1}$ by a power of γ. This makes it possible to base the fundamental groups as we did, without worrying about the fact that it is an orbifold point because the attached automorphism group is precisely generated by γ. The fact that β appears in $\pi_1(C_n)$ is actually shown in §§ 5, 7 of [**69**] and this concludes the proof of the proposition. For the convenience of the reader, however, and for the purpose of comparison with Proposition 4.11, we sketch the picture obtained in [**69**] and illustrate it in Figure 3 for $n = 5$ ($g = 2$), to be compared with Figure 1 in § 4.

Using the same notation as above, the curve X_n ($n = 2g + 1$) can be constructed by gluing two copies P and Q of the regular n-gon; Q is a copy of P rotated by π, that is obtained by applying the map $z \mapsto -z$ on the complex plane and sides are glued using parallel translations (see [**69**, § 4] for details). The holomorphic quadratic differential q on the hyperelliptic curve X_n is the lift of an integrable quadratic differential on the plane, whose singular points are the zeros of y, in other words the Weierstrass points of X_n. There are $n+1 = 2g+2$ such points, given by the n midpoints of the edges of P (and Q), which shape another regular n-gon, and the vertices of P (and Q), which are all identified because n is odd. The vertical foliation associated with q cuts X_n into g cylinders of closed leaves $V_1, \dots, V_g$, whose sides pass through the vertices of P and Q. A key point is that the moduli of these cylinders, namely the ratios of their heights to their widths, are all equal (see (5.2) in [**69**]). One can explicitly define the Dehn twists associated with these cylinders by specifying that the boundaries of the cylinders are fixed, increasing linearly the twist from 0 to 2π between the boundaries. Because the moduli coincide, these glue into an *affine* diffeomorphism which is thus an element of $\pi_1(C_n)$. It is in fact no other than β. Indeed one can number the Weierstrass points as in Figure 3 (but for any n), that is counterclockwise starting from the vertical boundary of the leftmost cylinder,

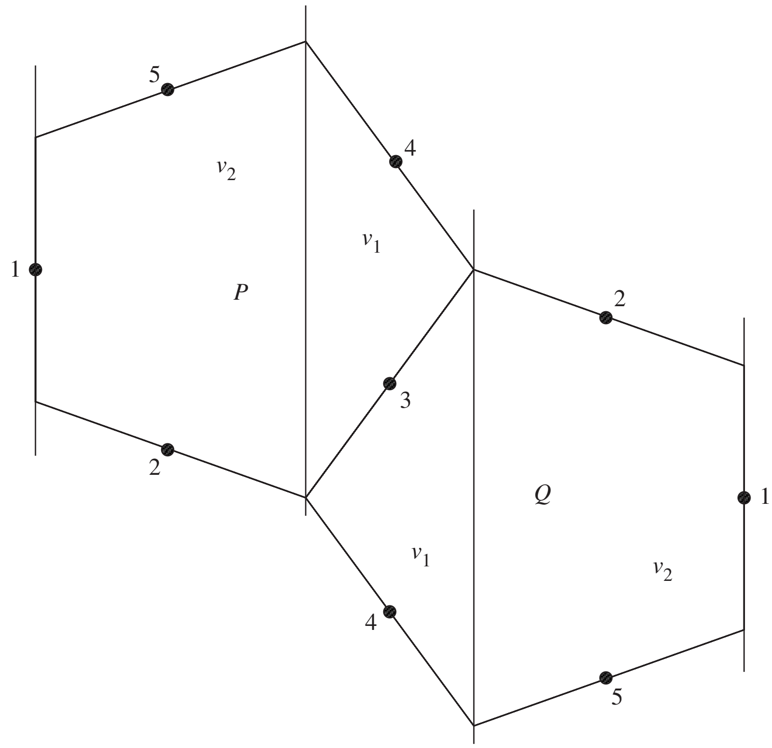

Figure 3.

plus the label $n+1$ for the vertices of P and Q (which are all one and the same point). Then the affine diffeomorphism constructed above, that is the product of the Dehn twists on the cylinders, permutes the points i and $n+2-i$ ($i = 2, \ldots, g+1$) leaving 1 and $n+1 = 2g+2$ fixed and it acts precisely like β.

Proposition 5.8 implies in particular that the rational curve described by (4.5) (or (4.7)) in § 4 with $\xi = \zeta^{-1}$ is geodesic and C_n gives an explicit example of a geodesic (algebraic) eigencurve. Treating the case of an even n requires only minor changes, mostly erasing the point ∞ in (4.5), since it is unramified for the hyperelliptic involution of the curve (4.7). The rest goes through essentially literally, but with a triangular group of different type. We have preferred to stick to the case n odd for simplicity and because a unified treatment tends to obscure the notation.

We remark that the curve C_5 (giving rise to a curve in $\mathcal{M}_2$) can be described in another (much more complicated) geometric way, originating in the work of G. Humbert [**32**] and tying it with abelian varieties (here Jacobians) with real multiplication, hence with the subject matter of [**49**]. It does not seem that this description can easily be extended to higher values of n. We refer the reader to [**73**] and [**2**] for recent sources of information.

Several papers have been devoted to the study of rational triangles (see in particular [**31**, **39**, **72**]). In the last of these papers the authors achieve a complete classification, that is they give a list of all the cases in which the disc associated with a rational triangle as in Proposition 5.7 projects to an algebraic curve in moduli space. It is interesting to note that looking at a triangle with angles $(k\pi/n, k\pi/n, (n-2k)\pi/n)$ with its canonical (quadratic) differential turns out to be equivalent to looking at X_n as above (coming from T_n which is the case $k = 1$) with another differential, namely $\omega_k = x^{k-1}\,\mathrm{d}x/y$ (and $q_k = \omega_k^2$). We refer to [**31**] for a detailed exposition and proof. These triangles do not give rise, for $k > 1$, to eigencurves. As mentioned already it would be interesting to investigate the situation more closely, comparing in particular with the algebraic eigencurves in (4.5) of § 4, which are by definition algebraic and whose lifts to Teichmüller space are tangent to the eigendiscs corresponding to the (X_n, q_k), letting ξ and k vary.

More generally one could try and generalize or apply the above to various situations, including those which are found in the literature on billiards; for instance one could investigate from this viewpoint the case of regular polygons, as treated in [**71**], which is closely related to that of triangles. One should also notice that prior to [**49**] and [**8**], all monodromy groups that had been computed were triangular. Those connected with origamis are commensurable with $PSL_2(\mathbb{Z}) \simeq \Delta(2,3,\infty)$ and are thus not terribly different from our present viewpoint.

We briefly return to (hypothetical) arithmetic, merely to remark that 'everything' that is algebraic here is actually arithmetic and indeed defined over $\mathbb{Q}^{\mathrm{ab}}$. These genus 0 objects should perhaps lend themselves to the construction of an elliptic lego, which moreover would be essentially cyclotomic in a genus 0 setting. How to implement this vague suggestion remains wholly to be seen (but see [**46**] for a start).

6. Conclusion in the form of an introduction

In this final section we will try to explain in plain words bits and pieces of what we are after, in part for the simple reason that *if* a road in that direction does exist, we surely stand only at its very beginning. Out of necessity we will be sketchy and will not even try to recall the definitions and properties of the objects we mention, as this would occupy far too much space. We also take full advantage of modern reference databases: recalling a theme and the name of an author makes it easy nowadays to retrieve the relevant references.

6.1. Glimpses of Grothendieck–Teichmüller theory

Grothendieck–Teichmüller theory was conceived or dreamt of by Grothendieck in his *Esquisse d'un programme* (now in [**63**]), following his *Longue marche à travers la théorie de Galois*. A few seminal papers, especially [**12**] and [**33**], started giving flesh to the vision. The theory is still very much in flux, so that there does not and probably cannot exist a satisfying global survey text. Moreover, there are several possible versions which are connected in subtle ways, for instance by all kinds of linearization processes. We will be concerned here with the version which we feel is closest to the spirit of the *Esquisse*, an assertion which should clearly be taken as nothing more than a personal opinion. It may perhaps be termed 'nonlinear', as opposed to 'motivic', although such terminology would deserve ample explanation. At any rate we refer to [**60**], [**43**], the introduction to [**30**], the more recent [**42**] and references therein for very imperfect and partial accounts. The papers [**34**] and [**26**] survey different but hopefully converging paths and the monograph [**1**] explores a related landscape.

Getting started in a nutshell, skipping the necessary motivating questions, we consider the collection of fine moduli spaces $\mathcal{M}_{g,n}$ for varying hyperbolic types (g,n), together with their stable completions $\bar{\mathcal{M}}_{g,n}$. These objects were constructed in [**11**] as algebraic stacks over $\mathbb{Z}$ but we confine attention to the generic fibre and view them as regular $\mathbb{Q}$-stacks. They fit together into a category or *modular tower* which we denote by $\underline{\mathcal{M}}$.

This means that one can define smooth morphisms of geometric origin between the $\mathcal{M}_{g,n}$. There are actually several (existing or not yet existing) versions, the simplest and most classical being derived from the stable stratification already briefly mentioned in §§ 2.6 and 3.2 above; the point is that the divisor at infinity $\bar{\mathcal{M}}_{g,n} \setminus \mathcal{M}_{g,n}$ is made of products of $\mathcal{M}_{g',n'}$ of lower dimensions, up to finite morphisms. These give rise to 'natural' morphisms, also-called Knudsen or 'clutching' morphisms (see the papers by F. Knudsen). By their very definition these morphisms 'live at infinity'. One can also add the point-erasing morphisms, that is the fibration $\mathcal{M}_{g,n+1} \to \mathcal{M}_{g,n}$ defining the universal curve over $\mathcal{M}_{g,n}$ and the corresponding universal monodromy map. Finally, one can take the action of the permutation groups into account and replace the $\mathcal{M}_{g,n}$ by the $\mathcal{M}_{g,[n]}$ or variants thereof. The long and the short is that the above defines, modulo 'technical details', a version of $\underline{\mathcal{M}}$ which essentially lives at infinity and is entirely defined over $\mathbb{Q}$. All the morphisms mentioned above are $\mathbb{Q}$-morphisms, essentially because they feature algebraic counterparts of familiar topological gestures: pasting, pinching, erasing, etc. It is this version of the modular tower $\underline{\mathcal{M}}$ which gives rise to the 'lego at infinity' (see [**30**] and [**55**] in this context). In the case of genus 0 as initiated in [**12**], it is sometimes referred to as the 'geometry of associativity' because it was Drinfeld's groundbreaking idea to, so to speak, not take associativity for granted (related themes: quasi-Hopf algebras, braided categories, McLane coherence relations, universal scattering matrix, Yang–Baxter equations, gravity operad, etc.).

Not much is known to date beyond this version of the modular tower which again (a) lives at infinity, and (b) is defined over $\mathbb{Q}$. In fact (a) and (b) are far from independent and it may well be that it is essentially the largest possible tower which is entirely defined over $\mathbb{Q}$. At any rate and for the time being, having built a more or less expensive version of $\underline{\mathcal{M}}$ over $\mathbb{Q}$, one applies the geometric fundamental group functor π_1^{geom}, from the category ($\mathbb{Q}$-Stacks) of $\mathbb{Q}$-stacks with $\mathbb{Q}$-morphisms to the category (Grps) of profinite groups with continuous outer homomorphisms. In other words, if $\mathcal{X}$ is a $\mathbb{Q}$-stack, one sets $\pi_1^{\mathrm{geom}}(\mathcal{X}) = \pi_1(\mathcal{X} \otimes \bar{\mathbb{Q}})$, which is a finitely generated profinite group. The theory of the fundamental group for stacks parallels, but with interesting differences, Grothendieck's classical [**22**] theory for schemes (see [**57**] and [**45**]); for the specific case of the moduli stacks of curves, which are classifying spaces, we also refer to the contribution of T. Oda in [**63**], which uses simplicial techniques. An extended version of the Lefschetz principle implies that, considering $\bar{\mathbb{Q}}$ as a subfield of $\mathbb{C}$, $\pi_1^{\mathrm{geom}}(\mathcal{X})$ is canonically isomorphic to the profinite completion of $\pi_1^{\mathrm{orb}}(\mathcal{X}^{\mathrm{an}})$, the orbifold fundamental group (as defined by Thurston) of the analytification of $\mathcal{X}$ viewed as a complex orbifold. In particular,

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

$\pi_1^{\rm geom}(\mathcal{M}_{g,n}) = \varGamma_{g,n}$, where $\varGamma_{g,n}$ denotes the *profinite* Teichmüller modular group. Note that in this section all groups are profinite by default. Regarding $\pi_1^{\rm geom}$ as a covariant functor from ($\mathbb{Q}$-Stacks) to (Grps) we can apply it to the subcategory $\underline{\mathcal{M}}$, getting the *Teichmüller tower* $\mathcal{T} = \pi_1^{\rm geom}(\underline{\mathcal{M}}) = \pi_1(\underline{\mathcal{M}} \otimes \bar{\mathbb{Q}})$. It is thus no more and no less than the collection of the $\varGamma_{g,n}$ with varying type and morphisms coming functorially from the morphisms in $\underline{\mathcal{M}}$. The name 'Teichmüller tower' for a tower of *groups* may not be so appropriate but comes directly from the *Esquisse*; it would be more consistent to denote it by $\underline{\varGamma}$, but there are already too many Gammas around.

The next step consists of considering the group $\mathrm{Out}(\pi_1^{\rm geom})$, the outer automorphism group of the functor $\pi_1^{\rm geom}$, meaning the group of automorphisms modulo inner automorphisms on the left-hand side, i.e. in (Grps). Let $G_{\mathbb{Q}}$ denote as usual the absolute Galois group of $\mathbb{Q}$. By a fairly easy extension to stacks of Grothendieck's short exact sequence [**22**, § IX.6], we find a map $G_{\mathbb{Q}} \to \mathrm{Out}(\pi_1^{\rm geom})$. This map is an injection, i.e. the outer Galois action is faithful as soon as we consider a big enough version of ($\mathbb{Q}$-Stacks). We were a little fuzzy above as to which geometric objects we wish to include but Belyi's theorem easily implies something much more drastic: as soon as ($\mathbb{Q}$-Stacks) contains the single object $\mathbb{P}^1_{\mathbb{Q}} \setminus \{0,1,\infty\}$, the action is faithful and the above map is an injection. Note that $\mathbb{P}^1_{\mathbb{Q}} \setminus \{0,1,\infty\}$ can be seen as the generic fibre of $\mathbb{P}^1_{\mathbb{Z}} \setminus \{0,1,\infty\}$ (indeed) and the latter can be characterized as the only smooth hyperbolic curve over $\mathbb{Z}$; this may serve to illustrate, although certainly not 'explain', the amazing universality of this object. We will sometimes write $\mathbb{P}^* = \mathbb{P}^1_{\mathbb{Q}} \setminus \{0,1,\infty\}$. Before going back to the modular tower, we mention an important recent and as yet unpublished result of Pop. Not only the above map is an injection, but under rather mild conditions it is an isomorphism! This means the following: consider $\mathcal{C} \subset$ ($\mathbb{Q}$-Stacks) a full subcategory and the restriction of $\pi_1^{\rm geom}(\mathcal{C})$ to $\mathcal{C}$. We still get a natural map $G_{\mathbb{Q}} \to \mathrm{Out}(\pi_1^{\rm geom}(\mathcal{C}))$ which is still an injection under very mild conditions (e.g. if $\mathcal{C}$ contains $\mathbb{P}^*$). Moreover, and somewhat informally, as $\mathcal{C}$ gets smaller, the target gets larger. Pop's result says that actually, even for a rather 'small' sample of geometric objects $\mathcal{C}$, the above map is an isomorphism. This is true for instance if one takes for $\mathcal{C}$ the quasi-projective varieties which are the complements in the projective plane of (not necessarily irreducible) curves; one can actually let $\mathcal{C}$ decrease still further. The upshot is that taking as the source category a (rather small) subcategory $\mathcal{C}$ of the quasi-projective $\mathbb{Q}$-schemes, one gets a natural isomorphism: $G_{\mathbb{Q}} \xrightarrow{\sim} \mathrm{Out}(\pi_1^{\rm geom}(\mathcal{C}))$. This is a striking and beautiful result, the first which connects a purely arithmetic object (on the left-hand side) to a purely geometric object (on the right-hand side). We note for further reference that the right-hand side is really unknown, apart from being proved to be isomorphic to the left-hand side.

Now let us go back to the modular tower $\underline{\mathcal{M}}$ and consider as above the group $\mathrm{Out}(\pi_1^{\rm geom}(\underline{\mathcal{M}})) = \mathrm{Out}(\mathcal{T})$ where the outer automorphisms are equivariant with respect to the morphisms in $\mathcal{T}$. One still has $G_{\mathbb{Q}} \subset \mathrm{Out}(\mathcal{T})$ since $\mathcal{M}_{0,4} \simeq \mathbb{P}^*$. In the colourful words of the *Esquisse*: 'l'action est déjà fidèle au premier étage'. Moreover, the action of $G_{\mathbb{Q}}$ enjoys a well-known property: it preserves inertia groups and the inertia groups in $\varGamma_{g,n}$ associated with the components of the divisors at infinity of the stable completion of $\mathcal{M}_{g,n}$ are nothing but Dehn twists. Concrete conclusion: the action of $G_{\mathbb{Q}}$ on the $\varGamma_{g,n}$

maps Dehn twists to conjugates of powers of themselves; in other words, they preserve the conjugacy classes of the procyclic groups they generate. This is folk wisdom but there is actually a lot more to say on the subject of inertia, in particular stack inertia; Proposition 4.6 and the short discussion above it provide a foretaste. Sticking here to the inertia groups classically (see especially [**23**]) associated with a divisor with (strict) normal crossings, we let $\mathbb{I}\Gamma = \mathrm{Out}^*(\mathcal{T}) \subset \mathrm{Out}(\mathcal{T})$ denote the subgroup of inertia-preserving outer automorphisms in that sense. This is by definition the *Grothendieck–Teichmüller group*, at least in its present, all genera, profinite version, and for the modular tower $\underline{\mathcal{M}}$ we outlined above.

As a first concrete approach, and in order to find 'coordinates' for $\mathbb{I}\Gamma$, one notes that

$$\mathbb{I}\Gamma = \mathrm{Out}^*(\mathcal{T}) \subset \mathrm{Out}^*(\pi_1^{\mathrm{geom}}(\mathcal{M}_{0,4})) = \mathrm{Out}^*(\pi_1^{\mathrm{geom}}(\mathbb{P}^*))$$

and

$$\pi_1^{\mathrm{geom}}(\mathbb{P}^*) = \pi_1(\mathbb{P}^* \otimes \mathbb{C}) = \pi_1(\mathbb{C} \setminus \{0,1\}) = \hat{F}_2,$$

the profinite completion of the free group on two generators, since obviously

$$\pi_1^{\mathrm{top}}(\mathbb{C} \setminus \{0,1\}) \simeq F_2.$$

We thus get $\mathbb{I}\Gamma \subset \mathrm{Out}^*(\hat{F}_2)$ and the latter group can be parametrized by pairs $(\lambda, F) \subset \hat{\mathbb{Z}}^* \times \hat{F}_2'$ where $\hat{F}_2'$ denotes the derived subgroup of $\hat{F}_2$ (see any paper on the subject, starting with [**33**] and [**12**]). We also note that the group $\mathbb{I}\Gamma$ is itself naturally endowed with the profinite topology because the $\Gamma_{g,n}$ are topologically finitely generated (hence characteristic sugbroups form a cofinal sequence). The amazing point, foreshadowed in the *Esquisse*, is that it is 'computable': it is given as a subgroup of $\mathrm{Out}^*(\hat{F}_2)$ by a small number of relations (say four) which translate into equations on the pair (λ, F). In fact $\mathbb{I}\Gamma$ has been computed in [**30**] and [**55**], adding one, perhaps not independent, relation to the genus 0 group introduced in [**12**]. Note that the term 'relation' which is commonly used here should not be misleading; $\mathbb{I}\Gamma$ is given as a subgroup, not a quotient of $\mathrm{Out}^*(\hat{F}_2)$. Essentially by definition there is a natural inclusion $G_{\mathbb{Q}} \subset \mathbb{I}\Gamma$. Whether or not this inclusion is strict is a main driving question of this young field. Note that the situation is in some sense opposite to that of Pop's result quoted above. Throwing in more, or say different, objects he proves the remarkable isomorphism: $G_{\mathbb{Q}} \simeq \mathrm{Out}(\pi_1^{\mathrm{geom}}(\mathcal{C}))$, giving in principle a geometric characterization of the arithmetic Galois group $G_{\mathbb{Q}}$. Yet as mentioned above we know nothing concrete about the right-hand side, so it does not immediately help study the left-hand side. On the other hand, using the Teichmüller tower as above, we get that $G_{\mathbb{Q}} \subset \mathbb{I}\Gamma = \mathrm{Out}^*(\mathcal{T})$. Here in some sense we are able to 'compute' the right-hand side, but we do not know whether or not the inclusion is an isomorphism.

Although it took quite some time to complete the picture sketched above, it should still be considered as a rather primitive stage of the theory. True we have used moduli spaces of curves of *all finite types*, especially all genera, and we have used the full *profinite* completions, two important positive features. But we have used essentially only the structure of the modular tower *at infinity.* So we get what can be called a (Grothendieck–Teichmüller) *lego at infinity* or parabolic lego to take up the terminology of the classification of diffeomorphisms. From this point of view, we are after a different and probably much more subtle sort of lego, connected in particular with the automorphisms of curves (so that it could be termed *elliptic* lego) which is actually the only one mentioned in the *Esquisse* and would encode a lot more arithmetic than the one at infinity. The present paper hopefully points to some relevant objects. In particular, one has to enrich the modular tower in a drastic way, probably throwing in morphisms which are defined over $\bar{\mathbb{Q}}$, not only over $\mathbb{Q}$. Each such morphism is actually defined over a finite extension K of $\mathbb{Q}$ because everything is finitely presented, and leads to an equivariant action of G_K, that is an open subgroup of $G_{\mathbb{Q}}$ which, however, effectively depends on the particular morphism one is looking at. A more general way to put it is to say that one should perhaps think in terms of outer actions modulo coverings, both on the arithmetic and the geometric sides. For example, Facts 4.4 and 4.5 above say that open subgroups of Teichmüller modular groups appear as subquotients of such. It may happen that the open subgroup is in fact the full group (the cover in Fact 4.5 has degree 1) and it may happen that the subquotient is a subgroup (sequence (4.1) after Fact 4.4 splits). Conditions for this to happen are given in [**61**,**62**] where simple cases of such situations are explored, which give rise to interesting compatibility conditions and relations. But in general it seems that 'anything is possible'; for instance the extension class of (4.1) is defined by a geometric cohomology class which may or may not vanish (cf. [**45**]) and special loci can probably be defined over any finite extension (i.e. covering) of $\mathbb{Q}$. At any rate, they give, up to finite covering (and normalization) copies of moduli spaces of curves inside such spaces, which is quite relevant for the theory (see [**61**]). Higher-dimensional algebraic eigenloci may also prove interesting objects in this respect. In the next and last subsection, however, we will concentrate on the potential relevance of one-dimensional objects, namely curves.

6.2. Mapping curves into moduli spaces of curves

Like anabelian geometry, Grothendieck–Teichmüller theory is concerned in the first place with fundamental groups and the arithmetic Galois action on their geometric part. Equivalently it deals with relative (or augmented) fundamental groups obtained by applying the fundamental group functor to structure morphisms, e.g. $\mathcal{M}_{g,n} \to \mathrm{Spec}(\mathbb{Q})$. This already tells us what the main characters should be, namely geometric objects which are effectively determined by their fundamental groups, roughly speaking points, curves, surfaces and classifying spaces. Indeed Lefschetz hyperplane section theorem precisely tells us that we cannot learn much in general from higher-dimensional objects by looking at the fundamental group only, except if these objects enjoy very special properties with respect to this invariant; the simplest and main such property consists of being a $K(\pi, 1)$-space.

Let us detail the above somewhat. We leave aside here the case of points, that is the birational side of the theory, which in anabelian geometry has given rise to a lot of activity (from J. Neukirch to Pop, Mochizuki and others). Curves enter in more than one way as should be plain already. In particular, $\mathbb{P}^1_{\mathbb{Q}} \setminus \{0, 1, \infty\} \simeq \mathcal{M}_{0,4}$ lies at the crossroads of hyperbolic curves and moduli spaces thereof, in a way akin to elliptic curves for curves and abelian varieties. Moreover, Belyi's theorem states that any hyperbolic arithmetic curve has an open set which is a finite étale covering of that object. This connects with the fact that curves enter through their moduli spaces, which is a key point and a defining original feature of Grothendieck's initial vision. The $\mathcal{M}_{g,n}$ are orbifold $K(\pi, 1)$-spaces, that is their orbifold universal covering spaces, alias Teichmüller spaces, are contractible (in cohomological terms they are called 'rationally' $K(\pi, 1)$). We remark that as partly demonstrated above, these spaces also contain essentially equivalent information to that conveyed by some closely related objects, Hurwitz spaces among them.

Now what about the last kind of objects in the list, namely surfaces? We will confine attention to those surfaces S, say connected smooth projective, which admit a stable fibration to a curve. In other words, there is a curve X with smooth projective completion $\bar{X}$ and a proper flat morphism $f : S \to \bar{X}$ such that for any $p \in X$ a (closed geometric) point, $f^{-1}(p)$ is a smooth hyperbolic curve of fixed type (g, n); above the finitely many exceptional points, the fibre is a stable curve, that is it has only normal crossing self-intersections (plus the stability condition). Such an object determines, and indeed is equivalent to, a classifying map $\bar{\phi} : \bar{X} \to \bar{\mathcal{M}}_{g,n}$. The points with non-smooth fibre land on the divisor at infinity and deleting these points from the basis, we recover a map $\phi : X \to \mathcal{M}_{g,n}$. Of course the fibration f, if it exists at all, is not determined by S and in any case it would be interesting to know how much information exactly we leave aside by confining attention to this class of surfaces. We just remark that blowing up a point on a surface does not alter the fundamental group, whereas applying the stable reduction theorem usually involves a base change which affects the fundamental group in a tractable way. The above may, or may not, have convinced the reader that it is at least natural to study maps from curves into moduli spaces of curves.

The fibration is locally trivial if and only if the base curve maps to a point, which is not especially interesting from our viewpoint, so we start from a non-constant (equivalently, generically injective) map $\phi : X \to \mathcal{M}_{g,n}$. Working over $\mathbb{C}$ with the analytic topology, it determines a topological monodromy representation:

$$\phi_*^{\mathrm{top}} = \mu_\phi^{\mathrm{top}} : \pi_1^{\mathrm{top}}(X^{\mathrm{an}}) \to \pi_1^{\mathrm{orb}}(\mathcal{M}_{g,n}^{\mathrm{an}}) = \mathit{\Gamma}_{g,n}^{\mathrm{top}}. \tag{6.1}$$

We apologize for the cumbersome and far from ideal notation here and in what follows; it just points to the fact that it is useful to keep the various structures in mind. To summarize again: this is the complex analytic or topological setting; X is regarded as a Riemann surface, $\mathcal{M}_{g,n}$ as a complex orbifold, π_1^{top} is the discrete or topological fundamental group and *idem* on the right-hand side, where $\mathit{\Gamma}_{g,n}^{\mathrm{top}}$ is just the discrete modular or mapping class group.

Rewriting the above viewing X as a $\mathbb{C}$-scheme $X_{\mathbb{C}}$ and $\mathcal{M}_{g,n}$ as a $\mathbb{C}$-stack $\mathcal{M}_{g,n,\mathbb{C}}$ ($= \mathcal{M}_{g,n,\mathbb{Q}} \otimes \mathbb{C}$) amounts to taking profinite completions of the fundamental groups: $\pi_1(X_{\mathbb{C}}) = \pi_1^{\mathrm{geom}}(X_{\mathbb{C}})$ is the profinite completion of $\pi_1^{\mathrm{top}}(X)$, *idem* for $\mathit{\Gamma}_{g,n}$ and $\mathit{\Gamma}_{g,n}^{\mathrm{top}}$. We get the map

$$\phi_* = \mu_\phi : \pi_1^{\mathrm{geom}}(X) \to \pi_1^{\mathrm{geom}}(\mathcal{M}_{g,n}) = \pi_1(\mathcal{M}_{g,n,\mathbb{C}}) = \mathit{\Gamma}_{g,n}. \tag{$\widehat{6.1}$}$$

https://doi.org/10.1017/S1474748005000101 Published online by Cambridge University Press

This all may look rather boring and formal but the fact is that there is a very serious rub. Assume for ease of notation that the topological monodromy map μ_ϕ^{top} is injective (if not, divide by the kernel) and recall that the modular groups $\Gamma_{g,n}$ are residually finite: in other words the completion map $\Gamma_{g,n}^{\mathrm{top}} \to \Gamma_{g,n}$ is injective. So the composed map $\pi_1^{\mathrm{top}}(X) \to \Gamma_{g,n}$ *is* injective. The geometric monodromy μ_ϕ is obtained by completing this map, using the universal property of the profinite completion. But there is no guarantee that the resulting map μ_ϕ is injective. It could be that μ_ϕ actually factors through a quotient of $\pi^{\mathrm{geom}}(X)$, which, however, has to be large enough so that $\pi_1^{\mathrm{top}}(X)$ injects into it. In other words, the topology induced by $\Gamma_{g,n}$ on $\pi_1^{\mathrm{top}}(X)$ may *a priori* very well be a quotient of the profinite topology. Geometric translation: is it true that any covering of X can be dominated by the pull-back of a covering of $\mathcal{M}_{g,n}$ via ϕ? This is equivalent to proving the injectivity of μ_ϕ and it seems very hard, because of our poor knowledge of the category of coverings of $\mathcal{M}_{g,n}$, in other words of the structure of the modular groups. The conclusion is that when we consider surface groups inside modular groups, it may be that the topologies on the completions do not coincide and that we will lose information when acting equivariantly with the arithmetic Galois group. One is tempted to conjecture that this does not happen but it is certainly good to keep this difficulty in mind. Note that it does happen, and in a most drastic way, if for instance one replaces the moduli of curves by moduli of polarized abelian varieties, whose fundamental groups displays rigidity in the form of the arithmetic congruence subgroup property.

Return to the curve X and the map ϕ and assume they are both defined over $\bar{\mathbb{Q}} \subset \mathbb{C}$, that is in fact over some number field K. We still get the geometric monodromy μ_ϕ, working over $\bar{\mathbb{Q}}$, or equivalently over $\mathbb{C}$. But we also get an action of the Galois group $G_K = \mathrm{Gal}(\bar{\mathbb{Q}}/K)$ on $\pi_1^{\mathrm{geom}}(X)$ and $\Gamma_{g,n}$ which is ϕ-equivariant, that is commute with $\mu_\phi = \phi_*$. Here one can as usual either be careful with basepoints or just think in terms of outer action, moding out by inner automorphisms. We choose the second option for simplicity and because we never mentioned basepoints, but the first one is also available. Finally, one can of course try and study the action of the finite group $\mathrm{Gal}(K/\mathbb{Q})$ (assume $K/\mathbb{Q}$ Galois) on the situation, mapping ϕ and perhaps X to different morphisms and perhaps different curves. Origamis provide a concrete class of examples, and so do arithmetic eigenloci (see [**53**] for a first case study).

By definition we have an action of the Grothendieck–Teichmüller group $\mathbb{\Gamma}$ on the Teichmüller tower, that is a coherent action on the groups $\Gamma_{g,n}$. On the other hand, Belyi's result says that by suitably puncturing an arithmetic curve X, we can realize the punctured curve $\check{X}$, that is a dense open subset of X, as an unramified covering of $\mathbb{P}^*$. Now $\mathbb{\Gamma}$ is a subgroup of $\mathrm{Out}^*(\pi_1^{\mathrm{geom}}(\mathbb{P}^*))$, so an open subgroup of it acts on $\pi_1^{\mathrm{geom}}(\check{X})$; moreover, $\pi_1^{\mathrm{geom}}(X)$ is the quotient of $\pi_1^{\mathrm{geom}}(\check{X})$ obtained by annihilating the inertia associated with the divisor $X \setminus \check{X}$, which sits over the points $(0, 1, \infty)$ in the covering $\check{X}/\mathbb{P}^*$. Since $\mathbb{\Gamma}$ is inertia preserving the action is compatible with that quotient. This is all very well, but essentially useless as is. The point is that there is nothing natural (functorial) in the above construction; one has to 'pick a Belyi function', and there is just no guarantee that these various actions are coherent (they are indeed, when restricting to the Galois action!). By mapping curves into moduli spaces of curves we are at least beginning to see an interplay between curves and their moduli spaces. Notwithstanding the difficulty mentioned above about induced topologies, we find copies of geometric fundamental groups of curves inside the modular groups and $\mathbb{I}\Gamma$ by definition acts coherently on the latter groups. Moreover, if the image $\phi(X) \subset \mathcal{M}_{g,n}$ happens to be geodesic for the Teichmüller metric (switching to a complex analytic setting), the plot thickens because of Proposition 2.4 above, which enables one to view the Fuchsian group uniformizing X inside the Teichmüller modular group, leading to the picture described in Proposition 2.10. Again origamis provide a particularly simple and attractive class of examples.

Let us recall a few relevant and perhaps helpful geometric properties before closing. First and only for the record, we recall that a lot of geometric work has been devoted to finding *complete* curves and higher-dimensional complete varieties inside the moduli spaces of curves. Here the curves and higher-dimensional objects we contemplate are not complete and it is not a problem to find a plentiful of them; the problems and interest lie elsewhere. Now at least two fundamental results may be relevant in our context: first, under the assumptions above, given the curve X and the hyperbolic type (g, n), the nonconstant map ϕ is entirely determined by its topological monodromy $\phi_*^{\text{top}} = \mu_\phi^{\text{top}}$; second, still fixing X up to isomorphism and the type (g, n), there are only finitely many nonconstant maps $\phi : X \to \mathcal{M}_{g,n}$. The second assertion is the celebrated Parshin–Arakelov theorem (ex-Shafarevich conjecture) and can be deduced from the first rigidity assertion. In turn it implies the Manin–Grauert theorem (ex-Mordell's conjecture for function fields). There is of course a large body of literature devoted to this circle of ideas; see [**36**] for a short analytic proof using Teichmüller spaces and [**48**] for an unusual survey, in the spirit of 'topological arithmetic'. Certainly both the rigidity and the finiteness assertions are striking from the point of view we are trying to build up and we used the first property in a crucial way in Proposition 5.8. Moreover, from the standpoint of Grothendieck–Teichmüller theory, it would surely be interesting to have an algebraic treatment of the rigidity of the monodromy available, perhaps in the style of [**66**]. Here 'algebraic' means as usual that it should not involve Teichmüller space and distance in an essential way. Rigidity of the monodromy is indeed in the spirit of Grothendieck's question or program in both Galois–Teichmüller theory (renamed 'Grothendieck–Teichmüller' in [**12**]) and anabelian geometry: how much geometry is encoded in the fundamental group functor? For results, references and a broad discussion of Galois and π_1-rigidity, we refer to [**54**], supplemented by [**56**] for more recent anabelian results.

So this is where we stand: again a wholly subjective and partial view. We have at our disposal a modular tower based essentially on considerations 'at infinity' that is on completely degenerate (Mumford) curves. It gives rise to a Teichmüller tower, a lego at infinity and a version of the Grothendieck–Teichmüller group; all this already gives flesh to Grothendieck's 'two levels' principle (see in particular [**30**] and the use of 'locality' there). The above situation is over $\mathbb{Q}$; by definition $G_{\mathbb{Q}} \subset \mathbb{I}\Gamma$ and equality remains obscure, not even necessarily conjectural. We now finds that moduli spaces of curves are populated by interesting subobjects, like special loci which themselves very much resemble the moduli spaces, or more generally arithmetic eigenloci which have hardly been studied at all. All this is in general defined over $\bar{\mathbb{Q}}$ but not over $\mathbb{Q}$. Some particular situations, like special loci in genus 0, are typically cyclotomic, that is defined over $\mathbb{Q}^{\mathrm{ab}}$. Finally, we find that it is also natural and important to map curves into their moduli spaces, and that geodesicity and algebraicity or arithmeticity mix up in an odd and hitherto mysterious way. Again this sort of thing is usually not defined over $\mathbb{Q}$.

We would now like to go after some sort of lego, assuming it exists at all, that is try to see if the information depends somehow on local data. Needless to say (using the familiar preterition) all this should be interesting not only in terms of putative Grothendieck–Teichmüller theory but also for the study of the Galois action in this highly non-abelian framework. Because the situation is not defined over $\mathbb{Q}$ anymore, one is tempted to take a small step and formalize it via the convenient notion of *virtual* group action. Given a profinite group G, we say that it acts virtually on an object (a variety, a group, etc.) if *some* open subgroup of G does. Of course one requires obvious coherence conditions, like the fact that if two subgroups act, the actions coincide on the intersection. Note that we use 'virtual' as group theorists do, that is for a property pertaining to an open subgroup (e.g. 'virtually torsion free'); arithmeticians use 'potential' for just the same notion. We also conform to the motto that things should sometimes be considered up to coverings. Clearly, this notion is of little or no interest when applied to just one object. But as a perhaps more interesting example, let k be a field, $\bar{k}$ an algebraic closure of k, and consider ($\bar{k}$-Var) the category of varieties over $\bar{k}$. We can apply again the geometric fundamental group functor $\pi_1^{\mathrm{geom}} : (\bar{k}\text{-Var}) \to (\mathrm{Grps})$ and we find that the absolute Galois group $G_k = \mathrm{Gal}(\bar{k}/k)$ defines a virtual outer action on the image. The long and the short is that given a situation, say, over $\bar{\mathbb{Q}}$, we may abstractly, that is without using the Galois correspondence, apply π_1^{geom} and define a virtual outer automorphism group (it is indeed a group). This applies to a host of situations we have encountered above and we leave it to the reader to think of others. In some of these situations at least, this virtual outer, say inertia-preserving, automorphism group again deserves the namesake 'Grothendieck–Teichmüller' and contains the absolute Galois group $G_{\mathbb{Q}}$. Moreover, it will usually also be contained in the enormous $\mathrm{Out}^*(\hat{F}_2)$. So there are two questions in essentially opposite directions, hopefully meeting at the midpoint of the tunnel. Is this group computable, that is defined by finitely many relations, or better said equations? A lego, that is a way of breaking information into local pieces, usually provides a positive answer. And is the group equal to the Galois group? The theorem of Pop quoted above gives a positive answer for a bona fide action in a setting where everything is defined over $\mathbb{Q}$. We have seen many ways of adding in maps between or into moduli spaces of curves. How much do we need to add in order to ensure that the corresponding 'Grothendieck–Teichmüller group' coincides with the Galois group?

Acknowledgements. It is a pleasure to thank the Mathematical Science Research Institute in Berkeley and the Max Planck Institut für Mathematik in Bonn for offering me ideal research conditions during the elaboration of the present paper. I cannot list all the persons who have enlightened me at some point concerning the matters discussed here; let me mention and warmly thank among many others E. Gutkin, D. Harbater, P. Hubert, C. McMullen, M. Möller, L. Schneps, T. A. Schmidt and L. Zapponi.

References

1. B. Bakalov and A. Kirillov Jr, *Lectures on tensor categories and modular functors*, University Lecture Series, vol. 21 (American Mathematical Society, Providence, RI, 2001).
2. P. Beazley Cohen, Humbert surfaces and transcendence properties of automorphic functions, *Rocky Mt. J. Math.* **26** (1996), 987–1001.
3. P. Beazley Cohen, C. Itzykson and J. Wolfart, Fuchsian triangle groups and Grothendieck dessins (variations on a theme of Belyi), *Commun. Math. Phys.* **163** (1994), 605–627.
4. J. Bertin, M. Imbert and M. Romagny, Compactification des schémas de Hurwitz et applications, preprint (2002).
5. J. Birman, *Braids, links, and mapping class groups*, Annals of Mathematics Studies, vol. 82 (Princeton University Press, 1974).
6. J.-B. Bost, Algebraic leaves of algebraic foliations over number fields, *Publ. Math. IHES* **93** (2001), 161–221.
7. X. Buff, J. Fehrenbach, P. Lochak, L. Schneps and P. Vogel, *Espaces de modules des courbes, groupes modulaires et théorie des champs*, Panoramas et Synthèses, vol. 7 (Société Mathématique de France, 1999) (English translation published by American Mathematical Society, Providence, RI, 2003).
8. K. Calta, Veech surfaces and complete periodicity in genus 2, *J. Am. Math. Soc.* **17** (2004), 871–908.
9. A. Chambert-Loir, Théorèmes d'algébrisation en géométrie diophantienne, Séminaire Bourbaki, Exposé 886, *Astérisque* **282** (2002), 175–209.
10. M. Cornalba, On the locus of curves with automorphisms, *Annli Mat. Pura Appl.* **149** (1987), 135–151.
11. P. Deligne and D. Mumford, The irreducibility of the space of curves of given genus, *Publ. Math. IHES* **36** (1969), 75–109.
12. V. G. Drinfel'd, On quasitriangular quasi-Hopf algebras and a group closely connected with Gal($\bar{\mathbb{Q}}/\mathbb{Q}$), *Leningrad Math. J.* **2** (1991), 829–860.
13. C. J. Earle and F. P. Gardiner, Teichmüller discs and Veech's $\mathcal{F}$-structures, *Contemp. Math.* **201** (1997), 165–189.
14. C. J. Earle and I. Kra, On holomorphic mappings between Teichmüller spaces, in *Contributions to analysis* (ed. L. V. Ahlfors *et al.*), pp. 107–124 (Academic, 1974).
15. G. Ellingsrud, K. Lønsted, An equivariant Lefschetz trace formula for finite reductive groups, *Math. Ann.* **251** (1980), 253–261.
16. A. Eskin, H. Masur and A. Zorich, Moduli spaces of abelian differentials: the principal boundary, counting problems and the Siegel–Veech constants, preprint (2002).
17. A. Fathi, F. Laudenbach and V. Poenaru, *Travaux de Thurston sur les surfaces*, Astérisque, vol. 66–67 (Société Mathématique de France, 1979).
18. L. Funar and R. Gelca, On the groupoid of transformations of rigid structure on surfaces, *J. Math. Sci. Univ. Tokyo* **6** (1999), 599–646.
19. F. P. Gardiner, *Teichmüller theory and quadratic differentials* (Wiley, 1987).
20. G. Gonzalez-Diez and W. J. Harvey, *Moduli of Riemann surfaces with symmetry*, London Mathematical Society Lecture Note Series, vol. 173, pp. 75–93 (Cambridge University Press, 1992).
21. G. Gonzalez-Diez and P. Lochak, On the fundamental groups at infinity of the moduli spaces of compact Riemann surfaces, *Michigan Math. J.* **49** (2001), 493–500.
22. A. Grothendieck, *Revêtements etales et groupe fondamental*, Lecture Notes in Mathematics, vol. 224 (Springer, 1971).
23. A. Grothendieck and J. P. Murre, *The tame fundamental group of a formal neighbourhood of a divisor with normal crossings on a scheme*, Lecture Notes in Mathematics, vol. 208 (Springer, 1971).

24. E. Gutkin and C. Judge, Affine mappings of translation surfaces: geometry and arithmetic, *Duke Math. J.* **103** (2000), 191–213.
25. E. Gutkin, P. Hubert and T. A. Schmidt, Affine diffeomorphisms of translation surfaces: periodic points, Fuchsian groups and arithmeticity, *Annls Scient. Éc. Norm. Sup.* **36** (2003), 847–866.
26. R. Hain and M. Matsumoto, Tannakian fundamental groups associated to Galois groups, in *Galois groups and fundamental groups*, Mathematical Sciences Research Institute Publications, vol. 41 (Cambridge University Press, 2003).
27. J. L. Harer, The cohomology of the moduli spaces of curves, in *Theory of moduli* (ed. E. Sernesi) Lecture Notes in Mathematics, vol. 1337, pp. 138–221 (Springer, 1988).
28. W. J. Harvey, On certain families of compact Riemann surfaces, *Contemp. Math.* **150** (1993), 137–148.
29. A. Hatcher and W. Thurston, A presentation for the mapping class group of a closed orientable surface, *Topology* **19** (1980), 221–237.
30. A. Hatcher, P. Lochak and L. Schneps, On the Teichmüller tower of mapping class groups, *J. Reine Angew. Math.* **521** (2000), 1–24.
31. P. Hubert and T. Schmidt, Veech groups and polygonal coverings, *J. Geom. Phys.* **35** (2000), 75–91.
32. G. Humbert, Sur les fonctions abéliennes singulières, *J. Math. Pures Appl.* **5** (1899), 233–350.
33. Y. Ihara, Profinite braid groups, Galois representations and complex multiplications, *Ann. Math.* **123** (1986), 43–106.
34. Y. Ihara, Some aspects of Galois actions on the pro-p fundamental group of $\mathbb{P}^1 \setminus \{0, 1, \infty\}$, in *Arithmetic fundamental groups and noncommutative geometry, Proc. Symp. Pure Math.*, vol. 70, pp. 247–273 (American Mathematical Society, 2002).
35. Y. Ihara and H. Nakamura, On deformation of maximally degenerate stable curves and Oda's problem, *J. Reine Angew. Math.* **487** (1997), 125–151.
36. Y. Imayoshi and H. Shiga, A finiteness theorem for holomorphic families of Riemann surfaces, in *Holomorphic functions and moduli*, vol. II, Mathematical Sciences Research Institute Series, vol. 11, pp. 207–219 (Springer, 1988).
37. Y. Imayoshi and M. Taniguchi, *An introduction to Teichmüller spaces* (Springer, 1992).
38. M. Imbert, Cellular decomposition of compactified Hurwitz spaces, in *Galois groups and fundamental groups*, Mathematical Sciences Research Institute Publications, vol. 41 (Cambridge University Press, 2003).
39. R. Kenyon and J. Smillie, Billiards in rational-angled triangles, *Comment. Math. Helv.* **75** (2000), 65–108.
40. G. Laumon and L. Moret-Bailly, *Champs algébriques*, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 39 (Springer, 2000).
41. P. Lochak, The fundamental groups at infinity of the moduli spaces of curves, in *Geometric Galois actions* (ed. L. Schneps and P. Lochak), pp. 139–158, London Mathematical Society Lecture Note Series, vol. 242 (Cambridge University Press, 1997).
42. P. Lochak, Fragments of nonlinear Grothendieck–Teichmüller theory, in *Woods Hole Mathematics*, pp. 225–262, Series on Knots and Everything, vol. 34 (World Scientific, 2004).
43. P. Lochak and L. Schneps, The Grothendieck–Teichmüller group and automorphisms of braid groups, in *The Grothendieck theory of dessins d'enfants* (ed. L. Schneps), pp. 323–358, London Mathematical Society Lecture Note Series, vol. 200 (Cambridge University Press, 1994).
44. P. Lochak and L. Schneps, A cohomological interpretation of the Grothendieck–Teichmüller group, *Invent. Math.* **127** (1997), 571–600.
45. P. Lochak and M. Vaquié, Inertie champêtre, inertie singulière et action galoisienne, in preparation.

46. P. Lochak, H. Nakamura and L. Schneps, Eigenloci of five point configurations on the Riemann sphere and the Grothendieck–Teichmüller group, preprint (2003).
47. C. MacLachlan and W. J. Harvey, On mapping class groups and Teichmüller spaces, *Proc. Lond. Math. Soc.* **30** (1975), 496–512.
48. C. T. McMullen, From dynamics on surfaces to rational points on curves, *Bull. Am. Math. Soc.* **37** (2000), 119–140.
49. C. T. McMullen, Billiards and Teichmüller curves on Hilbert modular surfaces, *J. Am. Math. Soc.* **16** (2003), 857–885.
50. H. A. Masur, On a class of geodesics in Teichmüller space, *Ann. Math.* **102** (1975), 205–221.
51. H. A. Masur and Y. N. Minsky, Geometry of the complex of curves, I, Hyperbolicity, *Invent. Math.* **138** (1999), 103–149.
52. H. A. Masur and S. Tabachnikov, Rational billiards and flat structures, in *Handbook on dynamical systems*, vol. 1A, pp. 1015–1089 (North-Holland, 2002).
53. M. Möller, Teichmüller curves, Galois actions and $\widehat{GT}$-relations, *Math. Nachr.*, in press.
54. H. Nakamura, Galois rigidity of algebraic mappings into some hyperbolic varieties, *Int. J. Math.* **4** (1993), 421–438.
55. H. Nakamura and L. Schneps, On a subgroup of the Grothendieck–Teichmüller group acting on the tower of profinite Teichmüller modular groups, *Invent. Math.* **141** (2000), 503–560.
56. H. Nakamura, A. Tamagawa and S. Mochizuki, Grothendieck's conjectures concerning fundamental groups of algebraic curves, *Sugaku Exposit.* **14** (2001), 31–53 (Japanese original: *Sugaku* **50** (1998), 113–129).
57. B. Noohi, Fundamental groups of algebraic stacks, *J. Inst. Math. Jussieu* **3** (2004), 69–103.
58. H. Popp, Stratifikation von quotientenmannigfaltigkeiten und insbesondere der modulmannigfaltigkeiten für kurven, *J. Reine Angew. Math.* **250** (1971), 12–41.
59. L. Schneps (ed.), *The Grothendieck theory of dessins d'enfants*, London Mathematical Society Lecture Note Series, vol. 200 (Cambridge University Press, 1994).
60. L. Schneps, The Grothendieck–Teichmüller group $\widehat{GT}$: a survey, in *Geometric Galois actions* (ed. L. Schneps and P. Lochak), pp. 183–203, London Mathematical Society Lecture Note Series, vol. 242 (Cambridge University Press, 1997).
61. L. Schneps, Automorphisms of curves and the Grothendieck–Teichmüller group, *Math. Nachr.*, in press.
62. L. Schneps, Special loci in moduli spaces of curves, in *Galois groups and fundamental groups*, Mathematical Sciences Research Institute Publications, vol. 41 (Cambridge University Press, 2003).
63. L. Schneps and P. Lochak (eds), *Geometric Galois actions*, London Mathematical Society Lecture Note Series, vol. 242 (Cambridge University Press, 1997).
64. G. Shimura, *Introduction to the arithmetic theory of automorphic functions* (Princeton University Press, 1971).
65. K. Strebel, *Quadratic differentials* (Springer, 1984).
66. L. Szpiro, *Sur le théorème de rigidité de Parshin et Arakelov*, Astérisque, vol. 64, pp. 169–202 (Société Mathématique de France, 1979).
67. S. Tabachnikov, *Billiards*, Panoramas et Synthèses, vol. 1 (Société Mathématique de France, 1995).
68. W. P. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces, *Bull. Am. Math. Soc.* **19** (1988), 417–431.
69. W. A. Veech, Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards, *Invent. Math.* **97** (1989), 553–583.
70. W. A. Veech, Moduli spaces of quadratic differentials, *J. Analyse Math.* **55** (1990), 117–171.

71. W. A. Veech, The billiard in a regular polygon, *Geom. Funct. Analysis* **2** (1992), 341–379.
72. C. C. Ward, Calculations of Fuchsian groups associated to billiards in a rational triangle, *Ergod. Theory Dynam. Syst.* **18** (1998), 1019–1042.
73. J. Wilson, Explicit moduli for curves of genus 2 with real multiplication by $\mathbb{Q}(\sqrt{5})$, *Acta Arith.* **93** (2000), 121–138.