GEOMETRY OF TEICHMÜLLER CURVES

MARTIN MÖLLER

Abstract. The study of polygonal billiard tables with simple dynamics led to a remarkable class of special subvarieties in the moduli of space of curves called Teichmüller curves, since they are totally geodesic submanifolds for the Teichmüller metric.

We survey the known methods to construct of Teichmüller curves and exhibit structure theorems that might eventually lead towards the complete classification of Teichmüller curves.

Contents

Introduction 1
1. Dynamically optimal billiard tables and flat surfaces 2
2. The list of known examples 4
3. Teichmüller curves and variations of Hodge structures 6
4. Constructing Veech surfaces and computing the Veech group 8
5. Affine invariant manifolds 9
6. Finiteness and classification results 10
7. Modular forms and Euler characteristics 12
8. Orbifold points and other connections to arithmetic geometry 13
References 14

Introduction

The origin of the notion Teichmüller curve goes back to a remarkable discovery of Veech [Vee89] who constructed billiard tables where the trajectories of a bouncing billiard ball have a remarkably simple dynamics, as simple as on a rectangular table. An unfolding construction of the billiard table yields a flat surface, that is a compact Riemann surface together with a flat metric and a finite number of cone-type singularities. Shearing such a flat surface by elements in $\text{GL}_2^+(\mathbb{R})$ provides a whole family of flat surfaces. Only rarely is such an orbit closed in the moduli space $\Omega\mathcal{M}_g$ parametrizing flat surfaces. In this case the image of the $\text{GL}_2^+(\mathbb{R})$-orbit in the moduli space of curves $\mathcal{M}_g$ is an algebraic curve, called a Teichmüller curve.

The moduli space of curves is not a locally homogeneous space and thus does not come naturally with a distinguished class of special algebraic subvarieties. Thanks

1991 Mathematics Subject Classification. Primary 32G15, 37D50, Secondary 32G20, 57M50.

Key words and phrases. Teichmüller curve, flat surfaces, real multiplication, torsion points, $\text{GL}_2^+(\mathbb{R})$-action, affine invariant manifolds.
to the $GL^+_2(\mathbb{R})$-action, the moduli space of flat surfaces $\Omega \mathcal{M}_g$ inherits quite a bit of the properties of a homogeneous space. The special subvarieties there are affine invariant submanifolds and the smallest of them are Teichmüller curves. Part of the beauty of studying their geometry is reflected in the fact that Teichmüller curves admit a variety of different characterizations that may roughly be phrased as follows.

i) Teichmüller curves are immersed curves in the moduli space of curves $\mathcal{M}_g$ that are totally geodesic for the Teichmüller metric.

ii) Teichmüller curves are the images in $\mathcal{M}_g$ of closed $GL^+_2(\mathbb{R})$-orbits in the moduli space $\Omega \mathcal{M}_g$ of flat surfaces.

iii) Teichmüller curves are curves in $\mathcal{M}_g$ whose variation of Hodge structures contains a rank two summand whose Kodaira-Spencer map is an isomorphism.

iv) Teichmüller curves are the images in $\mathcal{M}_g$ of two-dimensional subvarieties in strata of $\Omega \mathcal{M}_g$ cut out by torsion and real multiplication.

In all likelihood, Teichmüller curves should not exist, maybe except for low genus examples, and examples derived from them. The torsion and real multiplication conditions are just too restrictive. And yet they do exist!

The goal of this survey is to explain the above characterizations of Teichmüller curves and to summarize the current state of knowledge on the classification and geometry of Teichmüller curves.

1. Dynamically optimal billiard tables and flat surfaces

We start with a rational polygonal billiard table, that is, a planar polygon $P$ all whose angles are rational multiples of $\pi$. The trajectories of a single ball bouncing in such a $P$ might exhibit various types of long-term behavior. (If the trajectory hits a corner it just ends there and subsequently we disregard this measure zero set of cases.) The trajectory could be periodic. Second, trajectory might be dense, more precisely uniformly distributed all over the polygon, that is, visit every region with frequency proportional to the volume of the region. Last, the trajectory might be dense in some region strictly smaller than the whole polygon. For a rectangular table, the last possibility does not occur. Moreover which of the two first cases occurs depends on the initial direction only, not on the starting point. This simple trajectory behavior that rectangular tables exhibit is called Veech dichotomy or optimal dynamics. Polygons tiled by a rectangular table also exhibit this optimal dynamics.

Understanding whether a billiard table has optimal dynamics is simplified by performing the Katok-Zemlyakov unfolding construction [ZK75], as illustrated in Figure 1. Instead of reflecting the trajectory at the boundary of the polygon we reflect the polygon and continue the trajectory by a straight line. Since the polygon is rational this process ends with a finite number of reflection copies. Gluing them together gives a flat surface $(X, \omega)$, that is, a compact Riemann surface $X$ together with a holomorphic one-form $\omega$ that provides us with a flat metric $|\omega|$ outside a finite number of cone-type singularities where the angle is a multiple of $2\pi$. The set of all flat surfaces fit into a moduli space $\Omega \mathcal{M}_g$ with a natural forgetful map to the moduli space of curves $\mathcal{M}_g$. This moduli space is decomposed into strata

$$\Omega \mathcal{M}_g = \bigcup_{\mu} \Omega \mathcal{M}_g(\mu) \quad (\mu = (m_1, \ldots, m_n) \sum_{i=1}^n m_i = 2g-2), \quad (1)$$
The moduli space of flat surfaces $\Omega M_g$ carries a natural action of $GL^+_2(\mathbb{R})$ induced by the linear action on planar polygons, see Figure 3. This action preserves the stratification (1). Moreover, the straight line flow on $(X, \omega)$ is dynamically optimal if and only if it is dynamically optimal on $A \cdot (X, \omega)$ for any $A \in GL^+_2(\mathbb{R})$.

The initial observation of Veech [Vee89, Vee91] was that if the $GL^+_2(\mathbb{R})$-orbit of $(X, \omega)$ is closed in its stratum $\Omega M_g(\mu)$, then $(X, \omega)$ has optimal dynamics. This is to say that for each direction $\theta$ one of two cases happen: Either all trajectories in the direction $\theta$ are uniformly distributed (hence dense) or the Veech surface is foliated in direction $\theta$ by closed geodesics and saddle connections between the saddle points, the zeros of $\omega$. (The converse holds in low genus [McM05b], but it is false in general [SW08].) Such a $GL^+_2(\mathbb{R})$-orbit is closed if there is a lattice $\Gamma \subset SL_2(\mathbb{R})$ stabilizing $(X, \omega)$ and the converse also holds [SW04]. Since the rotation images and the homothety images of $(X, \omega)$ are in the same fiber of the projection to $M_g$, the images in $M_g$ of closed $GL^+_2(\mathbb{R})$-orbits are of the form $C = \Gamma \backslash \mathbb{H} \to M_g$. They are immersed algebraic (but non-complete) curves in the moduli space of curves.

The action of $GL^+_2(\mathbb{R})$ extends to an action on the moduli space $Q_g$ that parameterizes half-translation surfaces $(Y, q)$ consisting of a Riemann surface with a quadratic differential $q$ and the above statements about images of orbits in $M_g$ carry over verbatim. To summarize:

**Definition 1.1.** A flat surface $(X, \omega)$ or a half-translation surface $(Y, q)$ with closed $GL^+_2(\mathbb{R})$-orbit is called Veech surface and the image curve $C \to M_g$ of this orbit is called a Teichmüller curve.

The name Teichmüller curve reflects that images of $GL^+_2(\mathbb{R})$-orbits in the Teichmüller space $T_g$ of any (Teichmüller-) marked half-translation surface $(Y, q)$ are Teichmüller discs, i.e. discs $\Delta \to T_g$ that are totally geodesic for the Teichmüller metric. By Teichmüller’s fundamental theorems all such Teichmüller disc can be obtained as the orbits of half-translation surfaces $(Y, q)$.

In some sense it is not strictly necessary to discuss the case of half-translation surfaces. Associated with any half-translation surface $(Y, q)$ there is a canonical $GL^+_2(\mathbb{R})$-equivariant double cover construction $\pi : X \to Y$ on which the quadratic differential $\pi^* q = \omega^2$ admits a square root. Consequently to each Teichmüller curve $C \to M_g(Y)$ generated by $(Y, q)$ there is a Teichmüller curve $C' \to M_g(X)$ in a
moduli space of somewhat larger genus, generated by a flat surface. Since most of the classification of Teichmüller curves works using the cohomology of the Veech surfaces and hence abelian differentials, we restrict ourselves from now on to Teichmüller curves generated by flat Veech surfaces \((X, \omega)\). (The reader might then check in each case at hand if the surface admits an involution that makes those surfaces arise as double coverings.) Also the itemized characterizations in the introduction are equivalent only on this subclass of Teichmüller curves.

The uniformizing group \(\Gamma\) is also called the Veech group of the Veech surface \((X, \omega)\). It can be characterized as the group of orientation-preserving homeomorphisms of \(X\) that are affine when expressed in the flat charts of \(X \times \mathbb{Z}(\omega)\) provided by \(\omega\). An important invariant of \(\Gamma\) and thus of any Teichmüller curve is the trace field \(K = \mathbb{Q}[\text{tr}(\gamma), \gamma \in \Gamma]\).

2. The list of known examples

The known examples of Veech surfaces and Teichmüller curves consist of a short list of series, up to a natural notion of primitivity. We present all these series and come back in the subsequent sections to the ideas behind their discovery.

A Veech surface \((X, \omega)\) is called (geometrically) imprimitive if there is a (branched) covering \(\pi : X \to Y\) such that \(\omega = \pi^* \eta\) for some one-form \(\eta\) on \(Y\). Otherwise \((X, \omega)\) is called geometrically primitive. A Veech surface \((X, \omega)\) is called algebraically primitive if the trace field has degree \(g = [K : \mathbb{Q}]\). Theorem 3.1 implies that algebraically primitive implies geometrically primitive. All these properties are \(\text{GL}_2^+(\mathbb{R})\)-equivariant and we abuse the corresponding notions also for the Teichmüller curves the Veech surfaces generate.

The triangle group series. We currently know of a single series of primitive Teichmüller curves generated by Veech surfaces of unbounded genera, containing infinitely many algebraically primitive Teichmüller curves. This series is indexed by two parameters \(m, n \in \mathbb{N} \cup \infty\) and constructed so that the Veech groups are the triangle groups \(\Delta(m, n, \infty)\). The family was discovered in [BM10b] and contains the original examples of Veech \((n = 2\) and \(n = \infty\)) and those of his student Ward.
The Weierstraß family and the Prym family. The Weierstraß family is generated by (nearly) L-shaped flat surfaces in the stratum $\Omega \mathcal{M}_2(2)$ of genus two as in Figure 3. The family was discovered independently by Calta [Cal04] and McMullen [McM03]. Veech surfaces generating all the Teichmüller curves in this series are given by side length parameters $a = (0, \lambda), c = (\lambda, 0), b = (t, h), c + d = (w, 0)$ where $t \in \mathbb{N}, w, h \in \mathbb{N} > 0$ and where $\lambda = (e + \sqrt{D})/2$ for $D = e^2 + 4wh$ is a quadratic irrational number.

The Prym family is generated by the S-shaped genus three surfaces and the X-shaped genus four surfaces (in the strata $\Omega \mathcal{M}_3(4)$ and $\Omega \mathcal{M}_4(6)$) in Figure 4, discovered by McMullen in [McM06a]. The trace field has degree $r = 2$ in all cases.

The Gothic family. This is an infinite family of primitive Teichmüller curves discovered by McMullen, Mukamel and Wright [MMW17] in the stratum $\Omega \mathcal{M}_4(2, 2, 2)$. It is generated by Veech surfaces that resemble Gothic cathedrals (see Figure 5), again with trace field of degree $r = 2$. Another infinite series, generated by Veech surfaces in the stratum $\Omega \mathcal{M}_4(3, 3)$, has been announced by Eskin, McMullen, Mukamel and Wright.

The sporadic examples. There are two sporadic examples of Teichmüller curves. The Veech surfaces are constructed as the unfolding of the $(2\pi/9, \pi/3, 4\pi/9)$-triangle (in the stratum $\Omega \mathcal{M}_3(3, 1)$, see [KS00]) and as the unfolding of the $(\pi/3, \pi/5, 7\pi/15)$-triangle (in the stratum $\Omega \mathcal{M}_4(6)$, see [Vor96]).
3. Teichmüller curves and variations of Hodge structures

This section reveals the algebro-geometric nature of Teichmüller curves by two basic structure theorems. We also highlight similarities and differences to Shimura curves.

Theorem 3.1 ([Möl06b]). If \((X, \omega)\) is a Veech surface, then the Jacobian \(\text{Jac}(X)\) contains an abelian subvariety \(\text{Jac}(X, \omega)\) of dimension \(r = [K : \mathbb{Q}]\) with real multiplication by the trace field \(K\), i.e. the endomorphism ring of \(\text{Jac}(X, \omega)\) is an order in \(K\).

Here \(\text{Jac}(X, \omega)\) is the smallest abelian subvariety of \(\text{Jac}(X)\) whose tangent space contains \(\omega\) via the canonical identification \(T_{\text{Jac}(X)} \cong \Gamma(X, \Omega^1_X)\).

For a Veech surface \((X, \omega)\) we let \(z_1, \ldots, z_n\) be the zeros of \(\omega\), i.e. \(\text{div}(\omega) = \sum m_iz_i\).

Theorem 3.2 ([Möl06a]). For any \(i, j\) the divisor \([z_i - z_j]\) has finite order in \(\text{Jac}(X, \omega)\).

To sketch the proof of the two theorems we consider the Teichmüller curve \(C \to \mathcal{M}_g\) generated by the Veech surface. After passing to a finite unramified (in the orbifold sense) cover of \(C\) the universal curve over (a level cover of) \(\mathcal{M}_g\) pulls back to a family of curves \(f : X \to C\) that we may extend to a family of stable curves \(f : \tilde{X} \to \tilde{C}\) over a complete curve \(\tilde{C}\). The vector spaces \(H^1(X, \mathbb{Q})\) glue to a locally constant bundle \(V_\mathbb{Q}\) over \(C\). The Hodge bundle, the vector bundle with fiber \(H^0(X, \Omega^1_X)\), is a subbundle of (the extension to \(\tilde{C}\) of) the vector bundle \(V_{\tilde{C}}\). This vector bundle inclusion together with a polarization stemming from the symplectic pairings on the fibers of \(f\) defines a weight one variation of Hodge structures (VHS).

The starting point for all theorems in this section is the decomposition (as variation of Hodge structures)

\[
V_{\mathbb{Q}} = W_{\mathbb{Q}} \oplus M_{\mathbb{Q}}, \quad \text{where} \quad W_K = L_1 \oplus \cdots \oplus L_r
\]

over any Teichmüller curve. Here \(L_1\) is the rank-two locally constant subbundle generated by \(\text{Re}(\omega)\) and \(\text{Im}(\omega)\), the tautological plane, and the \(L_i\) are the Galois conjugates of \(L_1\) over \(K\). This decomposition follows from Deligne’s semisimplicity of VHS and since the tautological plane is a sub-VHS essentially by definition of a Teichmüller curve as \(\text{GL}^+_2(\mathbb{R})\)-orbit.

A \(\mathbb{Q}\)-decomposition of \(V_{\mathbb{Q}}\) defines a splitting of the family of Jacobian varieties. Moreover, for any \(\lambda \in K\) the endomorphism \(\oplus \sigma_i(\lambda)\), with \(\sigma_i\) running over the real embeddings of \(K\), defines a rational endomorphism of the family of abelian
The locally constant subbundle $L_1$ is special, since by construction its monodromy representation is the standard representation of the uniformizing Fuchsian group $\Gamma$ of $C$. As a consequence, the period map from the universal cover of $C$ to the upper half plane, the period domain for such a rank two subsheaf, is an isomorphism. In particular, if we let $\mathcal{L}$ be the $(1,0)$-part of (the Deligne extension to $\overline{C}$ of) $L_1$ then the Higgs field (also known as the Kodaira-Spencer map), that is the derivative
\[
\tau : \mathcal{L} \to \mathcal{L}^{-1} \otimes \Omega^1_{\overline{C}}(\Delta), \quad (\Delta = \overline{C} \setminus C)
\]
of the period map, is an isomorphism. Those subbundles are called maximal Higgs, since for those subbundles the degree of $\mathcal{L}$ attains its maximum value $\frac{1}{2} \text{deg} \Omega^1_{\overline{C}}(\Delta)$.

The above translation can also be read backwards: maximal Higgs subbundles have period maps that are isomorphisms. Consequently, suppressing the omnipresent passages to finite unramified covers, we can summarize the discussion by the following characterization of Teichmüller curves in the language of complex geometry.

**Proposition 3.3.** A Teichmüller curve is a curve $C \to \mathcal{M}_g$ such that the VHS of the family $f : \mathcal{X} \to C$ contains a rank two summand that is maximal Higgs. This maximal Higgs summand is unique.

To illustrate the idea behind Theorem 3.2 note that the zeros $z_1, \ldots, z_n$ on an individual Veech surface can be transported along the whole family $f : \mathcal{X} \to C$ without colliding, again by definition of $\text{GL}_2^+(\mathbb{R})$-action. Passing to an unramified cover of $C$ we may assume that they are the images of sections $z_i : C \to \mathcal{X}$. Using the theory of Néron models one can show that a finite index subgroup of the group of sections extends to the family of Jacobians. We may then project these sections to the family $\mathcal{A} \to \overline{C}$ of the abelian subvarieties whose fibers are $\text{Jac}(X, \omega)$. But this family does not have any non-zero sections. In fact, by the uniformization of the family $\mathcal{A} \to \overline{C}$, sections can be identified with elements of $H^1(\overline{C}, \mathbb{W}_\mathbb{Q})$. This cohomology group naturally has a weight two Hodge structure and the sections provide elements of type $(1,1)$. The maximal Higgs direct summand of $\mathbb{W}_\mathbb{Q}$ prohibits the existence of such non-zero elements.

The two theorems can be recast to characterize Teichmüller curves purely using terms from algebraic geometry, as observed by Alex Wright.

**Proposition 3.4.** A Teichmüller curve is the image in $\mathcal{M}_g$ of a two-dimensional suborbifold $\mathcal{M}$ of a stratum $\Omega \mathcal{M}_g(\mu)$ such that for each point $[(X, \omega)] \in \mathcal{M}$ the abelian variety $\text{Jac}(X, \omega)$ has real multiplication by an order in a field of degree $\dim(\text{Jac}(X, \omega))$ and such that for any two zeros $z_1, z_2$ of $\omega$ the difference $z_1 - z_2$ is torsion in $\text{Jac}(X, \omega)$.

Shimura curves are also defined as totally geodesic curves, but in the moduli space of Abelian varieties $\mathcal{A}_g$ (instead of $\mathcal{M}_g$) and for the Bergman-Siegel metric (instead of the Kobayashi metric). Usually Shimura curves are moreover required to have a CM point, and sometimes they are referred to as Kuga curves with this conditions relaxed. Shimura curves can also be defined as stemming from a homomorphism of a $\mathbb{Q}$-algebraic group into the symplectic group by quotienting the corresponding real groups by maximal compact subgroups and a lattice. Since $\mathcal{A}_g$ is
a locally homogeneous space and since Shimura curves are defined by group theory, there are plenty of Shimura curves. However, since the Torelli-image of $M_g$ in $A_g$ is of large codimension for $g \to \infty$, most of the Shimura curves don’t intersect the Torelli-image. The classification of Shimura curve in (the Torelli-image of) $M_g$ is an open problem that is morally similar to the classification of Teichmüller curves, see e.g. [LZ14] for one of the latest results.

Shimura curves can also be characterized by a decomposition of the VHS like in (2), but now the bundle $\mathcal{M}_\Omega$ has to have unitary monodromy and now all the bundles $L_i$ have to be maximal Higgs (rather than just one of them), but they are not necessarily all Galois conjugates.

4. Constructing Veech surfaces and computing the Veech group

We revisit the known examples of primitive Teichmüller curves in the light of the previous structure results and sketch their method of construction.

Veech’s and Ward’s original examples were constructed by exhibiting two elements in the Veech group that jointly generate a Fuchsian triangle group. In the example of the double pentagon in Figure 1 this is the triangle group generated by the rotation $\begin{pmatrix} \cos 2\pi/5 & \sin 2\pi/5 \\ -\sin 2\pi/5 & \cos 2\pi/5 \end{pmatrix}$ and the vertical shear $\begin{pmatrix} 1 & 2\cot \pi/5 \\ 0 & 1 \end{pmatrix}$. The latter element belongs to the Veech group since the straight line flow in the vertical direction is periodic and the periodic orbits come in two homotopy classes, each sweeping out a cylinder bounded by saddle connections. Both cylinders have the same modulus equal to $2\cot \pi/5$.

Expanding on the previous remark we note that Teichmüller curves are never compact since any direction on the Veech surface admitting a saddle connection provides a parabolic element in the Veech group that has this direction as an eigenvector. However, the Veech groups are in general not generated by elliptic and parabolic elements, as we will prove in Section 7. In fact, none of the series of Teichmüller curves besides the original examples of Veech and Ward was detected by computing the Veech group! Only recently Mukamel [Muk17] gave an algorithm to compute the Veech group for a general Veech surface. His basic idea is to associate to each Veech surface over a Teichmüller curve the number of girth directions that contain a shortest saddle connection. This number provides a stratification of the $GL_2(\mathbb{R})$-orbit of a Veech surface, since the generic number of girth directions is one. The algorithm proceeds by tracing along the spine of this stratification (consisting of surfaces with two girth directions) and testing if the Veech surfaces at vertices of the stratification are scissors congruent to each other.

The construction of the triangle group series started with the observation that families of cyclic coverings have rank two summands in their cohomology whose monodromy groups are triangle groups. However, these summands are not maximal Higgs in the sense of [3]. The problem is that at the points where the monodromy has finite order the family of curves degenerates, but the period map can be continued over these points (after passing to a finite cover). Luckily, if we consider triangle groups $(m, n, \infty)$, say with $m, n$ odd and coprime for simplicity, the group $(\mathbb{Z}/2)^2$ acts on the family of cyclic covers. The quotient family still has the rank two summand in cohomology. Moreover, the fibers over the orbifold points are now smooth and Proposition 3.3 applies.

The Weierstraß series consists of surfaces $(X, \omega) \in \Omega \mathcal{M}_2(2)$ whose cohomology admits a self-adjoint endomorphism $\phi \in \text{End}(H^1(X, \mathbb{Z}))$ such that $\phi^* \omega = \lambda \omega$ for
some \( \lambda \) generating a fixed real quadratic extension \( K \) of \( \mathbb{Q} \). Such a map \( \phi \) defines an endomorphism of \( \text{Jac}(X) \), since it preserves the period lattice and it induces a consistent map on the tangent space of \( \text{Jac}(X) \) as it preserves the line \( \mathbb{C} \cdot \omega \) by definition and another complex line by self-adjointness. The existence of such an endomorphism involves only the periods of \( \omega \) and can thus be checked to hold for surfaces of the form in Figure 3 with the explicitly given parameters of the saddle connection vectors on the boundary. In the minimal stratum the torsion condition is void and thus Proposition 3.4 implies that such \((X, \omega)\) are Veech surfaces.

For the Prym series the crucial observation is that the same argument as for the Weierstraß series can be made for a four-dimensional part of cohomology that is (anti-)invariant by an involution rather than for the whole \( H^1(X, \mathbb{Z}) \).

For the Gothic series this observation is refined to work for endomorphisms acting on an even smaller part of \( H^1(X, \mathbb{Z}) \), the kernel of two projections to the first cohomology of smaller genus curves, provided that the ambient variety without imposing the real multiplication endomorphism behaves like the stratum \( \Omega M_2(2) \) in a sense made precise in the next section.

5. Affine invariant manifolds

Recall that Teichmüller curves are images of closed \( \text{GL}_2^+(\mathbb{R}) \)-orbits. The groundbreaking results of Eskin-Mirzakhani [EM13] and Eskin-Mirzakhani-Mohammadi [EMM15] imply that all the non-closed \( \text{GL}_2^+(\mathbb{R}) \)-orbits have very nice closures: They are manifolds, affine and \( \mathbb{R} \)-linear in a natural ‘period’ coordinate system, and more precisely quasi-projective varieties by Filip’s results [Fil16]. These orbit closures are thus called affine invariant manifolds (AIM). Their classification is a very interesting question that created a lot of recent activity. We refer e.g. to [Api15], [AN16] and [EFW] for some of the latest results and highlight here only the aspects connected with the classification of Teichmüller curves.

Suppose some stratum \( \Omega M_g(\mu) \) contains an infinite number of algebraically primitive Teichmüller curves \( C_i \) for \( i \in \mathbb{N} \). The closure of their union is an AIM \( \mathcal{M} \) by [EMM15]. The main observation of Matheus-Wright [MW15] is that it is possible to spread out the decomposition information (2) from the union of the \( C_i \) to all of \( \mathcal{M} \). Namely, they define a Hodge-Teichmüller plane over the moduli point of \((X, \omega)\) to be a \( \mathbb{C} \)-rank-two subspace \( L \subset H^1(X, \mathbb{C}) \) defined over \( \mathbb{R} \) such that all its \( \text{GL}_2^+(\mathbb{R}) \)-translates intersect the \((1, 0)\)-part of the cohomology in a one-dimensional subspace. By (2) each point over each \( C_i \) has \( g \) orthogonal Hodge-Teichmüller planes and by a limiting argument each point of \( \mathcal{M} \) has them. This leads to an immediate contradiction in many cases, e.g. when the monodromy representation on \( H^1(X, \mathbb{C}) \) over Teichmüller curves generated by Veech surfaces that are torus covers can be shown to not have that many Hodge-Teichmüller planes.

This idea was subsequently refined (by working with relative cohomology and by computing the algebraic hull of \( \text{GL}_2^+(\mathbb{R}) \)-cocycle for general AIM, hence in particular for those containing an infinite number of Teichmüller curves) to yield the following optimal (though ineffective) finiteness result.

**Theorem 5.1 (EFW).** Each stratum \( \Omega M_g(\mu) \) contains only a finite number of Teichmüller curves with trace field of degree \( r > 2 \).

In each stratum \( \Omega M_g(\mu) \) there are only a finite number of AIMs \( \mathcal{M}_i \) 'like \( \Omega M_2(2) \)' such that all primitive Teichmüller curves with \( r = 2 \) are contained in
one of these $\mathcal{M}_1$. Conversely, any such AIM 'like $\Omega M_2(2)$' contains infinitely many Teichmüller curves with $r = 2$.

To give a precise definition of an AIM 'like $\Omega M_2(2)$' we recall that the tangent space of a stratum $\Omega M_2(m_1, \ldots, m_n)$ at $(X, \omega)$ is modelled on the relative cohomology $H^1(X, Z(\omega), \mathbb{C})$, where $Z(\omega) = \{z_1, \ldots, z_n\}$ is the zero set of $\omega$. The tangent space to an AIM $\mathcal{M}$ is by [EM13] and [Wri15] a linear subspace $TM \subseteq H^1(X, Z(\omega), \mathbb{C})$, defined over a real number field $K$ that generalizes the notion of the trace field. The rank of $\mathcal{M}$ is the integer \( \frac{1}{2} \dim p(TM) \) where $p : H^1(X, Z(\omega), \mathbb{C}) \to H^1(X, \mathbb{C})$. It measures (half of) the number of moduli of $\mathcal{M}$ discounting those that stem from moving the zeros relative to each other. An AIM is 'like $\Omega M_2(2)$' if it is rank two and $K = \mathbb{Q}$.

The quest for the classification of primitive Teichmüller curves is thus reduced to detecting the cases with exceptionally large trace field $r > 2$ (like most examples of the triangle series) and theAIMs 'like $\Omega M_2(2)$'. Currently, we only know of the few examples mentioned in the previous section.

6. Finiteness and classification results

All presently known finiteness and classification results for Teichmüller curves are based on the study of their cusps.

The classification of Teichmüller curves in the Weierstraß series and the Prym series starts with listing all possible cusps, which amounts to a finite list of possible combinatorics for the saddle connections and a list of possible length data compatible with real multiplication by the order of a given discriminant $D$. The main problem is to detect when two cusps lie on the same Teichmüller curve. Sometimes it is possible to spot this, like for the cusps belonging to the horizontal and vertical direction in Figure 4. Spotting enough of those direction changes to connect any pair of cusps is the tedious step in the proof of the following theorem.

\textbf{Theorem 6.1 (McM05a, LN14).} In $\Omega M_2(2)$ there is a unique primitive Teichmüller curve $W_D$ with real multiplication by the order of discriminant $D$ for each $D \equiv 1 \mod 8$ and two such Teichmüller curves $W_D^\pm$ for each $D \equiv 1 \mod 8$.

The Prym series in $g = 3$ consists of a unique primitive Teichmüller curve $W_D(4)$ with real multiplication by the order of discriminant $D$ for each $D \equiv 0, 4 \mod 8$, it has two components $W_D^\pm(4)$ for $D \equiv 1 \mod 8$ and is empty for $D \equiv 5 \mod 8$.

A similar result for the $g = 4$-series is known in some cases [LN14] and the classification is open for Gothic curves. It would be very interesting to find a more conceptual argument for the classification of connected components. This classification is currently the only property of Teichmüller curves not accessible through the viewpoint of modular forms, see Section 7 below.

Around the time of discovery, it was puzzling that $\Omega M_2(2)$ contains infinitely many primitive Teichmüller curves, while in $\Omega M_2(1,1)$ there was only a single such curve known, the decagon in Veech’s original family. Given Theorem 3.2 this should no longer come as a surprise: Finding two points $z_1$ and $z_2$ on a Riemann surface $X$ whose difference is torsion is a very rare pick. It is equivalent to finding a map $p : X \to \mathbb{F}^1$ with $p^{-1}(0) = \{z_1\}$ and $p^{-1}(\infty) = \{z_2\}$. Pushing this condition to the cusp of a primitive Teichmüller curve in $\Omega M_2(1,1)$ amounts to detecting when ratios of sines at rational multiples of $\pi$ belong to a quadratic number field. There
are only finitely many possibilities and examining these cases, McMullen showed in [McM06b] that the decagon is indeed the only example in $\Omega M_2(1,1)$.

Even if there are no torsion constraints, i.e. in the minimal strata $\Omega M_g(2g-2)$, there are many other constraints imposed by cusps on the existence of a primitive Teichmüller curve. We illustrate this in the smallest interesting case, the hyperelliptic connected component of the stratum $\Omega M_3(4)$. The fiber over the cusp of an algebraically primitive Teichmüller curve is a projective line (by the real multiplication condition) and the limit of the generating one-form is a stable differential $\omega_\infty$ that we may normalize due to the hyperelliptic involution to have simple poles at $\pm x_i$ for $i = 1, 2, 3$ with residues $\pm r_i$ and a four-fold zero at zero. This amounts to the conditions

\[ \sum_{i=1}^{3} r_i x_{i+1} x_{i+2} = 0 \quad \text{and} \quad \sum_{i=1}^{3} r_i x_i (x_{i+1}^2 + x_{i+2}^2) = 0, \]

where indices have to be read mod 3. We can moreover normalize $r_1 = 1$ and $x_1 = 1$. Algebraically primitive implies that the $r_i$ are integers such that $\sum_{i=1}^{3} b_i / s_i = 0$, where $\{s_1, s_2, s_3\}$ is the $\mathbb{Q}$-basis of $K$, the reciprocal of $x_i$ and $x_j$ at the cusps lie on the boundary of Hilbert modular threefolds, whose boundary has been computed in [BM12], the cross-ratio equation

\[ R_{23}^{d_1} R_{13}^{d_2} R_{23}^{d_3} = 1, \quad (R_{ij} = \left( \frac{x_i + x_j}{x_i - x_j} \right)^2) \]

holds, where $b_i$ are integers such that $\sum_{i=1}^{3} b_i / s_i = 0$, where $\{s_1, s_2, s_3\}$ is the $\mathbb{Q}$-basis of $K$ trace dual to $\{r_1, r_2, r_3\}$. Note that it is already a very restrictive property for a $\mathbb{Q}$-basis of $K$ that the reciprocals of dual basis are $\mathbb{Q}$-linearly dependent. Second, considerations of the Harder-Narasimhan filtration of the Hodge bundle over the Teichmüller curve imply that one of the two Galois conjugate forms also has to have a double zero in common with $\omega$. This implies

\[ \sum_{i=1}^{3} r_i^7 x_{i+1} x_{i+2} = 0. \]

We strongly suspect that the solution stemming from Veech’s 7-gon

\[ r_2 = v^2 + v - 2, \quad r_3 = v^2 - 2, \quad x_2 = -v^2 - v + 1, \quad x_3 = v^2 + v - 2, \quad (v = 2 \cos(2\pi/7)) \]

is the unique solution in some $K$ up to permutation of variables of the equations [3], [5], and [6]. However, the fact that equations are not algebraic, but involve Galois conjugates, makes the geometry of the solution set more interesting. Currently, we can only show that the set of solutions is finite. This is part of the following version of Theorem [5.1] for $g = 3$, that has the advantage to be at least in theory algorithmically implementable.

Theorem 6.2 ([BHM16]). There are finitely many algebraically primitive Teichmüller curves in genus three.

The proof in the minimal stratum uses a variant of the theory of just likely intersections. The rough statement of the main theorem of this theory is that all intersection points of an algebraic subvariety $Y$ of a multiplicative torus $\mathbb{G}_m^n$ with all subtori of dimension $n - \dim(Y) - 1$ have bounded height, except for the anomalous
locus \( Y^{an} \subseteq Y \) consisting of the subvarieties that intersect a translate of a subtorus of \( G_m \) in a larger subvariety than expected from the naive dimension count. The subtori this theory is applied to are those defined in [5], but the variant in \( BHM16 \) uses coupled equations in \( G_m \) and the additive group \( G_0^n \).

7. Modular forms and Euler characteristics

The locus of abelian surfaces with real multiplication by an order of discriminant \( D \) is the Hilbert modular surface \( X_D = \mathbb{H}^2/\langle \text{SL}(\mathfrak{o}_D \oplus \mathfrak{o}_D) \rangle \). By Theorem 3.1, the Torelli image of a Teichmüller curve with quadratic trace field lands in \( X_D \). We can thus use modular forms and other tools from number theory to approach the geometry of Teichmüller curves, e.g. in the Weierstrass, Prym and Gothic series.

The vanishing locus of a Hilbert modular form of weight \((k, \ell)\) descends to an algebraic curve in \( X_D \). Not all curves on a Hilbert modular surface are the vanishing locus of a Hilbert modular form, not even linearly equivalent to such a vanishing locus. However, all the Teichmüller curves in the Weierstrass and the Prym series can be described using Hilbert modular forms. Concretely, let \( \theta_{(m,m')}^{(m,m')} (z,u) \) be the restriction of the classical Riemann theta function with characteristic \( m,m' \in \frac{1}{2} \mathbb{Z}/\mathbb{Z}^2 \) to \( z = (z_1,z_2) \in \mathbb{H}^2 \), let \( u = (u_1,u_2) \) be coordinates of \( \mathbb{C}^2 \) that correspond to eigendirections of real multiplication, and let \( D_2 \theta_{(m,m')} = \frac{\partial}{\partial u_2} \theta_{(m,m')} (z,u) \).

**Theorem 7.1** ([Bai07], [MZ16]). The image of the Weierstrass Teichmüller curve \( W_D \) in \( X_D \) is the vanishing locus of the Hilbert modular form

\[
D \theta(z) = \prod_{(m,m') \text{ odd}} D_2 \theta_{(m,m')} (z)
\]

of weight \((3,9)\). The orbifold Euler characteristic of \( W_D \) is \( \chi(W_D) = -\frac{9}{2} \chi(X_D) \).

The Euler characteristic of Hilbert modular surfaces is explicitly computable, for fundamental discriminants \( D \) it is simply \( \chi(X_D) = \zeta_{\mathbb{Q}(\sqrt{D})}(-1) \).

Note that the weight of the modular form \( D \theta \) is non-parallel, while in the literature almost exclusively modular forms of parallel weight \((k,k)\) appear. The reason for this can be explained as follows. Teichmüller curves are geodesic for the Teichmüller metric on \( T_g \), which is the same as the Kobayashi metric. Images of Teichmüller curves are still geodesic for the Kobayashi metric on the Hilbert modular surface \( X_D \) and in fact also on the moduli space of abelian surfaces \( A_2 \). On \( X_D \), the Kobayashi metric is the supremum of the Poincaré metrics on the two factors. In each point of the Teichmüller curve this supremum is attained precisely for the first factor of \( \mathbb{H}^2 \). This is a restatement of the fact that the maximal Higgs summand in Proposition 3.3 is unique.

The proof of Theorem 7.1 recasts in terms of modular forms the fact that the Abel-Jacobi map based at a Weierstrass point embeds the Veech surface in its Jacobian as the vanishing locus of a translate of the theta divisor. Consequently, the eigenform for real multiplication has a zero at the Weierstrass point (i.e. the Veech surface belongs to the stratum \( \Omega \mathcal{M}_2(2) \)) if and only if the theta divisor has a tangent in an eigendirection for real multiplication. This is expressed by the right hand side of (7).

There is an analogous theorem that expresses the Teichmüller curves in the Prym series as the vanishing locus of a determinantal expression in derivatives of theta functions [Möl14]. It also yields an expression of the Euler characteristics...
\( \chi(W_D(4)) \) and \( \chi(W_D(6)) \) as a multiple (depending only on \( D \mod 8 \)) of \( \chi(X_D) \).

The proof refines the above argument, using that the Prym-Abel-Jacobi image of a Veech surface \((X, \omega)\) in the Prym series is immersed in \( \text{Jac}(X, \omega) \).

8. ORBIFOLD POINTS AND OTHER CONNECTIONS TO ARITHMETIC GEOMETRY

Orbifold points of Teichmüller curves are, besides cusps and the Euler characteristic, the last missing piece in determining their topology. Orbifold points provide an additional automorphism of the Jacobian, besides the real multiplication on \( \text{Jac}(X, \omega) \). Consequently, orbifold points give points of complex multiplication (in the general sense, allowing endomorphism rings that are matrix rings). Recall that by proven versions of the André-Oort conjecture (see [Edi01]) Shimura curves in Hilbert modular surfaces are characterized by having infinitely many CM points. On the other hand, a Teichmüller curve is a Shimura curve at most for those generated by some torus covering Veech surfaces in genus three and four. We will thus find only finitely many CM points and hence finitely many orbifold points on a primitive Teichmüller curve.

Orbifold points are loosely connected to billiards. The unfolding construction (Figure 1) exhibits a Veech surface \((X, \omega)\) arising from a dynamically optimal billiard table \( P \) as \( X = \bigcup_{g \in G} gP \) for some finite group \( G \) generated by reflections. The index two subgroup \( G' \subset G \) that preserves the orientation belongs to the Veech group of \((X, \omega)\). However, if the billiard table has right angles only, this group \( G' \) might just consist of an involution. E.g. for L-shaped billiard tables it gives the hyperelliptic involution, common to all Veech surfaces in genus two rather than to special orbifold points. The locus of unfoldings of billiard tables is a real codimension one submanifold of the Teichmüller curves in this case.

Orbifold points on all but the most recently discovered series of Teichmüller curves have been determined by Mukamel and by Torres-Teigell and Zachhuber.

Theorem 8.1 ([Muk14], [TZ15], [TZ16]). The orbifold points on the Weierstraß Teichmüller curve \( W_D \) are a point of order five on \( W_5 \) and \( \tilde{h}(-D) \) points of order two.

The orbifold points on the Prym Teichmüller curve \( W_D(4) \) for \( D > 12 \) are \( H_2(D) \) points of order two and \( H_3(D) \) points of order three.

Here \( \tilde{h}(-D) \) are generalized class numbers, e.g. \( \tilde{h}(-D) = h(-4D)/|\sigma_{-4D}| \) for odd discriminants \( D \) and \( H_2(D) \) and \( H_3(D) \) are representation number for \( D \) by quadratic forms, with \( H_2(D) = 0 \) if \( D \) is odd. For \( D \leq 12 \) there are a finite number of exceptional cases with orbifold points. A similar statement also holds for the Prym Teichmüller curves \( W_D(6) \), see [TZ16].

We conclude this survey by addressing various aspects that emphasize the arithmetic nature of Teichmüller curves. They are defined over number fields, since the existence of the maximal Higgs subbundle \( L_1 \) implies rigidity ([MV10], [McM09]). Since maximal Higgs is a numerical condition, the Galois conjugate of a Teichmüller curve is again a Teichmüller curve. This allows to search for natural integral models over number rings for Teichmüller curves and to study the primes of bad reduction of these models. Such models were computed in [BM10a] and many more in [KM14], providing an interesting conjectural picture of the bad primes.

Since Teichmüller curves are characterized by their uniformization, there is a natural notion of modular forms for the Veech group. Since the universal covering...
of the map $C \to X_D$ of a Teichmüller curve to the Hilbert modular surface can be written as $z \mapsto (z, \varphi(z))$ for some holomorphic map $\varphi$, there is, besides the usual automorphy factor $(cz + d)$ also the twisted automorphy factor $(e^\sigma \varphi(z) + d^\sigma)$ where $\sigma$ is a generator of $\text{Gal}(K/\mathbb{Q})$. This leads to a theory of twisted modular forms, studied in [MZ16]. However, since the Veech group of a primitive Teichmüller curve is not arithmetic there is no theory of Hecke operators on twisted modular forms. It is an open problem if there is any replacement of the pivotal role usually played by Hecke eigenforms in this context.

Acknowledgements: Considerable part of the work presented in this survey was obtained in collaboration. The author thanks his collaborators for the pleasure to work jointly. The author is grateful to C. Salzmann for her help in the preparation of the pictures.

References


Institut für Mathematik, Goethe-Universität Frankfurt, Robert-Mayer-Str. 6–8, 60325 Frankfurt am Main, Germany

E-mail address: moeller@math.uni-frankfurt.de