HOROCYCLE DYNAMICS: NEW INVARIANTS AND EIGENFORM LOCI IN THE STRATUM $\mathcal{H}(1,1)$

MATT BAINBRIDGE, JOHN SMILLIE, AND BARAK WEISS

Abstract. We study dynamics of the horocycle flow on strata of translation surfaces, introduce new invariants for ergodic measures, and analyze the interaction of the horocycle flow and real Rel surgeries. We use this analysis to complete and extend results of Calta and Wortman classifying horocycle-invariant measures in the eigenform loci. We classify the orbit-closures and prove that every orbit is equidistributed in its orbit-closure. We also prove equidistribution statements regarding limits of sequences of measures, some of which have applications to counting problems.

1. Introduction

Translation surfaces arise naturally in many different mathematical contexts, e.g. complex analysis, geometric group theory, geometry and billiards. See [MaTa, Zo] for surveys. A stratum is a moduli space of translation surfaces of a given topological type (detailed definitions will be given below). Ideas from renormalization theory have been useful in the study of translation surfaces, and these are encoded in the action of the group $G = \mathrm{SL}_2(\mathbb{R})$, and its subgroups, on strata. The study of the dynamics of these actions has led to significant advances in the understanding of the geometry of translation surfaces and the dynamics of flows on individual translation surfaces, and hence has been under intensive study for over 30 years. A fruitful analogy for understanding the dynamics, is an analogy between strata of translation surfaces and homogeneous spaces of Lie groups (in fact the simplest stratum, namely the stratum of tori with one marked point, is the homogeneous space $G/\mathrm{SL}_2(\mathbb{Z})$, but other strata are not homogeneous). A celebrated set of results in the homogeneous setting, due to Ratner, Margulis, Dani, and others, is the classification of invariant measures, orbit-closures, equidistribution of orbits, and additional equidistribution results. We refer to [KSS] for a detailed survey of results and methods of homogeneous dynamics. These results, and the techniques employed in their

Date: March 2, 2016.

proof, have been inspirational for many developments in the study of dynamics on strata.

A central problem is to classify the invariant measures and orbit-closures for the actions of G, and its subgroup U of unipotent matrices, on strata. For the G-action, measure and orbit-closure classification results were obtained by McMullen [McM3] in the genus two strata $\mathcal{H}(2)$ and $\mathcal{H}(1,1)$. In arbitrary genus, Eskin, Mirzakhani and Mohammadi [EMi, EMiMo] showed that invariant measures and orbit-closures are of a geometrical nature, sharing many of the features of the strata themselves, and subsequent and ongoing work of many authors (see e.g. [F, W2]) has revealed a rich connection to arithmetic and algebraic geometry. These results could be regarded as an analogue of Ratner's results for the action of G, and this analogy raises the question of whether similar phenomena are also valid for the U-action. However the U-action is not as well understood. The purpose of this paper is to make a contribution to the study of the U-action, building on previous work of Eskin, Marklof and Morris [EMaMo], and Calta and Wortman [CW], who adapted some of Ratner's arguments to the non-homogeneous setting.

In the first part of this paper we discuss general features of the U-action and in the second part, we give a complete picture in a restricted setting, namely in the eigenform loci. These spaces are G-invariant loci within the genus 2 stratum $\mathcal{H}(1,1)$, which were discovered by Calta and McMullen [C, McM2, McM3]. In these loci we extend the analogy with homogeneous spaces by proving analogues of results proved by Ratner, Dani, Margulis, and Shah.

First we describe our dynamical invariants for the U-action. Several of these involve the real Rel vector fields, which correspond to local motion in a direction in a stratum which moves one singularity horizontally with respect to another, while fixing absolute periods of the surface. As observed by Calta [C], these vector fields commute with the U-action, and in particular send U-orbits to U-orbits, preserving their parameterization. Combining these vector fields with the geodesic flow provides directions of motion which normalize the U-action, sending U-orbits to U-orbits, while inducing a time-change on the orbits. Thus the centralizer and normalizer of the U-action are locally modeled respectively on an abelian Lie group Z and a nilpotent Lie group N.

There is a fundamental difference between homogeneous spaces and strata of translation surfaces, in connection with these vector fields. On homogeneous spaces, natural vector fields can be integrated to define group actions on the space, and the centralizer and the normalizer subgroups of the U-action play an important role in studying the dynamics. In the case of strata, motion in the real Rel directions is not globally defined; i.e. the solution curves for the differential equation defined by these vector fields may not be defined for all times. This implies that the centralizer and normalizer of the horocycle flow make sense locally but do not correspond to globally defined actions of Lie groups (in [EMaMo] this is discussed using the terminology of "pseudo group-actions", which we will avoid favoring the language of geometric structures and vector fields). The domain of well-defined motion in the Z and N directions are invariants of a U-invariant ergodic measure μ, which we denote by $Z^{(\mu)}, N^{(\mu)}$ (see §4). Moreover within these domain of well-defined motion, are Lie groups Z_μ, N_μ which do act on $\operatorname{supp}\mu$ and are the stabilizer subgroups of μ within the centralizer and normalizer (see §4.1).

An additional invariant is a "horizontal data diagram" $\Xi(\mu)$ (see §5), recording topological and geometric structure left invariant by the U-action. This includes the number and length of horizontal saddle connections, the topology of the complements, and the incidence of these saddle connections at singularities.

Rigorous definitions of the invariants listed above require passing to certain finite covering spaces of strata. To this end we begin in §2 with a detailed discussion of "blow ups of translations surfaces" along with their corresponding moduli spaces and mapping class groups. These notions are useful throughout our discussion: for marking singularities, marking distinguished horizontal prongs, distinguishing horizontal data diagrams, resolving orbifold issues for both the stratum and for the structure of individual rel leaves, and for discussing surgeries involving a stratum and nearby boundary strata obtained from it as limits of rel operations. An interesting novelty of our approach is that certain finite covers $\hat{G}$ of G appear as the groups which acts naturally on our covers (see §2.10). The finite covers we consider have arisen in topological contexts (see [Boi]), in the study of interval exchange transformations (see [Y]) and in computations of monodromy representations (see e.g. [MYZ]), as well as in our forthcoming [SmWe3]. We believe that the terminology we introduce will be useful in future work on translation surfaces.

Recall that a surface M is *generic for* μ if for any continuous compactly supported function f on $\operatorname{supp}\mu$, $\lim_{T\to\infty}\frac{1}{T}\int_0^T f(u_sM)\,ds = \int f\,d\mu$, where $U = \{u_s : s \in \mathbb{R}\}$. The analysis of generic points plays a major role in the analysis of ergodic measures on homogeneous spaces. The Rel surgery depends on a parameter T and as remarked above, the Rel surgery is not globally defined. As a consequence, the set of surfaces for which it is defined for all T is not locally compact and it is not a-priori clear that the centralizer of the U-action should map U-generic points to U-generic points. We clarify this in §4.2. In §6 we explicitly identify the domain of definition of real Rel surgeries, continuing work done in [MW2, McM8, B2].

In the second part of the paper we specialize to eigenform loci. Namely, let $D \geqslant 4$ be an integer congruent to either 0 or 1 mod 4, and let $\mathcal{E}_D(1,1)$ be the space of translation surfaces in $\mathcal{H}(1,1)$ which are eigenforms for real multiplication by a quadratic order of discriminant D. The following is our measure classification result.

Theorem 1.1. *Let D be as above and let μ be a U-invariant U-ergodic Borel probability measure on $\mathcal{E}_D(1,1)$. Then one of the following holds:*

1. *Every surface in* $\operatorname{supp}\mu$ *has a horizontal cylinder decomposition and μ is the length measure on a periodic U-orbit.*
2. *Every surface in* $\operatorname{supp}\mu$ *has a horizontal cylinder decomposition into three cylinders and μ is the area measure on a 2-dimensional minimal set for the U-action. In this case μ is invariant under the real Rel operation.*
3. *For every $M \in \operatorname{supp}\mu$, the horizontal data diagram $\Xi(M)$ contains two saddle connections, joining distinct singularities, whose union divides M into two isogenous tori glued along a slit. In this case μ is the image of the G-invariant measure on a quotient G/Γ for some lattice Γ, via a Borel U-equivariant map.*
4. *For every $M \in \operatorname{supp}\mu$, $\Xi(M)$ contains one saddle connection joining distinct singularities, and μ is the image of the $\widehat{G}$-invariant measure on a quotient $\widehat{G}/\Gamma$ for some lattice Γ in the 3-fold connected cover $\widehat{G}$ of G, via a Borel U-equivariant map.*
5. *The set of surfaces with horizontal saddle connections has μ measure zero and μ is the image of the G-invariant measure on a closed G-orbit in $\mathcal{H}(1,1)$ via a Borel U-equivariant map. In this case D is either a square or is equal to 5.*
6. *For every $M \in \operatorname{supp}\mu$, $\Xi(M)$ contains two saddle connections joining distinct singularities, whose complement in M is a torus with two parallel slits of equal length, which are images of each other under a translation by an exactly d-torsion element of the torus. In this case $D = d^2$ is a square and μ is the image, via a Borel U-equivariant map, of the G-invariant measure on the space of tori.*
7. *μ is the canonical flat measure on $\mathcal{E}_D(1,1)$ obtained from period coordinates.*

We describe these measures in detail in §8. The construction of most of the measures in this list involves the real Rel operation on $\mathcal{H}(1,1)$. Any U-invariant measure μ, which is not preserved by real Rel, gives rise to a one-parameter family of U-invariant translates of μ by real Rel. This observation of Calta [C, CW] is crucial to our analysis, and the measures (3)–(6) all arise in this way as real Rel translates of G-invariant measures. The measures (5) are the real Rel translates of the natural measures on closed G-orbits in $\mathcal{H}(1,1)$. Loosely speaking, the measures (3),(4), and (6) are all pushforwards of measures on closed G-orbits in a suitable boundary component in a bordification of $\mathcal{H}(1,1)$, where the maps pushing the measure consist of the composition of real Rel with a map passing from the boundary component to $\mathcal{H}(1,1)$. In (4), the closed G-orbit belongs to the stratum $\mathcal{H}(2)$, while in (3) it belongs to $\mathcal{H}(0) \times \mathcal{H}(0)$, and in (6) it belongs to $\mathcal{H}(0,0)$ (where $\mathcal{H}(0)$ and $\mathcal{H}(0,0)$ denote respectively the moduli spaces of genus one translation surfaces with one or two marked points).

The equivariant maps in cases (3)–(6) arise from the real Rel operation and are not defined for every point in G/Γ. They define Borel isomorphisms between the supports of these measures and homogeneous spaces (that is, quotients G/Γ for lattices $\Gamma \subset G$). We emphasize however that these maps are not everywhere defined, but rather on a dense open set of full measure, and thus their existence does not imply the existence of homeomorphisms of the supports of these measures with homogeneous spaces of G. In fact such homeomorphisms do not exist in general; in a forthcoming paper [SmWe3], we show that the supports of these measures are not homeomorphic to homogeneous spaces. In particular the supports of the measures appearing in case (5) above are manifolds with nonempty boundaries and infinitely generated fundamental groups. It is likely that the same is true in cases (3), (4) and (6).

Theorem 1.1 extends a theorem of Calta and Wortman [CW]. In [CW] it was assumed that D is not a square, and thus case (6) did not arise. Also [CW] assumed that μ is not invariant under the real Rel operation, and case (2) did not arise. Theorem 1.1 is proved in a more general form in §9. The statement is inspired by Ratner's measure classification theorem [KSS, Thm. 3.3.2], and its proof is inspired by [EMaMo, CW], which in turn employs arguments of Ratner. We provide some technical shortcuts and clarify some delicate steps. Our treatment relies on the general results worked out in §4.1, §4.2 and §6.

The measures in cases (1)–(6) of Theorem 1.1 are naturally organized in continuous families depending on some real parameters (e.g. in case

(5) the continuous parameter involves a real Rel perturbation, and in case (2) there is a natural two parameter family of perturbations). We call such continuous families of U-invariant measures "beds". Within them, the invariants $\Xi(\mu)$, $Z^{(\mu)}$ and Z_μ are constant, while $N^{(\mu)}$ and N_μ vary in a natural way with respect to the real Rel action. In §11, we develop an analogue of the "linearization technique" developed for homogeneous spaces (see [KSS, §3.4]) to analyze the behavior of U-orbits which are near beds. Using these ideas we classify all orbit-closures for the U-action on $\mathcal{E}_D(1,1)$.

In fact we prove a stronger statement. The assertion that M is generic for μ constitutes a quantitative strengthening of the assertion that $\overline{UM} = \operatorname{supp} \mu$. The following statement is an analogue of Ratner's genericity theorem [KSS, Thm. 3.3.10]:

Theorem 1.2. *For any D as above and any $M \in \mathcal{E}_D(1,1)$, there is a measure μ as in Theorem 1.1 such that M is generic for μ, and belongs to the support of μ.*

In §11 we deduce Theorem 1.2 from a more explicit result which explains, given M, how to construct the measure μ for which M is generic.

Continuing the analogy with related results for homogeneous flows, we prove several equidistribution results. We mention three of our results of this type here, and refer the reader to §12 for more results along these lines.

Theorem 1.3. *Let μ be the length measure on a periodic U-orbit in $\mathcal{E}_D(1,1)$, let $\{g_t\}$ denote the geodesic flow, and suppose $\operatorname{supp} \mu$ is not contained in a closed G-orbit. Then the measures $g_{t*}\mu$ converge to the flat measure on $\mathcal{E}_D(1,1)$ as $t \to \infty$.*

This could be viewed as a non-homogeneous analogue of a theorem that Shah proves in a homogeneous setting [KSS, Thm. 3.7.6]. The following result also has an analogue in the homogeneous setting (see [KSS, §3.7]):

Theorem 1.4. *Let μ be the G-invariant measure on a closed G-orbit in $\mathcal{E}_D(1,1)$. Let $t \in \mathbb{R}$ and let μ_t be the measure obtained by applying the real Rel operation to surfaces in $\operatorname{supp} \mu$, with rel parameter t. Then as $t \to \infty$ or $t \to -\infty$, μ_t converges to the flat measure on $\mathcal{E}_D(1,1)$.*

In response to a question of Giovanni Forni, we prove:

Theorem 1.5. *Let μ be any ergodic U-invariant measure on $\mathcal{E}_D(1,1)$, then as $t \to +\infty$, $g_{t*}\mu$ converges to a G-invariant measure, and as*

$t \to -\infty$, either $g_{t*}\mu$ converges to a G-invariant measure or $g_{t*}\mu$ is divergent in the space of probability measures on $\mathcal{H}(1,1)$.

We also prove an equidistribution result for large circles in Theorem 12.7. Following a strategy of Eskin and Masur [EM], we use this result to solve a counting problem. We obtain:

Theorem 1.6. *For any $M \in \mathcal{E}_D(1,1)$, the limit*

$$C_D = \lim_{T\to\infty} \frac{\#\{\textit{saddle connections on } M \textit{ of length } \leqslant T\}}{T^2}$$

exists.

Theorem 1.6 was proved in [EMS] for the case that D is a square, and in [B2] for the case that D is not a square. In both of these papers precise formulae for the constants C_D were given. In fact, when D is not square, $C_D = 4\pi$, which agrees with the value of this constant for a generic surface in $\mathcal{H}(1,1)$. The proof given in [B2] contained a gap in the case $D = 5$ which our results bridge.

We expect that the results obtained in §8–§12 for the eigenform loci in genus 2, can be extended to the Prym loci in genera 3, 4, 5 (see [McM6]), and more generally, to the rank-one loci defined by Wright [W1]. We hope to return to this topic in future work.

1.1. **Acknowledgements.** The authors are grateful to Uri Bader, Elon Lindenstrauss and Saul Schleimer for useful conversations. This work is supported by BSF grant 2010428, Royal Society Wolfson Research Merit Award, ERC starter grant DLGAPS 279893, and Simons Foundation grant 359821.

2. Basics

In this section we define our objects of study: translation surfaces, moduli spaces of translation surfaces and dynamics on moduli spaces. For background and alternate treatments we refer the reader to [EMZ, MaTa, MaSm, Zo]. Our treatment expands on previous work by discussing in detail blow-ups of translation surfaces and associated orbifold covering spaces of strata, which are needed in later sections. Our discussion also has interesting consequences of indendent interest about fundamental groups of strata (see §2.11).

2.1. **Translation surfaces.** A *translation surface,* or a *surface with translation structure* can be defined in several equivalent ways. We will describe it by gluing polygons, in terms of an atlas using the language of (G, X) structures, or as a holomorphic differential.

Let M be a surface obtained from a collection of polygons in $\mathbb{R}^2$ which are glued together by isometries of the edges which are restrictions of translations. Each polygon edge receives an orientation when viewed as part of the boundary of a polygon and we assume that the gluing isometries reverse these boundary orientations. This construction typically produces a finite set of *singular points* corresponding to the vertices of the polygons at which the cone angle is larger than 2π. The cone angle at a singularity is $2\pi(n+1)$ for some natural number n. We call n the order of the singularity. In addition to the singular points that we have defined it is often useful to mark a finite set of points at which the cone angles are 2π. Let Σ be a finite subset of M which contains all of the singular points and possibly some additional "marked" points.

A translation structure on a surface determines an atlas of charts for the surface $M \smallsetminus \Sigma$ taking values in $\mathbb{R}^2$ where the transition maps are restrictions of translations. A translation structure can be determined by specifying this atlas. If we have a space X equipped with an action of a group G, a (G, X)*-structure* is an atlas of charts with overlap functions in G. Thus a translation surface produces a $(G, X) = (\mathbb{R}^2, \mathbb{R}^2)$ structure on M where the first $\mathbb{R}^2$ represents a Lie group acting on the second $\mathbb{R}^2$ by translation. See [Th] for more information about (G, X)-structures.

We can use the atlas of charts on M to define geometric structures on M which are naturally associated with the translation structure. Since the one-forms dx and dy are translation invariant on $\mathbb{R}^2$ these charts allow us to build globally defined one-forms dx and dy on M. Similarly we can use the planar charts to define a metric, an area form, and an orientation on M. A *saddle connection* on M is a path with endpoints in Σ, which is a straight line in each chart of this atlas, and does not contain singularities in its interior. The one-forms dx and dy are closed and represent cohomology classes in $H^1(M, \Sigma; \mathbb{R})$. For an oriented path γ connecting points in Σ write $\mathrm{hol}(M, \gamma)$ for $\left(\int_\gamma dx, \int_\gamma dy\right)$. We can think of $\mathrm{hol}(M, \cdot)$ as giving a homomorphism from $\pi_1(M)$ to $\mathbb{R}^2$ or as determining an element of $H^1(M, \Sigma; \mathbb{R}^2)$. This is the *holonomy homomorphism.* If $\widetilde{M}$ is the universal cover of M then a map from $\widetilde{M}$ to $\mathbb{R}^2$ with derivative equal to the identity is a *developing map.* Developing maps exist and are unique up to translation. The developing map is equivariant with respect to the holonomy homomorphism. This is an instance of the general principle that for any (G, X)-structure on a manifold M, there is a developing map $\mathrm{dev} : \widetilde{M} \to X$ and a holonomy homomorphism $\pi_1(M) \to G$.

The atlas for a translation surface gives a trivialization of the tangent bundle at non-singular points. Let us say that $f : M \to N$ is a smooth map between translation surfaces if it preserves singular sets, and is smooth away from the singular set of M. Using the trivialization of the tangent space, we can view the derivative of a smooth map f as a 2×2 real matrix-valued function. We say that a smooth map f is a *translation equivalence* if it is a homeomorphism and its derivative is the identity matrix. The notion of translation equivalence gives a natural equivalence relation on translation surfaces. We say that M and N are *affinely equivalent* if there is an orientation-preserving smooth map f between them which is a homeomorphism and for which Df is constant but not necessarily equal to the identity. The affine equivalence classes are orbits of a $\mathrm{GL}_2^\circ(\mathbb{R})$-action we will discuss shortly.

If we identify $\mathbb{R}^2$ with $\mathbb{C}$ then the coordinate charts of a translation structure induce a conformal structure on $M \smallsetminus \Sigma$. This conformal structure extends to M so that the points in Σ correspond to punctures. We can define the complex-valued one-form $dz = dx + i\,dy$. This is a holomorphic one-form or Abelian differential with zeros at the singular points where the order of the zero is the order of the singular point. If we are given a surface with a conformal structure X and we wish to define a compatible translation surface then this is determined by a nonzero holomorphic one-form ω. We use the notation (X, ω) to denote this pair. For the relations between these three points of view, see [Zo] and the references cited there.

2.2. **Strata as sets.** Strata are moduli spaces of translation surfaces. In this section we describe the set of surfaces corresponding to a given stratum and the relevant notion of equivalence. In section 2.4 we will show that the set of these equivalence classes of surfaces can be given natural orbifold structures. Let $k \in \mathbb{N}$ and let $a_1, \ldots, a_k$ be a sequence of non-negative integers. Let M be a translation surface and let Σ be a finite subset of M which consists of k points labeled $\xi_1, \ldots, \xi_k$ and contains all singular points of M. (Note that our conventions imply that Σ has at least one point.) We say that M is a translation surface of type $(a_1, \ldots, a_k)$ if the cone angle at ξ_j is $2\pi(a_j + 1)$. We refer to a_j as the order of ξ_j. The points ξ_j with order zero are not singular and we will refer to them as marked points. We want to construct a version of the stratum in which singular points have well defined labels. To this end we say that two surfaces of type $(a_1, \ldots, a_k)$ are label preserving translation equivalent if they are translation equivalent by means of a translation equivalence preserving the labelling of the singular points. (When no confusion will result we will drop the expression "label preserving".)

Let $\mathcal{H} = \mathcal{H}(a_1, \ldots, a_k)$ denote the *stratum* of translation surfaces of type $(a_1, \ldots, a_k)$, considered up to label preserving translation equivalence. The surfaces in $\mathcal{H}$ have genus g where $2g - 2 = \sum a_j$. It is sometimes convenient to denote this stratum as $\mathcal{H}_g(a_1, \ldots, a_k)$. In particular note that $g \geqslant 1$ since the a_j are non-negative.

Caution should be exercised in dealing with surfaces with non-trivial translation automorphisms. Since there is no *canonical* equivalence between two equivalent surfaces with non-trivial translation automorphisms it is dangerous to think of them as being 'the same'. When we build the stratum as a topological space these surfaces will correspond to orbifold points.

2.3. **Strata of marked surfaces.** Let S be an oriented surface of genus g and let Σ be a subset consisting of k points labeled $\xi_1, \ldots, \xi_k$. The pair (S, Σ) will serve as a topological model for a translation surface. A *marked translation surface* is a translation surface M of type $(a_1, \ldots, a_k)$, equipped with an orientation-preserving homeomorphism $f : S \to M$ where Σ maps to the appropriate distinguished points of the translation structure, respecting the labels. We will consider S as fixed and sometimes write maps $f : S \to M$ as pairs (f, M). Two marked translation surfaces (f_1, M_1) and (f_2, M_2) are considered to be equivalent as marked translation surfaces if there is a label preserving translation equivalence $g : M_1 \to M_2$ so that $g \circ f_1$ and f_2 are isotopic via an isotopy that fixes the points in Σ.

Let $\mathcal{H}_{\mathrm{m}} = \mathcal{H}_{\mathrm{m}}(a_1, \ldots, a_k)$ denote the set of marked translation surfaces, up to the equivalence relation described above. As discussed above, a surface $M \in \mathcal{H}$ determines a cohomology class $\mathrm{hol}(M, \cdot) \in H^1(M, \Sigma; \mathbb{R}^2)$. A marked translation surface $f : S \to M$ in $\mathcal{H}_{\mathrm{m}}$ determines a pullback cohomology class $f^*(\mathrm{hol}(M, \cdot)) \in H^1(S, \Sigma; \mathbb{R}^2)$ where $f^*(\mathrm{hol}(M, \gamma)) = \mathrm{hol}(M, f_*(\gamma))$. We denote by

$$\mathrm{dev} : \mathcal{H}_{\mathrm{m}} \to H^1(S, \Sigma; \mathbb{R}^2) \tag{1}$$

the map which takes a marked translation surface to the corresponding element of $H^1(S, \Sigma; \mathbb{R}^2)$. This map is often called the *period map.*

2.4. **Strata as spaces.** We now define a topology on $\mathcal{H}_{\mathrm{m}}$ using an atlas of charts which is defined via triangulations of translation surfaces. Let τ be a triangulation of S so that the vertices are points in Σ. We do not require that edges have distinct endpoints in S though we do require that they have distinct endpoints in the universal cover of

S or equivalently that no two edges are homotopic relative to their endpoints. This triangulation gives S the structure of a Δ-complex (for a definition see [H, p. 102]). Let $\mathcal{U}_\tau \subset \mathcal{H}_{\mathrm{m}}$ be the set of marked translation surfaces containing a representative $f : S \to M$ which takes the edges of τ to saddle connections in M.

The function that maps an oriented edge to its holonomy vector is a 1-cochain in the cochain complex associated to the triangulation τ. The condition that the sums of vectors on the boundary of a triangle is zero means that this 1-cochain is a cocycle. Using the fact that τ is a Δ-complex we can identify the space of such cocycles with $H^1(S, \Sigma; \mathbb{R}^2)$ (see [H]).

Using the triangulation τ allows us to define *comparison maps* between two marked translation surfaces lying in a given chart $\mathcal{U}_\tau$. Namely, suppose (f, M) and (f', M') are two representatives of marked translation surfaces in the same $\mathcal{U}_\tau$. Then $f' \circ f^{-1} : M \to M'$ is a homeomorphism but it is only well defined up to isotopy. We define a map F in this isotopy class by requiring that for each triangle Δ in τ, F is an affine map when restricted to $f(\Delta)$. This requirement and the triangulation τ determine $F = F(\tau, M, M')$ uniquely, and we call it the *comparison map* between M and M'.

Proposition 2.1. *The map* $\mathrm{dev}|_{\mathcal{U}_\tau} : \mathcal{U}_\tau \to H^1(S, \Sigma; \mathbb{R}^2)$ *is injective and its image is open.*

Proof. Suppose (f, M) and (f', M') are in $\mathcal{U}_\tau$ and have the same image under dev. The comparison map $F(\tau, M, M') : M \to M'$ is a translation equivalence since the holonomy of corresponding edges are equal. By changing f and f' by isotopies we can assume that the image $f(\tau)$ and $f'(\tau)$ are geodesic triangulations. So the maps $F \circ f$ and f' are equal. We conclude that (f, M) and (f', M') are equivalent and the developing map is injective on $\mathcal{U}_\tau$.

A cohomology class ϕ is in the image of $\mathrm{dev}|_{\mathcal{U}_\tau}$ if the values of ϕ on the sides of every triangle Δ in τ correspond to the coordinates of a non-degenerate triangle in $\mathbb{R}^2$ with the appropriate orientation — an open condition in $H^1(S, \Sigma; \mathbb{R}^2)$. If this condition is satisfied then an appropriate translation surface can be built by gluing together triangles in $\mathbb{R}^2$ with edge coordinates given by ϕ. In particular the image $\mathrm{dev}(\mathcal{U}_\tau)$ is open in $H^1(S, \Sigma; \mathbb{R}^2)$. $\square$

It was shown in [MaSm] that every translation surface M with Σ nonempty admits a triangulation of this type. Thus the charts $\mathcal{U}_\tau$ cover $\mathcal{H}_{\mathrm{m}}$ as τ ranges over triangulations of S. The change of coordinate maps for this system of charts are linear. These charts on the sets $\mathcal{U}_\tau$ give

$\mathcal{H}_{\mathrm{m}}$ an affine manifold structure. This affine manifold structure can be discussed using the terminology of (G, X)-structures. Specifically we can take X to be $H^1(S, \Sigma; \mathbb{R}^2)$ and the structure group G can be taken to be the group of linear automorphisms of $H^1(S, \Sigma; \mathbb{R}^2)$. Our choice of dev as notation for the map in (1) is motivated by the fact that this map is the developing map for the affine structure.

We have described triangulations of (S, Σ) in terms of a homeomorphism with a Δ-complex. We note that such a homeomorphism is determined up to isotopy by knowing the relative homotopy classes of the edges of the triangulation (see [FaMa, Lemma 2.9]). Furthermore distinct edges are not homotopic to each other relative to their endpoints. In the sequel we will use τ to denote the homotopy classes of edges of a triangulation.

We now make use of the marked stratum to put a topology on the stratum. Consider the group of isotopy classes of orientation-preserving homeomorphisms of S fixing Σ pointwise. This is sometimes called the pure mapping class group (see e.g. [I]). We will simply refer to it as the mapping class group and we will denote it by $\mathrm{Mod}(S, \Sigma)$. Up to the action of $\mathrm{Mod}(S, \Sigma)$ there are finitely many charts $\mathcal{U}_\tau$. It can be checked that the $\mathrm{Mod}(S, \Sigma)$-action on $\mathcal{H}_{\mathrm{m}}$ is properly discontinuous. This equips the quotient $\mathcal{H} = \mathcal{H}_{\mathrm{m}} / \mathrm{Mod}(S, \Sigma)$ with the structure of an affine orbifold, with respect to which the map $\mathcal{H}_{\mathrm{m}} \to \mathcal{H}$ is an orbifold covering map (see [Th] for the definition and basic properties of properly discontinuous actions and orbifolds).

2.5. **Blow ups of translation surfaces.** We now discuss a "blow-up" construction that replaces a singularity ξ by a boundary circle. This is a special case of a more general construction of a "real oriented blow-up" (see e.g. [HPV]). The real oriented blowup of a point ξ in $\mathbb{R}^2$ is a new space $\mathrm{Blo}_\xi(\mathbb{R}^2)$ together with a collapsing map $c : \mathrm{Blo}_\xi(\mathbb{R}^2) \to \mathbb{R}^2$ with the property that the inverse image of any point other than ξ is a single point while the inverse image of ξ is the circle of directions at ξ which we can identify with $(\mathbb{R}^2 \smallsetminus \{0\})/\mathbb{R}^+$ (or with the unit circle) and denote by S^1. The space $\mathrm{Blo}_\xi(\mathbb{R}^2)$ has the property that a smooth path in $\mathbb{R}^2$ landing at ξ and with non-zero derivative at ξ has a lift to the space $\mathrm{Blo}_\xi(\mathbb{R}^2)$ which takes the endpoint of the path to a point in the circle of directions.

If we blow up the vertex ψ of a polygon P in $\mathbb{R}^2$ then we obtain the space $\mathrm{Blo}_\psi(P)$ which is the result of replacing the vertex ψ in P by an interval. We describe this construction explicitly. Applying a translation, assume ψ is at the origin. Say that the two edges of the polygon incident to ψ are in directions θ_1, θ_2 with $0 < \theta_2 - \theta_1 < 2\pi$, and there is $\varepsilon > 0$ such that

$$\{r(\cos\theta, \sin\theta) : \theta \in [\theta_1, \theta_2],\ r \in [0, \varepsilon)\}$$

parametrizes a neighborhood of ψ in P. The blow-up of this neighborhood in $\mathrm{Blo}_\psi(P)$ corresponds to the rectangle

$$\{(r, \theta) : \theta \in [\theta_1, \theta_2],\ r \in [0, \varepsilon]\}$$

and the collapsing map c from $\mathrm{Blo}_\psi(P)$ to P takes (r, θ) to $r(\cos\theta, \sin\theta)$. In particular the interval $\{(0, \theta) : \theta \in [\theta_1, \theta_2]\}$ is collapsed to the point ψ. Note that this interval has a natural "angular coordinate" with values in the unit circle.

Given a translation surface M with singular points Σ we construct a surface with boundary $\check{M}$ by blowing up all of the points of Σ. We can do this as follows. Choose a triangulation of M. Blow up each vertex of each triangle thereby creating a family of hexagons where each hexagon contains edges of two types: those corresponding to edges of triangles and those corresponding to vertices of triangles. Glue the (triangle edge) sides of the hexagons together according to the gluing pattern of the original triangles. The result is the surface $\check{M}$ which is in fact independent of the particular choice of triangulation. At a singular point ξ_j of the surface the intervals mapping to ξ_j glue together to form a circle which we call $\partial_j \check{M}$. The angular coordinates glue together to give us a map $p_j : \partial_j \check{M} \to S^1$. The total angular measure of $\partial_j \check{M}$ is $2\pi(a_j + 1)$ which is the cone angle at ξ_j. We can choose an identification of the circle $\partial_j \check{M}$ with the circle $\mathbb{R}/(2\pi(a_j + 1))\mathbb{Z}$ so that the angular coordinate of a point is equal to its circle coordinate modulo 2π. This identification of $\partial_j \check{M}$ with the circle is well-defined up to translation by 2π while the map p_j is defined independently of any choices.

If we associate a point ν in the j-th boundary component $\partial_j \check{M}$ with a short ray heading away from ξ_j, then $p_j(\nu)$ is the direction of the ray. We define the *prongs* to be points on boundary circles corresponding to horizontal rays, i.e. points whose angular parameter is πk with $k \in \mathbb{Z}$. We will call a prong *right-* or *left-pointing*, if k is even (resp. odd), that is, according to the orientation that the prong inherits from the plane. A particular choice of an identification of $\partial_j \check{M}$ with $\mathbb{R}/2\pi(a_j + 1)\mathbb{Z}$ is equivalent to the choice of a right-pointing prong. The boundary circles of $\check{M}$ inherit boundary orientations as boundary components of the oriented manifold M. With respect to the boundary orientation the maps p_j are covering maps of degree $-(a_j + 1)$ where the negative sign reflects the fact that moving in the direction of the boundary orientation corresponds to *decreasing* the angular coordinate.

2.6. **Strata of boundary marked surfaces.** In this subsection we define a notion of marked surface appropriate to surfaces with boundary and a corresponding mapping class group. To this end we define a "model surface" $\check{S}$, which will capture some of the structure common to the surfaces $\check{M}$ for $M \in \mathcal{H}(a_1, \ldots, a_k)$. Let $\check{S}$ be a surface with boundary which has genus g where $2g - 2 = \sum a_j$ and k boundary components labeled $\partial_1 \check{S}, \ldots, \partial_k \check{S}$. The circles inherit a boundary orientation from S. We equip each boundary circle $\partial_j \check{S}$ with orientation reversing homeomorphisms $q_j : \partial_j \check{S} \to \mathbb{R}/(2\pi(a_j + 1)\mathbb{Z})$ which give angular coordinates on the boundaries. A *marked translation surface rel boundary* is a surface $\check{M}$ which is a blow up of a translation surface M of type $(a_1, \ldots, a_k)$, equipped with an orientation preserving homeomorphism $\check{f} : \check{S} \to \check{M}$ respecting the labels, and such that on each boundary circle $\partial_j \check{S}$ we have $p_j \circ \check{f} \equiv q_j \bmod 2\pi$. Note that a boundary marking of M induces an explicit coordinate on $\partial \check{M}$ and an explicit choice of a prong, namely the image under $\check{f}$ of the prong on $\partial_j \check{S}$ corresponding to angular parameter zero.

We say that two boundary marked translation surfaces rel boundary $(\check{f}_1, \check{M}_1)$ and $(\check{f}_2, \check{M}_2)$ are *equivalent* if there is a translation equivalence $g : M_1 \to M_2$ such that $\check{g} \circ \check{f}_1$ and $\check{f}_2$ are equal on $\partial \check{S}$ and are isotopic via an isotopy that fixes the boundary. Let $\widetilde{\mathcal{H}} = \widetilde{\mathcal{H}}(a_1, \ldots, a_k)$ denote the set of boundary marked translation surfaces, up to equivalence. We call $\widetilde{\mathcal{H}}$ a stratum of boundary marked translation surfaces.

There is a natural collapsing map $c : \check{M} \to M$ which collapses each boundary component to a single point. A map $\check{f} : \check{S} \to \check{M}$ induces a map $f : S \to M$, where $c \circ \check{f} = f \circ c$. If $\check{f}_1$ and $\check{f}_2$ are equivalent (in the sense of §2.6), then so are f_1 and f_2 (in the sense of §2.2). We say that f is *obtained from* $\check{f}$ *by projection.* The forgetful map which takes $(\check{f}, \check{M})$ to the pair (f, M) where f is obtained by projection gives a map $\mathrm{pr} : \widetilde{\mathcal{H}} \to \mathcal{H}_{\mathrm{m}}$.

We now define a *mapping class group rel boundary.* We say a homeomorphism $\check{f} : \check{S} \to \check{S}$ is *admissible* if $\check{f}$ takes each boundary circle $\partial_j \check{S}$ to itself and the restriction of $\check{f}$ to $\partial_j \check{S}$ is a rotation by a multiple of 2π with respect to the circle coordinate given by q_j. We say that two admissible homeomorphisms are *isotopic rel boundary* if they agree on $\partial \check{S}$ and if the are isotopic by an isotopy which is the identity when restricted to $\partial \check{S}$. We denote by $\mathrm{Mod}(\check{S}, a_1, \ldots, a_k)$ the group of isotopy classes rel boundary of admissible homeomorphisms of $\check{S}$. When there is no chance of confusion we will abbreviate this to $\mathrm{Mod}(\check{S}, \partial \check{S})$. There is a natural right action of the mapping class group $\mathrm{Mod}(\check{S}, a_1, \ldots, a_k)$ on the stratum $\widetilde{\mathcal{H}}(a_1, \ldots, a_k)$ by precomposition.

2.7. **Relative homotopy classes of paths.** A useful tool in the study of mapping class groups is the study of their action on curves. In order to analyze $\widetilde{\mathcal{H}}$ and $\mathrm{Mod}(\check{S}, \partial\check{S})$ we will consider the action on homotopy classes of paths with endpoints in the boundary. We consider paths in $\check{S}$ with endpoints in $\partial\check{S}$. We say that two paths are *homotopic* if there is a homotopy between them that keeps the endpoints in $\partial\check{S}$. We say that two homotopic paths are *relatively homotopic* if the homotopy can be chosen to fix the endpoints of the path. The collection of relative homotopy classes of paths is a natural set to consider but, unlike homotopy classes of paths in (M, Σ), it is not discrete. We analyze its topology below. There is a well-defined action of $\mathrm{Mod}(\check{S}, \partial\check{S})$ on relative homotopy classes of paths. We say that a path α is *peripheral* if both endpoints lie in the same boundary component and α is homotopic to a path contained in that boundary component. Let $\sigma : [0,1] \to \check{S}$ be a non-peripheral oriented path with $\sigma(0) \in \partial\check{S}_j$ and $\sigma(1) \in \partial\check{S}_k$. Let $\mathcal{C}_\sigma(\check{S})$ be the set of relative homotopy classes of oriented paths in $\check{S}$ which are homotopic to σ.

Definition 2.2. . *Let $\check{\sigma}$ and $\check{\sigma}'$ be elements of $\mathcal{C}_\sigma(\check{S})$ going from $\partial_j\check{S}$ to $\partial_k\check{S}$, and let $\varepsilon > 0$. We say that $\check{\sigma}'$ is* ε-close to $\check{\sigma}$ *if there are intervals $I_1 \subset \partial_j\check{S}$ and $I_2 \subset \partial_k\check{S}$ of length ε, containing the endpoints of $\check{\sigma}$ and $\check{\sigma}'$, such that $\check{\sigma}'$ is homotopic to $\check{\sigma}$ through a family of paths each of which has one endpoint in I_1 and one endpoint in I_2.*

Sets of ε-close homotopy classes of curves for given intervals I_1 and I_2 form the basis for a topology on $\mathcal{C}_\sigma(\check{S})$. We have an endpoint map $\epsilon : \mathcal{C}_\sigma(\check{S}) \to \partial_j\check{S} \times \partial_k\check{S}$ which takes a relative homotopy class of curves to its endpoints. With respect to the topology on $\mathcal{C}_\sigma(\check{S})$ the endpoint map is continuous.

Lemma 2.3. *Say that S is a surface with boundary with negative Euler characteristic. The endpoint map $\epsilon : \mathcal{C}_\sigma(\check{S}) \to \partial_j\check{S} \times \partial_k\check{S}$ is a covering map and $\mathcal{C}_\sigma(\check{S})$ is the universal cover of $\partial_j\check{S} \times \partial_k\check{S}$. It follows that the space of relative homotopy classes of paths in this homotopy class is homeomorphic to the product of the universal covers $\widetilde{\partial_j\check{S}}$ and $\widetilde{\partial_k\check{S}}$.*

We will apply this result to the surfaces with boundary $\check{S}$ arising as 'model surfaces' corresponding to some family of translation surfaces. According to our conventions these surfaces always have negative Euler characteristic. This result captures the idea that the distinction between homotopy and relative homotopy for paths is measured by the amount of twisting around each boundary component.

Proof. Since the Euler characteristic of $\check{S}$ is negative we can give $\check{S}$ a hyperbolic structure so that the boundaries are geodesics. The universal cover $\tilde{\check{S}}$ is isometric to a convex subset of the hyperbolic plane with geodesic boundary. Choose a boundary component B_1 of $\tilde{\check{S}}$ corresponding to $\partial_j \check{S}$. Identify B_1 with $\widetilde{\partial_j \check{S}}$. Choose a lift of σ to a path $\tilde{\sigma}$ in $\tilde{\check{S}}$ starting in B_1. The other endpoint of $\tilde{\sigma}$ lands in a component B_2 which maps to $\partial_k \check{S}$. Identify B_2 with $\widetilde{\partial_k \check{S}}$. Since the path is non-peripheral, B_1 and B_2 are distinct. Any path homotopic to σ has a lift to a path from B_1 to B_2 and this lift is unique since the subgroup of the deck group that stabilizes B_1 and B_2 is trivial. This follows from the fact that a hyperbolic isometry that fixes four points is the identity.

We get a map from the space of relative homotopy classes of paths homotopic to σ to $B_1 \times B_2 = \widetilde{\partial_j \check{S}} \times \widetilde{\partial_k \check{S}}$ as follows. Given a path homotopic to σ we lift to a path from $B_1 \times B_2$ and we associate this path to its endpoints. Given a pair of points $(p_1, p_2) \in B_1 \times B_2$ we associate the projection to $\check{S}$ of the unique geodesic from p_1 to p_2. The fact that any path between $\widetilde{\partial_j \check{S}}$ and $\widetilde{\partial_k \check{S}}$ is relatively homotopic to a unique geodesic implies that these maps are inverses. This map is a continuous bijection with respect to the natural topology on $\mathcal{C}_\sigma(\check{S})$. $\square$

We now describe explicitly some elements of $\mathrm{Mod}(\check{S}, \partial\check{S})$ which correspond to partial Dehn twists around boundary components. Let A_j be an annular neighborhood of $\partial_j \check{S}$ where we choose coordinates $\{(t, \theta) : t \in [0,1],\, \theta \in \mathbb{R}/2\pi(a_j+1)\mathbb{Z}\}$. Here $t = 0$ corresponds to the boundary circle $\partial_j \check{S}$, and the θ coordinate of A_j is compatible at $t = 0$ with the θ coordinate of the boundary circle. We will define a particular homeomorphism $\tau_j \in \mathrm{Mod}(\check{S}, \partial\check{S})$ as follows. On A_j we define $\tau_j(t, \theta) = (t, \theta + 2\pi(1-t))$ so that τ_j rotates $\partial_j \check{S}$ by 2π and is the identity on the other boundary of A_j. We extend τ_j to a map $\tau_j : \check{S} \to \check{S}$ by setting it to be the identity outside of A_j. The map τ_j represents an element of $\mathrm{Mod}(\check{S}, \partial\check{S})$ which we call a *fractional Dehn twist* by angle 2π. In particular $\tau_j^{a_j+1}$ is a full Dehn twist around the boundary curve $\partial_j \check{S}$. Note that $\tau_j^{a_j+1}$ is a right Dehn twist and that it also makes sense to describe τ_j as a right fractional Dehn twist (see [FaMa]). The collapsing map $c : \check{S} \to S$ induces a map $c_* : \mathrm{Mod}(\check{S}, \partial\check{S}) \to \mathrm{Mod}(S, \Sigma)$. Denote by FT the group generated by the fractional Dehn twists $\tau_1, \ldots, \tau_k$. Note that FT depends on $(a_1, \ldots, a_k)$.

Lemma 2.4. *We have a short exact sequence*

$$1 \to FT \to \mathrm{Mod}(\check{S}, \partial\check{S}) \xrightarrow{c_*} \mathrm{Mod}(S, \Sigma) \to 1, \tag{2}$$

where the group FT is the free Abelian group generated by the τ_j and is central in $\mathrm{Mod}(\check{S}, \partial\check{S})$.

Proof. We can see from the definition of a fractional Dehn twist that an element of FT is isotopic to the identity by an isotopy which moves points in the boundary of S. These isotopies descend to isotopies of (S, Σ). It follows that FT belongs to the kernel of c_*. The fact that the kernel of c_* is exactly FT follows from the arguments used in [FaMa, Prop. 3.20].

To see the surjectivity of c_*, let h be a homeomorphism of S fixing points of Σ. The homeomorphism h is isotopic to a diffeomorphism (see [FaMa, Thm. 1.13]) which, using the properties of the real blow-up, has a lift to a homeomorphism h' from $\check{S}$ to itself. This homeomorphism induces a homeomorphism of $\partial_j\check{S}$ for each j. By applying an isotopy in the annular neighborhood A_j of each $\partial_j S$ as above, we can replace h' by a map h'', which is the identity on A_j. In particular h and h'' represent the same element of $\mathrm{Mod}(S, \Sigma)$, and h'' fixes each point of $\partial\check{S}$. This shows that the equivalence class of h contains a representative which is in the image of c_*, proving surjectivity.

Since h'' is the identity on each A_j and τ_j is the identity on the complement of A_j we see that h'' and τ_j have disjoint support so they commute which implies that FT is central.

We now show that there are no relations between elements of FT. Consider a word in the collection of twists that represents a relation. Since FT is abelian we can write it as $w = \tau_1^{m_1} \cdots \tau_j^{m_j} \cdots \tau_k^{m_k}$. Now since $\check{S}$ has negative Euler characteristic we can find a non-peripheral path from a boundary component $\partial_j\check{S}$ to itself. The effect of w on this path is to shift both endpoints by $2\pi m_j$. By Lemma 2.3, since w acts trivially, $m_j = 0$. Since j was arbitrary, $w = 1$. □

2.8. **Strata of boundary marked surfaces as topological spaces.** In this subsection we define a topology on $\widetilde{\mathcal{H}}$ in a manner somewhat analogous to the method used for defining the topology on $\mathcal{H}_{\mathrm{m}}$. We construct a cover of $\widetilde{\mathcal{H}}$ by sets on which the developing map is an injection into $H^1(S, \Sigma; \mathbb{R}^2)$ and then we use these maps to endow each such set with the topology induced by the developing map. We show that with respect to this topology the map pr is a covering map, thus we conclude that these charts give not only a topology on $\widetilde{\mathcal{H}}$ but a compatible affine structure.

To construct these charts fix a point $(\check{f}, \check{M}) \in \widetilde{\mathcal{H}}$ and a geodesic triangulation τ of M. We caution the reader that in contrast to §2.4, here τ denotes a geodesic triangulation of M rather than a topological triangulation of S. We have canonical lifts of the edges σ of the triangulation τ to edges $\check{\sigma}$ in $\check{M}$ so that the endpoints of $\check{\sigma}$ lie in $\partial\check{M}$. Let $\check{\tau}$ be the collection of paths in $\check{S}$ of the form $\check{f}^{-1}(\check{\sigma})$ for σ an edge of τ. These paths are embedded and we refer to them as arcs. These arcs of $\check{\tau}$ decompose $\check{S}$ into hexagons where edges of the hexagons consist alternately of arcs and intervals in boundary circles.

We now define what it means for two hexagon decompositions to be *ε-close*. Firstly we require that the arcs in the two decompositions are pairwise homotopic. Secondly we require that these pairs of homotopic arcs are ε-close in the sense of Definition 2.2. Note that any triangulation of a translation surface has the property that distinct edges lie in distinct homotopy classes, thus there is no ambiguity in comparing homotopy classes of arcs in the two decompositions.

Given a geodesic triangulation τ of M, let $\check{\tau}$ be its pullback under a marking of blown up translation surfaces $\check{S} \to \check{M}$, and let $\mathcal{U}_{\check{\tau},\varepsilon}$ consist of $(\check{f}, \check{M}') \in \widetilde{\mathcal{H}}$ for which there is a geodesic triangulation σ of M' which lifts to a hexagon decomposition of $\check{\sigma}$ of $\check{M}'$ so that the pullback of $\check{\sigma}$ under $\check{f}'$ is ε-close to $\check{\tau}$.

Lemma 2.5. *The developing map is injective on the set $\mathcal{U}_{\check{\tau},\pi/2}$.*

Remark: The developing map for $\widetilde{\mathcal{H}}$ is most naturally defined to take values in $H^1(S, \partial\check{S}; \mathbb{R}^2)$ but the the collapsing map $c : \check{S} \to S$ induces a map $c^* : H^1(S, \Sigma; \mathbb{R}^2) \to H^1(\check{S}, \partial\check{S}; \mathbb{R}^2)$ which is an isomorphism. In the sequel we will make use of this isomorphism to identify the two spaces.

Proof. Assume that $(\check{f}_1, \check{M}_1)$ and $(\check{f}_2, \check{M}_2)$ map to the same point in $H^1(S, \Sigma; \mathbb{R}^2)$ and both lie in $\mathcal{U}_{\check{\tau},\pi/2}$. Let τ_1 and τ_2 be triangulations of M_1 and M_2 so that $\check{\tau}_1 = f_1^*(\check{\tau}_1)$ and $\check{\tau}_2 = {f_2}^*(\check{\tau}_2)$ are $\pi/2$-close to $\check{\tau}$, and hence π-close to each other. Since M_1 and M_2 have geodesic triangulations such that corresponding edges have the same image in $H^1(S, \Sigma; \mathbb{R}^2)$, the comparison map $F(\tau, M_1, M_2)$ is a translation equivalence, and we denote it by $g : M_1 \to M_2$. In order to show that $(\check{f}_1, \check{M}_1)$ and $(\check{f}_2, \check{M}_2)$ represent the same element of $\widetilde{\mathcal{H}}$ we need to show that $\check{g} \circ \check{f}_1$ and $\check{f}_2$ agree on $\partial\check{S}$ and that they are isotopic via an isotopy which fixes $\partial\check{S}$. Both $\check{\tau}_1$ and $\check{\tau}_2$ produce collections of arcs in $\check{S}$. There is a unique correspondence between these collections of arcs so that corresponding arcs are homotopic. We want to show that corresponding arcs are not just homotopic but in fact relatively homotopic.

We begin by working with a single edge. Let σ be an edge of $\check{\tau}$ and let σ_1 and σ_2 be corresponding oriented edges of $\check{\tau}_1$ and $\check{\tau}_2$. Our first objective is to show that σ_1 and σ_2 have the same endpoints and are relatively homotopic. The relative homotopy classes of σ_1 and σ_2 determine points $[\sigma_1]$ and $[\sigma_2]$ in $\mathcal{C}_\sigma(\check{S})$. It suffices to show that these two points are the same.

Say that σ_1 and σ_2 run from $\partial_i\check{S}$ to $\partial_j\check{S}$. We have the endpoint map $\epsilon : \mathcal{C}_\sigma(\check{S}) \to \partial_i\check{S} \times \partial_j\check{S}$. According to Lemma 2.3 this is a covering map. We also have projection maps $p_k : \partial_k\check{S} \to \mathbb{R}/2\pi\mathbb{Z}$ which are covering maps (in either S or M). Consider the composition $\Pi = (p_i \times p_j) \circ \epsilon : \mathcal{C}_\sigma(\check{S}) \to \mathbb{R}/2\pi\mathbb{Z} \times \mathbb{R}/2\pi\mathbb{Z}$. This is a covering map. The oriented segments σ_1 and σ_2 have the same holonomy so they point in the same direction in $\mathbb{R}^2$ hence $\Pi([\sigma_1]) = \Pi([\sigma_2])$. Since both triangulations lie in $\mathcal{U}_{\check{\tau},\pi/2}$ there are intervals $I \subset \partial_i\check{S}$ and $J \subset \partial_j\check{S}$ of length π and a homotopy from σ_1 to σ_2 for which the endpoints of the paths remain in I and J. This homotopy gives a path ρ from $[\sigma_1]$ to $[\sigma_2]$ in $\mathcal{C}_\sigma(\check{S})$ which projects to $I \times J \subset \mathbb{R}/2\pi\mathbb{Z} \times \mathbb{R}/2\pi\mathbb{Z}$ under the covering map Π. The image of ρ is a loop, which is contractible since it is contained in the contractible set $I \times J$. Covering theory implies that ρ itself must be a loop so $[\sigma_1] = [\sigma_2]$. It follows that $\check{g} \circ \check{f}_1(\sigma_1)$ and $\check{f}_2(\sigma_2)$ are relatively homotopic.

Proposition 2.1 shows that $g \circ \check{f}_1$ and $\check{f}_2$ induce the isotopic maps on (S, Σ). It remains to show that they agree as maps on $(\check{S}, \partial\check{S})$. According to Lemma 2.4 this means that there is no boundary twisting but boundary twisting is detected by the effect on relative homotopy classes of the paths σ. We conclude that $g \circ \check{f}_1$ and $\check{f}_2$ are isotopic relative to their boundaries. $\square$

We can use the injectivity of the developing maps on the sets $\mathcal{U}_{\check{\tau},\pi/2}$ to define a topology on $\widetilde{\mathcal{H}}$. With respect to this topology the developing map becomes a local homeomorphism. The next proposition refers to this topology.

Proposition 2.6. *The projection map* pr *from $\widetilde{\mathcal{H}}$ to $\mathcal{H}_{\mathrm{m}}$ has the path lifting property.*

Recall that the path lifting property for the map pr means that if we are given a path $f : I \to \mathcal{H}_{\mathrm{m}}$ and a lift $\tilde{x}_0$ of the endpoint $f(0) = x_0$ then there is a unique path $\tilde{f} : I \to \tilde{X}$ with $\tilde{f}(0) = \tilde{x}_0$ which satisfies $\mathrm{pr} \circ \tilde{f} = f$ (see [H, p. 60] for more information).

Proof. It suffices to construct lifts of paths locally. Consider an open set $\mathcal{U}_\tau \subset \mathcal{H}_{\mathrm{m}}$ corresponding to a triangulation τ as in §2.4. Let $\phi_t = (f_t, M_t)$ for $t \in [0, 1]$ be a path taking values in $\mathcal{U}_\tau$ and let $(\check{f}_0, \check{M}_0)$ represent a point in $\widetilde{\mathcal{H}}$ such that $\mathrm{pr}(\check{f}_0, \check{M}_0) = (f_0, M_0)$. We define a path $\check{\phi}_t = (\check{f}_t, \check{M}_t)$ for $t \in [0,1]$, satisfying

$$\check{\phi}_0 = (\check{f}_0, \check{M}_0) \text{ and } \mathrm{pr}(\check{\phi}_t) = \phi_t, \tag{3}$$

as follows.

Let $F(\tau, M_0, M_t) : M_0 \to M_t$ be the comparison maps defined in §2.4. These maps are piecewise linear and hence they have unique extensions to the blow-ups, that is there are homeomorphisms $\check{F}(\tau, M_0, M_t) : \check{M}_0 \to \check{M}_t$ of the blown-up surfaces, satisfying

$$c \circ \check{F}(\tau, M_0, M_t) = F(\tau, M_0, M_t) \circ c.$$

We define $\check{F}_t = \check{F}(\tau, M_0, M_t) \circ \check{f}_0 : \check{S} \to \check{M}_t$, and denote the restriction of $\check{F}_t$ to $\partial_j \check{S}$ by $\partial_j \check{F}_t$. We would like to set $\check{\phi}_t = (\check{F}_t, \check{M}_t)$ but there is a problem in that the maps $\check{F}_t$ need not preserve the boundary coordinates given by the maps $p_j : \partial_j \check{M}_t \to S^1$. In other words, it will not generally be the case that $p_j \circ \partial_j \check{F}_t = p_j$. We claim that there is a unique (up to isotopy) way to modify the maps $\check{F}_t$ by precomposition with a continuous family of maps of $\check{S}$, so that they do satisfy the boundary coordinate condition, and so that (3) holds. We will do this by precomposing the maps $\check{F}_t$ with homeomorphisms $H_t : \check{S} \to \check{S}$ supported in a neighborhood of the boundary and then set $\check{f}_t = \check{F}_t \circ H_t$.

In order to prove the existence of the required homeomorphisms H_t, we consider the condition that they will need to satisfy. For each boundary component $\partial_j \check{S}$ the restriction $\partial_j H_t$ should satisfy

$$p_j \circ \partial_j \check{F}_t \circ \partial_j H_t = p_j \text{ and } \partial_j H_0 = \mathrm{Id}. \tag{4}$$

We rewrite this as

$$p_j \circ \partial_j \check{F}_t = p_j \circ (\partial_j H_t)^{-1}. \tag{5}$$

Setting $\ell_{j,t} = (\partial_j H_t)^{-1}$, we see that $\ell_{j,t} : \partial_j \check{S} \to \partial_j \check{S}$ is a solution to the homotopy lifting problem

$$p_j \circ \ell_{j,t} = p_j \circ \partial_j \check{F}_t \quad \text{and} \quad \ell_{j,0} = \mathrm{Id},$$

see Figure 1.

Since p_j is a covering map the Homotopy Lifting Theorem [H, Prop 1.30] asserts that lifts $\ell_{j,t}$ exist and are unique. Thus $H_t = (\ell_{j,t})^{-1}$ is defined on the boundary of $\check{S}$. It remains to extend H_t to annular neighborhoods of the boundaries. As in §2.6 we have a family of disjoint annuli A_j in $\check{S}$ parametrized by $\{(r, \theta) : r \in [0,1], \theta \in \mathbb{R}/2\pi(a_j+1)\mathbb{Z}\}$ where in these coordinates, $\{r = 0\}$ is the j-th boundary component of $\check{S}$. Then we define H_t to be the identity outside the union of annuli, to be equal to the prescribed map $\partial_j H_t$ on $\{r = 0\}$, and be given by the formula $H_t(r, \theta) = (r, \psi_{(1-r)t}(\theta))$ for $0 \leqslant r \leqslant 1$.

$$\begin{tikzcd}
& \partial_j \check{S} \arrow[d, "p_j"] \\
\partial_j \check{S} \arrow[ur, dashed, "\ell_{j,t}"] \arrow[r, "p_j \circ \partial_j \check{F}_t"] & S^1
\end{tikzcd}$$

FIGURE 1. The boundary map $\ell_{j,t} = (\partial_j H_t)^{-1}$ rectifies the discrepancies of $\partial_j \hat{F}_t$ along boundary circles.

It is clear that with these definitions, the path $\check{\phi}_t = (\check{f}_t, \check{M}_t)$ with $\check{f}_t = \check{F}_t \circ H_t$ satisfies (3).

It is not hard to verify the homotopy invariance of the lift, that is, if we have two homotopic paths ϕ_t and ϕ'_t (for $t \in [0,1]$), one can lift the homotopy to obtain a one-parameter family of maps $\check{f}_1^{\,s} : \check{S} \to \check{M}_1$, which gives us an isotopy from $\check{f}_1^0$ to $\check{f}_1^1$, fixing the boundary. That is, the lifts of homotopic paths have the same endpoint in $\widetilde{\mathcal{H}}$.

The uniqueness (up to isotopy) of the maps H_t (where H_0 is the identity and (4) holds), follows from the uniqueness of the lifts $\ell_{j,t}$ and standard properties of annuli, and is left as an exercise. □

Corollary 2.7. *The map* $\mathrm{pr} : \widetilde{\mathcal{H}} \to \mathcal{H}_{\mathrm{m}}$ *is a covering map and* $\widetilde{\mathcal{H}}$ *has an affine structure given by period coordinates.*

Proof. A local homeomorphism to a semilocally simply connected space with the path lifting property is a covering map. □

The group $\mathrm{Mod}(\check{S}, \partial\check{S})$ acts on $\widetilde{\mathcal{H}}$ by precomposition. It acts continuously and properly discontinuously on $\widetilde{\mathcal{H}}$, the quotient is $\mathcal{H}$ and according to Lemma 2.4 the subgroup FT acts simply transitively on each fiber of pr. We warn the reader that the spaces $\widetilde{\mathcal{H}}$ and $\mathcal{H}_{\mathrm{m}}$ are not in general connected. Typically they have infinitely many components.

2.9. **The space of framed surfaces.** Our next objective is to define and analyze the covering space of framed surfaces. Since $\mathcal{H}_{\mathrm{m}}$ is an affine manifold and $\mathrm{pr} : \widetilde{\mathcal{H}} \to \mathcal{H}_{\mathrm{m}}$ is a covering map, we have equipped $\widetilde{\mathcal{H}}$ with the structure of an affine manifold. Since the action of $\mathrm{Mod}(\check{S}, \partial\check{S})$ is properly discontinuous, for each subgroup Γ of $\mathrm{Mod}(\check{S}, \partial\check{S})$, we can form the quotient $\widetilde{\mathcal{H}}/\Gamma$. By Lemma 2.4 we have $\mathcal{H} = \mathcal{H}/\mathrm{Mod}(\check{S}, \partial\check{S})$ and $\mathcal{H}_{\mathrm{m}} = \mathcal{H}/FT$. Moreover each $\widetilde{\mathcal{H}}/\Gamma$ is an orbifold cover of $\mathcal{H}$. Note that this is not the Galois correspondence relating connected covers to subgroups of the fundamental group, since $\widetilde{\mathcal{H}}$ is not connected. Nevertheless we can define the space of framed surfaces via a group-theoretic approach.

We start with a discussion of subgroups of $\mathrm{Mod}(\check{S}, \partial\check{S})$. While an element of the group $\mathrm{Mod}(\check{S}, \partial\check{S})$ is only defined up to isotopy on the interior of $\check{S}$, it is well-defined on the boundary circles and acts on each circle $\partial_j\check{S}$ by rotations which are multiples of 2π. Let $\mathrm{Mod}(\check{S})$ be the subgroup of $\mathrm{Mod}(\check{S}, \partial\check{S})$ represented by homeomorphisms that fix the boundary pointwise. Let PR be the *prong rotation group*, consisting of homeomorphisms of the boundary $\partial\check{S}$ which, on each boundary component $\partial_j\check{S}$, are rotations by an integral multiple of 2π. Since $\partial_j\check{S}$ is parametrized by a circle of length $2\pi(a_j+1)$, as a group PR is isomorphic to $\prod_{j=1}^{k}\mathbb{Z}/(a_j+1)\mathbb{Z}$. We have a short exact sequence

$$1 \to \mathrm{Mod}(\check{S}) \to \mathrm{Mod}(\check{S}, \partial\check{S}) \to PR \to 1. \tag{6}$$

Surjectivity in (6) follows from the fact that the fractional twists $\tau_j \in \mathrm{Mod}(\check{S}, \partial\check{S})$ map to a collection of generators for PR.

A *framed* translation surface is a translation surface M equipped with a right-pointing horizontal prong at each singular point. We will call this prong (considered as an element of $\partial_j\check{M}$) the *selected prong*. Equivalently a framed surface is equipped with a choice of boundary coordinate on C_j taking values in the circle $\mathbb{R}/2\pi(a_j+1)\mathbb{Z}$ so that the map p_j is reduction modulo 2π and the selected prong corresponds to the angle 0.

The space of framed translation surfaces is naturally a finite cover of $\mathcal{H}$ and we will denote it by $\mathcal{H}_{\mathrm{f}}$. The connected components of $\mathcal{H}_{\mathrm{f}}$ were classified by Boissy [Boi]. We recover $\mathcal{H}_{\mathrm{f}}$ as the quotient of $\widetilde{\mathcal{H}}$ by the group $\mathrm{Mod}(\check{S})$ in (6):

Proposition 2.8. *We have $\widetilde{\mathcal{H}}/\mathrm{Mod}(\check{S}) = \mathcal{H}_{\mathrm{f}}$.*

Proof. We define a map from $\widetilde{\mathcal{H}}$ to $\mathcal{H}_{\mathrm{f}}$ as follows. For each boundary component $\partial_j\check{S}$, let 0_j denote the point in $\partial_j\check{S}$ with angular coordinate 0. Given a marked blown-up surface $(\check{f}, \check{M})$, define a framed surface by letting the point $\check{f}(0_j) \in \partial_j\check{M}$ be the selected prong. Say that two surfaces $(\check{f}_1, \check{M}_1)$ and $(\check{f}_2, \check{M}_2)$ map to the same surface in $\mathcal{H}_{\mathrm{f}}$. Thus there is a translation equivalence $h : M_1 \to M_2$ so that $\check{h}$ takes selected prongs in each $\partial_j\check{M}_1$ to selected prongs in $\partial_j\check{M}_2$. This means that the restriction of the map $(\check{f}_2)^{-1} \circ \check{h} \circ \check{f}_1$ to $\partial_j\check{S}$ fixes 0_j. Since this map is a rotation of the circle with a fixed point it is the identity map. Thus $(\check{f}_2)^{-1} \circ \check{h} \circ \check{f}_1 \in \mathrm{Mod}(\check{S})$ so $(\check{f}_1, \check{M}_1)$ and $(\check{f}_2, \check{M}_2)$ are $\mathrm{Mod}(\check{S})$-equivalent. $\square$

We summarize our constructions in Figure 2.

$$\begin{tikzcd}
 & \widetilde{\mathcal{H}} \arrow[dl] \arrow[dr] \arrow[dd] & \\
\mathcal{H}_{\mathrm{m}} \arrow[dr] & & \mathcal{H}_{\mathrm{f}} \arrow[dl] \\
 & \mathcal{H} &
\end{tikzcd}$$

FIGURE 2. The vertical arrow corresponds to factoring by the action of $\mathrm{Mod}(\check{S}, \partial\check{S})$, and the pair of arrows on the left and right correspond to factoring by the groups appearing in the sequences (2) and (6) respectively.

2.10. **Action of G and its covers on covering spaces of $\mathcal{H}$.** The affine equivalence classes of translation structures are orbits of a group actions which we now define. Recall that $\mathrm{GL}_2^\circ(\mathbb{R})$ and $\widetilde{\mathrm{GL}_2^\circ(\mathbb{R})}$ denote respectively the group of orientation preserving invertible 2×2 real matrices, and its universal cover group. Given $g \in \mathrm{GL}_2^\circ(\mathbb{R})$ and a translation surface M we construct a new surface gM as follows. As discussed in §2.1, in the language of (G, X)-structures, a translation surface can be given by an atlas of charts on M with overlap functions taking values in the group of translations $\mathbb{R}^2$. The element g is a linear map $\mathbb{R}^2 \to \mathbb{R}^2$, and postcomposing each chart in an atlas, we obtain a new translation atlas. Let gM be this new translation surface and let ϕ_g be the identity map from M to gM. The map ϕ_g is an affine map with derivative g.

The group $g \in \mathrm{GL}_2^\circ(\mathbb{R})$ acts on $\widetilde{\mathcal{H}}_{\mathrm{m}}$ as follows. If $(f, M) \in \widetilde{\mathcal{H}}_{\mathrm{m}}$ then define $g(f, M)$ to be $(\phi_g \circ f, gM)$. The condition that $g \in \mathrm{GL}_2^\circ(\mathbb{R})$ insures that $\phi_g \circ f$ is orientation preserving. Since the action of $\mathrm{Mod}(S, \Sigma)$ on marked surfaces is by pre-composition, this action induces a well-defined action on $\mathcal{H}_{\mathrm{m}}$. If we let $\mathrm{GL}_2^\circ(\mathbb{R})$ act on $H^1(S, \Sigma; \mathbb{R}^2)$ by acting on the coefficients via the linear action on the plane, then the map $\mathrm{dev} : \mathcal{H}_{\mathrm{m}} \to H^1(S, \Sigma; \mathbb{R}^2)$ will be equivariant.

It is a general principle that if a connected topological group acts on a topological space X, then its universal cover acts on any cover of X. Since $\mathrm{GL}_2^\circ(\mathbb{R})$ and its subgroup $G = \mathrm{SL}_2(\mathbb{R})$ act on $\mathcal{H}_{\mathrm{m}}$, and $\widetilde{\mathcal{H}} \to \mathcal{H}_{\mathrm{m}}$ is a covering map, we conclude that their universal covering groups $\widetilde{\mathrm{GL}_2^\circ}(\mathbb{R})$ and $\widetilde{G}$ act on $\widetilde{\mathcal{H}}$. For related discussions in the case of strata of meromorphic quadratic differentials see [BS] and [HKK].

It will be useful for us to not only know that this action exists but to have an explicit description of the action. An element of $\widetilde{g} \in \widetilde{\mathrm{GL}_2^\circ}(\mathbb{R})$ can be represented by a pair (ρ, g) where $g \in \mathrm{GL}_2^\circ(\mathbb{R})$ and $\rho : [0,1] \to \mathrm{GL}_2^\circ(\mathbb{R})$ is a path with $\rho(0) = \mathrm{Id}$ and $\rho(1) = g$. Two such representations are equivalent if the corresponding paths are homotopic relative to their endpoints. Given an element of $\widetilde{\mathcal{H}}$, $(\check{f}, \check{M})$ and an element $\widetilde{g} = (\rho, g)$ of $\widetilde{\mathrm{GL}_2^\circ}(\mathbb{R})$ we have a path $\alpha(t) = \rho(t)(f, M)$ in $\mathcal{H}_{\mathrm{m}}$ and a lift of the initial point of this path (f, M) to $(\check{f}, \check{M}) \in \widetilde{\mathcal{H}}$. According to Proposition 2.6 this path lifts uniquely to a path $\tilde{\alpha}(t) \in \widetilde{\mathcal{H}}$ and we set $\tilde{g}(\check{f}, \check{M}) = \tilde{\alpha}(1)$. The resulting element is independent of the choice of path ρ.

Let

$$u_s \stackrel{\mathrm{def}}{=} \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \ g_t \stackrel{\mathrm{def}}{=} \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} \text{ and } r_\theta \stackrel{\mathrm{def}}{=} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \quad (7)$$

We will refer to the action of u_s as the *horocycle flow* and g_t as the *geodesic* flow and write

$$U = \{u_s : s \in \mathbb{R}\}, \quad A = \{g_t : t \in \mathbb{R}\}, \ \mathrm{SO}_2(\mathbb{R}) = \{r_\theta : \theta \in \mathbb{R}\}, \ B = AU.$$

Note that B is the connected component of the identity in the group of upper-triangular matrices and it normalizes U. Since U is simply connected, the connected component of the identity in the pre-image of U in $\widetilde{G}$ is a subgroup isomorphic to U. Thus we can identify U with a subgroup of $\widetilde{G}$, and the same is true for the groups B and A. The groups G and $\mathrm{GL}_2^\circ(\mathbb{R})$ are both homotopy equivalent to $\mathrm{SO}_2(\mathbb{R}) \cong S^1$, and $\widetilde{G}$ and $\widetilde{\mathrm{GL}_2^\circ}(\mathbb{R})$ are both homotopy equivalent to $\widetilde{\mathrm{SO}_2}(\mathbb{R}) \cong \mathbb{R}$.

As we have seen the fundamental groups of both groups $\widetilde{G}$ and $\widetilde{\mathrm{GL}_2^\circ}(\mathbb{R})$ are isomorphic to $\mathbb{Z}$. Since $\mathrm{SO}_2(\mathbb{R}) \subset G \subset \mathrm{GL}_2^\circ(\mathbb{R})$ and the induced maps on fundamental groups are isomorphisms it will cause no confusion to identify the fundamental groups of these three groups. This group is infinite cyclic and we denote it by C. We have three short exact sequences:

$$\begin{gathered} 1 \to C \to \widetilde{\mathrm{GL}_2^\circ}(\mathbb{R}) \to \mathrm{GL}_2^\circ(\mathbb{R}) \to 1 \\ 1 \to C \to \widetilde{G} \to G \to 1 \\ 1 \to C \to \widetilde{\mathrm{SO}_2}(\mathbb{R}) \to \mathrm{SO}_2(\mathbb{R}) \to 1. \end{gathered} \quad (8)$$

We will write the element of $\widetilde{\mathrm{SO}_2}(\mathbb{R})$ corresponding to $\theta \in \mathbb{R}$ as $\tilde{r}_\theta$, so that $\tilde{r}_\theta \mapsto r_{\theta \bmod 2\pi}$ is the projection $\widetilde{\mathrm{SO}_2}(\mathbb{R}) \to \mathrm{SO}_2(\mathbb{R})$. The group

C is central in $\widetilde{\mathrm{GL}}_2^\circ(\mathbb{R})$ and in these coordinates it is identified with $\{\tilde{r}_{2\pi n} : n \in \mathbb{Z}\}$.

We will explicitly describe the action of $\widetilde{\mathrm{SO}}_2(\mathbb{R})$ on $\widetilde{\mathcal{H}}$.

Proposition 2.9. *The left action of $\tilde{r}_{2\pi}$ on $\widetilde{\mathcal{H}}$ is equal to the right action of τ^{-1} where $\tau = \tau_1 \cdots \tau_k \in FT$. That is, for $(\check{f}, \check{M}) \in \widetilde{\mathcal{H}}$ we have $\tilde{r}_{2\pi}(\check{f}, \check{M}) = (\check{f} \circ \tau^{-1}, \check{M})$.*

Proof. Given a marked translation surface rel boundary $\check{f} : \check{S} \to \check{M}$ we follow the definition of the action of $\widetilde{\mathrm{GL}}_2^\circ(\mathbb{R})$ on $\widetilde{\mathcal{H}}$ given above. Let D_θ be the map supported in annuli around boundary components of $\check{S}$ that performs a Dehn twist by $\theta \in \mathbb{R}$ on each boundary component. Let $\theta_0 \in \mathbb{R}$, and assume for concreteness that $\theta_0 > 0$. We will describe the action of $\tilde{r}_{\theta_0}$ by explicitly lifting the action of $\{r_\theta : \theta \in [0, \theta_0]\}$ using the procedure described in Proposition 2.6.

Let $(f, M) \in \mathcal{H}_{\mathrm{m}}$ be a marked translation surface, and let $(\check{f}, \check{M})$ be an element of $\widetilde{\mathcal{H}}$ projecting to (f, M). We want to lift the path $\theta \mapsto r_\theta(M, f) = (f, r_\theta M)$ to $\widetilde{\mathcal{H}}$. As in §2.8, fix a triangulation τ of S. By a compactness argument it suffices to analyze the lift of the path $\theta \mapsto r_\theta(M, f)$, for the subset

$$\{\theta \in [0, \theta_0] : r_\theta(M, f) \in \mathcal{U}_\tau\},$$

and we can assume with no loss of generality that $(f, M) \in \mathcal{U}_\tau$.

Let $F(\tau, M, r_\theta M)$ be the comparison map defined in §2.4. An important observation, which follows immediately from the definition of the comparison maps, is that the derivative of $F(\tau, M, r_\theta M)$ is everywhere equal to the matrix r_θ, and hence the comparison map is in fact independent of the triangulation τ. Let $\check{F}(\tau, M, r_\theta M)$ denote the extension of $F(\tau, M, r_\theta M)$ to blown-up surfaces, and let $\check{F}_\theta = \check{F}(\tau, M, r_\theta M) \circ f$. Then as discussed in §2.8, $\check{F}_\theta$ does not preserve boundary coordinates, but the composition $D_{-\theta} \circ \check{F}_\theta$ does. Thus the path $\theta \mapsto (\check{F}_\theta \circ D_{-\theta}, r_\theta \check{M})$ satisfies (3), so by uniqueness, is the desired lift of the path $(f, r_\theta M)$ to $\widetilde{\mathcal{H}}$.

In particular, setting $\theta_0 = 2\pi$ and using the fact that $r_{\theta_0} = \mathrm{Id}$, we get $(\check{f} \circ D_{-2\pi}, \check{M})$ and $D_{-2\pi} = \tau^{-1}$ with $\tau = \tau_1 \cdots \tau_k$. □

2.11. **Consequences.** We note some consequences of the above discussion. These will not be needed in this paper but are of independent interest.

Corollary 2.10. *Every path component of $\mathcal{H}_{\mathrm{m}}$ has an infinite fundamental group. In particular it is never the case that the space of marked surfaces is the universal cover of the stratum.*

Proof. Let $\mathcal{C}$ be a path component of $\mathcal{H}_{\mathrm{m}}$ and let $M \in \mathcal{C}$. Proposition 2.9 shows that lifting the closed path $\{r_\theta M : \theta \in [0, 2\pi]\}$ to $\widetilde{\mathcal{H}}$, we get a non-closed path whose endpoints differ by an application of the element $\tau^{-1} \in FT$. This element has infinite order in FT. Since the group FT acts freely on the fiber $\mathrm{pr}^{-1}(M)$, none of the lifted paths are closed. It follows that $\mathcal{C}$ has an infinite fundamental group. $\square$

An instructive example is the case $\mathcal{H} = \mathcal{H}(0)$ (where the model surface S is the torus with one marked point and $\check{S}$ is the torus with an open disk removed). The covering space $\mathcal{H}_{\mathrm{m}}$ can be identified with $\mathrm{GL}_2^\circ(\mathbb{R})$ (the connected component of the identity in $\mathrm{GL}_2(\mathbb{R})$), and is not simply connected, whereas $\widetilde{\mathcal{H}}$ is identified with its universal cover group $\widetilde{\mathrm{GL}_2^\circ}(\mathbb{R})$ and is simply connected. A generator of the fundamental group of $\mathrm{GL}_2^\circ(\mathbb{R})$ acts by a boundary Dehn twist on $\check{S}$.

We present another useful consequence of our lifting construction.

Let Γ_0 be a subgroup of $\mathrm{Mod}(\check{S}, \partial\check{S})$, let $\mathcal{H}_{\Gamma_0}$ denote the quotient $\widetilde{\mathcal{H}}/\Gamma_0$, let $\mathcal{C}$ be a path component of $\mathcal{H}_{\Gamma_0}$ and let $p \in \mathcal{C}$ and $q \in \widetilde{\mathcal{H}}$ such that q projects to p. We define a homomorphism $\rho_q : \pi_1(\mathcal{C}, p) \to \Gamma_0$ as follows. Let ϕ be a loop based at p. Let $[\phi]$ be the element of π_1 that it represents. We can lift ϕ to a path starting at q. The endpoint of this lifted path maps to p so it has the form $p\gamma$ for some $\gamma \in \Gamma_0$ (where our notation reflects the fact that $\mathrm{Mod}(\check{S}, \partial\check{S})$ acts by precomposition, so defines a right-action). Define $\rho_q([\phi])$ to be γ. The homotopy lifting property shows that $\rho_q([\phi])$ depends only on q and on the homotopy class $[\phi]$, and not the particular loop ϕ chosen to represent it.

The construction of ρ_q depends on the choice of the point q. If we were to choose a different point q' mapping to p then $q' = q\alpha$ for some $\alpha \in \Gamma_0$. In this case the lift of ϕ starting at $p\alpha$ is the path $\phi\alpha$ and the other endpoint is $p\gamma\alpha = p\alpha(\alpha^{-1}\gamma\alpha)$. Thus $\rho_{q'}(\phi) = \alpha^{-1}\rho_q\alpha(\phi)$ so $\rho_{q'}$ differs from ρ_q by an inner automorphism of Γ_0. In other words we have constructed a preferred homomorphism $\rho : \pi_1(\mathcal{C}) \to \Gamma_0$, well-defined up to a choice of an inner automorphism of Γ_0.

Let $n \in \mathbb{N}$ and let C_n denote the subgroup of C generated by $r_{2\pi n}$. Then C_n is central in G and $\hat{G}_n = \widetilde{G}/C_n$ is the unique connected n-fold cover of G.

Corollary 2.11. *Let Γ_0 be a subgroup of $\mathrm{Mod}(\check{S}, \partial\check{S})$. Then $\hat{G}_n$ acts on $\widetilde{\mathcal{H}}/\Gamma_0$ if and only if $\tau^n \in \Gamma_0$. In particular, suppose n is the least common multiple of the numbers $\{a_i + 1 : i = 1, \ldots, k\}$. Then $\hat{G}_n$ acts on $\mathcal{H}_{\mathrm{f}}$, but $\hat{G}_m$ does not act when $m < n$.*

Proof. Let $\mathcal{H}_{\Gamma_0} = \widetilde{\mathcal{H}}/\Gamma_0$, and let $(\check{f}, \check{M}) \in \widetilde{\mathcal{H}}$. If the action of $\tilde{r}_{2\pi n}$ is well-defined on $\mathcal{H}_{\Gamma_0}$ then $\tilde{r}_{2\pi n}(\check{f}, \check{M})$ and $(\check{f}, \check{M})$ are equivalent in $\mathcal{H}_{\Gamma_0}$ so by Proposition 2.9,

$$(\check{f}, \check{M})\tau^{-n} = \tilde{r}_{2\pi n}(\check{f}, \check{M}) = (\check{f}, \check{M})\gamma$$

for some $\gamma \in \Gamma_0$. That is, $\check{f}\circ\gamma$ and $\check{f}\circ\tau^{-n}$ are isotopic via an isotopy fixing $\partial\check{S}$. In particular they represent the same elements in $\mathrm{Mod}(\check{S}, \partial\check{S})$ and $\tau^{-n} \in \Gamma_0$. Conversely if $\tau^{-n} \in \Gamma_0$ then $(\check{f}, \check{M})$ and $\tilde{r}_{2\pi n}(\check{f}, \check{M})$ represent the same surface in $\mathcal{H}_{\Gamma_0}$. $\square$

2.12. **Area form.** Let $\widetilde{\mathcal{H}}^{(1)}$ and $\mathcal{H}^{(1)}$ denote the subset of area-one surfaces in $\widetilde{\mathcal{H}}$ and $\mathcal{H}$ respectively. With respect to the charts afforded by the map dev, $\widetilde{\mathcal{H}}^{(1)}$ is a submanifold cut out by a quadratic equation.

We can give these quadratic equations explicitly. Identify the coefficients $\mathbb{R}^2$ with $\mathbb{C}$ and define a Hermitian form on $H^1(S, \Sigma; \mathbb{R}^2)$ by

$$\langle \alpha, \beta \rangle = \frac{1}{2i} \int_M \alpha \wedge \bar{\beta}. \tag{9}$$

The area of M is $\langle \omega, \omega \rangle$, where $\omega = \mathrm{dev}(M)$.

We now explain how this Hermitian form can be obtained in more topological terms, from the cup product and a particular choice of coefficient pairing. First note that if we take z and w to represent two sides of a triangle, then the signed area of the triangle is equal to $\Re(z\bar{w}/2i)$. Now recall that the cup product of two $\mathbb{R}$-valued simplicial cochains is defined on a simplex by taking the product of values of the cochains on the simplex, and is then extended by linearity to chains. If we have coefficients which are not in $\mathbb{R}$, we can replace the operation of multiplying values of cochains by a bilinear pairing of the values of cochains. Motivated by the above observation, we will use the coefficient pairing

$$\mathbb{C} \ni (z, w) \mapsto z\bar{w}/2i \in \mathbb{C}. \tag{10}$$

If we interpret the 1-forms in (9) as complex valued cohomology classes, interpret integration as evaluation on the fundamental class of M, and use evaluation on the fundamental class to identify $H^2(S, \Sigma; \mathbb{C})$ with $\mathbb{C}$, then our Hermitian form is the cup product

$$H^1(S, \Sigma; \mathbb{C}) \otimes H^1(S, \Sigma; \mathbb{C}) \to H^2(S, \Sigma; \mathbb{C}) \cong \mathbb{C},$$

and the particular choice of coefficient pairing (10) is responsible for the connection with the area of translation surfaces.

Since the Hermitian form on $\widetilde{\mathcal{H}}$ was defined purely topologically, it is preserved by the $\mathrm{Mod}(S, \Sigma)$-action. Thus $\widetilde{\mathcal{H}}^{(1)}$ and $\mathcal{H}^{(1)}$ have (G, X)-structures where X is a quadric in $H^1(S, \Sigma; \mathbb{C})$ and G can be taken to be the subgroup of the general linear group which preserves this quadric.

3. Rel foliation and Rel vector-fields

It is a general principle that geometric structures on $H^1(S, \Sigma; \mathbb{R}^2)$ which are invariant under the action of the mapping class group induce geometric structures on the stratum. We will now see an example of this principle. Consider the following exact sequence:

$$0 \longrightarrow H^0(S) \longrightarrow H^0(\Sigma) \longrightarrow H^1(S, \Sigma) \xrightarrow{\mathrm{Res}} H^1(S) \longrightarrow 0 \quad (11)$$

Let us take the coefficients in cohomology groups to be $\mathbb{R}^2$. Let $\mathfrak{R}$ denote the image of $H^0(\Sigma; \mathbb{R}^2)$ in $H^1(S, \Sigma; \mathbb{R}^2)$ which we can also identify with the kernel of the restriction map $\mathrm{Res} : H^1(S, \Sigma; \mathbb{R}^2) \to H^1(S; \mathbb{R}^2)$. We can identify $H^0(\Sigma; \mathbb{R}^2)$ with the set of functions from Σ to $\mathbb{R}^2$ and we can identify the image of $H^0(S; \mathbb{R}^2)$ in $H^0(\Sigma; \mathbb{R}^2)$ with the subspace of constant functions. Thus $\mathfrak{R}$ can be seen as $\mathbb{R}^2$-valued functions on Σ modulo constant functions. If k is the cardinality of Σ then the real dimension of $\mathfrak{R}$ is $2(k-1)$. In the second part of this paper we will be concerned with the case when Σ consists of two points, so that $\dim_{\mathbb{R}} \mathfrak{R} = 2$.

We will explicitly describe the action of $\mathfrak{R}$ on $H^1(S, \Sigma; \mathbb{R}^2)$. Pick $v \in \mathfrak{R}$ and let $\gamma \in H^1(S, \Sigma; \mathbb{R}^2)$. We will define $\gamma + v \in H^1(S, \Sigma; \mathbb{R}^2)$. Explicitly, the elements $\gamma, \gamma + v$ are determined by their values on oriented paths in S with endpoints in Σ. Let σ be one such oriented path starting at ξ_i and ending at ξ_j. Since $\mathfrak{R} \cong H^0(\Sigma; \mathbb{R}^2)/H^0(S; \mathbb{R}^2)$, v is an equivalence class of functions $\widetilde{v} : \Sigma \to \mathbb{R}^2$, where functions are equivalent if they differ by a constant. We define $(\gamma + v)(\sigma) = \gamma(\sigma) + \widetilde{v}(\xi_j) - \widetilde{v}(\xi_i)$. Since representatives of v differ by constants, the preceding formula does not depend on the choice of $\widetilde{v}$. Also $v(\sigma)$ gives the same value for any σ from ξ_i to ξ_j.

The group G acts equivariantly on the terms of the exact sequence (11). If we think of the terms as vector spaces of $\mathbb{R}^2$-valued functions then G acts on these functions by acting on their values. In particular there is a natural action of G on $\mathfrak{R}$ since it is the quotient of the first two terms.

A subspace W of a vector space V defines a linear foliation of V where the leaves are the translates of W. In this way the subspace $\mathfrak{R}$ defines a foliation of $H^1(S, \Sigma)$. Since the mapping class group $\mathrm{Mod}(S, \Sigma)$ preserves the short exact sequence it preserves this foliation and thus the foliation descends to a well-defined foliation on $\mathcal{H}$. We call this the Rel foliation. The names 'kernel foliation' and 'absolute period foliation' have also been used in the study of this foliation, see [Zo, Sch, McM7].

Proposition 3.1. *Two surfaces in the same* Rel *leaf have the same area.*

Proof. As we have seen in (9) the area of a surface M can be written as $\langle \omega, \omega \rangle$ where $\omega = \mathrm{dev}(M)$ and the bilinear form is the cup product with a certain choice of coefficient pairing. So it suffices to show that the cup product of two classes in $H^1(S, \Sigma)$ depends only on the image of the classes in absolute cohomology $H^1(S)$; the latter statement follows from the fact that the cup product is natural with respect to the inclusion $(M, \varnothing) \to (M, \Sigma)$, i.e. diagram (12) below commutes. We refer to [H] for the definition of the cup product in the relative case, and to [H, Prop. 3.10] for a proof of naturality.

$$\begin{tikzcd} H^1(S,\Sigma) \times H^1(S,\Sigma) \arrow[r, "\cup"] \arrow[d] & H^2(S,\Sigma) \arrow[d, "\simeq"] \\ H^1(S) \times H^1(S) \arrow[r, "\cup"] & H^2(S) \end{tikzcd} \tag{12}$$

□

We will define a notion of parallel translation on the leaves of the Rel foliation. We can identify the elements of $\mathfrak{R}$ with constant vector fields on the vector space $H^1(S, \Sigma; \mathbb{R}^2)$. Recall that we have insisted on labeling the points in Σ, and that the mapping class group fixes Σ pointwise. With these conventions, $\mathrm{Mod}(S, \Sigma)$ acts trivially on $\mathfrak{R}$. Thus the vector fields corresponding to $\mathfrak{R}$ are invariant under $\mathrm{Mod}(S, \Sigma)$ and induce well-defined vector fields on $\mathcal{H}_{\mathrm{m}}$ and $\mathcal{H}$. The leaves of the Rel foliation have natural translation structures and these are the coordinate vector fields.

The constant vector field associated with $v \in \mathfrak{R}$ can be integrated on the vector space $H^1(S, \Sigma; \mathbb{R}^2)$ to give a one-parameter flow. Our next objective is to lift this flow, to the extent possible, to $\mathcal{H}_{\mathrm{m}}$.

Definition 3.2. *Let M_0 be point in $\mathcal{H}$ and let $v \in \mathfrak{R}$. Let $\vec{v}$ denote the rel vector field on $\mathcal{H}$ corresponding to v. We say that $\mathrm{Rel}_v(M_0)$ is defined and equal to M_1 if there is a smooth path $\phi(t)$ in $\mathcal{H}$ with $\phi(0) = M_0$, $\frac{d}{dt}\phi(t) = \vec{v}(\phi(t))$ and $\phi(1) = M_1$.*

The translation structures on the leaves of the Rel foliation are not complete in general and this means that the trajectories of the vector fields cannot always be defined for all time.

Proposition 3.3. *Let $\Omega \subset \mathcal{H} \times \mathfrak{R}$ be the set of pairs (M, v) for which $\mathrm{Rel}_v(M)$ is defined. Then Ω is open and the map $(M, v) \mapsto \mathrm{Rel}_v(M)$ is continuous when viewed as a map from Ω to $\mathcal{H}$.*

Proof. This follows from properties of solutions of first order ordinary differential equations. □

Our next result deals with the interaction between the natural actions of G on $\mathcal{H}$ and $\mathfrak{R}$, and the partially defined maps Rel_v.

Proposition 3.4. *Let $M \in \mathcal{H}$ and $v \in \mathfrak{R}$. If $\mathrm{Rel}_v(M)$ is defined and $g \in G$ then $\mathrm{Rel}_{gv}(gM)$ is defined and $g(\mathrm{Rel}_v(M)) = \mathrm{Rel}_{gv}(gM)$.*

Proof. Let $\vec{v}$ denote the vector field corresponding to $v \in \mathfrak{R}$. To say that $\mathrm{Rel}_v(M)$ is defined means that there is a smooth path $\phi(t)$ with $\phi(0) = M$, $\frac{d}{dt}\phi(t) = \vec{v}(\phi(t))$, and in this case $\phi(1) = \mathrm{Rel}_v(M)$. Consider the path $t \mapsto g(\phi(t))$. It has the property that $g(\phi(0)) = gM$, $\frac{d}{dt}g(\phi(t)) = g(\frac{d}{dt}\phi(t)) = g\vec{v}(\phi(t)) = (\overrightarrow{gv})(\phi(t))$ and $g(\phi(1)) = g(\mathrm{Rel}_v(M))$. The existence of this path shows that $\mathrm{Rel}_{gv}(gM)$ is defined and $g(\mathrm{Rel}_v(M)) = \mathrm{Rel}_{gv}(gM)$. □

We can think of $\mathfrak{R}$ as a Lie group acting on $H^1(S, \Sigma; \mathbb{R}^2)$. The fact that we can lift elements of the Lie group action on $H^1(S, \Sigma; \mathbb{R}^2)$ to $\mathcal{H}_{\mathrm{m}}$ does not imply that the relations in the Lie group necessarily lift. For example the transformations Rel_v and Rel_w acting on $H^1(S, \Sigma; \mathbb{R}^2)$ commute but the corresponding lifted transformations of $\mathcal{H}_{\mathrm{m}}$ need not commute where they are defined. The following result gives criteria for a composition law and for commutation.

Proposition 3.5. *Let*

$$\square = [0,1]^2 \text{ and } \triangle = \{(s,t) \in \mathbb{R}^2 : 0 \leqslant s \leqslant t \leqslant 1\},$$

and let $v, w \in \mathfrak{R}$. Then:

(i) *If $\mathrm{Rel}_{sv+tw}(M)$ exists for all $(s,t) \in \triangle$ then*

$$\mathrm{Rel}_v \circ \mathrm{Rel}_w(M) = \mathrm{Rel}_{v+w}(M). \tag{13}$$

(ii) *If $\mathrm{Rel}_{sv+tw}(M)$ exists for all $(s,t) \in \square$ then*

$$\mathrm{Rel}_v \circ \mathrm{Rel}_w(M) = \mathrm{Rel}_w \circ \mathrm{Rel}_v(M) = \mathrm{Rel}_{v+w}(M). \tag{14}$$

Proof. It suffices to prove the result in $\mathcal{H}_{\mathrm{m}}$, since the maps Rel_u are $\mathrm{Mod}(S, \Sigma)$-equivariant. Note that (ii) follows immediately from (i), so we prove (i). Define

$$\sigma : \triangle \to H^1(S, \Sigma; \mathbb{R}^2) \quad \text{by} \quad \sigma(s,t) = \mathrm{dev}(M) + sv + tw.$$

Recall that we have a developing map $\mathrm{dev} : \mathcal{H}_{\mathrm{m}} \to H^1(S, \Sigma; \mathbb{R}^2)$. The developing map is a local homeomorphism. This does not imply that paths can be lifted but it does mean that when paths can be lifted the lifts are unique (see [E] for more information). The hypothesis that $\mathrm{Rel}_{sv+tw}(M)$ is defined for all $(s,t) \in \triangle$ means that every path $r \mapsto \sigma(rs, rt)$ for $0 \leqslant r \leqslant 1$ lifts to $\mathcal{H}_{\mathrm{m}}$. Let $\tilde{\sigma}(s,t) = \mathrm{Rel}_{sv+tw}(M)$ be the lift of σ to $\mathcal{H}_{\mathrm{m}}$. Arguing as in Proposition 1.11 in [E] we see that $\tilde{\sigma}$ is a continuous map. By construction $\tilde{\sigma}(0,1) = \mathrm{Rel}_w(M)$ and $\tilde{\sigma}(1,1) = \mathrm{Rel}_{v+w}(M)$. The path $\rho_0(r) = \tilde{\sigma}(r,1)$ satisfies $\rho_0(0) = \mathrm{Rel}_w(M)$ and $\rho_0' = v$, so by the definition of Rel, $\rho_0(1) = \mathrm{Rel}_v(\mathrm{Rel}_w(M))$. Also $\rho_0(1) = \tilde{\sigma}(1,1) = \mathrm{Rel}_{v+w}(M)$, and (13) follows. □

3.1. **Real Rel.** Let us write $\mathbb{R}^2$ as $\mathbb{R}_x \oplus \mathbb{R}_y$. We then write

$$H^1(S,\Sigma;\mathbb{R}^2) \cong H^1(S,\Sigma;\mathbb{R}_x) \oplus H^1(S,\Sigma;\mathbb{R}_y). \tag{15}$$

and we refer to $H^1(S,\Sigma;\mathbb{R}_x)$ as the *horizontal space*. Let Z denote the intersection of $\mathfrak{R}$ and the horizontal space. We will refer to Z as *real Rel*. Since the subgroup B of Z preserves the horizontal directions $\mathbb{R}_x$, its action on $\mathfrak{R}$ leaves Z invariant.

A special case which will concern us here are strata with two singularities. In this case $\mathfrak{R}$ can be identified with $\mathbb{R}^2$, and we make the identification explicit. Label the singularities of the model surface S by ξ_1 and ξ_2, we will identify $\mathfrak{R}$ with $\mathbb{R}^2$ as follows: a cochain $v : H_1(S,\Sigma) \to \mathbb{R}^2$ which vanishes on cycles represented by closed curves is identified with the vector $v(\delta)$ for some (any) directed path δ from ξ_1 to ξ_2. In this case Z is one dimensional, and we write $\mathrm{Rel}_t(M)$ for $\mathrm{Rel}_v(M)$, where $v = (t,0) \in \mathbb{R}^2 \cong \mathfrak{R}$ via the identification above.

Figures 3 and 4 show the effect of flowing along the real Rel vector field on a decagon with opposite sides identified. In Figure 3 the flow has the effect of shortening the top saddle connection. The flow cannot be continued past the point at which the length of the top saddle connection shrinks to zero. In Figure 4 the lengths of saddle connections are preserved since they connect vertices of the same color, and hence represent a saddle connection from a singularity to itself. In this case the flow can be continued for all time.

Definition 3.6. *Let $v \in \mathfrak{R}$. We denote by $\mathcal{H}'_v$ be the set of $M \in \mathcal{H}$ for which* $\mathrm{Rel}_v(M)$ *is defined.*

Proposition 3.5(i), with $w = -v$, implies $\mathrm{Rel}_{-v} \circ \mathrm{Rel}_v(M) = M$ for $M \in \mathcal{H}'_v$, and this yields a useful equivariance property:

$$\mathrm{Rel}_v(\mathcal{H}'_v) = \mathcal{H}'_{-v}. \tag{16}$$

In the case of two singularities keeping in mind our convention regarding the identification $\mathfrak{R} \cong \mathbb{R}^2$, we will continue to use the notation $\mathcal{H}'_v$ for $v \in \mathbb{R}^2$. Then Proposition 3.4 implies that for $v \in \mathbb{R}^2$ and $g \in G$

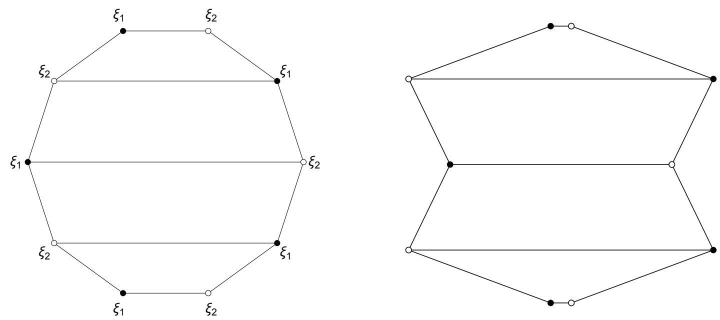


FIGURE 3. Applying Rel_t (with $t < 0$) to the decagon. When $t < -a$ or $t > b$ then $\mathrm{Rel_t}$ fails to be defined, where a is the length of the top segment and b is the length of the second segment from the top.

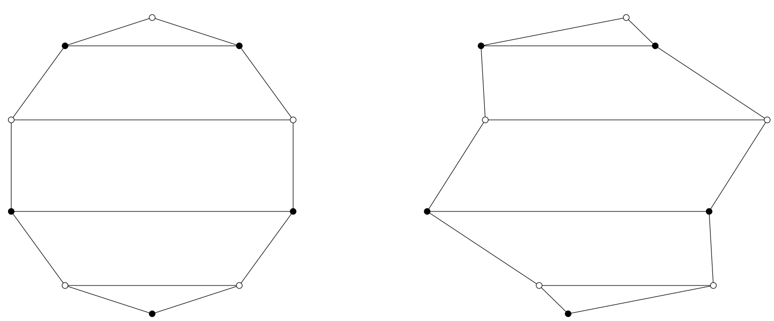


FIGURE 4. Applying Rel_t (with $t > 0$) to the tipped decagon.

we have $g(\mathcal{H}'_v) = \mathcal{H}'_{gv}$, where gv is the image of v under the linear action of g on $\mathbb{R}^2$. We now introduce some more notation for discussing directions belonging to Z.

Definition 3.7. *Let $\mathcal{H}'_\infty = \bigcap_{z \in Z} \mathcal{H}'_z$ (i.e. the subset of M on which Rel_z is defined for all $z \in Z$).*

In the case of two singularities, for a real number t let $\mathcal{H}'_t$ denote the set $\mathcal{H}'_v$ with $v = (t, 0)$ (i.e. the subset of M on which Rel_t is defined).

For a fixed $M \in \mathcal{H}$, let

$$Z^{(M)} = \{z \in Z : M \in \mathcal{H}'_z\}$$

(i.e., the subset of Z corresponding to surgeries which are defined on M). Thus $\mathcal{H}'_\infty = \{M \in \mathcal{H} : Z^{(M)} = Z\}$.

Recall that if V is a vector space and $V_0 \subset V$, we will say that V_0 is a *star body* if

$$v \in V_0, s \in [0,1] \implies sv \in V_0.$$

We denote the convex hull of a subset $W \subset V$ by $\operatorname{conv} W$.

Proposition 3.8. *The set $Z^{(M)}$ has the following properties:*

(i) *It is an open star body in Z.*
(ii) *If $b \in B$ then $Z^{(bM)} = b(Z^{(M)})$.*
(iii) *If $z \in Z^{(M)}$ then $-z \in Z^{(\mathrm{Rel}_z(M))}$.*
(iv) *If $z, z' \in Z$ and $\operatorname{conv}\{0, z, z + z'\} \subset Z^{(M)}$, then $\mathrm{Rel}_z(M)$ and $\mathrm{Rel}_{z'}(\mathrm{Rel}_z(M))$ are defined and $\mathrm{Rel}_{z'}(\mathrm{Rel}_z(M)) = \mathrm{Rel}_{z+z'}(M)$.*

Proof. The fact that $Z^{(M)}$ is open follows from Proposition 3.3. The fact that it is a star body is immediate from Definition 3.2. This proves (i). Assertions (ii), (iii), (iv) follow respectively from Propositions 3.4, (16), and 3.5(i). □

We will need a significant strengthening of Proposition 3.8:

Proposition 3.9. *For any M, $Z^{(M)}$ is convex.*

Whereas the proof of Proposition 3.8 relies only on general principles, Proposition 3.9 relies on additional information about Rel and will be proved in §6.

4. The centralizer and normalizer of the horocycle flow

Recall that G acts on $H^1(S, \Sigma; \mathbb{R}^2)$ via its linear action on $\mathbb{R}^2$. Since the linear action of U on $\mathbb{R}^2$ preserves horizontal vectors, it fixes elements of $\mathbb{R}_x$. This implies that real Rel commutes with the horocycle flow. Namely, by Proposition 3.4, if $z \in Z$ and $u \in U$ then $uz = z$, and hence $u(\mathcal{H}'_z) = \mathcal{H}'_z$, and

$$u(\mathrm{Rel}_z(M)) = \mathrm{Rel}_{u(z)}(uM) = \mathrm{Rel}_z(uM), \quad \text{for } M \in \mathcal{H}'_z.$$

Now define

$$N \stackrel{\text{def}}{=} \{(b, z) : b \in B, z \in Z\} \quad \text{and } L = \{(g, v) : g \in G, v \in \mathfrak{R}\}.$$

We equip N and L with the natural group structures as semi-direct products $N = B \ltimes Z$ and $L = G \ltimes \mathfrak{R}$, which are compatible with their actions on period coordinates. For the action of the first factor on the second in this semi-direct product, we take the natural action of G on $\mathfrak{R}$, and its restriction to the B-invariant subspace Z. In particular in the case of two singularities, G acts on $\mathfrak{R} \cong \mathbb{R}^2$ via its standard linear action, and B acts on $Z \cong \mathbb{R}_x$ via the restriction of its linear action on

$\mathbb{R}^2$, to the horizontal axis. We will write this semidirect product group law explicitly as

$$(g_1, v_1) \cdot (g_2, v_2) = (g_1 g_2, v_1 + g_1 v_2) \text{ for } \ell_i = (g_i, v_i) \in L,\ i = 1, 2,$$

where in the expression $g_1 v_2$ we mean the action of G on $\mathfrak{R}$ described above. We can associate a partially defined transformation of $\mathcal{H}$ to each element of L as follows:

$$gM \mathbin{\oplus} v = \mathrm{Rel}_v(gM) \ \text{ when } (g, v) \in L \ \text{ and } gM \in \mathcal{H}'_v. \tag{17}$$

We will use a different notation for the restriction to N. Namely we write

$$nM = \mathrm{Rel}_z(bM) \ \text{ when } n = (b, z) \in N \ \text{ and } bM \in \mathcal{H}'_z. \tag{18}$$

Note that (17) and (18) give two different notations for the same transformations. The reason for this is a fundamental difference between the behavior of the same operation on N and on L. When dealing with all of L, these operations do not obey a group action law. Indeed, it may happen that $(M \mathbin{\oplus} v) \mathbin{\oplus} w \neq (M \mathbin{\oplus} w) \mathbin{\oplus} v$. However, for N we have the following:

Proposition 4.1. *Let $n_1 = (b_1, z_1)$, $n_2 = (b_2, z_2)$ be two elements of N. Suppose that $n_2 M$ and $n_1(n_2 M)$ are defined. Then $(n_1 n_2)M$ is defined, and $n_1(n_2 M) = (n_1 n_2)M$.*

Proof. Let $n_1 = (b_1, z_1)$ and $n_2 = (b_2, z_2)$. Since $n_2 M = \mathrm{Rel}_{z_2}(b_2 M)$ is defined we have $z_2 \in Z^{(b_2 M)}$, and since $n_1(n_2 M)$ is defined, we have $z_1 \in Z^{(b_1 n_2 M)}$. Using Proposition 3.8 we find that $Z^{(n_2 M)}$ contains $-z_2$ and $b_1^{-1} z_1$. By Proposition 3.9 we see that

$$\mathrm{conv}\{0, -z_2, b_1^{-1} z_1\} \subset Z^{(n_2 M)}. \tag{19}$$

Then

$$\begin{aligned} n_1(n_2(M)) =& \mathrm{Rel}_{z_1}(b_1 n_2(M)) && (20)\\ =& \mathrm{Rel}_{b_1^{-1} z_1}(n_2(M)) && (21)\\ =& \mathrm{Rel}_{b_1^{-1} z_1 + z_2} \circ \mathrm{Rel}_{-z_2}(n_2(M)) && (22)\\ =& \mathrm{Rel}_{b_1^{-1} z_1 + z_2}(b_2(M)) && (23)\\ =& \mathrm{Rel}_{z_1 + b_1 z_2}(b_1 b_2(M)) && (24)\\ =& (n_1 n_2)(M). \end{aligned}$$

Here Proposition 3.4 is used to derive (21) from (20) and to derive (24) from (23), and (22) is obtained from (21) by Proposition 3.8 and (19). $\square$

The following is immediate from Propositions 3.3, 3.8 and 4.1 (and was proved previously in [MW2]):

Corollary 4.2. *The set $\mathcal{H}'_\infty$ is N-invariant. The map $N \times \mathcal{H}'_\infty \to \mathcal{H}'_\infty$ defined by $(\ell, M) \mapsto \ell M$ defines a continuous action of N on $\mathcal{H}'_\infty$. The Z-orbits in this action are the real Rel leaves in $\mathcal{H}'_\infty$.*

The following will be useful. For $u \in U$ and $n = (b_n, z_n) \in N$, write $u^{(b_n)} = bub^{-1}$, so that $nu = u^{(b_n)}n$ as elements of N. Then we have:

Corollary 4.3. *For any $u \in U$, $n \in N$ and $M \in \mathcal{H}$, if $n(uM)$ is defined then nM is defined and $n(uM) = u^{(b_n)}(nM)$.*

Proof. We apply Proposition 4.1 with $b_2 = u$ and $z_2 = 0$. □

4.1. **The stabilizer group of a measure.** Proposition 3.8 implies the invariance property $Z^{(uM)} = Z^{(M)}$. We now extend this to semi-direct products and use it to define the stabilizer of a measure, within a collection of partially defined transformations.

For fixed $M \in \mathcal{H}$ we write

$$N^{(M)} = \left\{(b, z) \in N : z \in Z^{(bM)}\right\}.$$

(i.e. the set of $n \in N$ for which nM is defined). Then it follows from Proposition 3.3 that $N^{(M)}$ is open for each M, and Corollary 4.3 implies $N^{(uM)} = N^{(M)}$.

Proposition 4.4. *Suppose M is in the U orbit-closure of M'. Then $N^{(M)} \subset N^{(M')}$.*

Proof. If $n \in N^{(M)}$ then by Proposition 3.3 there is a neighborhood $\mathcal{W}$ of M in $\mathcal{H}$ such that $n \in N^{(M_1)}$ for any $M_1 \in \mathcal{W}$. Let $u \in U$ such that $uM' \in \mathcal{W}$. Then $z \in N^{(uM')} = N^{(M')}$. □

We also need the following:

Proposition 4.5. *Given an ergodic U-invariant measure μ there is a subset $\Omega \subset \mathcal{H}$ such that $\mu(\Omega) = 1$ and for any $M_1, M_2 \in \Omega$, $Z^{(M_1)} = Z^{(M_2)}$ and $N^{(M_1)} = N^{(M_2)}$.*

Proof. To explain the idea, we first prove the assertion for $Z^{(M)}$ in case $\dim Z = 1$. In this case, $Z^{(M)}$ is an open interval for every $M \in \mathcal{H}$, which we can write as

$$Z^{(M)} = (a_M, b_M), \text{ for some } -\infty \leqslant a_M < 0 < b_M \leqslant \infty.$$

The maps $M \mapsto a_M, M \mapsto b_M$ are measurable maps with values in the extended real line so by ergodicity, are constant μ-a.e., and the statement follows.

In the general case the proofs for $Z^{(M)}$ and $N^{(M)}$ are identical, so we discuss the case of $Z^{(M)}$. The map $M \mapsto Z^{(M)}$ is a map from $\mathcal{H}$ to the collection of open subsets of Z. Rather than worry about measurability issues, we proceed as follows. Let Z_0 be a dense countable subset of Z. For each $z \in Z_0$, the set

$$\Omega_z = \left\{ M \in \mathcal{H} : z \in Z^{(M)} \right\}$$

is a measurable invariant set, so has measure 0 or 1 by ergodicity. We define

$$\Omega = \bigcap_{\mu(\Omega_z)=1} \Omega_z \smallsetminus \bigcup_{\mu(\Omega_{z'})=0} \Omega_{z'},$$

where z, z' range over the countable set Z_0. Then $\mu(\Omega) = 1$, and for any $M \in \Omega$,

$$Z^{(M)} \cap Z_0 = \{z \in Z_0 : \mu(\Omega_z) = 1\}.$$

Since an open set is the interior of the closure of its intersection with a dense set, we see that $Z^{(M_1)} = Z^{(M_2)}$ for any $M_1, M_2 \in \Omega$. $\square$

We will denote by $Z^{(\mu)}$ and $N^{(\mu)}$ the sets $Z^{(M)}, N^{(M)}$ which appear in Proposition 4.5 for $M \in \Omega$. If $z \in Z^{(\mu)}$ we define a pushforward

$$\mathrm{Rel}_{z*}\mu(X) = \mu(\mathrm{Rel}_{-z}(X \cap \mathcal{H}'_{-z})), \text{ for all measurable } X \subset \mathcal{H}.$$

Note that $\mu(\mathcal{H}'_z) = 1$, and now it follows from (16) that $\mathrm{Rel}_{z*}\mu$ is a probability measure. Moreover, the partially defined map Rel_z is a measurable conjugacy between $(\mathcal{H}, \mu)$ and $(\mathcal{H}, \mathrm{Rel}_{z*}\mu)$, thought of as dynamical systems for the U-action. Thus $\mathrm{Rel}_{z*}\mu$ is again an ergodic U-invariant measure. The same observations are valid for $n \in N^{(\mu)}$. Namely, if we denote

$$\mathcal{H}'_n = \{M \in \mathcal{H} : n(M) \text{ is defined}\} = \left\{ M \in \mathcal{H} : z \in Z^{(bM)} \right\} \tag{25}$$

(where $n = (b, z)$), then we have an equivariance property

$$n(\mathcal{H}'_n) = \mathcal{H}'_{n^{-1}} \tag{26}$$

and we can define an ergodic U-invariant measure

$$n_*\mu(X) = \mu(n^{-1}(X \cap \mathcal{H}'_{n^{-1}})), \text{ for all measurable } X \subset \mathcal{H}.$$

Corollary 4.3 now implies that $n_*\mu$ is a U-invariant measure, and the partially defined map $M \mapsto n(M)$ is equivariant for the action of U on $(\mathcal{H}, \mu)$ and the "time-changed" action of U on $(\mathcal{H}, n_*\mu)$ via

$$u \cdot M = u^{(n)}M, \text{ for } n_*\mu \text{ a.e. } M$$

(where $u^{(n)} = nun^{-1}$).

The collection of Borel probability measures on a locally compact space X can be given the weak-$*$ topology by embedding it in the dual space of the space $C_c(X)$ of continuous functions with compact support.

Proposition 4.6. *Let μ be an ergodic U-invariant probability measure. The map which takes $n \in N^{(\mu)}$ to $n_*\mu$ is a continuous with respect to the weak-$*$ topology.*

Proof. Let n_j be a sequence of elements of $N^{(\mu)}$ converging to $n_\infty \in N^{(\mu)}$. In order to show that $n_{j*}\mu$ converges to $n_{\infty *}\mu$ we need to show that for any continuous compactly supported function φ on $\mathcal{H}$, we have

$$\lim_{j\to\infty} \int_{\mathcal{H}} \varphi \, d(n_{j*}\mu) = \int_{\mathcal{H}} \varphi \, d(n_{\infty *}\mu).$$

For each j, the set $\mathcal{H}'_{n_j}$ has full μ-measure, and hence so does $\mathcal{H}'_0 = \bigcap_{1\leqslant j\leqslant\infty} \mathcal{H}'_{n_j}$. Now we have:

$$\begin{aligned}
\lim_{j\to\infty} \int_{\mathcal{H}} \varphi \, d(n_{j*}\mu) &= \lim_{j\to\infty} \int_{n_j(\mathcal{H}'_0)} \varphi \, d(n_{j*}\mu) && (27)\\
&= \lim_{j\to\infty} \int_{\mathcal{H}'_0} \varphi \circ n_j \, d\mu && (28)\\
&= \int_{\mathcal{H}'_0} \lim_{j\to\infty} \varphi \circ n_j \, d\mu && (29)\\
&= \int_{\mathcal{H}'_0} \varphi \circ n_\infty \, d\mu && (30)\\
&= \int_{n_\infty(\mathcal{H}'_0)} \varphi \, d(n_{\infty *}\mu) && (31)\\
&= \int_{\mathcal{H}} \varphi \, d(n_{\infty *}\mu). && (32)
\end{aligned}$$

The equalities (27) and (32) follow from the fact that each $n_{j*}\mu$ assigns full measure to $n_j(\mathcal{H}'_0)$ respectively. The lines (28) and (31) follow from the definition of the pushforward of a measure. Line (29) follows from Lebesgue's Dominated Convergence Theorem using the fact that, since f has compact support, it is bounded and hence the family of functions $f \circ n_j$ is uniformly bounded and also using the fact that constant functions are in $L^1(\mu)$ since μ is a probability measure. Line (30) follows from Proposition 3.3 and the continuity of φ. $\square$

Definition 4.7. *For any ergodic U-invariant measure μ we define*

$$N_\mu \overset{\text{def}}{=} \left\{ n \in N^{(\mu)} : n_*\mu = \mu \right\}.$$

Corollary 4.8. *N_μ is a closed subgroup of N.*

Proof. The fact that N_μ is closed follows from Proposition 4.6, and in order to prove that N_μ is closed under compositions, it suffices to show that for μ-almost every M, $n_1(n_2M) = (n_1n_2)M$. This follows from Proposition 4.1. ☐

Proposition 4.9. *If N_μ contains a non-unipotent element (i.e. an element in $N \smallsetminus UZ$) and $\dim Z = 1$ then $N_\mu \cap Z$ is connected. In particular, if N_μ contains a non-unipotent element of N and a nontrivial element of Z then it contains all of Z.*

Proof. Write $Z_1 = UZ \cong \mathbb{R}^2$ and $N_1 = N_\mu \cap Z_1$, and let $a \in N_\mu \smallsetminus N_1$. By Corollary 4.8, N_μ is a closed subgroup of N, and hence N_1 is a closed subgroup of Z_1 containing U. If it is not connected then there is a minimal positive distance between two distinct cosets for N_1 in Z_1. However N_1 is invariant under conjugation by the elements a and a^{-1}, and one of these acts on Z_1 by contractions. This contradiction proves the claim. ☐

4.2. **Generic points and real Rel.** Recall that if μ is a U-invariant ergodic probability measure on a closed subset $\mathcal{L}$ of a stratum $\mathcal{H}$, then $M \in \mathcal{L}$ is said to be *generic for μ* if for any continuous compactly supported function f on $\mathcal{L}$, we have

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(u_sM)\, ds = \int f\, d\mu. \tag{33}$$

Definition 4.10. *Let μ be an invariant U-ergodic measure on $\mathcal{H}$. Then Ω_μ denotes the set of generic points for μ.*

We collect some well-known facts about generic points.

Proposition 4.11. *If μ is ergodic then Ω_μ has full μ measure. If $\mu \neq \nu$ then Ω_μ and Ω_ν are disjoint. If $M \in \Omega_\mu$ then the support of μ is contained in the orbit closure of M.*

In light of Proposition 4.4, this implies:

Corollary 4.12. *If $M \in \Omega_\mu$ then $N^{(\mu)} \subset N^{(M)}$.*

It will prove useful to reformulate the notion of genericity in terms of convergence of measures. It is important to keep in mind that we are dealing with spaces which are locally compact but not compact. The weak-$*$ topology on the space of Borel measures is defined in terms of integrals of continuous functions of compact support. In this topology the total mass of a measure need not be a continuous function on the space of measures since the constant function 1 need not have compact support.

Definition 4.13. *Let $\nu(M,T)$ be the probability measure defined by*

$$\int \varphi \, d\nu(M,T) = \frac{1}{T}\int_0^T \varphi(u_s M)\, ds, \quad \textit{for all } \varphi \in C_c(\mathcal{H}).$$

Then M is generic for μ if

$$\lim_{T\to\infty} \nu(M,T) = \mu \tag{34}$$

where the limit is taken with respect to the weak-$*$ topology on measures on $\mathcal{H}$.

In preparation for Proposition 4.15 below, we collect some results about integrals of continuous bounded functions along orbits. The space of probability measures on a locally compact space can be given another topology by embedding it in the dual space of $C_b(X)$, the space of continuous bounded functions on X. This topology is called the strict topology (see [Bo] for more information). Clearly convergence in the strict topology implies convergence in the weak-$*$ topology. The following result gives a simple criterion for showing that weak-$*$ convergence implies strict convergence.

Proposition 4.14 ([Bo] Prop. 9, p. 61)**.** *Suppose X is a locally compact space, μ_j is a sequence of probability measures on X, and μ is also a measure on X. If $\mu_j \to \mu$ with respect to the weak-$*$ topology, and if μ is a probability measure, then $\mu_j \to \mu$ with respect to the strict topology.*

If the limit measure μ is not a probability measure then it has total mass less than one. This phenomenon is referred to as "loss of mass". In §10 we will give conditions that establish that loss of mass does not occur for measures $\nu(M,T)$.

We will repeatedly use the following:

Proposition 4.15. *Let μ be an ergodic U-invariant measure. If $n \in N^{(\mu)}$ then $n(\Omega_\mu)$ is the set of generic points for $n_*\mu$.*

Proof. Let $M \in \Omega_\mu$. According to Corollary 4.12, nM is defined. We use the reformulation of genericity in terms of weak-$*$ convergence of measures given above. Thus we are given that $\lim_{T\to\infty} \nu(M,T) = \mu$ with respect to the weak-$*$ topology on measures on $\mathcal{H}$ and we want to show that $\lim_{T\to\infty} \nu(nM,T) = n_*\mu$ with respect to the same topology.

Since N normalizes U, there is $c > 0$ such that $nu_s = u_{cs}n$. We have $n_*\nu(M,T) = \nu(nM,cT)$. Thus it suffices to show:

$$\lim_{T\to\infty} n_*\nu(M,T) = n_*\mu.$$

Let $\mathcal{H}'_1 = \mathcal{H}'_n$ as in (25) and let $\mathcal{H}'_2 = n(\mathcal{H}'_1) = \mathcal{H}'_{n^{-1}}$. Let $\mathfrak{M}$ denote the set of probability measures on $\mathcal{H}$ which assign mass 1 to $\mathcal{H}'_1$, and let $\mathfrak{N}$ denote the set of probability measures on $\mathcal{H}$ which assign mass 1 to $\mathcal{H}'_2$. By Proposition 3.3, $\mathcal{H}'_1$ and $\mathcal{H}'_2$ are open subsets of $\mathcal{H}$, and in particular are locally compact. Moreover any continuous compactly supported function on $\mathcal{H}'_1$ extends to a continuous function on $\mathcal{H}$ by setting it equal to zero on $\mathcal{H} \smallsetminus \mathcal{H}'_1$. It follows that $\lim_{T\to\infty} \nu(M,T) = \mu$ with respect to the weak-$*$ topology on $\mathfrak{M}$.

Since n is a homeomorphism from $\mathcal{H}'_1$ to $\mathcal{H}'_2$, the map $n_* : \mathfrak{M} \to \mathfrak{N}$ is continuous and $\lim_{T\to\infty} n_*\nu(M,T) = n_*\mu$ with respect to the weak-$*$ topology on $\mathfrak{N}$. Since $n_*\mu$ is a probability measures which assigns full measure to $\mathcal{H}'_2$, and since $\mathcal{H}'_2$ is locally compact, we can use Proposition 4.14 to conclude that $\lim_{T\to\infty} n_*\nu(M,T) = n_*\mu$ in the strict topology on measures on $\mathcal{H}'_2$.

We need to show that convergence also holds with respect to the weak-$*$ topology on measures on $\mathcal{H}$. Given a function φ on $\mathcal{H}$ which is continuous and compactly supported, its restriction to $\mathcal{H}'_2$ is a bounded continuous function. Therefore

$$\lim_{T\to\infty} \int_{\mathcal{H}'_2} \varphi \, dn_*\nu(M,T) = \int_{\mathcal{H}'_2} \varphi \, dn_*\mu$$

by strict convergence which gives

$$\lim_{T\to\infty} \int_{\mathcal{H}} \varphi \, dn_*\nu(M,T) = \int_{\mathcal{H}} \varphi \, dn_*\mu,$$

and establishes convergence with respect to the weak-$*$ topology on measures on $\mathcal{H}$. □

Corollary 4.16. *If μ is an ergodic U-invariant probability measure and if M_1 and M_2 are in Ω_μ and there is an element $n \in N^{(\mu)}$ of N such that $M_2 = nM_1$, then $n \in N_\mu$.*

Proof. Let M_1 and M_2 be generic for μ and $M_2 = nM_1$ where $n \in N$. According to Proposition 4.15, μ and $n_*\mu$ share a generic point and hence, by Proposition 4.11, they must coincide. □

5. Horizontal equivalence of surfaces

Given $M \in \mathcal{H}$ with singularity set Σ, we denote by $\Xi(M)$ the set of horizontal saddle connections for M. We would like to use $\Xi(M)$ to define two equivalence relations on surfaces: *topological horizontal equivalence* and *geometrical horizontal equivalence*. Analogously to the sets $Z^{(\mu)}, N^{(\mu)}$ appearing in Proposition 4.5, they will serve as invariants of ergodic U-invariant measures.

Let $\check{M}$ be the blowup of M and let $(\check{f}, \check{M}) \in \widetilde{\mathcal{H}}$ be a marking of $\check{M}$ rel boundary. Then $\check{f}^{-1}(\Xi(M))$ is a subset of $\check{S}$, which we denote by $\check{f}^*(\Xi)$. We take $\Xi(M)$ to include all of the points of Σ and hence $\check{f}^*(\Xi)$ contains the boundary $\partial\check{S}$. In addition, for each edge of $\Xi(M)$, $\check{f}^*(\Xi)$ contains an edge in the interior of S, which intersects the boundary $\partial\check{S}$ at points with angular parameters which are multiples of π, even or odd according as the point is the initial or terminal point of the edge.

Definition 5.1. *We will say that M_1 and M_2 are* topologically horizontally equivalent *if there are markings rel boundary $(\check{f}_i, \check{M}_i)$ such that $\check{f}_2^*(\Xi)$ and $\check{f}_1^*(\Xi)$ can be obtained from each other by an isotopy of $\check{S}$ that does not move points of $\partial\check{S}$. We will say that M_1 and M_2 are* geometrically horizontally equivalent *if they are topologically horizontally equivalent, and if for any edge δ in $\check{f}_i^*(\Xi(M_i))$, $\mathrm{hol}(M_1, \check{f}_1(\delta)) = \mathrm{hol}(M_2, \check{f}_2(\delta))$, where $\check{f}_1$ and $\check{f}_2$ are as in the definition of topological equivalence.*

Proposition 5.2. *If $M \in \mathcal{H}$, $b \in B$ and $u \in U$ then M and bM are topologically horizontally equivalent, and M and uM are geometrically horizontally equivalent.*

Proof. Since B acts by postcomposition on the charts, and preserves the horizontal direction, it preserves the set $\Xi(M)$ (possibly changing the lengths of edges), and hence M and bM are topologically equivalent. Moreover if $b = u \in U$ then the lengths of edges are unaffected and so M and uM are geometrically horizontally equivalent. □

Corollary 5.3. *If μ is an ergodic U-invariant measure on $\mathcal{H}$, then there is a subset $X \subset \mathcal{H}$ of full μ-measure such that for any $M_1, M_2 \in X$, M_1 and M_2 are geometrically horizontally equivalent.*

We will define a combinatorial invariant of topological horizontal equivalence – the *horizontal data diagram*. This diagram is an analogue of the separatrix diagram introduced by Kontsevich and Zorich [KoZo] and captures some of the properties of $\Xi(M)$ which depend only on the class of M. Note that a graph embedded on a translation surface is a ribbon graph, namely at each vertex it inherits from the surface a cyclic order of edges incident to the vertex. Its vertices are labeled by the labels $\xi_1, \ldots, \xi_k$ of the corresponding singularities. Additionally, if the graph consists of horizontal saddle connection, each edge inherits an orientation from the translation structure mapping it to a horizontal edge in the plane.

For any $(\check{f}, \check{M}) \in \widetilde{\mathcal{H}}$, $\check{f}^*(\Xi)$ is a graph embedded in $\check{S}$. By projecting it to S we thus obtain a ribbon graph with labeled singularities and oriented edges. The graphs $\check{f}^*(\Xi)$ carry additional information, namely the angular distance separating consecutive edges incident at a vertex. This angle is an integer multiple of π, and where the orientations of consecutive edges agree if and only if this distance is an even multiple of π. It is clear that all of this information is an invariant of geometric horizontal equivalence.

The horizontal data diagram of M records the graph structure induced by $\check{f}^*(\Xi)$, as well as the orientation of edges, labeling of vertices, and cyclic structure at each singularity. In order to record the information of angular separation of edges at vertices, we indicate as dashed lines additional left- and right-pointing horizontal. Thus at the singularity ξ_i there are $2(a_i + 1)$ prongs, some of which correspond to edges of the graph. Figures 5, 6 and 7 depict graphs which can occur as $\Xi(M)$ for some M in $\mathcal{H}(1,1)$.

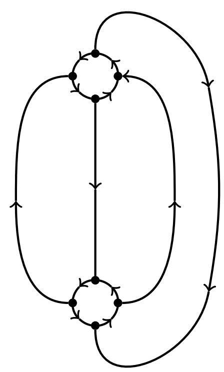


FIGURE 5. This figure represents a horizontal data diagram on a blown up surface.

In Figure 6 the angle between any two adjacent ends is exactly π, but this is not the case in Figure 7. We capture this angle information in the diagrams in Figure 7 by inserting additional dotted lines at each vertex indicating the ends of horizontal separatrices which are not saddle connections. We can determine the angle between two prongs by counting the number of ends of separatrices between them. Note that it makes sense to give orientations to separatrices so that the orientations of ends at a given vertex alternate with respect to the cyclic ordering.

It follows from Corollary 5.3 that for any ergodic U-invariant measure μ, there is a subset X of $\mathcal{H}$ such that $\mu(X) = 1$ and every $M \in X$ has the same horizontal data diagram. We call it the *horizontal data diagram of* μ and denote it by $\Xi(\mu)$.

If a horizontal data diagram is *maximal*, i.e. at each vertex, all prongs are initial or terminal prongs of edges, then the horizontal data diagram coincides with the separatrix diagram of [KoZo].

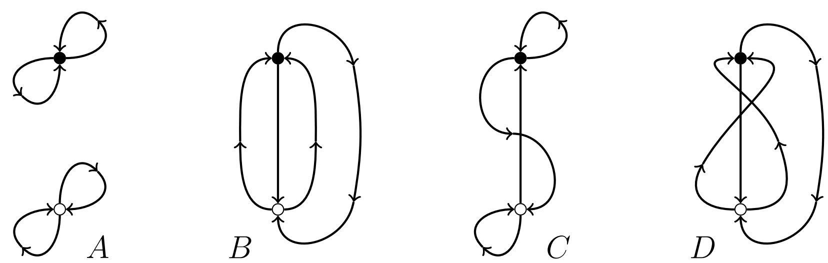


FIGURE 6. These figures represent all maximal horizontal data diagrams in $\mathcal{H}(1,1)$ up to switching labels and orientations. See Figure 4 for a polygonal presentation of a cylinder decomposition of type A.

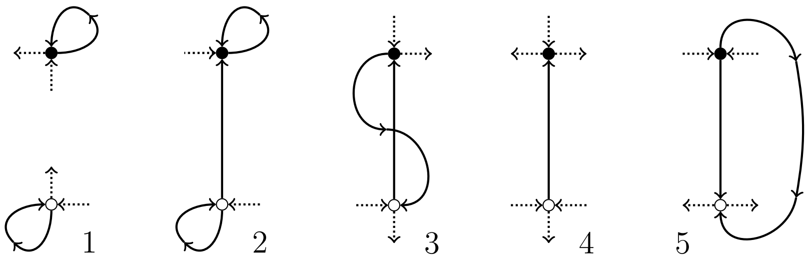


FIGURE 7. These figures represent some non-maximal horizontal data diagrams in $\mathcal{H}(1,1)$.

We will sometimes need an extension of $\Xi(M)$. Recall that at each $\xi_i \in \Sigma$ there are $2(a_i+1)$ prongs. Some of these are part of saddle connections in $\Xi(M)$, and we call the others the *unoccupied prongs* of $\Xi(M)$. Let $\Xi_-(M)$ denote the graph which contains $\Xi(M)$ and has one additional *prong edge* for every unoccupied prong, with one vertex at a singular point and one vertex at a point of $M \smallsetminus \Sigma$, where the edge is realized by a path in M which lies on the horizontal separatrix issuing at the corresponding prong.

We will sometimes need to be specific about the length of edges. In that case, given $L > 0$, we will let $\Xi_-(M, L)$ denote the graph $\Xi_-(M)$ described above, where the prong edges all have length L. Note that since all of the prong edges are part of an infinite separatrix in the horizontal direction, the graph $\Xi_-(M, L)$ can be embedded in M for any L.

6. An explicit surgery for real Rel

In this section, for fixed $M \in \mathcal{H}$ and $z \in Z^{(M)}$, we will present an explicit presentation of $\mathrm{Rel}_z M$ in terms of glued polygons. Our explicit surgery generalizes the special cases treated in [McM8] and [B2, §3]. As a by-product, it will enable us to determine $Z^{(M)}$ explicitly from the geometry of M. This analysis makes it possible to analyze limits $\lim_{j\to\infty} \mathrm{Rel}_{z_j} M$, for $z_j \in Z^{(M)}$ with $\lim_{j\to\infty} z_j \in \partial Z^{(M)}$. Such limits do not exist as elements of $\mathcal{H}$; loosely speaking they belong to a bordification of $\mathcal{H}$ obtained by adjoining boundary strata. We will not construct this bordification in this paper but hope to return to it in future work. A particularly simple case of this bordification arises when one takes limits of surfaces in $\mathcal{H}(1,1)$ for which a segment joining the two singularities collapses. Even in this relatively simple case, which will arise in Theorem 6.5, the construction we use differs from earlier related work (see [KoZo, EMZ]) and leads naturally to the use of framed surfaces.

We will establish some terminology and make a construction which will be used in stating Theorem 6.1. Let L and ε be positive numbers. Given M, we let $\Xi(M)$, $\Xi_-(M)$ and $\Xi_-(M,L)$ be as in §5. We say that $\mathcal{N} = \mathcal{N}(L,\varepsilon) \subset M$ is the *(L,ε)-rectangle thickening of* $\Xi_-(M)$ if $\mathcal{N}$ is a union of rectangles R_e^+ and R_e^- in M, with sides parallel to the coordinate axes, where e ranges over the edges of $\Xi_-(M,L)$, R_e has vertical sides of length ε, and the edge e runs along the bottom of R_e^+ and the top of R_e^-. Here the words 'bottom' and 'top' refer to the orientation provided by the translation surface structure and need not correspond to the directions shown in our figures. The edge identification maps of the rectangles are inherited from $\Xi_-(M)$, namely they are as follows. Each R_e^- is attached to R_e^+ along e, and the bottom of R_e^- (resp. top of R_e^+) is unattached. If the right end of e is not a singularity (i.e. e is right-pointing prong edge) then the right hand boundaries of $R_e^\pm$ are unattached. If the right end of e is the singularity ξ then the right hand boundary of R_e^- (resp. R_e^+) is attached to the left hand boundary of R_f^-, where f is the edge of $\Xi_-(M)$ which is counterclockwise (resp. clockwise) from e at ξ. For the gluing rule for the left edges, replace left with right and clockwise with counterclockwise in the above description. See Figure 8.

Note that for any L there is $\varepsilon_0 = \varepsilon_0(M,L)$ such that for all $\varepsilon < \varepsilon_0$, the (L,ε)-rectangle thickening of $\Xi_-(M)$ exists and is embedded in M. Moreover the constant $\varepsilon_0(M,L)$ can be chosen to be independent of M and L, for L in a bounded set of numbers and M in a compact set of topologically horizontally equivalent surfaces.

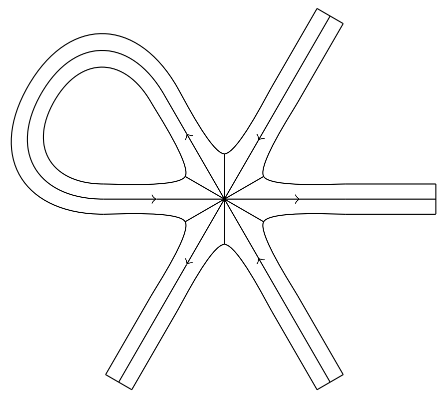


FIGURE 8. A rectangle thickening of $\Xi_-(M)$ in $\mathcal{H}(2)$ with 10 rectangles.

Given translation surfaces M, M', and a subset $M_0 \subset M$ we say that *M' can be obtained from M by modifying M_0* if there are polygon representations of M and M', and a subset $M_0' \subset M'$, such that M_0, M_0' are unions of polygons, and there is a homeomorphism $f : (M \smallsetminus M_0) \to (M' \smallsetminus M_0')$, which is a translation in each coordinate chart of the translation surface structures on M and M'.

Theorem 6.1. *Let $M \in \mathcal{H}$. For each $\delta \in \Xi(M)$, let $Z^{(M,\delta)}$ denote the connected component of 0 in $\{z \in Z : \mathrm{hol}(M,\delta) + z(\delta) \neq 0\}$. Then*

$$Z^{(M)} = \bigcap_{\delta \in \Xi(M)} Z^{(M,\delta)}. \tag{35}$$

For any $z \in Z^{(M)}$ there is $L = L_z > 0$, such that for any $\varepsilon < \varepsilon_0(M, L)$, $\mathrm{Rel}_z M$ can be obtained from M by modifying $\mathcal{N}(L,\varepsilon)$. Moreover, the function $z \mapsto L_z$ can be taken to be bounded when z varies in a bounded subset of $Z^{(M)}$.

Proof. First note that $Z^{(M)} \subset \bigcap_{\delta \in \Xi(M)} Z^{(M,\delta)}$. Indeed, if $z \in Z^{(M)}$ then the straight line path $\{\mathrm{Rel}_{tz} M : t \in [0,1]\}$ is embedded in $\mathcal{H}$. In particular, for any saddle connection $\delta \in \Xi(M)$,

$$\mathrm{hol}(M,\delta) + tz(\delta) = \mathrm{hol}(\mathrm{Rel}_{tz} M, \delta) \neq 0.$$

So the path $\{tz : t \in [0,1]\}$ is contained in $Z^{(M,\delta)}$ and hence $z \in Z^{(M,\delta)}$.

Now let $z \in \bigcap_{\delta \in \Xi(M)} Z^{(M,\delta)}$. Recall that we may think of an element z of $Z \cong H^0(\Sigma;\mathbb{R})/H^0(S;\mathbb{R})$ as being represented by a function $\bar{z} : \Sigma \to \mathbb{R}$. All such representatives $\bar{z}$ differ by constants, and

$$z_{ij} = \bar{z}(\xi_j) - \bar{z}(\xi_i) = z(\delta_{ij}), \tag{36}$$

for any path $\delta_{ij} : [0,1] \to S$ with $\delta_{ij}(0) = \xi_i$, $\delta_{ij}(1) = \xi_j$. We fix one representative $\bar{z}$ of z, and let

$$L > \max_{i=1,\dots,k} |\bar{z}(\xi_i)|. \tag{37}$$

Our assumption about z implies that

$$\text{if } \delta \text{ is from } \xi_i \text{ to } \xi_j \text{ and } \mathrm{hol}(M,\delta) > 0, \text{ then } \mathrm{hol}(M,\delta) + z_{ij} > 0. \tag{38}$$

Let $\varepsilon < \varepsilon_0(M, L)$ and let $\mathcal{N}_0 = \mathcal{N}(L, \varepsilon)$. Choose a marking $(f, M) \in \mathcal{H}_{\mathrm{m}}$ of M. For each $t \in [0,1]$, we will define a translation surface M_t and a marking $f_t : S \to M_t$, as follows. To each rectangle $R = R_e^{\pm}$ in $\mathcal{N}_0$ we define a trapezoid $R_t = R_{t,e}^{\pm}$, by choosing a plane development $\bar{R}$ of R and moving the vertices of $\bar{R}$ horizontally, where points of $\bar{R}$ which come from $\partial\mathcal{N}_0$ are not moved, and points which correspond to the singularity ξ_i are moved by adding $t\bar{z}(\xi_i)$ to their horizontal component. See Figures 9 and 10.

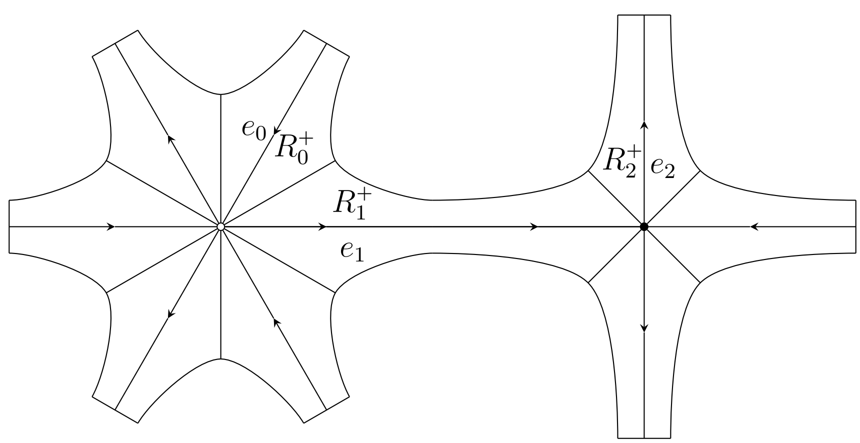
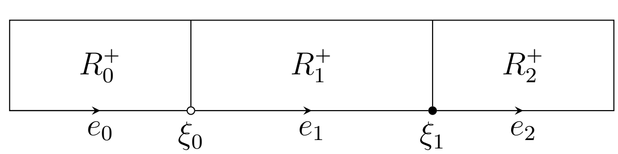


FIGURE 9. The complex $\mathcal{N}_0$, presented topologically, showing three adjacent rectangles. The rectangles are presented with correct geometries below.



Note that for any $t \in [0,1]$, the lengths of edges of R_t do not vanish. Indeed, $R_e^{\pm}$ has either one or two vertices which are in Σ, depending on whether e is a prong edge or an edge of $\Xi(M)$. In case it is a prong edge, the length of horizontal sides of R is L and (37) implies that the sidelength of R_t is positive, and in case it is an edge of $\Xi(M)$,

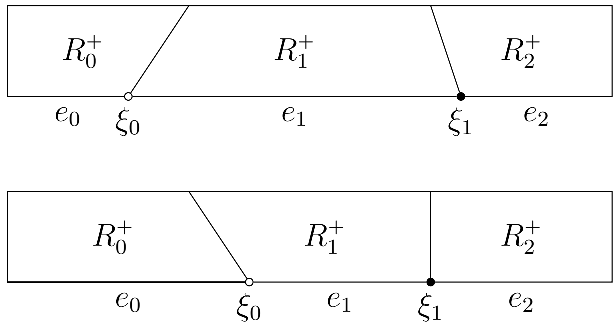


FIGURE 10. The effect of different deformations, transforming the rectangles into trapezoids.

(38) implies that the sidelength of R_t is positive. We glue the different trapezoids $\{R_t\}$ to each other along their edges, using the same gluing that defines $\mathcal{N}_0$, to obtain a complex $\mathcal{N}_t$. Since the sides of R corresponding to $\partial\mathcal{N}_0$ have the same length in R_t, their boundaries $\partial\mathcal{N}_t$ and $\partial\mathcal{N}_0$ can be identified by a translation, and so we can glue $\mathcal{N}_t$ to $M \smallsetminus \mathcal{N}_0$ along $\partial\mathcal{N}_0$. We denote the resulting translation surface by M_t; clearly it is obtained from M by modifying $\mathcal{N}_0$.

On each rectangle R we choose a homeomorphism $\bar{f}_t : R \to R_t$ which sends vertices of R to the corresponding vertices of R_t, and acts affinely on each boundary edge of ∂R_t. This choice ensures that $\bar{f}_t$ can be extended consistently from rectangles to their union, defining a map $\bar{f}_t : \mathcal{N}_0 \to \mathcal{N}_t$, and then extended to a homeomorphism $\bar{f}_t : M \to M_t$. We set $f_t = \bar{f}_t \circ f$. With this choice (f_t, M_t) is a path in $\mathcal{H}_{\mathrm{m}}$, and the maps $f_t \circ f^{-1}$ are isotopic to the identity via an isotopy fixing Σ.

We claim that for each t, the pullback $\mathrm{dev}(f_t, M_t) = f_t^*\mathrm{hol}(M_t, \cdot)$ is the cohomology class $\mathrm{dev}(f, M) + tz$. We verify this formula on each path $\gamma : [0,1] \to M$ between singularities ξ_i and ξ_j. The path γ is homotopic to a concatenation of segments $\delta_1, \ldots, \delta_\ell$ which are completely contained in $\mathcal{N}_0$, and segments $\delta'_1, \ldots, \delta'_m$ which are not completely contained in $\mathcal{N}_0$. Each of the segments $\delta = \delta_i$ is homotopic to a saddle connection in $\Xi(M)$. If δ is a saddle connection of $\Xi(M)$ then it is one of the edges e of the rectangles R_e, and by construction $\mathrm{hol}(f_t^* M_t, \delta) = \mathrm{hol}(f^*M, \delta) + tz(\delta)$. If $\delta' = \delta'_\ell$ is a path not contained in $\mathcal{N}_0$ we can subdivide it into a concatention of paths $\sigma_1, \sigma_2, \sigma_3, \sigma_4$, where σ_1 goes from ξ_1 to $\partial\mathcal{N}_0$, σ_2 goes from $\partial\mathcal{N}_0$ to ξ_j, σ_3 (resp. σ_4) is a union of segments inside (resp. outside) $\mathcal{N}$ from $\partial\mathcal{N}_0$ to $\partial\mathcal{N}_0$. Moreover by applying a homotopy we can assume that on the initial surface M, each of the segments in σ_1, σ_2 and σ_3 proceeds along a vertical line in the rectangles $R_e^\pm$. Now we compute the difference $\mathrm{hol}(f_t^* M_t, \sigma) - \mathrm{hol}(f^*M, \sigma)$

in each case. We have $\mathrm{hol}(f_t^*M_t, \sigma_1) - \mathrm{hol}(f^*M, \sigma_1) = -\bar{z}(\xi_i)$ and $\mathrm{hol}(f_t^*M_t, \sigma_2) - \mathrm{hol}(f^*M, \sigma_2) = \bar{z}(\xi_j)$ by definition of the deformed flat structure on $R_{t,e}^{\pm}$. We also have $\mathrm{hol}(f_t^*M_t, \sigma_4) - \mathrm{hol}(f^*M, \sigma_4) = 0$, since σ_4 is in the complement of $\mathcal{N}_0$ where the two flat structures are the same, and we have $\mathrm{hol}(f_t^*M_t, \sigma_3) - \mathrm{hol}(f^*M, \sigma_3) = 0$ since each of the segments of σ_3 passes through both R_e^+ and R_e^- for some e, leaving and exiting at symmetric points, and the change to the holonomy in these two rectangles cancel each other. All together we have $\mathrm{hol}(f^*M, \delta') - \mathrm{hol}(f_t^*M_t, \delta') = \bar{z}(\xi_j) - \bar{z}(\xi_i)$, as required. □

The following result says that the only obstruction to defining the Rel flow is the one illustrated in Figure 3. Let z_{ij} be defined as in equation (36).

Corollary 6.2. *Suppose Z is the real Rel subspace for a stratum $\mathcal{H}$ as above and let $z \in Z$, $M \in \mathcal{H}$. Then $M \in \mathcal{H}'_z$ exactly when there is no horizontal saddle connection δ on M from singularity ξ_i to singularity ξ_j, and $t \in [0,1]$ such that $\mathrm{hol}(M, \delta) + tz_{ij} = 0$. In particular $\mathcal{H}'_\infty$ is the set of surfaces which have no horizontal saddle connections joining distinct singularities.*

Furthermore, if $\mathcal{H}$ has two singularities, and $v \in \mathfrak{R} \cong \mathbb{R}^2$, then $M \notin \mathcal{H}'_v$ if and only if there is a saddle connection δ on M from ξ_2 to singularity ξ_1, with $\mathrm{hol}(M, \delta) = tv$ for some $t \in [0,1]$.

Proof. The first assertion is a restatement of (35), and the second assertion follows immediately from the first. For the third assertion, let r_θ be the rotation matrix for which $r_\theta v$ is horizontal. We obtain the assertion by applying the first assertion to the surface $r_\theta M$ and using Proposition 3.4 with $g = r_\theta$. □

Corollary 6.2 was proved in [MW2, Thm. 11.2]. See also [McM8].

Figures 3 and 4 illustrate the meaning of Corollary 6.2. In Figure 3 the saddle connection at the top and bottom of the decagon violates the first condition, when $v = (T, 0)$ for T which is at least as large as the length of this segment. In Figure 4 there are no horizontal saddle connections joining distinct singularities, and as a consequence of Corollary 6.2, $\mathrm{Rel}_{(T,0)}M$ is defined for all T.

We now derive some consequences. As a first consequence we have:

Proof of Proposition 3.9. Each $Z^{(M,\delta)}$ is a half-space, and in particular is convex. Thus Proposition 3.9 follows immediately from (35). □

An immediate consequence of the explicit surgery we have presented in the proof of 6.1 is the following useful statement:

Corollary 6.3. *For any M and any t for which $\mathrm{Rel}_t M$ is defined, there is a natural bijection between horizontal saddle connections on M and on $\mathrm{Rel}_t M$, and for each saddle connection δ directed from ξ_2 to ξ_1, $\mathrm{hol}(\mathrm{Rel}_t M, \delta) = \mathrm{hol}(M, \delta) - (t, 0)$. In particular M and $\mathrm{Rel}_t M$ are topologically horizontally equivalent.*

The following will be crucial for analyzing U-invariant measures in §8.

Definition 6.4. *Suppose $\mathcal{H} = \mathcal{H}(1,1)$, and let $T \in Z \cong \mathbb{R}$. Let $\mathcal{H}''_T$ denote the set of surfaces $M \in \mathcal{H}$ for which*

(i) *M contains exactly one directed saddle connection δ' from ξ_2 to ξ_1 with $\mathrm{hol}(M, \delta') = (T, 0)$;*
(ii) *M contains no directed saddle connection δ from ξ_2 to ξ_1, such that $\mathrm{hol}(M, \delta) = (c, 0)$ with c between 0 and T.*

Theorem 6.5. *There is a map*

$$\Phi : \mathcal{H}''_T \to \mathcal{H}(2)$$

which is affine in charts (hence continuous) and U-equivariant. For each $M \in \mathcal{H}''_T$, $\Phi(M)$ is obtained by modifying $\mathcal{N}(\varepsilon, L)$ for some $\varepsilon > 0, L > T$ depending on M. There is a map

$$\Phi_{\mathrm{f}} = \Phi_{\mathrm{f}}^{(T)} : \mathcal{H}''_T \to \mathcal{H}_{\mathrm{f}}(2),$$

which is a lift of Φ (i.e. $\Phi = P \circ \Phi_{\mathrm{f}}$ where $P : \mathcal{H}_{\mathrm{f}}(2) \to \mathcal{H}(2)$ is the projection), and Φ_{f} is a homeomorphism onto its image.

Suppose $T > 0$. Note that assumption (ii) implies that $[0, T) \subset Z^{(M)}$, and assumption (i) implies that $T \notin Z^{(M)}$; that is $T \in \partial Z^{(M)}$. A topology on $\mathcal{H}(1,1) \cup \mathcal{H}_{\mathrm{f}}(2)$ can be constructed in which the map Φ can be recovered as $\Phi(M) = \lim_{s \to T-} \mathrm{Rel}_s(M)$. We will not construct this topology here.

Proof. We will assume throughout the proof that $T > 0$. The case in which $T < 0$ can be dealt with by repeating the arguments below, switching the labels of the two singularities. Let δ' be as in (i) in the definition of $\mathcal{H}''_T$ and let e' be the corresponding edge of $\Xi_-(M)$. Let $L > |T|$ and define $\mathcal{N}(L, \varepsilon)$ as in the discussion prior to the statement of Theorem 6.1, where $\varepsilon > 0$ is small enough so that the rectangles $R_e^{\pm}$ are all embedded in M. Define the polygons $R_{T,e}^{\pm}$ as in the proof of Theorem 6.1. For all $e \neq e'$ condition (ii) ensures that $R_{T,e}^{\pm}$ is a nondegenerate trapezoid. For e', the length of the edge corresponding to e' is zero and so it is a triangle. See Figure 11. We glue the two triangles $R_{T,e'}^{\pm}$ to the trapezoids $R_{T,e}^{\pm}, e \neq e'$ to obtain the complex $\mathcal{N}_T$

which we glue to $M \smallsetminus \mathcal{N}_0$ as before, to obtain M_T. We note that $\mathcal{N}_T$ is a thickening of a graph obtained from $\Xi_-(M, L)$ by collapsing the edge e'. One can compute explicitly that there is one singular point in $\mathcal{N}_T$ and that it has cone angle 6π; that is $M_T \in \mathcal{H}(2)$. We set $\Phi(M) = M_T$.

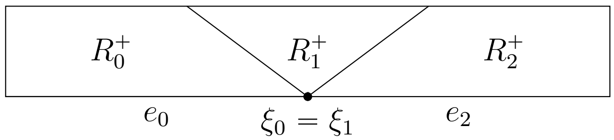


FIGURE 11. The developing image of degenerate rectangles. The trapezoid R_1^+ has degenerated to a triangle, the edge $e' = e_1$ has disappeared, and the two singular points have coalesced.

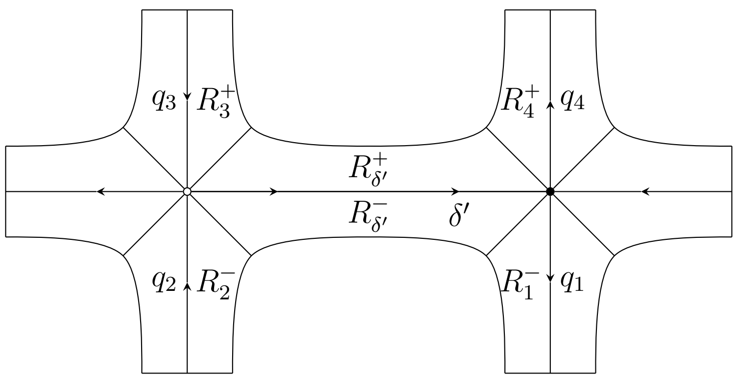


FIGURE 12. The complex $\mathcal{N}_0$, shown topologically, with δ' connecting the two singularities.

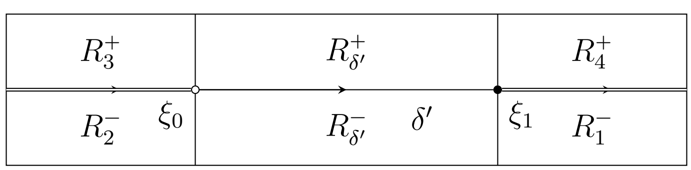


FIGURE 13. The developing image of rectangles. The double lines represent two different edges on the surface.

We now show that Φ is affine in charts. We first explain what this means. Let $\mathcal{H}''_{\mathrm{m},T}$ be the pre-image of $\mathcal{H}''_T$ in $\mathcal{H}_{\mathrm{m}}(1,1)$, and let $\mathrm{dev}_{(1,1)} : \mathcal{H}_{\mathrm{m}}(1,1) \to H^1(S, \{\xi_1, \xi_2\}; \mathbb{R}^2)$ be the developing map as in (1). Condition (i) in Definition 6.4 can be expressed as a linear condition on the image of $\mathrm{dev}_{(1,1)}$ and condition (ii) is an open

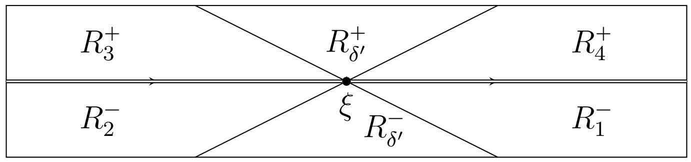


FIGURE 14. The developing image of rectangles when δ' shrinks to a point and the two singularities coalesce.

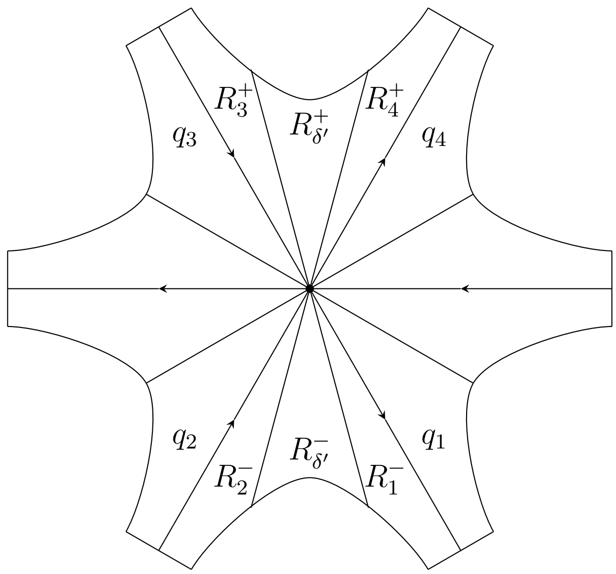


FIGURE 15. The corresponding topological correct picture. The chosen prong edge is labeled q_1.

condition on the image of $\mathrm{dev}_{(1,1)}$, and thus $\mathcal{H}''_{\mathrm{m},T}$ is an affine submanifold of $\mathcal{H}_{\mathrm{m}}(1,1)$. Also let $\mathrm{dev}_{(2)} : \mathcal{H}_{\mathrm{m}}(2) \to H^1(S', \{\xi'\}; \mathbb{R}^2)$, for some model surface S' of genus 2 with one distinguished point ξ'. To say that Φ is affine in charts is to say that that there is a map $\Phi_{\mathrm{m}} : \mathcal{H}''_{\mathrm{m},T} \to \mathcal{H}_{\mathrm{m}}(2)$, which is a lift of Φ, and a linear map $L : H^1(S, \{\xi_1, \xi_2\}; \mathbb{R}^2) \to H^1(S', \{\xi'\}; \mathbb{R}^2)$ such that

$$\mathrm{dev}_{(2)} \circ \Phi_{\mathrm{m}} = L \circ \mathrm{dev}_{(1,1)}. \tag{39}$$

Let $(f, M) \in \mathcal{H}''_{\mathrm{m},T}$ be a marked surface projecting to $M \in \mathcal{H}''_T$, and define the map $\bar{f}_T : M \to M_T$ as in the proof of Theorem 6.1. Let $f_T = f \circ \bar{f}_T : S \to M_T$. The map $\bar{f}_T$ is injective on the complement of e' and maps all points in e', including its endpoints ξ_1 and ξ_2, to the unique singularity of M_T, which we denote by ξ. Note that $f^{-1}(e')$ is a simple path connecting the two points ξ_1 and ξ_2. Let S' be the surface obtained from S by collapsing $f^{-1}(e')$ to a point ξ', and let

$p : S \to S'$ be the quotient map. Since $f^{-1}(e')$ is contractible, S' is also a genus 2 surface, and since f_T is constant on $f^{-1}(e')$, it descends to a homeomorphism $S' \to M_T$, which we continue to denote by f_T. We see that

$$\Phi_{\mathrm{m}} : \mathcal{H}''_{\mathrm{m},T} \to \mathcal{H}_{\mathrm{m}}(2), \quad \Phi_{\mathrm{m}}(f, M) = (f_T, M_T)$$

is a lift of Φ. Since $f^{-1}(e')$ is contractible, the pullback $p^* : H^1(S'; \mathbb{R}^2) \to H^1(S; \mathbb{R}^2)$ is an isomorphism. Let $\mathrm{Res}_{(1,1)}$ and $\mathrm{Res}_{(2)}$ denote the restriction maps in (11), in the two cases corresponding respectively to $\mathcal{H}(1,1)$ and $\mathcal{H}(2)$. Since ξ_1, ξ_2 are contained in $f^{-1}(e')$, and the holonomies of absolute periods are on the same on (f, M) and (f_T, M_T), we find that $\mathrm{Res}_{(2)} \circ \mathrm{dev}_{(2)} \circ \Phi_{\mathrm{m}} = p^* \circ \mathrm{Res}_{(1,1)} \circ \mathrm{dev}_{(1,1)}$. Since $\mathrm{Res}_{(2)}$ is an isomorphism, this yields (39) with $L = \mathrm{Res}_{(2)}^{-1} \circ p^* \circ \mathrm{Res}_{(1,1)}$. This computation, and the fact that the action of U does not move the points of e', also shows that Φ is U-equivariant.

We now show that Φ lifts to a continuous map $\Phi_{\mathrm{f}} : \mathcal{H}''_T \to \mathcal{H}_{\mathrm{f}}(2)$. In view of the discussion in §2.9, in order to lift Φ to a map to $\mathcal{H}_{\mathrm{f}}(2)$ we need to equip M_T with a right-pointing horizontal prong at the singular point ξ. Let $\delta' = \delta'(M)$ be as in the definition of $\mathcal{H}''_T$ and let $q(M)$ be the prong which is obtained by moving an angle π in the counterclockwise direction from the terminal prong of δ', at the singularity ξ_2. Then $q(M)$ is in the complement of δ' and so is mapped by $\bar{f}_T$ to a horizontal prong on M_T. See Figures 12 and 15, where $q(M)$ is marked as q_1. We need to show that with this choice of selected prong, Φ_{f} is continuous. In light of Proposition 2.8, it is enough to show two things: (i) that the choice of prong $M \mapsto q(M)$ is continuous with respect to the coordinates given by the developing map, for any fixed triangulation; and (ii), that $M \mapsto q(M)$ is $\mathrm{Mod}(\breve{S})$-invariant. It is clear from our description of $q(M)$ that (i) and (ii) are satisfied.

Finally, since Φ_{f} is a lift of a locally affine map in charts, in order to show that it is a homeomorphism onto its image we need only verify that it is injective, and for this we explicitly construct its inverse. We first pick one basepoint $M_0' = \Phi_{\mathrm{f}}(M_0)$ in each connected component of the image of Φ, choose a horizontal prong at M_0' using the direction of δ in M_0 as in the preceding paragraph, and extend this choice continuously to all surfaces in the image of Φ. Now for any M' in the image of Φ, let $\mathcal{N}'$ be the (L, ε)-rectangle thickening of $\Xi_-(M')$ for $L > |T|$ and small enough ε. We consider $\mathcal{N}'$ with its decomposition into rectangles as in the preceding discussion, and we now modify this decomposition. Let q_1 be the chosen right-pointing prong on M' and let q_2 be the left-pointing prong which is clockwise from q_1. Similarly let q_3, q_4 be the left- and right-pointing prongs which are at angular distance 3π

from q_1, q_2 respectively. Let σ_1 (resp. σ_2) be the two vertical segment of lengths ε between q_1, q_2 (resp., between q_3 and q_4) connecting ξ to the boundary of $\partial N'$. For $i = 1, 2$, the segments σ_1, σ_2 are boundary segments of two rectangles $R_i^{\pm}$ of the complex $\mathcal{N}'$, one on each side. Let Δ_i be a triangle which is embedded in the union of $R_i^+ \cup R_i^-$, has an apex at ξ, and has σ_i as an altitude contained in its interior. We now replace each $R_i^{\pm}$ with $R_i^{\pm} \smallsetminus \Delta_i$, and add the Δ_i to the polygonal decomposition of $\mathcal{N}'$. Thus we have a decomposition of $\mathcal{N}'$ into rectangles, trapezoids, and two triangles. To each of them we apply the map described in the proof of Theorem 6.1, with $-T$ instead of T. That is, we do not move points on $\partial\mathcal{N}'$ and the non-boundary edges. The two triangles are thought of as degenerate trapezoids. The choice of the prongs at ξ, and the fact that M' is in the image of Φ, ensure that these operations are well-defined, that is for all t strictly between 0 and T, the deformed shapes are nondegenerate trapezoids. Gluing them to each other using the gluing map of $\mathcal{N}'$ and gluing the resulting complex to $M \smallsetminus \mathcal{N}'$ completes the definition of the inverse of Φ_{f}. $\square$

The image of Φ in Theorem 6.5 can be described explicitly in terms of the choice of horizontal prong at the singularity. Namely suppose again that $T > 0$ and that q_1 is the chosen prong. Let $q_2, \ldots, q_6$ be the additional prongs at ξ in counterclockwise order. Then the image of Φ is the set of $M \in \mathcal{H}_{\mathrm{f}}(2)$ which have no horizontal saddle connections of length at most T from q_1 or q_3 to q_4 or q_6.

The inverse of Φ appearing in Theorem 6.5 is the operation of 'splitting open a singularity' which was discussed in [EMZ].

7. The eigenform locus

In this section we will define the eigenform locus $\mathcal{E}_D$. We describe its intersection with $\mathcal{H}(1,1)$ and $\mathcal{H}(2)$ and describe how it meets some boundary strata. We summarize some properties of surfaces in the eigenform locus.

7.1. **Definition of the eigenform loci.**

The eigenform loci were defined by Calta [C] and McMullen [McM2]. Calta made use of the J invariant and McMullen used properties of real multiplication on Jacobians. Here we follow the approach of McMullen.

For every positive integer $D \equiv 0, 1 \pmod 4$ with $D \geqslant 4$ there is a closed, connected, G-invariant locus $\mathcal{E}_D \subset \mathcal{H}(2) \cup \mathcal{H}(1,1)$, called the *eigenform locus*, which we now describe.

An *order* in a number field F is a subring $\mathcal{O}$ of the ring of integers $\mathcal{O}_F$ which is finite index as an abelian group. Orders in quadratic fields are particularly simple as they can be classified by a single integer D, the *discriminant*. More precisely, for every positive integer $D \equiv 0, 1 \pmod 4$, we consider the real quadratic order

$$\mathcal{O}_D = \mathbb{Z}[T]/(T^2 + bT + c),$$

where $b, c \in \mathbb{Z}$ are such that $b^2 - 4c = D$. If D is not square, it is a subring of the real quadratic field, $F_D = \mathbb{Q}[T]/(T^2 + bT + c)$. We also allow D to be square, in which case F_D is isomorphic to $\mathbb{Q} \oplus \mathbb{Q}$ as a $\mathbb{Q}$-algebra. In either case the isomorphism classes of $\mathcal{O}_D$ and F_D depend only on D. We fix a choice of a ring homomorphism $\iota\colon F_D \to \mathbb{R}$. When D is not square, ι is a field embedding. If D is square, ι is obtained from the projection of $\mathbb{Q} \oplus \mathbb{Q}$ onto its first factor. A more detailed discussion of orders in number fields appears in [BoSh].

Consider a genus two Riemann surface X with Jacobian variety $\mathrm{Jac}(X) = \Omega(X)^*/H_1(X;\mathbb{Z})$, where $\Omega(X)$ is the space of holomorphic one-forms on X. *Real multiplication* by $\mathcal{O}_D$ on $\mathrm{Jac}(X)$ is a ring monomorphism $\rho\colon \mathcal{O}_D \to \mathrm{End}^0(\mathrm{Jac}(X))$, where $\mathrm{End}^0(\mathrm{Jac}(X))$ is the ring of endomorphisms of $\mathrm{Jac}(X)$ which are self-adjoint with respect to the intersection form on $H_1(X;\mathbb{Z})$. We also require ρ to be *proper*, in the sense that it does not extend to $\mathcal{O}_E \supsetneq \mathcal{O}_D$ for some $E|D$.

Real multiplication by $\mathcal{O}_D$ induces a representation of $\mathcal{O}_D$ on $\Omega(X)$, and by self-adjointness, a decomposition of $\Omega(X)$ into complementary eigenspaces. A nonzero holomorphic one-form on X is an eigenform if it lies in one of these eigenspaces. We say that a pair (X, ω) is an eigenform for real multiplication if $\mathrm{Jac}(X)$ has real multiplication with ω an eigenform.

Real multiplication for curves in genus two is very special, as it can be detected from knowledge of the absolute periods of a single one-form on the curve. That is to say, real multiplication in genus two has a "purely flat" description. More precisely:

Proposition 7.1 ([B2]). *A genus two translation surface M is an eigenform for real multiplication by $\mathcal{O}_D$ if and only if there is a proper monomorphism $\rho_0\colon \mathcal{O}_D \to \mathrm{End}^0(H_1(M;\mathbb{Z}))$ such that*

$$\mathrm{hol}(M, \rho_0(\lambda) \cdot \gamma) = \iota(\lambda)\mathrm{hol}(M, \gamma) \tag{40}$$

for each $\lambda \in \mathcal{O}_D$ and $\gamma \in H_1(M;\mathbb{Z})$.

The ρ_0 in this Proposition is simply the action on homology induced by the real multiplication $\rho\colon \mathcal{O}_D \to \mathrm{End}^0(\mathrm{Jac}(M))$. See also [McM2, Lemma 7.4], and see [CaSm] for an alternative approach.

We define the eigenform locus $\mathcal{E}_D \subset \mathcal{H}(2) \cup \mathcal{H}(1,1)$ to be the locus of eigenforms for real multiplication by $\mathcal{O}_D$, and we define $\mathcal{E}_D(2)$

and $\mathcal{E}_D(1,1)$ to be the intersections of $\mathcal{E}_D$ with the respective strata. Similarly we denote the corresponding subsets of area-one surfaces by $\mathcal{E}_D^{(1)}(2)$ and $\mathcal{E}_D^{(1)}(1,1)$.

The locus $\mathcal{E}_D(1,1)$ is $\mathrm{GL}_2(\mathbb{R})$-invariant, as it can be easily seen that the condition of Proposition 7.1 is G-invariant (this was first proved in [McM2] and [C]). It is also Rel invariant since this condition only involves absolute periods. Moreover $\mathcal{E}_D(1,1)$ is a six dimensional linear submanifold of $\mathcal{H}(1,1)$ with respect to the period coordinates from §2.2. To see this explicitly, choose two generators γ_1, γ_2 of $H_1(M;\mathbb{Z})$ (as an $\mathcal{O}_D$-module) and complete to a set of four generators as a $\mathbb{Z}$-module, e.g. by adjoining a multiple of each γ_i by a generator of $\mathcal{O}_D$ over $\mathbb{Z}$. Equation (40) now gives linear equations which the vectors $\mathrm{hol}(M,\gamma)$ must satisfy, and these equations define $\mathcal{E}_D(1,1)$ locally. As a consequence $\mathcal{E}_D^{(1)}(1,1)$ is a five-dimensional manifold locally defined in periodic coordinates by linear equations and one quadratic equation.

This dimension count easily implies that $\mathrm{GL}_2(\mathbb{R})$-orbits and Rel leaves locally fill out $\mathcal{E}_D(1,1)$ and yields:

Proposition 7.2. *For any $M \in \mathcal{E}_D^{(1)}(1,1)$ there is a neighborhood $\mathcal{U}$ of the identity in L and a neighborhood $\mathcal{U}'$ of M in $\mathcal{E}_D^{(1)}(1,1)$ such that the map $p\colon \mathcal{U} \to \mathcal{U}'$ defined by*

$$p(g,v) = gM \mathbin{\text{⨁}} v$$

is the restriction of an affine homeomorphism to $G \ltimes \mathbb{R}^2$.

Proof. Consider a precompact neighborhood $\mathcal{W}$ of the identity in $\mathrm{GL}_2(\mathbb{R})$. For some $\varepsilon > 0$, no surface in $\mathcal{W} \cdot M$ has saddle connections of length less than ε. By Corollary 6.2, p is well-defined on $\mathcal{V} = \mathcal{W} \times B_\varepsilon(0) \subset \mathrm{GL}_2(\mathbb{R}) \ltimes \mathbb{R}^2$. Possibly decreasing ε so that the image of p is contained in an affine coordinate chart as defined above, p is a homeomorphism onto its image. Since the action of Rel preserves the area of surfaces, p sends L into the locus of area-one surfaces. Intersecting $\mathcal{V}$ with L and the image of p with the locus of area-one surfaces, we obtain $\mathcal{U}$ and $\mathcal{U}'$ with the required properties. $\square$

The eigenform locus has a more elementary description in the case where D is a square. A translation surface X is a *torus cover* if there is a branched cover $p\colon X \to T$ which is a local translation for some flat torus T (note that the branch points of p are not required to lie over a single point of T). We say p is *primitive* if it does not factor through a torus cover of smaller degree, equivalently if the map on homology $p_*\colon H_1(X;\mathbb{Z}) \to H_1(T;\mathbb{Z})$ is onto. McMullen established in [McM3] that $\mathcal{E}_{d^2}$ is the locus of primitive degree d torus covers.

7.2. **G-invariant measures in genus two.** We now discuss the G-invariant measures in genus two. These were classified by McMullen in [McM3]. Measures supported on the full strata were constructed by Masur [Ma] and Veech [Ve1] using period coordinates on these strata. In [McM3] McMullen constructed measures on the eigenform loci in an analogous way using period coordinates.

We may use Proposition 7.2 to define a measure on $\mathcal{E}_D(1,1)$ by locally pushing forward Haar measure on L. More precisely, given E in the image $\mathcal{U}'$ of p, we assign to E the Haar measure of $p^{-1}(E) \subset L$ (the L invariance of Haar measure implies that the measure of E doesn't depend on the choice of basepoint). We call this the *flat measure* on $\mathcal{E}_D(1,1)$. McMullen proved that this measure is finite and G-invariant [McM3].

Here is an alternative description of the flat measure which can be generalized to define a measure on any linear submanifold of a stratum. Suppose $f\colon S \to M$ is a marked translation surface, write $M = (X, \omega)$, and suppose that $\mathrm{Jac}(X)$ admits real multiplication by $\mathcal{O}_D$. The real multiplication on $\mathrm{Jac}(X)$ gives $H_1(S;\mathbb{Z})$ the structure of an $\mathcal{O}_D$-module. A choice of embedding $\mathcal{O}_D \to \mathbb{R}$ makes $\mathbb{R}^2$ an $\mathcal{O}$-module as well. We define $H^1_{\mathcal{O}_D}(S,\Sigma;\mathbb{R}^2) \subset H^1(S,\Sigma;\mathbb{R}^2)$ to be the subspace of cocycles for which the induced period map $H_1(S;\mathbb{Z}) \to \mathbb{R}^2$ is $\mathcal{O}_D$-linear. This is in other words the space of cocycles satisfying (40). By Proposition 7.1, the linear subspace $H^1_{\mathcal{O}_D}(S,\Sigma;\mathbb{R}^2)$ parameterizes the eigenform locus in $\mathcal{H}_{\mathrm{m}}$ near M.

We have the commutative diagram of homology groups (all coefficients in $\mathbb{R}^2$):

$$\begin{tikzcd}
0 \arrow[r] & H^0(\Sigma)/H^0(S) \arrow[r] \arrow[d] & H^1_{\mathcal{O}_D}(S,\Sigma) \arrow[r] \arrow[d] & H^1_{\mathcal{O}_D}(S) \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & H^0(\Sigma)/H^0(S) \arrow[r] & H^1(S,\Sigma) \arrow[r] & H^1(S) \arrow[r] & 0
\end{tikzcd}$$

We give $\mathcal{E}_D(1,1)$ a measure by defining a measure on the linear space $H^1_{\mathcal{O}_D}(S,\Sigma)$ on which it is modeled which is invariant under the monodromy action. To define such a measure, we give a measure on the other two terms of the short exact sequence. The left term is canonically $\mathbb{R}^2$, and the monodromy action is trivial. We give it the usual Euclidean area form. The $H^1(S)$ term has a symplectic form arising from the intersection form, which is preserved by the action of the mapping class group. This descends to a symplectic form on $H^1_{\mathcal{O}_D}(S)$ which is preserved by monodromy. This form is non-degenerate; this could be checked by direct computation, and was proved in complete generality by Avila, Eskin and Möller [AEM]. Therefore the symplectic form defines a volume form on $H^1_{\mathcal{O}_D}(S)$ which is also monodromy invariant. The product of these volume forms induces one on $H^1_{\mathcal{O}_D}(S,\Sigma)$ which defines a measure on $\mathcal{E}_D(1,1)$. Finally, we apply the standard cone construction (meaning we push forward the restriction of the measure to surfaces in $\mathcal{E}_D(1,1)$ of area at most 1, by the canonical projection onto the locus of surfaces of area one; see [Zo] for details) to obtain a finite G-invariant measure on $\mathcal{E}_D(1,1)$ which is supported on the area one forms.

The eigenform locus $\mathcal{E}_D(1,1)$ is nonempty and connected for $D \geqslant 4$ and $D \equiv 0$ or $1 \pmod 4$ by [McM2].

In the stratum $\mathcal{H}(2)$, the locus of eigenforms $\mathcal{E}_D(2)$, is called the Weierstrass curve in McMullen's papers. By [McM5], $\mathcal{E}_D(2)$ consists of a single G-orbit if $D \not\equiv 1 \pmod 8$ and $D \geqslant 5$ (note $\mathcal{E}_D(2)$ is empty for $D < 5$), or if $D = 9$. Otherwise $\mathcal{E}_D(2)$ consists of two orbits. It is equipped with a finite measure coming from Haar measure on G.

The square-discriminant eigenform locus $\mathcal{E}_{d^2}$ also contains a countable, dense collection of closed G-orbits. A translation surface X is called a *square-tiled surface* if it is a branched cover of the standard square torus with all of the branching lying over a single point. Every square-tiled surface has a closed G-orbit, and the square-tiled surfaces are dense in each $\mathcal{E}_{d^2}$.

Closed G-orbits inherit a measure from the Haar measure on G, and this measure is finite by a result of Smillie (see [SmWe2]). We will refer to this measure as the Haar measure on the closed G-orbit. The *decagon surface* is the surface obtained by identifying opposite sides of the regular 10-gon. It was shown by Veech [Ve2] that it has a closed G-orbit. It belongs to $\mathcal{H}(1,1)$ and in fact to the eigenform locus $\mathcal{E}_5(1,1)$. We write $\mathcal{L}_{\mathrm{dec}} \subset \mathcal{E}_5(1,1)$ for its G-orbit. The closed G-orbits in genus two were classified by McMullen in [McM2] and [McM4] (these closed G-orbits were constructed independently by Calta [C] using different methods):

Theorem 7.3. *Each connected component of $\mathcal{E}_D(2)$ is a closed G-orbit, and every closed G-orbit in $\mathcal{H}(2)$ is of this form.*

Every closed G-orbit in $\mathcal{H}(1,1)$ is either the G-orbit of a square-tiled surface or is $\mathcal{L}_{\mathrm{dec}}$.

In [McM3], McMullen showed that the measures defined above are the full list of ergodic G-invariant measures in genus two:

Theorem 7.4. *Every ergodic G-invariant measure on $\mathcal{H}(2)$ is either the flat measure on the full stratum or Haar measure on a closed G-orbit.*

Every ergodic G-invariant measure on $\mathcal{H}(1,1)$ is either the flat measure on the full stratum, the flat measure on some $\mathcal{E}_D(1,1)$, or the Haar measure on a closed G-orbit.

7.3. **Degenerate eigenform surfaces.** We will also be interested in eigenforms which are not genus two surfaces but can be thought of as surfaces lying in a bordification of $\mathcal{H}(1,1)$. We will consider two cases, where the role of "boundary strata" will be played respectively by $\mathcal{H}(0) \times \mathcal{H}(0)$ and $\mathcal{H}(0,0)$.

Given a pair of genus one translation surfaces E_1 and E_2, we may consider the one-point connected sum $X = E_1 \# E_2$. These degenerations arise from families of genus two surfaces where a separating curve has been pinched. As we have the direct sum decompositions $H_1(X) = H_1(E_1) \oplus H_1(E_2)$ and $\Omega(X) = \Omega(E_1) \oplus \Omega(E_2)$, the Jacobian of X is simply the product of E_1 and E_2:

$$\begin{aligned}\mathrm{Jac}(X) &= \Omega^*(X)/H_1(X;\mathbb{Z}) \\ &\cong \Omega^*(E_1)/H_1(E_1;\mathbb{Z}) \oplus \Omega^*(E_2)/H_1(E_2;\mathbb{Z}) \\ &\cong E_1 \times E_2\end{aligned}$$

Just as for the smooth case, we say that X is an eigenform for real multiplication if $\mathrm{Jac}(X)$ has real multiplication, with the one-form defining the translation structure belonging to one of the eigenspaces in $\Omega(X)$. McMullen gave a more explicit description of real multiplication for these degenerate surfaces in terms of isogenies of the E_i.

Recall that E_1 and E_2 are *isogenous* if there is a holomorphic covering map $p\colon E_1 \to E_2$. The isogeny p is *primitive* if it cannot be written as a composition of an isogeny of lower degree with a self-covering of E_2. Existence of p yields a dual isogeny $\bar{p}\colon E_2 \to E_1$, so isogeny is an equivalence relation. In translation coordinates, an isogeny p is of the form $p(z) = \lambda z + c$ for some complex number λ which we call the *scaling factor* of p. In our setting, λ will always be real, in which case p preserves the horizontal direction but scales the metric by a factor of λ.

Proposition 7.5 ([McM5]). *The surface $E_1 \# E_2$ is an eigenform for real multiplication by $\mathcal{O}_D$ if and only if there exists a primitive degree m isogeny $p\colon E_1 \to E_2$, together with an integral solution (e,ℓ) to the equation $e^2 + 4m\ell^2 = D$ with $\ell > 0$ and $\gcd(e,\ell) = 1$, such that the scaling factor λ is the unique real positive root to the equation $\lambda^2 - e\lambda - \ell^2 m = 0$.*

We define $\mathcal{P}_D \subset \mathcal{H}(0) \times \mathcal{H}(0)$ to be the locus of pairs (E_1, E_2) such that $E_1 \# E_2$ is an eigenform for real multiplication by $\mathcal{O}_D$. With the diagonal G-action on $\mathcal{H}(0) \times \mathcal{H}(0)$, the locus $\mathcal{P}_D$ is G-invariant by Proposition 7.5.

By [McM5], the locus $\mathcal{P}_D$ consists of finitely many closed G-orbits. We recall McMullen's classification of these G-orbits. A *prototype* for real multiplication by $\mathcal{O}_D$ is a triple of integers (e, ℓ, m) such that $D = e^2 + 4\ell^2 m$, with $\ell, m > 0$ and $\gcd(\ell, m) = 1$. A prototype (e, ℓ, m) determines a prototypical form in $\mathcal{P}_D$ as follows. Let λ be the unique positive solution of $\lambda^2 = e\lambda + \ell^2 m$. We define a pair of lattices in $\mathbb{C}$:

$$\Lambda_1 = \mathbb{Z}(\lambda, 0) \oplus \mathbb{Z}(0, \lambda) \quad \Lambda_2 = \mathbb{Z}(\ell m, 0) \oplus \mathbb{Z}(0, \ell),$$

and associated genus one translation surfaces $E_i = (\mathbb{C}/\Lambda_i, dz)$. Multiplication by λ defines an isogeny $p\colon E_1 \to E_2$ of degree $l^2 m$.

Proposition 7.6 ([McM5]). *Each G-orbit of $\mathcal{P}_D$ contains a unique prototypical form.*

Given a prototype (e, ℓ, m), we define $\mathcal{P}_D(e, \ell, m)$ to be the closed G-orbit containing the prototypical form associated to (e, ℓ, m). By the above proposition, this gives a bijection between prototypes and components of $\mathcal{P}_D$. We say that two pairs of genus-one forms *have the same combinatorial type* if they lie on the same component of $\mathcal{P}_D$. By [McM5, Theorem 2.1], the G-orbit corresponding to the prototype (e, ℓ, m) is isomorphic to the modular curve $G/\Gamma_0(m)$, where $\Gamma_0(m) \subset \mathrm{SL}(2, \mathbb{Z})$ is the group of matrices which are upper-triangular mod m.

Finally, in the case when $D = d^2$ with $d > 1$, we consider the moduli space $\mathcal{H}(0, 0)$ of genus one translation surfaces E with two marked points p and q. Again, there is a natural G-action on this space. We define $\mathcal{S}_{d^2} \subset \mathcal{H}(0, 0)$ to be the locus of (E, p, q) such that $p - q$ is exactly d-torsion in the group law on E (that is $d(p - q) = 0$ and $d'(p - q) \neq 0$ for any $d' < d$, in particular this implies $p \neq q$). By [B1], $\mathcal{S}_{d^2}$ is a closed G-orbit isomorphic to $G/\Gamma_1(d)$, where

$$\Gamma_1(d) = \left\{ \begin{pmatrix} a & b \\ c & e \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}) : a \equiv e \equiv 1 \pmod{d} \text{ and } c \equiv 0 \pmod{d} \right\}.$$

We can regard (E, p, q) as a degenerate degree d torus cover obtained by pinching a nonseparating curve. Since $p - q$ is a point of order d in the group law on E, we have a degree d map $\pi\colon E \to F$, where F is the quotient of E by the order d subgroup generated by $p - q$. This π is the torus cover of minimal degree sending p and q to the same point.

The precise sense in which surfaces in the loci $\mathcal{P}_D$, $\mathcal{S}_D$, and $\mathcal{E}_D(2)$ can all be regarded as lying in the boundary of $\mathcal{E}_D(1,1)$, is explained in [B1]. We will not explicitly use this point of view.

7.4. **Properties of eigenform surfaces.** We will require the following properties of surfaces in $\mathcal{E}_D(1,1)$. Recall that the *modulus* of a flat cylinder is defined to be its height divided by its circumference.

Proposition 7.7. *In any cylinder decomposition of a surface in* $\mathcal{E}_D(1,1)$ *with more than one cylinder, there is at least one rational relation among the moduli of horizontal cylinders.*

Proof. Given a horizontally periodic surface M with n cylinders, the orbit-closure $\overline{UM}$ is a torus of dimension $n-r$, where r is the number of independent rational relations among the moduli of the cylinders (see [SmWe1, Proposition 4]). We must then show that in any case the dimension of $\overline{UM}$ is smaller than n.

First, suppose $n = \dim \overline{UM} = 3$. We may parameterize a neighborhood of M in $\mathcal{H}(1,1)$ by the three holonomy vectors (x_i, y_i) of a saddle connection joining cone points on the two boundary components of the ith cylinder C_i. Since the y_i are constant on $\overline{UM}$ and $\dim \overline{UM} = 3$, the x_i are arbitrary on $\overline{UM}$. However, since $\overline{UM}$ is contained in the eigenform locus, the x_i satisfy some nontrivial real-linear equation, a contradiction.

Now suppose $n = \dim \overline{UM} = 2$. There must then be a horizontal saddle connection joining distinct zeros. Applying the real-rel flow so that the length of this saddle connection tends to 0, yields a surface M' in $\mathcal{H}(2)$ with two horizontal cylinders with the same moduli. A neighborhood of M' in $\mathcal{H}(2)$ is parameterized by two holonomy vectors (x_i, y_i). Again, if $\dim \overline{UM} = 2$, the x_i would be arbitrary on $\overline{UM'}$, which contradicts the real-linear equation imposed by the eigenform locus. (Note in the two-cylinder case, the claim also follows directly from [McM3, Theorem 9.1].) $\square$

Proposition 7.8. *The U-action on each* $\mathcal{E}_D(1,1)$, *with respect to the flat measure, is ergodic.*

Proof. In a G-action, ergodicity of the geodesic flow implies ergodicity of the U-action by the "Mautner phenomenon" (see e.g. [EW]). So it suffices to prove ergodicity of the G-action. This can be proved by applying the Hopf argument to the geodesic flow as in [Ma]. $\square$

Recall that a *periodic direction* of a translation surface is a direction in which the surface is a union of parallel cylinders and saddle connections. A translation surface is *completely periodic* if for any direction which contains a cylinder, the surface is a union of parallel cylinders and saddle connections.

Theorem 7.9 ([McM4, C])**.** *Every genus two surface M which is an eigenform is completely periodic. In particular for $M \in \mathcal{E}_D(1,1)$, if $\Xi(M)$ contains a saddle connection joining a zero to itself, then the horizontal direction of M is periodic.*

Given two genus one translation surfaces E_1 and E_2, consider a horizontal segment $I \subset \mathbb{R}^2$ which embeds into each E_i. We may form the *connected sum* $X = E_1 \#_I E_2$ by removing the image of I from each E_I, and then gluing the resulting boundary components. The result is a surface $X \in \mathcal{H}(1,1)$ with two horizontal saddle connections whose union is a loop separating X into two tori (see [McM3] for details).

Theorem 7.10. *Suppose that D is not square. Let $M \in \mathcal{E}_D$, and suppose M contains a loop which is a union of horizontal saddle connections. Then either*

- *The horizontal direction of M is periodic, or*
- *M is obtained by gluing two genus one translation surfaces with uniquely ergodic horizontal directions along a horizontal slit.*

In particular, in the second case $\Xi(M)$ consists of two horizontal saddle connections joining distinct zeros which are interchanged by the hyperelliptic involution (for the corresponding horizontal data diagram, see diagram 5 in Figure 7).

Proof. We regard M as a Riemann surface X equipped with a holomorphic one-form ω. By [McM1], if $\mathrm{Re}\,\omega$ has zero "Galois flux" and $\Xi(M)$ contains a loop, then we are in one of these two cases. Zero flux follows from equation (4.2) of [McM1] and Theorem 5.1 of [McM3].

The final statement is proved in the following lemma. $\square$

Lemma 7.11. *Let $M \in \mathcal{H}(1,1)$ be a surface with two horizontal saddle connections I and I' joining distinct zeros. The loop $\gamma = I \cup I'$ disconnects M if and only if the hyperelliptic involution interchanges I and I'.*

Proof. Let $\eta\colon M \to M$ denote the hyperelliptic involution and ω the holomorphic one-form on M induced by the translation structure.

Suppose that we are in the case where η interchanges I and I'. Since $\eta^*\omega = -\omega$, the map η sends γ to itself preserving the orientation. Thus $[\eta_*\gamma] = [\gamma] \in H_1(M;\mathbb{Z})$. But the hyperelliptic involution acts on $H_1(M;\mathbb{Z})$ as minus the identity. It follows that γ is homologous to zero, so it separates M.

Suppose now that η fixes I and I'. Since η reverses the orientation of each of these saddle connections, each must contain a single fixed point of η (i.e. one of the six Weierstrass points of M). If $f\colon M \to S^2$ is the quotient map induced by the hyperelliptic involution, then in this case the union $f(I) \cup f(I')$ is a smooth embedded path joining two branch points of f. Since an embedded path does not disconnect the sphere, and f has branch points disjoint from $f(I) \cup f(I')$, the complement $M \backslash (I \cup I')$ is also connected, a contradiction. ☐

When $D = d^2$, there is one more possible configuration of horizontal saddle connections. Suppose (E, p, q) is a genus one translation surface with two marked points p, q whose difference $p - q$ is exactly d-torsion in the group law of E. There is then a genus one surface F and degree d cover $\pi\colon E \to F$ which identifies p and q. Let $I \subset \mathbb{R}^2$ be a segment which may be embedded in E by a translation to yield disjoint parallel segments I' and I'' beginning at p and q respectively. We may then form the self-connected-sum $M = (E, p, q)\#_I$ by cutting along I' and I'' and then regluing to obtain a genus two surface. Since the gluings are compatible with the covering $\pi\colon E \to F$, the surface M is a primitive degree d branched cover of F, so $M \in \mathcal{E}_{d^2}(1,1)$. If the slope of I is not a periodic direction on E, the surface M has exactly two saddle connections of this slope of the same length, and the complement of this pair of saddle connections is a genus one surface. As in the second case in Theorem 7.10, the horizontal diagram is once again diagram 5 of Figure 7, but the complement of the slit has a different topology than in the case that D is not a square.

The following theorem says when D is square, this is the only additional configuration of saddle connections.

Theorem 7.12. *Suppose that $D = d^2$. Let $M \in \mathcal{E}_D$, and suppose M contains a loop which is a union of horizontal saddle connections. Then either*

- *The horizontal direction of M is periodic, or*
- *M is obtained by gluing two genus one translation surfaces with uniquely ergodic horizontal directions along a horizontal slit,*
- *M is a self-connected sum of (E, p, q) along two horizontal slits of the same length based at p and q respectively, where E a is genus one translation surface with uniquely ergodic horizontal direction, and $p - q$ is exactly d-torsion in the group law of E.*

In particular, in the second case $\Xi(M)$ consists of two horizontal saddle connections joining distinct zeros which are interchanged by the hyperelliptic involution. In the third case, $\Xi(M)$ consists of two horizontal *saddle connections of the same length fixed by the hyperelliptic involution.*

Proof. Suppose that M is not periodic. If M has a saddle connection joining a singularity to itself, this saddle connection is the boundary of a cylinder, which implies that the horizontal direction of M is periodic by Theorem 7.9, a contradiction. Therefore $\Xi(M)$ consists of two horizontal saddle connections I and I' joining singularity ξ_1 to singularity ξ_2.

Let $\pi\colon M \to F$ be a primitive degree d cover of a genus one translation surface. The horizontal direction of F is not periodic, otherwise M would be periodic as well. The singularities ξ_i are also the branch points of π, and the images $\pi(I)$ and $\pi(I')$ are horizontal segments on F joining $\pi(\xi_1)$ to $\pi(\xi_2)$. Since the horizontal direction of F is not periodic, the images $\pi(\xi_1)$ and $\pi(\xi_2)$ must be distinct. There is then a unique horizontal segment on F joining these points which I and I' are both mapped to injectively by π. It follows that I and I' have the same length.

By Lemma 7.11, the loop $I \cup I'$ disconnects M if and only if I and I' are interchanged by the hyperelliptic involution.

Suppose we are in the case where I and I' are fixed by the hyperelliptic involution. The complement $M \backslash (I \cup I')$ is then a genus one translation surface with two boundary components, each a union of two horizontal segments of the same length. The holonomy around each boundary component is trivial, so we have exhibited M as a self-connected-sum $(E, p, q)\#_I$. The branch covering π must arise from a primitive covering $E \to F$ which identifies p and q. This implies that $p - q$ is exactly d-torsion in the group law of E. $\square$

8. Construction of ergodic measures for U

In the previous section we discussed G-invariant measures on $\mathcal{E}_D$. In this section we will discuss U-invariant measures which are not G-invariant.

8.1. Minimal sets.

A *minimal set* for a flow is a minimal (with respect to inclusion) nonempty closed invariant subset. An example is a closed orbit, but more complicated examples may arise. The minimal sets for the horocycle flow in any stratum were classified in [SmWe1], where the following was shown.

Proposition 8.1. *For $M \in \mathcal{H}$, the following are equivalent:*
- *$\overline{UM}$ is minimal.*
- *$\overline{UM}$ is compact.*

- *M has a horizontal cylinder decomposition, with $\Xi(M)$ consisting of the boundaries of these cylinders.*

In case these hold, $\overline{UM}$ is topologically conjugate to a torus, whose dimension is the dimension over $\mathbb{Q}$ of the span of the moduli of the cylinders in the horizontal direction of M, and under this conjugacy, the U-action becomes conjugate to a straight-line linear flow on this torus. The closure of any U-orbit in $\mathcal{H}$ contains some M satisfying the above conditions.

Corollary 8.2. *Each minimal set supports a unique U-invariant measure (which in particular is ergodic). The measure is linear with respect to the linear structure on the torus which is the minimal set.*

Proof. These are all well-known properties of straight line flows on tori, and follow from the isomorphism above. □

Corollary 8.3. *For cylinder decompositions of type (A) of Figure 6, the corresponding orbit-closure is a minimal set which is one-dimensional or two-dimensional. In the former case it is a closed horocycle, and in the latter, the corresponding measure is invariant under UZ, and this group acts transitively on the support of the measure. For cylinder decompositions of types (B), (C) and (D), U-orbits are periodic.*

Proof. By Proposition 8.1 the dimension of the U-orbit closure is the dimension of the $\mathbb{Q}$-span of the moduli, which by Proposition 7.7 is either 1 or 2. Suppose that it is 2. Again by Proposition 7.7, there must be three cylinders, so we are in case (A). The description in [SmWe1] shows that the additional dimension of the orbit closure is given by changing only the twists around cylinders, and it is straightforward to check using the explicit coordinates used in [SmWe1], that the Z-action also affects only the twists, changing them linearly in a way which is distinct from the U action. Thus the UZ-orbits are 2 dimensional, and hence UZ is transitive on the orbit-closure and the corresponding measure is UZ-invariant.

In cases (B) and (C) there are only two cylinders which, by Proposition 7.7, have rationally related moduli. In case (D) there is only one cylinder. So in all of these cases the minimal set is 1-dimensional. □

8.2. **Parameter spaces of minimal sets.** For every discriminant D, there are examples in $\mathcal{E}_D$ of surfaces belonging to minimal sets of type (A), for which the U-orbit closure is two-dimensional, i.e. is an orbit of UZ. These minimal spaces are naturally grouped in continuous families, which we will refer to as 'beds'. Their structure was studied in detail in [B2, §6] when D is not square. We summarize this structure here. This analysis will not be required in the sequel. Contrary to the rest of the paper, this discussion will be in the context of surfaces with unlabeled singularities as the combinatorial structure is simpler.

Each $\mathcal{E}_D(1,1)$ contains finitely many four (real) dimensional beds which parameterize unit area surfaces with periodic horizontal directions having three cylinders, where the ratios of circumferences of these cylinders are fixed, while twist parameters and heights are allowed to vary, subject to the linear condition which defines the eigenform locus. These beds are invariant under the horocycle flow, geodesic flow and real and imaginary rel. Each bed is foliated by a family of 2-tori defined by fixing the heights as well as the circumferences of the cylinders, and varying the twist parameters. Each torus is a closed UZ-orbit and is either a U-minimal set or a union of closed U-orbits, depending on whether or not moduli of the cylinders have rational ratios. We will see in Theorem 12.3 that applying the geodesic flow to one of these tori gives a family of measures which equidistribute in $\mathcal{E}_D(1,1)$.

As we vary the heights of the cylinders in one of these families of three-cylinder surfaces (see Figure 16), the height of a cylinder may tend to 0, yielding a family of two-cylinder surfaces in the boundary of the family of three-cylinder surfaces (see Figure 17). For a given three cylinder surface there are two cylinders which are candidates for degeneration and for each candidate cylinder there is a path in the family of three cylinder surfaces leading to a surface in which this cylinder is degenerate. Thus there are two distinct families of two-cylinder surfaces lying in the boundary of each three-cylinder family. These two-cylinder families are three-dimensional and composed of surfaces having horizontal data diagram of type (B) or (C). By Corollary 8.3, each two-cylinder family is a union of closed U-orbits.

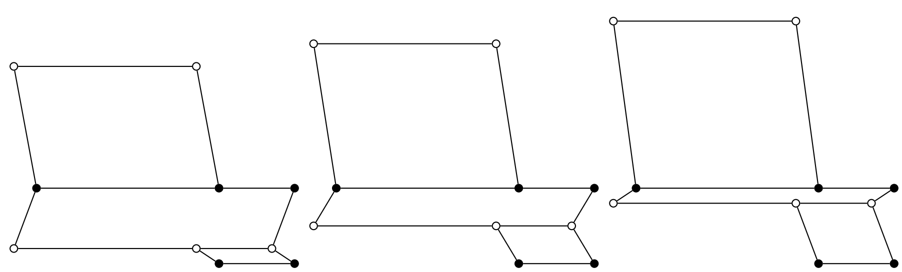


FIGURE 16. Moving in the bed of minimal sets: 3 cylinder surfaces in $\mathcal{E}_8(1,1)$, deformed by imaginary rel.

Each family of two-cylinder surfaces is in turn in the boundary of exactly two families of three-cylinder surfaces. Starting from a given

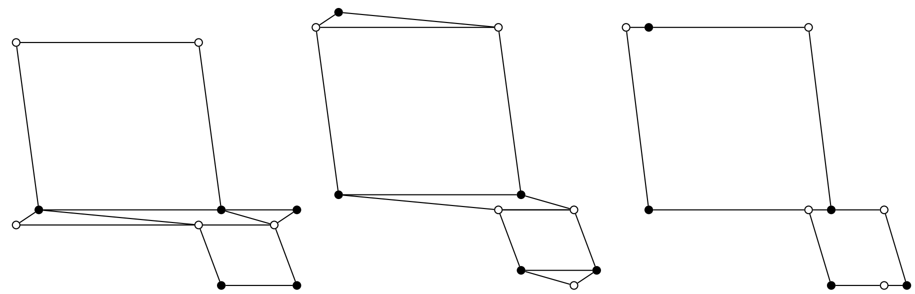


FIGURE 17. Applying a cut and paste surgery, one can continue the imaginary rel motion and arrive at a 2-cylinder surface.

family of three-cylinder surfaces, we may continue this process of moving in the family so that the cylinders degenerate to one of the boundary two-cylinder families and then by undoing a different degeneration we get to a different three-cylinder family. After passing through a finite sequence of three-cylinder families, we eventually return to the original one. These families of minimal sets then comprise finitely many disjoint cycles.

If $\mathcal{O}_D$ is a maximal order, the set of cycles is naturally in bijection with the ideal class group of $\mathcal{O}_D$. The bijection is given by associating to a periodic surface in one of these families the fractional ideal in F_D obtained by taking the $\mathbb{Z}$-span of the circumferences of its cylinders. These circumferences are constant (up to scale) in each two- or three-cylinder family of minimal sets, and when a three-cylinder family degenerates to a two-cylinder family, one cylinder is lost but the fractional ideal generate by their circumferences is unchanged. If $\mathcal{O}_D$ is not maximal the analysis becomes more involved. In the case where D is square, the structure of the collection of minimal sets is more complicated due to the presence of minimal sets consisting of one-cylinder surfaces, in addition to two and three cylinder decompositions.

8.3. **Framing and splitting.** Let $\mathcal{H}_{\mathrm{f}}(2) \to \mathcal{H}(2)$ be the threefold cover corresponding to choosing a right-pointing horizontal prong at the singularity, as in §2.9. This cover is connected, since as we have seen in Proposition 2.9, the $\widetilde{\mathrm{SO}}_2(\mathbb{R})$-orbits of a point contain all of its pre-images under the map $\mathcal{H}_{\mathrm{f}}(2) \to \mathcal{H}(2)$. We denote by $\widehat{\mathcal{E}}_D(2)$ the pre-image of $\mathcal{E}_D(2)$ in $\mathcal{H}_{\mathrm{f}}(2)$. By the same reasoning as above, each connected component of $\mathcal{E}_D(2)$ has connected inverse image. Recall that since U is simply connected it lifts to $\widehat{G}$, so it acts on $\widehat{\mathcal{E}}_D(2)$.

Proposition 8.4. *The action of U is ergodic on any connected component of $\widehat{\mathcal{E}}_D(2)$.*

Proof. Since any connected component of $\mathcal{E}_D(2)$ is a G-orbit (see Theorem 7.3), $\widehat{G}$ acts transitively on each connected component of $\widehat{\mathcal{E}}_D(2)$, and in particular, ergodically. The inclusion $U \subset \widehat{G}$ has the Mautner property (see e.g. [EW]) and hence any ergodic $\widehat{G}$-action is also ergodic for U. $\square$

8.4. **Other U-invariant measures.** We now discuss the remaining ergodic U-invariant measures. Each of these arise from applying the rel flow to a closed G-orbit, possibly lying in a boundary stratum of $\mathcal{E}_D$. To avoid working with compactifications, we will not always take this point of view explicitly, in favor of more elementary constructions.

We classify the remaining measures in terms of their configuration of horizontal saddle connections $\Xi(\mu)$. We have the following possibilities.

Proposition 8.5. *Let μ be an ergodic U-invariant measure, which is not supported on a minimal set. There are four possibilities for $\Xi(\mu)$:*

(i) $\Xi(\mu) = \varnothing$.
(ii) $\Xi(\mu)$ *consists of one saddle connection joining distinct singularities. For μ-a.e. surface, this saddle connection is fixed by the hyperelliptic involution.*
(iii) $\Xi(\mu)$ *consists of two saddle connections joining distinct singularities which are interchanged by the hyperelliptic involution. For μ-a.e. surface M, the union of these saddle connections disconnects M into a union of two tori with slits removed.*
(iv) $\Xi(\mu)$ *consists of two saddle connections joining distinct singularities which are fixed by the hyperelliptic involution. For μ-a.e. surface M, the union of these saddle connections disconnects M into a torus with two slits removed, and D is a square.*

Proof. This follows directly from Theorems 7.10 and 7.12. $\square$

In the case where M belongs to a minimal set, $\Xi(M)$ is one of configurations pictured in Figure 6.

Our goal in this section will be to provide examples of U-ergodic U-invariant measures in each case.

For fixed discriminant D, let $\mathcal{C}$ denote one of the following:

(i) The G-orbit of the regular decagon $\mathcal{L}_{\mathrm{dec}} \subset \mathcal{E}_5(1,1)$, or if D is square, the G-orbit of a square-tiled surface in $\mathcal{E}_D(1,1)$.
(ii) A component of $\widehat{\mathcal{E}}_D(2)$ in $\mathcal{H}_{\mathrm{f}}(2)$.
(iii) A component of $\mathcal{P}_D \subset \mathcal{H}(0) \times \mathcal{H}(0)$.

(iv) If $D = d^2$, the d-torsion locus $\mathcal{S}_D \subset \mathcal{H}(0,0)$.

Note that in this case and subsquently we use the same symbol to denote subsets of different spaces; that is, more formally we should replace $\mathcal{C}$ with one of $\mathcal{C}$(i)–$\mathcal{C}$(iv) depending on the cases above. We will continue with this slight inaccuracy with the set $\mathcal{B}$ and map Ψ below, and this should cause no confusion.

In each case, $\mathcal{C}$ has a natural G or $\widehat{G}$ action, and $\mathcal{C}$ is isomorphic to G/Γ or $\widehat{G}/\Gamma$ for some lattice Γ. Let $\mathcal{B} \subset \mathcal{C}$ denote the subset consisting of surfaces without horizontal saddle connections. In case (iii) or (iv), we interpret a saddle connection to be either a closed geodesic or a segment joining two marked points.

Proposition 8.6. *The set $\mathcal{B} \subset \mathcal{C}$ is the complement of the set of closed horocycles.*

Proof. In cases (i) and (ii) this follows from the Veech alternative [Ve2].

In case (iii), recall that every point of $\mathcal{C}$ represents a pair of isogenous genus one translation surfaces, so the horizontal direction is periodic for one if and only if it is periodic for the other. A torus with periodic horizontal direction is stabilized by an infinite cyclic subgroup of U, and for two isogenous tori, these subgroups are commensurable. Thus a point of $\mathcal{C}$ representing a pair of tori with periodic horizontal direction lies on a closed horocycle.

In case (iv), every point of $\mathcal{C}$ represents a torus with two marked points (E, p, q) where $p - q$ is d-torsion. It has a horizontal saddle connection exactly when the horizontal direction of E is periodic, in which case E is stabilized by an infinite cyclic subgroup of U. Given E, there are finitely many choices of p, q which differ by d-torsion, up to isomorphism of E. Thus (E, p, q) is also stabilized by an infinite cyclic subgroup of U, so lies on a closed horocycle. □

Note that $\mathcal{B}$ is a dense G_δ subset of a locally compact metrizable space, and hence the Borel σ-algebra structure on $\mathcal{B}$ is a standard Borel space (see [K]).

Corollary 8.7. *In each case the action of U on $\mathcal{B}$ is uniquely ergodic, i.e. there is a unique U-invariant regular Borel measure on $\mathcal{B}$.*

Proof. By Proposition 8.6, the action of U on $\mathcal{B}$ is measurably conjugate to the horocycle flow on the complement of the set of horocycles in G/Γ or $\widehat{G}/\Gamma$ for some lattice Γ. Dani [D] classified the U-invariant measures on G/Γ and $\widehat{G}/\Gamma$, showing that they are either supported on closed horocycles or are the global measure induced by Haar measure.

Thus Dani's theorem implies that the U-action on each $\mathcal{B}$ is uniquely ergodic. $\square$

In each case, we define a map $\Psi\colon \mathcal{B} \times (0, \infty) \to \mathcal{H}(1,1)$ as follows:

(i) $\Psi(M,T) = \mathrm{Rel}_T M$. This is well-defined by Corollary 6.2.

(ii) $\Psi(M,T)$ is the inverse of the map $\Phi_{\mathrm{f}}^{(-T)}$ of Theorem 6.5. Informally, $\Psi(M,T)$ is the surface obtained by splitting the singularity of angle 6π to two singularities of angle 4π each, and moving them apart using Rel_T. The way in which the two singularities are to be split apart is made explicit by the framing.

(iii) Given a pair of genus one translation surfaces $(E_1, E_2) \in \mathcal{B} \subset \mathcal{P}_D$, we define $\Psi((E_1,E_2),T)$ to be the surface obtained by removing a horizontal slit of length T from each E_i and then gluing the resulting surfaces along their boundaries to obtain a genus two surface in $\mathcal{H}(1,1)$. The resulting surface is an eigenform, since real multiplication is preserved by deformations which leave absolute periods constant by Proposition 7.1.

(iv) Given a genus one surface E with marked points p, q which differ by d-torsion, we define $\Psi((E,p,q),T)$ to be the surface obtained by cutting E along two horizontal slits of length T with left endpoints p and q and then gluing the resulting two boundary components to obtain a genus two surface in $\mathcal{H}(1,1)$. We saw above that there is a primitive degree d torus covering $\pi\colon E \to F$ sending p and q to the same point. Our gluing identifies only pairs of points which have the same image, so the resulting surface is also a primitive degree d torus cover belonging to $\mathcal{E}_{d^2}(1,1)$.

In each case, we label the singularities of the resulting surface so that singularity ξ_1 is on the left hand side of the horizontal saddle connection of length T, and singularity ξ_2 is on the right hand side.

We extend the definition of $\Psi(M,T)$ to $T < 0$ as follows. In case (i), we simply apply Rel_T. In the remaining cases, we repeat the same construction using $|T|$ in place of T, but choose the opposite labeling for the singularities.

For fixed $T \neq 0$, let $\Psi_T : \mathcal{B} \to \mathcal{E}_D(1,1)$ be defined by $\Psi_T(M) = \Psi(M,T)$. In cases (ii), (iii), and (iv), we denote by $\mathcal{B}_T$ the image of Ψ_T.

We summarize this discussion in the following:

Proposition 8.8. *For each T, Ψ_T is continuous on $\mathcal{B}$. In case (i), the image is a set of surfaces with no horizontal saddle connections. In the remaining cases, the image $\mathcal{B}_T$ is the set of surfaces in $\mathcal{E}_D(1,1)$ with*

- *in case (ii) there is exactly one horizontal saddle connection δ, it has length $|T|$, joins distinct singularities.*
- *in case (iii) there are exactly two horizontal saddle connections of length $|T|$ joining distinct singularities which are interchanged by the hyperelliptic involution.*
- *in case (iv) there are exactly two horizontal saddle connections of length $|T|$ joining distinct singularities which have the same length and are both fixed by the hyperelliptic involution.*

These saddle connections are all positively oriented in the case $T > 0$ and negatively oriented in the case $T < 0$.

In each case, the pushforward under Ψ_T of the G-invariant (or $\widehat{G}$-invariant) Haar measure on $\mathcal{B}$ is the unique U-invariant and U-ergodic measure μ on $\mathcal{E}_D(1,1)$, such that $\Xi(\mu)$ is as in cases (i)-(iv) above.

Proof. Most of the statements in the Proposition were established in the preceding discussion. To obtain continuity of Ψ_T, in case (i), continuity follows from Proposition 3.3, and in case (ii), from Theorem 6.5. In the other cases, it follows immediately from the explicit definition of the surgeries described above that continuity holds in a neighborhood of M, for all T sufficiently small (depending on this neighborhood). Continuity for arbitrary T follows by combining this with Proposition 3.3.

In order to prove uniqueness, we show that for each of the cases (i)–(iv), each surface M with $\Xi(M) = \Xi(\mu)$ as described in the list, is in the image of the map Ψ_T. Then the claim follows using the unique ergodicity of the U-action on the corresponding set $\mathcal{B}$. To show that M is in the image of Ψ_T, we define the inverse of Ψ_T by inverting the surgery construction used to define it. Thus for case (ii), we use Theorem 6.5. In case (iii), Ψ_T associates to a pair of isogenous tori (E, F), the connected sum $E\#_I F$. Given a surface M with two horizontal saddle connections interchanged by the hyperelliptic involution, we cut M along these saddle connections to obtain a pair of isogenous tori (E, F), and so $M = \Psi_T(E, F)$. The other cases are similar. □

9. Classification of ergodic measures

In this section we show that the only U-invariant U-ergodic measures in $\mathcal{E}_D(1,1)$ are those described in Section 8. The following is our full measure classification result, which is an expanded version of Theorem 1.1.

Theorem 9.1. *Let μ be a U-invariant U-ergodic Borel probability measure on $\mathcal{E}_D$. Then exactly one of the following cases holds:*

(1) *μ is length measure on a periodic U-orbit, $\Xi(\mu)$ is a complete separatrix diagram of type (A), (B), (C) or (D), and all the cylinders in the corresponding cylinder decomposition have commensurable moduli.*
(2) *μ is the flat measure on a 2-dimensional torus which is a U-minimal set, $\Xi(\mu)$ is a complete separatrix diagram of type (A), and the cylinders in the corresponding cylinder decomposition do not have commensurable moduli. The stabilizer of μ is $U \times Z$, and $U \times Z$ acts transitively on* $\mathrm{supp}\,\mu$.
(3) *$\Xi(\mu)$ consists of two saddle connections joining distinct singularities, which disconnect M. The complement consists of two isogenous tori glued along a slit as in case (iii) of §8.4, and there is $T \in \mathbb{R} \smallsetminus \{0\}$ such that μ is the pushforward of Haar measure on a connected component of $\mathcal{P}_D$, via the map Ψ_T.*
(4) *$\Xi(\mu)$ consists of one saddle connection joining distinct singularities, as in case (ii) of §8.4, and there is $T \in \mathbb{R} \smallsetminus \{0\}$ such that μ is the image under Ψ_T of Haar measure on a finite volume $\widehat{G}$-orbit in $\mathcal{H}_{\mathrm{f}}(2)$, where $\widehat{G}$ and $\mathcal{H}_{\mathrm{f}}(2)$ are the threefold covers of G and $\mathcal{H}(2)$ described in §2.10.*
(5) *$\Xi(\mu) = \varnothing$ and there is $T \in \mathbb{R}$ such that μ is the image under* Rel_T, *of Haar measure on a closed G-orbit in $\mathcal{E}_D(1,1)$. In this case either D is a square or is equal to 5.*
(6) *$\Xi(\mu)$ contains two saddle connections joining distinct singularities, whose complement in M is a torus with two parallel slits of equal length, as in case (iv) of §8.4, and there is $T \in \mathbb{R} \smallsetminus \{0\}$ such that μ is the image under Ψ_T of a G-invariant measure on the space of tori with two marked points. In this case D is a square.*
(7) *μ is the flat measure on $\mathcal{E}_D(1,1)$.*

Definition 9.2. *The numbering in Theorem 9.1 corresponds to that in Theorem 1.1. If μ is a measure satisfying the conditions of item n in this theorem then we will refer to n as the* type *of the measure μ.*

Remark 9.3. *Recall from Proposition 7.2 that $\mathcal{E}_D(1,1)$ is an affine manifold in period coordinates, and the tangent space to $\mathcal{E}_D(1,1)$ at a point M is naturally identified with the Lie algebra $\mathfrak{l} \cong \mathfrak{sl}_2(\mathbb{R}) \ltimes \mathbb{R}^2$. Hence we can discuss the Lie subalgebra $\mathfrak{l}_\mu$ which fixes the measure μ. The following list identifies $\mathfrak{l}_\mu$ in each case:*

(1) $\mathfrak{l}_\mu = \mathfrak{u} = \mathrm{Lie}(U)$.
(2) $\mathfrak{l}_\mu = \mathfrak{u} \oplus \mathfrak{z}$.

(3) $\mathfrak{l}_\mu$ *is the Lie algebra of the subgroup of* L *fixing the holonomy of the two saddle connections in* $\Xi(\mu)$ *(note that these vectors have the same holonomy).*
(4) $\mathfrak{l}_\mu$ *is the Lie algebra of the subgroup of* L *fixing the holonomy of the vector in* $\Xi(\mu)$.
(5) $\mathfrak{l}_\mu$ *is the Lie algebra of the subgroup of* L *fixing the vector* $(T, 0) \in \mathbb{R}^2$.
(6) $\mathfrak{l}_\mu$ *is the Lie algebra of the subgroup of* L *fixing the holonomy of the two vectors in* $\Xi(\mu)$ *(note that these vectors have the same holonomy).*
(7) $\mathfrak{l}_\mu = \mathfrak{l}$.

In all cases above it is easy to show using the maps defined in §8, that for μ*-a.e. point* M*, the intersection of* $\operatorname{supp} \mu$ *with a neighborhood of* M *is an affine submanifold whose dimension equals* $\dim \mathfrak{l}_\mu$*. So we may say that* $\operatorname{supp} \mu$ *is almost everywhere an affine manifold locally modeled on the Lie algebra* $\mathfrak{l}_\mu$.

The proof of Theorem 9.1 occupies the rest of this section. It is modeled on arguments of Ratner dealing with horocycle orbits in homogeneous spaces (see [R] for a survey), which were first applied to spaces of translation surfaces in [EMaMo] and applied to the eigenform locus by Calta and Wortman in [CW]. We follow the outline of [CW] but make several improvements. These enable us to bypass some entropy arguments used in [CW] and give a clearer justification of Proposition 9.4 below.

Ratner's argument hinges on the analysis of transverse divergence of nearby trajectories for the U-action which she calls the *R-property* (see [R] p. 22). Suppose we want to compare two orbits u_tM and u_tM' where M and M' are close. We can write $M' = gM \Phi v$. The case when $(g, v) \in N$ is somewhat special and we consider it first. In this case g normalizes U and satisfies $u_tg = gu_{ct}$ for some $c > 0$ independent of t, and therefore

$$u_tM' = u_t(g, v)M = (u_tg, v)M = (gu_{ct}, v)M = (g, v)u_{ct}M.$$

The divergence is caused by two factors. Since (g, v) is independent of t and small we think of u_tM' and $u_{ct}M$ as being close. In particular if c is not 1 then the primary divergence of these two points is in the horocycle direction. Rescaling time along the second orbit by replacing t with ct has the effect of removing this divergence. We describe the remaining orbit divergence as transverse divergence.

Now let us consider the general case where $M' = gM \oplus v$ with $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$, and v arbitrary. In order to pick out the divergence in the directions transverse to the horocycle direction we will rescale the time factor along the second orbit. Specifically let u_tM be the first orbit and write the second orbit as u_sM' with $s = s(t)$. Unlike the previous case, the time parameter $s(t)$ for the second orbit will depend on the time parameter t in the first orbit in a nonlinear way and the rescaling will only work in a finite interval of time. We will show that the time change is not too far from being linear. In order to do this we choose a family of transversals to the horocycle orbits and choose the function $s(t)$ so that u_tM and u_sM' stay in the same transversal. Recall from Proposition 7.2 that locally we can model $\mathcal{E}_D(1,1)$ on L. We choose as our family of transversals the elements of L corresponding to elements (g, v) where g is lower triangular and v is in $\mathbb{R}^2$, and both g, v are small. This defines a submanifold transverse to U of complementary dimension.

Since $M \oplus v$ is defined, by Proposition 3.4, $u_sgM \oplus u_sv$ is also defined and is equal to $u_s(gM \oplus v) = u_sM'$. Since $u_sgM = (u_sgu_{-t})u_tM$, the pair (u_sgu_{-t}, u_sv), considered as an element of L, gives a well-defined map which moves u_tM to u_sM'. We calculate:

$$u_sgu_{-t} = \begin{pmatrix} a + sc & b - at + s(d - ct) \\ c & d - ct \end{pmatrix}.$$

In order for u_tM and u_sM' to lie on the same transversal, we require that the matrix u_sgu_{-t} is close to the identity and lower triangular. For $t \in \mathbb{R}$ we set

$$s = s(t, g) \stackrel{\text{def}}{=} \frac{at - b}{d - ct}. \tag{41}$$

With this choice we have:

$$u_{s(t)}gu_{-t} = \begin{pmatrix} \frac{1}{d-ct} & 0 \\ c & d - ct \end{pmatrix} \tag{42}$$

(here $s(t) = s(t, g)$), so that u_tM and $u_{s(t)}M'$ lie in the same transverse leaf as long as t is chosen such that $d - tc$ is close to 1.

Note that while the initial displacement may consist of displacement in the Rel direction as well as the G-direction, the time change only depends on the initial displacement in the G direction.

Let μ be a U-invariant ergodic measure on $\mathcal{E}_D(1,1)$ with $\Xi(\mu) = \varnothing$. Since N acts on the set of surfaces with $\Xi(\mu) = \varnothing$ we can consider the subgroup of N that preserves μ. Let N_μ° to be the connected component of the identity in N_μ, where N_μ is the stabilizer of μ as in Definition 4.7.

Proposition 9.4. *For any $\varepsilon > 0$ there is a compact subset K of $\mathcal{E}_D(1,1)$ with $\mu(K) \geqslant 1-\varepsilon$ such that if $M \in K$ is a limit of $M_k = h_k M \oplus v_k \in K$, with $h_k \in G$ and $v_k = (x_k, y_k) \in \mathbb{R}^2$, and $(h_k, v_k) \to 1_L$, $(h_k, v_k) \notin U$ for infinitely many k, then* $\dim N_\mu^\circ \geqslant 2$.

Proof. Let Ω_μ be the set of μ-generic points for the U-action. The set Ω_μ has full μ measure. Given ε, let Ω_1 be a compact subset of Ω_μ such that $\mu(\Omega_1) \geqslant 1-\varepsilon$. By the Birkhoff ergodic theorem there is a compact subset K of $\mathcal{E}_D(1,1)$ of measure $\mu(K) \geqslant 1-\varepsilon$, and a $T_0 > 0$, such that for all $M \in K$ and all $T > T_0$,

$$\frac{1}{T}\left|\{t \in [0,T] : u_t M \in \Omega_1\}\right| \geqslant 1 - 2\varepsilon. \tag{43}$$

We will show that this choice of K satisfies the conclusion of the Proposition. So let M_k, M, h_k and $v_k = (x_k, y_k)$ be as in the statement of the Proposition. Write $c_k \stackrel{\mathrm{def}}{=} c(h_k)$ and $d_k \stackrel{\mathrm{def}}{=} d(h_k)$.

By Corollary 4.8, N_μ is a closed subgroup of N and hence a Lie group. Thus N_μ° is a connected Lie group that contains U. In order to show that $\dim N_\mu^\circ \geqslant 2$ it suffices to rule out the possibility $U = N_\mu^\circ$, that is, in any neighborhood of the identity in N, to find an element in $N \smallsetminus U$ preserving μ. If the sequence (h_k, v_k) contains infinitely many elements of $N \smallsetminus U$ then it follows from Corollary 4.16 that we have such a sequence of elements. We may thus assume that $(h_k, v_k) \notin N$ for infinitely many k.

Write elements of Z as Rel_s (where defined). Assume first that along a subsequence

$$y_k = o(c_k). \tag{44}$$

In this case we will show that for each sufficiently small $\delta > 0$, there is an element

$$\ell = \ell(\delta) = g_\tau \mathrm{Rel}_\sigma \tag{45}$$

preserving μ, such that

$$\frac{\varepsilon\delta}{2} \leqslant \tau \leqslant 2\delta \quad \text{and } \sigma = O(\delta). \tag{46}$$

The fact that ℓ can be written as $g_\tau \mathrm{Rel}_\sigma$ means that it lies in N. The fact that $\tau \neq 0$ means that ℓ does not lie in U.

In light of Corollary 4.16, in order to show that $\ell \in N$ belongs to N_μ, it suffices to prove that there are surfaces $M^{(1)}, M^{(2)} \in \Omega_\mu$ such that $\ell M^{(1)} = M^{(2)}$. We will find $M^{(1)}$ and $M^{(2)}$ as limits of convergent subsequences in Ω_1 and ℓ as a limit of a convergent subsequence in L.

Recall that in order for u_tM and $u_{s(t)}M'$ to be close it is necessary that t be chosen so that $d - tc$ is close to 1. For $g \in G$ set

$$I_g \stackrel{\text{def}}{=} \{t \in \mathbb{R} : |\ln(d - tc)| \leqslant \delta\}, \text{ where } d = d(g), c = c(g).$$

An elementary computation shows that there is a neighborhood $\mathcal{W}$ of 1_G such that if $g \in \mathcal{W}$ then $0 \in I_g$, and for any t_1, t_2 in the connected component of I_g containing 0, we have

$$\frac{|t_1 - t_2|}{2} \leqslant |s(t_1, g) - s(t_2, g)| \leqslant 2|t_1 - t_2|. \tag{47}$$

Here the $s(t_i, g)$ are defined by equation (41) and δ is assumed to be sufficiently small.

Let $P(t)$ be any linear function, that is P is of the form $P(t) = At+B$ for some A, B, and for $a < b$, let $\|P\| = \max_{x\in[a,b]} |P(x)|$. Then it is easy to check that

$$\frac{|\{t \in [a, b] : |P(t)| < \varepsilon\|P\|\}|}{b - a} < 2\varepsilon. \tag{48}$$

The definition of I_g and equation (42) ensure that at the endpoints of I_g, the g_t-component of the displacing element $u_s g u_{-t}$ is 2δ. Namely, given $c \neq 0$ and d, the formula $T = T(c, d) = \frac{d - e^{\pm\delta}}{c}$ gives the two solutions to the equation $|\ln(d - cT)| = \delta$. Since $d_k \to 1$ and $c_k \to 0$, for all sufficiently large k there is a unique such $T_k = T(c_k, d_k)$ with $T_k > 0$, and moreover for all sufficiently large k, we will have $T_k \geqslant T_0$. Here we take δ sufficiently small and k sufficiently large so that

$$\frac{1}{2} \leqslant \frac{|d_k - c_k t - 1|}{|\ln(d_k - c_k t)|} \leqslant 2. \tag{49}$$

We apply the above estimates for $T = T_k$ and $P(t) = d_k - c_k t - 1$ and obtain the existence of a subset of $t \in [0, T]$, of measure at least $1-8\varepsilon$, for which all of the following hold:

- $u_tM \in \Omega_1$ (by (43)).
- $u_sM_k \in \Omega_1$, where $s = s(h_k, t)$ (by (47), (43) and since $M_k \in K$).
- $|d_k - c_k t - 1| \geqslant \frac{\varepsilon\delta}{2}$, and hence $|\ln(d_k - c_k t)| \geqslant \frac{\varepsilon\delta}{4}$ (by (48) and (49)).

In particular, if we choose $\varepsilon < 1/8$ (which we may with no loss of generality), then the set of t satisfying all these properties is nonempty. Fixing such a t, it follows from (42) and Proposition 3.4 that

$$\begin{aligned} u_sM_k &= u_s(h_kM \mathbin{+\!\!\!\!\circ} v_k) = u_sh_ku_{-t}u_tM \mathbin{+\!\!\!\!\circ} u_sv_k \\ &= \begin{pmatrix} \frac{1}{d_k - c_k t} & 0 \\ c_k & d_k - c_k t \end{pmatrix} u_tM \mathbin{+\!\!\!\!\circ} (x_k + sy_k, y_k)\,. \end{aligned} \tag{50}$$

For each k we let t_k and $s_k = s(h_k, t_k)$ satisfy (50). Using the compactness of Ω_1 we can take a subsequence along which we have convergence

$$u_{t_k} M \to M^{(1)}, \ u_{s_k} M_k \to M^{(2)}.$$

By (50), we have $u_{s_k} M_k = \ell_k u_{t_k} M$ where

$$\ell_k = \left[\begin{pmatrix} \frac{1}{d_k - c_k t_k} & 0 \\ c_k & d_k - c_k t_k \end{pmatrix}, (x_k + s(t_k, h_k) y_k, y_k) \right].$$

We claim that $M^{(2)} = \ell M^{(1)}$, for $\ell \in N$ satisfying the bounds (45), (46). It is enough to show that after passing to a further subsequence, the sequence ℓ_k converges to an element ℓ satisfying these bounds. Since $|\ln(d_k - c_k t)| \geqslant \frac{\varepsilon\delta}{4}$, since $c_k \to 0$, and by definition of the interval I_{h_k}, the matrices $\begin{pmatrix} \frac{1}{d_k - c_k t_k} & 0 \\ c_k & d_k - c_k t_k \end{pmatrix}$ converge to g_τ with τ as in (46). Since $x_k, y_k \to 0$, in order to obtain the required bound on σ we need to show that $s_k y_k = O(\delta)$. But

$$\begin{aligned} s_k y_k = & y_k \frac{a_k t_k - b_k}{d_k - c_k t_k} = \frac{y_k}{c_k} \frac{a_k t_k - b_k}{d_k / c_k - t_k} \\ & \overset{(44)}{=} o(1) \frac{a_k t_k - b_k}{c_k / d_k - t_k} \overset{g_k \to 1_G}{=} o(1) = O(\delta), \end{aligned}$$

as required.

In case the condition of equation (44) does not hold, we have

$$c_k = O(y_k).$$

In this case we can employ a similar argument, and find an element $\ell(\delta)$ as in (45) which satisfies (instead of (46)) the estimates

$$\sigma \geqslant c\delta \text{ for some } c > 0, \text{ and } \tau = O(\delta), \ \sigma = O(\delta).$$

□

Remark 9.5. *The proof of* [CW, Lemma 2]*, which is the statement analogous to Proposition 9.4, is incomplete. The formula preceding the final paragraph of the proof in* [CW] *would hold for an honest group action, but in this case requires justification.*

Proof of Theorem 9.1. To each ergodic horocycle invariant measure μ we associate a horizontal data diagram via Corollary 5.3. The possibilities for $\Xi(\mu)$ given in the statement of Theorem 9.1 cover all cases by Proposition 8.5. In §8, for each possibility for $\Xi(\mu)$, a measure was constructed. If it is a complete separatrix diagram, then by Proposition 8.1 and Corollary 8.2, we must be in cases (1) or (2). In all other cases for which $\Xi(\mu) \neq \varnothing$, it was shown in Proposition 8.8 that there are no additional measures. This shows that if $\Xi(\mu) \neq \varnothing$, then μ must be one of the measures described in items (1), (2), (3), (4) or (6).

We now treat the case $\Xi(\mu) = \varnothing$. Recall from Definition 3.7 and Corollary 6.2 that $\mathcal{H}'_\infty$ is the set of surfaces in $\mathcal{H}(1,1)$ with no horizontal saddle connections joining distinct singular points. Then $\mu(\mathcal{H}'_\infty) = 1$ and $N = B \ltimes Z$ acts on $\mathcal{H}'_\infty$. Recall that N°_μ is the connected component of the stabiliser of μ so that $U \subset N^\circ_\mu$. We first claim that there is no U-orbit with positive measure. Indeed, if this were the case then by ergodicity, μ is supported on one orbit which, by Poincaré recurrence, must be a periodic orbit. A periodic orbit is a special case of a minimal set so by Proposition 8.1, $\Xi(\mu)$ consists of a full separatrix diagram. This contradicts our assumption that $\Xi(\mu) = \varnothing$.

We now claim that we cannot have $N^\circ_\mu = U$, i.e. that $\dim N^\circ_\mu \geqslant 2$. To see this, let $\Omega_\mu \subset \mathcal{H}'_\infty$ be the set of generic points as in Proposition 4.15, and let K be as in Proposition 9.4. We have $K \subset \Omega_\mu$. Since μ assigns zero measure to individual U-orbits and K is a compact set of positive measure, K contains a sequence (M_k) such that $M_k \to M \in K$ and none of the M_k are on the U-orbit of M. According to Proposition 9.4 $\dim N^\circ_\mu \geqslant 2$.

Now suppose $Z \subset N^\circ_\mu$. We will show that in this case μ is the flat measure on $\mathcal{E}_D(1,1)$. Let $\mathcal{L}$ be an affine G-invariant subspace of $\mathcal{H}(1,1)$, and let $\mathcal{L}^{(1)}$ be its subset of area-one surfaces. We define the *horospherical foliation for* $\{g_t\}$ to be the foliation into leaves locally modelled on the intersection of the tangent space of $\mathcal{L}^{(1)}$ with the horizontal space $H^1(S,\Sigma;\mathbb{R}_x)$ (the first summand in the splitting (15)). This terminology is motivated by the fact that the $\{g_t\}$-flow is non-uniformly hyperbolic and the above leaves are generically its strong unstable leaves. For the case $\mathcal{L} = \mathcal{E}_D(1,1)$, the leaves of the horospherical foliation are just the orbits of UZ, so in the case we are now considering, the measure μ is invariant under the horospherical foliation, and must be flat measure in light of the following:

Claim 1. *Suppose that:*

- $\mathcal{L}$ *is* $\mathrm{GL}_2(\mathbb{R})$*-invariant and affine in period coordinates, and is defined by real-linear equations in* $H^1(S,\Sigma;\mathbb{R}^2) \cong H^1(S,\Sigma;\mathbb{C})$.
- *The flat measure on* $\mathcal{L}^{(1)}$ *obtained from applying the cone construction to* $\mathcal{L}$ *is* G*-invariant and ergodic.*

Then any strong-stable invariant measure μ *on* $\mathcal{L}^{(1)}$, *with* $\Xi(\mu) = \varnothing$ *coincides with the flat measure.*

This result was proved by Lindenstrauss and Mirzakhani in [LM] under the additional hypothesis that $\mathcal{L}$ the principal stratum, i.e. $\mathcal{L} =$

$\mathcal{H}(1,\ldots,1)$. In [SmWe4] we adapt the argument of Lindenstrauss and Mirzakhani to prove the more general statement given above. The special case we require here, namely $\mathcal{L} = \mathcal{E}_D(1,1)$, was explained by Calta and Wortman, see [CW, §6].

Now suppose $\dim N_\mu^\circ \geqslant 2$ but N_μ° does not contain Z. Since Z is a normal subgroup of N, N_μ° does not contain a conjugate of Z. Let A be the diagonal group in N. Up to conjugacy, the only three one-parameter subgroups of N are U, Z and A, so we find that N_μ° contains a conjugate of A. We can write this conjugate as $A' = nAn^{-1}$ where $n \in N$, and since $N = B \ltimes Z$, we can write $n = \mathrm{Rel}_T ua$ where $a \in A$, $u \in U$ and $T \in \mathbb{R}$. Since N_μ° contains U and $A = aAa^{-1}$ we find that N_μ° contains $(\mathrm{Rel}_T)A(\mathrm{Rel}_{-T})$. Therefore the measure $\mu' \stackrel{\mathrm{def}}{=} (\mathrm{Rel}_{-T})_*\mu$ is B-invariant. By a result of Eskin, Mirzakhani and Mohammadi [EMiMo], μ' is G-invariant, and thus by McMullen's classification of G invariant measures [McM3, McM4], the following are the only possibilities:

- μ' is the G-invariant measure on a closed G-orbit, and D is either a square or $D = 5$. This implies that $\mu = \mathrm{Rel}_{T*}\mu'$ is as described in case (5).
- μ' is the flat measure on $\mathcal{E}_D(1,1)$. Since the flat measure is Z-invariant, $\mu = \mu'$ and we are in case (7).

□

10. Injectivity and nondivergence

In this section we will prove some results which will be used in the proof of Theorem 11.1. The strategy of proof involves showing that typical horocycle orbits do not spend too much time close to a closed G-orbit $\mathcal{L}$ or translations of closed G-orbits by the Rel flow. The argument depends on the fact that a point in an eigenform locus which is close to the Rel_t orbit of $\mathcal{L}$ but not in the Rel_t orbit of $\mathcal{L}$ drifts slowly in the direction of the Rel_t vector field. In order to exploit this phenomenon it is useful to have coordinates close to the Rel_t orbit of $\mathcal{L}$. This argument is similar in spirit to the linearization method of Dani-Margulis, see [KSS, §3.4].

For a stratum $\mathcal{H}$ and $r > 0$, let $\mathcal{H}_r$ denote the set of surfaces M in $\mathcal{H}$ for which there are no horizontal saddle connections of length less than r, and let

$$\mathcal{H}_\infty \stackrel{\mathrm{def}}{=} \{M \in \mathcal{H} : \Xi(M) = \varnothing\} = \bigcap_{r>0} \mathcal{H}_r.$$

The following is an analogue of [KSS, Cor. 3.4.8].

Lemma 10.1. *Let $\mathcal{L}$ be a closed G-orbit in $\mathcal{E}_D(1,1)$. Then for any $T > 0$ there is $r > 0$ such that the map*

$$(M, x, y) \mapsto M + (x, y) \tag{51}$$

is well-defined and injective on $\mathcal{L}_r \times [-T, T] \times \{0\}$, where $\mathcal{L}_r = \mathcal{L} \cap \mathcal{H}_r$. In particular, if we set

$$\mathcal{L}_\infty \stackrel{\mathrm{def}}{=} \mathcal{L} \cap \mathcal{H}_\infty \tag{52}$$

then the map (51) *is well-defined and injective on $\mathcal{L}_\infty \times \mathbb{R} \times \{0\}$.*

Proof. We begin by proving that the map (51) is well-defined and injective on $\mathcal{L}_\infty \times \mathbb{R} \times \{0\}$. The assertion that the map is well defined on $\mathcal{L}_\infty$ follows from Corollary 6.2. We prove injectivity. Suppose that $M_1, M_2 \in \mathcal{L}_\infty$ and $x_1, x_2 \in \mathbb{R}$ are such that $\mathrm{Rel}_{x_1} M_1 = M_1 + (x_1, 0) = M_2 + (x_2, 0) = \mathrm{Rel}_{x_2} M_2$. If $x_1 = x_2$ then by applying Rel_{-x_1} we find that $M_1 = M_2$, so we can assume $x_1 \neq x_2$. Let $x_3 \stackrel{\mathrm{def}}{=} x_1 - x_2$. Then by Proposition 3.5

$$\mathrm{Rel}_{x_3}(\mathrm{Rel}_{x_1} M_1) = \mathrm{Rel}_{x_3}(\mathrm{Rel}_{x_2} M_2) = \mathrm{Rel}_{x_1} M_2, \tag{53}$$

and Rel_{x_3} is not the identity element of Z.

Let ν denote the G-invariant measure on $\mathcal{L}$ coming from the Haar measure on G, and let $\nu_1 = \mathrm{Rel}_{x_1 *}\nu$. By [DS], the set of U-generic points for ν is the set of surfaces which are not on periodic U-orbits, and by [Ve2], this is the set of surfaces without horizontal saddle connections. By Proposition 4.15, the same holds for ν_1, that is the set of U-generic points for ν_1 are the surfaces in $\mathrm{Rel}_{x_1}\mathcal{L}$ without horizontal saddle connections. By (53), Rel_{x_3} maps $\mathrm{Rel}_{x_1} M_1$, which is a generic point for ν_1, to $\mathrm{Rel}_{x_1} M_2$, which is another generic point. In light of Corollary 4.16, Rel_{x_3} preserves ν_1. The stabilizer of ν in N contains the diagonal subgroup A and hence the stabilizer of ν_1 in N contains a non-unipotent element (take for example a nontrivial element of the conjugate $\mathrm{Rel}_{x_1} A \mathrm{Rel}_{-x_1}$ in $B \ltimes Z$). Proposition 4.9 implies that the stabilizer of ν_1 properly contains the group ZU. According to Theorem 9.1, the only U-invariant measure on $\mathcal{E}_D(1,1)$ whose stabilizer properly contains ZU is the flat measure, so ν_1 is the flat measure. Hence so is $\nu = \mathrm{Rel}_{-x_1 *}\nu_1$, a contradiction.

We now prove the first assertion of the Lemma. According to Corollary 6.2, the map (51) is well-defined on $\mathcal{L}_r \times [-T, T] \times \{0\}$ if $r > T$. It remains to show that this map is injective when r is sufficiently large. Suppose that $M' \stackrel{\mathrm{def}}{=} \mathrm{Rel}_{x_1} M_1 = \mathrm{Rel}_{x_2} M_2$ for $M_1, M_2 \in \mathcal{L}$ and $x_1, x_2 \in \mathbb{R}$. According to the first part of the proof, at least one of the surfaces M_i has horizontal saddle connections. Since the Z-action preserves the property of having horizontal saddle connections (by Corollary 6.3), both surfaces have horizontal saddle connections, and since $\mathcal{L}$ is a closed G-orbit, they lie on closed U-orbits and have a horizontal cylinder decomposition. Since circumferences and heights of cylinders are the same for M_i and M', they are the same for M_1 and M_2 and this ensures that M_1 and M_2 lie on the union of one of finitely many closed horocycles $Ux_1, \ldots, Ux_\ell$ of equal length. Suppose first that this length is 1, and denote the union of the horocycles of length 1 by A_1. Let $r_1 > 0$ be the length of the shortest horizonal saddle connection on any surface in A_1, and choose $T_1 \in (0, r_1)$ for which the map (51) is injective on $A_1 \times [-T_1, T_1] \times \{0\}$. Now suppose the length of the closed horocycles is L. Then for $t_0 \stackrel{\mathrm{def}}{=} \log L$, the union of the closed horocycles is $A_L \stackrel{\mathrm{def}}{=} g_{t_0} A_1$, the length of the shortest saddle connections on these horocycles is $r_L \stackrel{\mathrm{def}}{=} e^{t_0/2} r_1$ and the map (51) is injective on $[-T_L, T_L]$ where $T_L \stackrel{\mathrm{def}}{=} e^{t_0/2} T_1$. That is, the ratio r_L/T_L is independent of L. In particular, for any $T > 0$, we can choose L so that $T = T_L$ and choose $r > r_L$. Since

$$\mathcal{L}(r) \subset \mathcal{L} \smallsetminus \bigcup_{r_L \leqslant r} A_L,$$

for these choices, the map (51) will be injective on $\mathcal{L}(r) \times [-T, T] \times \{0\}$. $\square$

Remark 10.2. *The condition that $M \in \mathcal{L}_r$ is necessary for the validity of Lemma 10.1. When D is a square and $\mathcal{L}$ is the closed G-orbit of a square-tiled surface, there are surfaces $M \in \mathcal{L}$, with a horizontal cylinder decomposition of type (A), whose real Rel orbit is compact. When $D = 5$ the ZU-orbit of a surface in $\mathcal{L}$ with a cylinder decomposition of type (A) is also not injectively embedded. That is, in both cases, injectivity will fail if M has short horizontal saddle connections.*

We will also need the following quantitative nondivergence results for the horocycle flow. The following results are proved in [MW1]:

Theorem 10.3. *For any stratum $\mathcal{H}$ there are constants C and α such that for any $M \in \mathcal{H}$, any $\rho \in (0, 1]$ and any $T > 0$, if for any saddle connection δ for M, $\max_{s \in [0,T]} \|\mathrm{hol}(u_s M, \delta)\| \geqslant \rho$ then*

$$|\{s \in [0, T] : u_s M \text{ has saddle connections of length} < \varepsilon\}| < CT \left(\frac{\varepsilon}{\rho}\right)^{\alpha}.$$

In particular:

(I) *For any $\varepsilon > 0$ and any compact $K' \subset \mathcal{H}$ there is a compact $K \subset \mathcal{H}$ such that for any $T > 0$ and any $M \in K'$,*

$$\frac{1}{T} \left|\{s \in [0,T] : u_s M \notin K\}\right| < \varepsilon. \tag{54}$$

(II) *For any $\varepsilon > 0$ there is a compact $K \subset \mathcal{H}$ such that for any $M \in \mathcal{H}_\infty$ there is $T_0 > 0$ such that for all $T > T_0$, (54) holds.*

We will need a refinement of statement (I).

Proposition 10.4. *For any positive constants ε and r, and any compact $L \subset \mathcal{H}_\infty$, there is an open neighborhood $\mathcal{W}$ of L and a compact $\Omega \subset \mathcal{H}$ containing $\mathcal{W}$, such that surfaces in Ω have no horizontal saddle connections shorter than r, and such that for any surface M and any interval $I \subset \mathbb{R}$ for which $u_s M \in \mathcal{W}$ for some $s \in I$, we have*

$$\frac{\left|\{s \in I : u_s M \notin \Omega\}\right|}{|I|} < \varepsilon. \tag{55}$$

Proof. If $s_0 \in I$ is such that $u_{s_0} M \in \mathcal{W}$ then by making the change of variables $s \mapsto s - s_0$ we can assume $M \in \mathcal{W}$ and $I = [a,b]$ with $a \leqslant 0 \leqslant b$. By considering separately the subintervals $[a,0]$ and $[0,b]$, we can assume that $I = [0,T]$ for some $T > 0$. Given a surface M we let $\mathrm{hol}(M)$ be the subset of $\mathbb{R}^2$ consisting of the holonomies of the saddle connections of M, and set

$$K_{r_1,r_2} \stackrel{\mathrm{def}}{=} \{M \in \mathcal{H} : \mathrm{hol}(M) \cap ((-r_1,r_1) \times (-r_2,r_2)) = \varnothing\}. \tag{56}$$

Then K_{r_1,r_2} is compact for any r_1, r_2.

Given ε, r and L as in the statement of the proposition, let η satisfy $C\eta^\alpha < \varepsilon$, where C, α are as in Theorem 10.3, and let $t_0 \stackrel{\mathrm{def}}{=} 2\log\left(\frac{\eta}{\sqrt{2}r}\right)$ and $\sigma' = \frac{\eta^2}{2r}$. These choices guarantee

$$e^{t_0/2} r = \frac{\eta}{\sqrt{2}} \quad \text{and} \quad e^{-t_0/2}\sigma' = \frac{\eta}{\sqrt{2}}. \tag{57}$$

Let $L'' = g_{t_0}(L)$. Then L'' is a compact subset of $\mathcal{H}_\infty$. Compactness implies that there is a constant $\theta > 0$ with the property that for any $M \in L''$ with a saddle connection δ of length less than 1, the vertical component $y = y(M,\delta)$ satisfies $|y| \geqslant \theta$. Moreover by making θ smaller if necessary, we can ensure that the same property holds for all surfaces in a neighborhood $\mathcal{W}''$ of L''. Let $T_1 \stackrel{\mathrm{def}}{=} \frac{2}{\theta}$. Then for any $M \in \mathcal{W}''$ and any saddle connection δ for M, either $\|\mathrm{hol}(M,\delta)\| \geqslant 1$ or

$$\begin{aligned}\|\mathrm{hol}(u_{T_1}M,\delta)\| &\geqslant |x(u_{T_1}M,\delta)| \\ &= |x(M,\delta) + T_1 y(M,\delta)| \\ &\geqslant \frac{2}{\theta}|y(M,\delta)| - 1 \geqslant 1.\end{aligned}$$

That is, $\max_{s\in[0,T_1]} \|\mathrm{hol}(u_sM,\delta)\| \geqslant 1$, so we can apply Theorem 10.3 with $\rho = 1$. We obtain that for all $T \geqslant T_1$, and all $M \in \mathcal{W}''$,

$$\frac{1}{T}\left|\{s \in [0,T] : u_sM \text{ has saddle connections of length} < \eta\}\right| < \varepsilon. \quad (58)$$

We now claim that if we set $\Omega = K_{r,\sigma}$ for any $\sigma \leqslant \sigma'$ then (55) holds when $M \in \mathcal{W}' \stackrel{\mathrm{def}}{=} g_{-t_0}(\mathcal{W}'')$ and $I = [0,T]$, $T > e^{-t_0}T_1$. Indeed suppose $M \in \mathcal{W}'$, δ is a saddle connection for M, and $s \in [0,T]$ such that $\mathrm{hol}(u_sM,\delta) \in (-r,r) \times (-\sigma,\sigma)$. Let $M'' \stackrel{\mathrm{def}}{=} g_{t_0}M \in \mathcal{W}''$, $s' \stackrel{\mathrm{def}}{=} e^{t_0}s$ so that $g_{t_0}u_sM = u_{s'}M''$. By (57),

$$\mathrm{hol}(u_{s'}M'',\delta) = g_{t_0}\mathrm{hol}(u_sM,\delta) \in \left(-\frac{\eta}{\sqrt{2}}, \frac{\eta}{\sqrt{2}}\right) \times \left(-\frac{\eta}{\sqrt{2}}, \frac{\eta}{\sqrt{2}}\right)$$

and in particular $\|\mathrm{hol}(u_{s'}M'',\delta)\| < \eta$. Since the change of variables $s \mapsto s'$ is linear, (55) follows from (58).

Finally, since $L \subset \mathcal{H}_\infty$ is compact, by making σ small enough, we see that Ω contains a neighborhood of L, so that taking a small enough neighborhood $\mathcal{W}$ of L contained in $\mathcal{W}'$, we can ensure that $\mathcal{W} \subset \Omega$ and also that for any $s \in [0, e^{-t_0}T_1]$ and any $M \in \mathcal{W}$, $\mathrm{hol}(u_sM) \cap (-r,r) \times (-\sigma,\sigma) = \varnothing$. $\square$

In light of Proposition 3.3, the following is an immediate consequence of Lemma 10.1 and Proposition 10.4:

Corollary 10.5. *Given positive numbers T and ε, and a compact subset $L \subset \mathcal{L}_\infty$, there is $\delta > 0$, a neighborhood $\mathcal{W}$ of L and a compact set $\Omega \subset \mathcal{L}$, containing $\mathcal{W}$, such that the map* (51) *is well-defined, continuous and injective on $\Omega \times [-T,T] \times [-\delta,\delta]$, and for any interval I and any $M \in \mathcal{H}$, if there is $s_0 \in I$ such that $u_{s_0}M \in \mathcal{W}$ then equation* (55) *holds.*

$\square$

We take this opportunity to record a consequence of Proposition 10.4, which will not be used in this paper but is of independent interest.

Corollary 10.6. *For any $\varepsilon > 0$ and any compact $L \subset \mathcal{H}_\infty$, there is a compact $\Omega \subset \mathcal{H}_\infty$ containing L such that for any $M \in L$ and any interval $I \subset \mathbb{R}$ containing 0,* (55) *holds.*

Proof. For any $\varepsilon > 0$ and $j = 1, 2, \ldots$, let $\varepsilon_j \stackrel{\text{def}}{=} \frac{\varepsilon}{2^j}$. By Proposition 10.4, we can find Ω_j containing L such that all surfaces in Ω_j have no horizontal saddle connections of length less than j, and (55) holds with ε_j in place of ε for any interval I containing 0 and any $M \in L$. Then $\Omega \stackrel{\text{def}}{=} \bigcap_j \Omega_j$ has the required properties. □

11. All horocycle orbits are generic

The goal of this section is to prove the following more detailed version of Theorem 1.2.

Theorem 11.1. *Let $M \in \mathcal{E}_D(1,1)$. Then M is generic for a U-invariant ergodic measure μ_M, and the type of μ_M (as described in Definition 9.2) is as follows:*

- *μ_M has type 1 when $\Xi(M)$ is a complete separatrix diagram of types (B), (C) or (D), or a complete separatrix diagram of type (A) with commensurable moduli.*
- *μ_M has type 2 when $\Xi(M)$ is a complete separatrix diagram of type (A) with two cylinders of incommensurable moduli.*
- *μ_M has type 3 when $\Xi(M)$ consists of two saddle connections joining distinct singularities, which disconnect M into two isogenous tori glued along a slit, as in case (iii) of §8.4.*
- *μ_M has type 4 when $\Xi(M)$ consists of one saddle connection joining distinct singularities, as in case (ii) of §8.4.*
- *μ_M has type 5 when $\Xi(M) = \varnothing$ and there is $s \in \mathbb{R}$ so that* $\mathrm{Rel}_s M$ *is a lattice surface.*
- *μ_M has type 6 when $\Xi(M)$ is a pair of saddle connections which do not disconnect M, as in case (iv) of §8.4.*
- *μ_M has type 7 if it does not correspond to one of the previous cases i.e. $\Xi(M) = \varnothing$ and M is not the result of applying the* Rel *flow to a lattice surface.*

Remark 11.2. *It would be interesting to characterize case (5) explicitly. That is, give a geometric characterization of those surfaces $M \in \mathcal{E}_D(1,1)$ with $\Xi(M) = \varnothing$, for which there is $s \in \mathbb{R}$ such that* $\mathrm{Rel}_s M$ *is a lattice surface.*

The proof of Theorem 11.1 relies on an analogue of the "linearization method", see [KSS, §3.4]. We need the following notion.

Definition 11.3. *The support of a U-invariant U-ergodic measure on $\mathcal{E}_D(1,1)$ is called a* sheet. *The* type *of the sheet is the type of the corresponding measure.*

Sheets of the same type typically appear in families. It will be convenient to partition the sheets into families, called beds. In the sequel, a *bed* will be a measurable set which is a union of sheets in $\mathcal{E}_D(1,1)$. We further require the sheets to be of a fixed type j, for $j \in \{3,4,5,6,7\}$, or to be of one of the two types 1 or 2. A bed corresponding to sheets of type 1 or 2 will be called a *minimal sets bed* and a bed corresponding to sheets of type $j \in \{3,\ldots,7\}$ will be called a *bed of type j*. If $\mathcal{B}$ is a bed corresponding to measures μ of a certain type, we say that μ *belongs to* $\mathcal{B}$, and we define $\Xi(\mathcal{B})$ to be $\Xi(\mu)$ for μ belonging to $\mathcal{B}$.

The reason for combining sheets of type 1 and 2 into the same bed of minimal sets, is that an arbitrarily small perturbation of a sheet of type 1 can be a sheet of type 2 and conversely (the condition that the moduli of the cylinders are rationally related is not stable under small perturbations).

Depending on the application, we will use different partitions of the sheets into beds. For example, for measures of type 1 or 2 (U-minimal sets), one could take one bed consisting of all U-minimal sets, or take one bed each for each type (A, B, C, D) of complete separatrix diagram.

Similarly for type 5, if $D = 5$ then the bed could consist of all surfaces $\mathrm{Rel}_s M$ for M in the G-orbit of the regular decagon. For type 5 when D is a square, note that $\mathcal{E}_D(1,1)$ contains countably many G-orbits of arithmetic lattice surfaces. In this case one could either take one bed consisting of all $\mathrm{Rel}_s M$ for all these lattice surfaces, or alternatively, for each G-orbit GM_0, one could take the bed

$$\bigcup_{s\in\mathbb{R}, M\in GM_0} \mathrm{Rel}_s M.$$

In our definition of beds we only required them to be measurable but in fact they can be chosen so that they have a nicer structure. Continuing with Remark 9.3, one can partition the sheets into beds which are almost everywhere locally modeled on affine subspaces of the Lie algebra $\mathfrak{l}$. However the corresponding affine subspaces of $\mathfrak{l}$ need not correspond to Lie subalgebras. Additionally beds may have complicated topology (e.g. boundary). This will be discussed further in [SmWe3].

For our analysis, the following property of a bed will be helpful:

Definition 11.4. *We will say that a sequence of sets $K_1, K_2, \ldots$* fills out *$\mathcal{B}$ if the K_i are compact, and $\mu\left(\bigcup_i K_i\right) = 1$ for any measure μ which belongs to $\mathcal{B}$.*

Proposition 11.5. *For each type $j \leqslant 6$, all sheets in $\mathcal{E}_D(1,1)$ of type j are contained in a finite or countable union of beds, and for each* *bed $\mathcal{B}$ there is a countable sequence of compact sets which fills out $\mathcal{B}$. Explicitly these properties are satisfied by the following choices:*

- *For types $j \in \{3, 4, 6, 7\}$, we let $\mathcal{B}$ be the union of all sheets of type j, and we let K_i be the closure of the set of $M \in \mathcal{E}_D(1,1)$ which have no saddle connections shorter than $1/i$, and such that $\{\delta \in \Xi(M) : \|\mathrm{hol}(M,\delta)\| \leqslant i\}$ and $\Xi(\mathcal{B})$ are the same (as horizontal data diagrams).*
- *For types $j \in \{1, 2\}$ we let $\mathcal{B}$ be the union of all sheets of type 1 or 2, and we define K_i as before.*
- *If $j = 5$, for each closed G-orbit $\mathcal{L}$ we take $\mathcal{B} = \bigcup_{t\in\mathbb{R}} \mathrm{Rel}_t(\mathcal{L})$. Let ν_0 be the Haar measure on $\mathcal{L}$, let $\mathcal{L}_\infty$ be as in (52), let $L_1, L_2, \ldots$ be a nested sequence of compact subsets of $\mathcal{L}_\infty$ with $\nu_0\left(\bigcup_i L_i\right) = 1$, and let*

$$K_i \stackrel{\mathrm{def}}{=} \bigcup_{|t|\leqslant i} \mathrm{Rel}_t(L_i).$$

Proof. It is clear that the beds listed above contain all sheets of type j. In all cases except $j = 5$ there is just one bed. In the case $j = 5$ the number of beds is at most countable, since each $\mathcal{E}_D(1,1)$ contains at most finitely many closed G-orbits.

We now show that in all cases, the sets K_i listed above fill out the bed. Suppose first that $\Xi(\mathcal{B}) \neq \varnothing$. It is clear that each K_i as in the statement of the proposition is compact. The set $\bigcup_i K_i$ contains all $M \in \mathcal{B}$ for which $\Xi(M) = \Xi(\mathcal{B})$, so Corollary 5.3 implies that $\mu\left(\bigcup_i K_i\right) = 1$ for any measure μ which belongs to $\mathcal{B}$.

Now suppose $\mathcal{B}$ is a bed of type (5). Let $\mathcal{L}$, ν_0, L_i, K_i be as above. Each K_i is compact by Proposition 3.3. Also, for any measure μ belonging to $\mathcal{B}$, there is $t \in \mathbb{R}$ such that $\mu = \nu_t \stackrel{\mathrm{def}}{=} \mathrm{Rel}_{t*}\nu_0$. Since $\nu_0\left(\bigcup_i L_i\right) = 1$, we have $\nu_t\left(\bigcup_i \mathrm{Rel}_t(L_i)\right) = 1$ for all t. For each t, since $\bigcup_i K_i$ contains $\bigcup_i \mathrm{Rel}_t(L_i)$ for each t, we have $\nu_t\left(\bigcup_i K_i\right) = 1$ for each t. $\square$

The following result summarizes a strategy for proving equidistribution results which we will use repeatedly.

Proposition 11.6. *Let $\{\mu_t\}$ be a collection of measures where t ranges over either the positive integers or non-negative real numbers. Suppose the following hold:*

(a) *The sequence (μ_t) as $t \to \infty$ has no escape of mass; i.e. for any $\varepsilon > 0$ there is a compact $K \subset \mathcal{E}_D(1,1)$ and t_0 such that for all $t \geqslant t_0$, $\mu_t(K) \geqslant 1 - \varepsilon$.*

(b) *Any convergent subsequence* (μ_{t_k}) *of* (μ_t) *with* $t_k \to \infty$ *converges to a measure which is invariant under a conjugate of* U *by an element of* G.

(c) *For any bed* $\mathcal{B} \subsetneq \mathcal{E}_D(1,1)$ *there is a sequence* $K_1, K_2, \ldots$ *of sets which fill out* $\mathcal{B}$, *and for any* i *and any* $\varepsilon > 0$, *there is* t_0 *and an open set* $\mathcal{U}$ *containing* K_i *such that for all* $t \geqslant t_0$, $\mu_t(\mathcal{U}) < \varepsilon$.

Then the sequence μ_t *converges to the flat measure on* $\mathcal{E}_D(1,1)$ *as* $t \to \infty$.

Proof. It is enough to show that any subsequence (μ_{t_k}) of (μ_t), with $t_k \to \infty$ contains a further subsequence converging to the flat measure. So re-indexing, suppose (μ_t) is already a subsequence. Since there is no escape of mass, the set (μ_t) is precompact with respect to the weak-$*$ topology. So passing to a subsequence we can assume that μ_t converges weak-$*$ to a probability measure ν and we need to show that ν is the flat measure. By assumption (b), ν is invariant under a conjugate $U' = g^{-1}Ug$ of U, and hence $g_*\nu$ is U-invariant. We need to show that $g_*\nu$ is the flat measure, as this will imply that ν is the flat measure as well. To simplify notation we therefore replace $g_*\nu$ with ν to assume that ν is U-invariant. By ergodic decomposition, and since there are countably many beds, we can write $\nu = \sum_1^7 \nu_j$ where each ν_j is supported on the beds of type j, and is a convex combination of the measures belonging to these beds. We must show that $\nu_1 = \cdots = \nu_6 = 0$. We derive this from assumption (c), as follows.

Fix a type $j \leqslant 6$ and let $\mathcal{B} \subsetneq \mathcal{E}_D(1,1)$ be a bed of type j. Let $K_1, K_2, \ldots$ be compact sets as described in assumption (c). Since $\mu(\bigcup K_i) = 1$ for every U-invariant ergodic measure μ belonging to $\mathcal{B}$, and each μ belonging to $\mathcal{B}$ satisfies $\mu\left(\bigcup_i K_i\right) = 1$, in order to show $\nu_j = 0$ it suffices to show that $\nu_j(\bigcup K_i) = 0$. Suppose by contradiction that $a \stackrel{\text{def}}{=} \nu_j(\mathcal{E}_D(1,1))$ is strictly positive. For any i, and any $\varepsilon > 0$, let t_0 and $\mathcal{U}$ be as in (c). There is a continuous compactly supported function $\varphi : \mathcal{E}_D(1,1) \to [0,1]$ which is identically 1 on K_i and vanishes outside $\mathcal{U}$. The definition of the weak-$*$ topology and condition (c) now ensure that $\nu_j(K_i) \leqslant \int_{\mathcal{E}_D(1,1)} \varphi \, d\nu_j \leqslant \frac{1}{a} \lim_k \int_{\mathcal{E}_D(1,1)} \varphi \, d\mu_{t_k} \leqslant \frac{\varepsilon}{a}$. Since $\varepsilon > 0$ was arbitrary we must have $\nu_j(K_i) = 0$ for each i, and hence $\nu_j\left(\bigcup K_i\right) = 0$. ◻

Proof of Theorem 11.1. Step 1: M is of types (1–6). If $\Xi(M)$ decomposes M into horizontal cylinders then the U-action on the orbit-closure of M is conjugate to an irrational straightline flow on a torus and so every orbit is equidistributed. This is what happens when $\Xi(M)$ is a complete separatrix diagram, i.e. in cases (1) or (2). When $\Xi(M) \neq \varnothing$

but $\Xi(M)$ is not a complete separatrix diagram, the U-action on $\overline{UM}$ is obtained from the U-action on a simpler space via a U-equivariant map Ψ_T. Namely, by Propositions 8.8, in each of the cases (3), (4) and (6) this simpler space is a finite volume homogeneous space G/Γ, and $M = \Psi_T(M_0)$ where $M_0 \in G/\Gamma$ does not lie on a periodic U-orbit, since $\Xi(M)$ is not a complete separatrix diagram. The equidistribution of UM_0 in G/Γ follows from [DS], and the equidistribution of M follows from the fact that Ψ_T is U-equivariant. The same argument applies when $\Xi(M) = \varnothing$ and M belongs to $\mathrm{Rel}_s\mathcal{L}$, for a closed G-orbit $\mathcal{L}$ and $s \in \mathbb{R}$.

Step 2: M is of type (7) and $\Xi(\mathcal{B}) \neq \varnothing$: When M is of type (7) we apply Proposition 11.6 to the collection of measures $\{\mu_t : t > 0\}$ defined by averaging along the orbit of M; that is, $\mu_t = \nu(t, M)$, where $\nu(M,t)$ is as in (4.13). Property (a) of Proposition 11.6 follows from Theorem 10.3 (I) and property (b) is immediate from the definition of μ_t. To verify (c) we first discuss beds $\mathcal{B}$ for which $\Xi(\mathcal{B}) \neq \varnothing$, adapting an argument of [EMaMo]. Let $K_1, K_2, \ldots$ be the sequence filling out the bed $\mathcal{B}$, as in Proposition 11.5. Fix i and let $\varepsilon > 0$. Then any surface in K_i contains a horizontal saddle connection of length at most i. Let $C(\delta)$ be the open set of surfaces whose shortest saddle connection is shorter than δ. By Theorem 10.3 (II) there is $\delta > 0$ (depending only on the stratum $\mathcal{H}(1,1)$) such that every surface M' without horizontal saddle connections satisfies

$$\limsup_{t\to\infty} \mu_{M',t}(C(\delta)) < \varepsilon \tag{59}$$

(where $\mu_{M',t} = \nu(M', t)$ is as in Definition 4.13). Let

$$t_0 < 2(\log \delta - \log i), \tag{60}$$

so that if M_1 has a horizontal saddle connection of length at most i, then $g_{t_0}M_1$ has a saddle connection of length less than δ; in other words, $K_i \subset \mathcal{U} \stackrel{\text{def}}{=} g_{-t_0}(C(\delta))$. Since $g_{t_0}u_s g_{-t_0} = u_{e^{t_0}s}$, setting $M' = g_{t_0}M$, we have $g_{t_0*}\mu_{M,t} = \mu_{M',e^{t_0}t}$ and (59) implies that

$$\limsup_{t\to\infty} \mu_{M,t}(\mathcal{U}) = \limsup_{t\to\infty} \mu_{M',t}(C(\delta)) < \varepsilon,$$

so (c) is satisfied.

Step 3: M is of type (7) and $\Xi(\mathcal{B}) = \varnothing$. That is $\mathcal{B}$ is a bed of type (5). The measures belonging to $\mathcal{B}$ are of the form $\mathrm{Rel}_{s*}\mu$, where $s \in \mathbb{R}$ and μ is the G-invariant measure on a closed G-orbit $\mathcal{L}$. Let $\mathcal{L}_\infty$ be as in (52). Set

$$F(M,x) \stackrel{\text{def}}{=} M \mathbin{+\!\!\!\!\!\!\;\bigcirc} (x,0) = \mathrm{Rel}_x M, \quad F(M,x,y) \stackrel{\text{def}}{=} M \mathbin{+\!\!\!\!\!\!\;\bigcirc} (x,y), \tag{61}$$

postponing for the moment questions of domain of definition of F. Let L_i and K_i be as in Proposition 11.5, so that the K_i fill out $\mathcal{B}$. Note that $L_i \subset \mathcal{L}_\infty$ for each i and

$$K_i = F\big(L_i \times [-i, i]\big).$$

We will verify (c) for the sets K_i. Given i and $\varepsilon > 0$, choose

$$j > \frac{(8+\varepsilon)(i+1)}{\varepsilon}. \tag{62}$$

Corollary 10.5 gives us a constant $\delta > 0$, an open set $\mathcal{W} \subset \mathcal{L}$ containing L_i, and a compact set Ω containing $\mathcal{W}$ such that: (1) F is well-defined, continuous and injective on $\Omega \times [-j, j] \times [-\delta, \delta]$, and (2) whenever $M' \in \mathcal{L}$ and $I \subset \mathbb{R}$ is an interval such that $u_s M' \in \mathcal{W}$ for some $s \in I$, we have

$$\frac{1}{|I|} \left|\{s \in I : u_s M' \notin \Omega\}\right| < \frac{\varepsilon}{4}. \tag{63}$$

Now set

$$\mathcal{U} \stackrel{\text{def}}{=} F\big(\mathcal{W} \times (-(i+1), i+1) \times (-\delta, \delta)\big). \tag{64}$$

Then $\mathcal{U}$ is a neighborhood of K_i, and we need to show that $\mu_{M,t}(\mathcal{U}) < \varepsilon$ for all sufficiently large t. That is, we need to find t_0 so that for $t > t_0$,

$$\frac{|\widehat{\mathcal{I}} \cap [0,t]|}{t} < \varepsilon, \quad \text{where } \widehat{\mathcal{I}} \stackrel{\text{def}}{=} \{s \in \mathbb{R} : u_s M \in \mathcal{U}\}. \tag{65}$$

Whenever $s_0 \in \widehat{\mathcal{I}}$ there are $M' = M'(s_0) \in \mathcal{W}$ and $(x, y) = (x(s_0), y(s_0)) \in (-(i+1), i+1) \times (-\delta, \delta)$ such that $u_{s_0} M = F(M', x, y)$. Since M is not of type (5) we have $y(s_0) \neq 0$. We define the following intervals:

$$\mathcal{I} = \mathcal{I}(s_0) \stackrel{\text{def}}{=} \{s : |x + sy| \leqslant i+1\}$$
$$\mathcal{J} = \mathcal{J}(s_0) \stackrel{\text{def}}{=} \{s : |x + sy| \leqslant j\}.$$

These are nested intervals with common centerpoint $s = -\dfrac{x(s_0)}{y(s_0)}$. If an interval I contains an endpoint of $\mathcal{J}$ and intersects $\mathcal{I}$ then it contains one of the two connected components of $\mathcal{J} \smallsetminus \mathcal{I}$. These two components have equal lengths since the two intervals share the same centerpoint and by (62),

$$\frac{|\mathcal{I} \cap I|}{|\mathcal{J} \cap I|} \leqslant \frac{2(i+1)|y_1|}{(j-i-1)|y_1|} < \frac{\varepsilon}{4}. \tag{66}$$

We claim that if $s \in \mathcal{J} \smallsetminus \mathcal{I}$ then either $u_s M' \notin \Omega$ or $u_{s+s_0} M \notin \mathcal{U}$. Indeed, suppose $s \in \mathcal{J} \smallsetminus \mathcal{I}$ and $u_s M' \in \Omega$. According to Proposition

3.4,

$$\begin{aligned} u_{s+s_0}M &= u_sF(M',x,y) = u_s(M' \mathbin{+\!\!\!\!\bigcirc} (x,y)) \\ &= u_sM' \mathbin{+\!\!\!\!\bigcirc} (x+sy,y) = F(u_sM', x+sy, y). \end{aligned}$$

The injectivity of F on $\Omega \times [-j,j] \times [-\delta,\delta]$ and $|x+sy| > i+1$ imply that $u_{s+s_0}M \notin \mathcal{U}$. This proves the claim.

For each $s_0 \in \mathcal{I}$, denote the translates by

$$\mathcal{J}'(s_0) \overset{\text{def}}{=} \mathcal{J}(s_0) + s_0, \quad \mathcal{I}'(s_0) \overset{\text{def}}{=} \mathcal{I}(s_0) + s_0.$$

The claim, combined with (63) and (66), imply that if $[0,t]$ contains an endpoint of $\mathcal{J}' = \mathcal{J}'(s_0)$, then

$$\frac{|[0,t] \cap \mathcal{J}' \cap \widehat{\mathcal{I}}|}{|[0,t] \cap \mathcal{J}'|} \leqslant \frac{|[0,t] \cap \mathcal{I}'|}{|[0,t] \cap \mathcal{J}'|} + \frac{|\{s \in [0,t] \cap \mathcal{J}' : u_{s-s_0}M' \notin \Omega|}{|[0,t] \cap \mathcal{J}'|} < \frac{\varepsilon}{2}.$$

Since $\widehat{\mathcal{I}}$ is covered by the intervals $\{\mathcal{J}'(s_0) : s_0 \in \widehat{\mathcal{I}}\}$, a standard covering argument shows that we can take a countable subcover $\mathcal{J}_1', \mathcal{J}_2', \dots$ satisfying

$$s \in \widehat{\mathcal{I}} \implies 1 \leqslant \#\{\ell : s \in \mathcal{J}_\ell'\} \leqslant 2.$$

In particular 0 is contained in at most two of the intervals $\mathcal{J}'(s_\ell)$ and if we take t_0 to be larger than the right endpoint of these two intervals, and $t > t_0$, $[0,t]$ will contain at least one of the endpoints of any $\mathcal{J}'(s_\ell)$ which intersects $[0,t]$. Then for any $t > t_0$ we will have:

$$\left|[0,t] \cap \widehat{\mathcal{I}}\right| \leqslant \sum_\ell \left|[0,t] \cap \mathcal{J}_\ell' \cap \widehat{\mathcal{I}}\right| \leqslant \frac{\varepsilon}{2} \sum_\ell |[0,t] \cap \mathcal{J}_\ell'| \leqslant \frac{\varepsilon}{2}\, 2\, |[0,t]|$$

and this proves (65). □

12. Additional equidistribution results

In this section we prove equidistribution for periodic horocycles in Theorem 1.3, for translates of G-invariant measures by the Rel flow in Theorem 1.4, for circle orbits of increasing radius in Theorem 1.6, and for pushforwards of invariant measures of minimal sets by the geodesic flow in Theorem 12.3.

Proof of Theorem 1.4. Let $\mathcal{L}$ be a closed G-orbit in $\mathcal{E}_D(1,1)$, let μ be the Haar measure on $\mathcal{L}$, and let $\mu_t = \mathrm{Rel}_{t*}\mu$. We will prove that as $t \to +\infty$ or $t \to -\infty$, the measure μ_t converges to the flat measure on $\mathcal{E}_D(1,1)$. For definiteness we discuss the case $t \to +\infty$, the second case being similar. It is enough to show that if $t_n \to \infty$ is any sequence of real numbers for which μ_{t_n} converges to some ν (where ν is not necessarily a probability measure), then ν is the flat measure (and in particular is a probability measure).

First we show that ν is a probability measure, i.e. that there is no escape of mass. This follows from statement (II) of Theorem 10.3 as follows. We need to show that for any $\varepsilon > 0$ there is a compact subset K_0 of $\mathcal{E}_D(1,1)$ such that for all t, $\mu_t(K_0) \geqslant 1-\varepsilon$. Given ε let K be the intersection of $\mathcal{E}_D(1,1)$ with the compact set in statement (II) of Theorem 10.3. Let φ be a continuous compactly supported function on $\mathcal{E}_D(1,1)$, with values in $[0,1]$, which is identically equal to 1 on K, and let $K_0 \stackrel{\mathrm{def}}{=} \operatorname{supp}\varphi$. Let M be a generic point for μ_t. According to Proposition 4.15, $M = \mathrm{Rel}_t(M')$ for $M' \in \mathcal{L}$, such that M' is generic for μ. In particular M' has no horizontal saddle connections, and hence neither does M. According to (II),

$$\liminf_{T\to\infty} \frac{1}{T} \left|\{s \in [0,T] : u_s M \in K\}\right| \geqslant 1-\varepsilon,$$

and therefore

$$\begin{aligned}\mu_t(K_0) \geqslant \int \varphi\, d\mu_t &= \lim_{T\to\infty} \frac{1}{T}\int_0^T \varphi(u_s M) ds \\ &\geqslant \liminf_{T\to\infty} \frac{1}{T}\int_0^T 1_K(u_s M) ds \geqslant 1-\varepsilon.\end{aligned}$$

Now we claim that $\nu(\mathcal{H}_\infty) = 1$, that is ν gives no mass to the set of surfaces with horizontal saddle connections. This follows by again expressing μ_t as the limit of an integral along a generic horocycle orbit, and repeating the argument of Step 2 of the proof of Theorem 11.1. Thus, in view of Claim 1 (see the proof of Theorem 9.1), it suffices to prove that ν is invariant under the 'horospherical foliation' UZ. It is clear that ν is U-invariant, and it remains to show that it is invariant under Rel_s for any $s \in \mathbb{R}$.

Let $\varphi \in C_c(\mathcal{E}_D(1,1))$. Since φ is uniformly continuous, for any $\varepsilon > 0$ there is a neighborhood $\mathcal{U}$ of the identity in G so that

$$|\varphi(M) - \varphi(gM)| < \varepsilon \quad \text{for any } M \in \mathcal{E}_D(1,1),\ g \in \mathcal{U}. \tag{67}$$

Now define

$$\tau(t,s) \stackrel{\mathrm{def}}{=} 2\log\left(1+\frac{s}{t}\right) \text{ and } \tau_n \stackrel{\mathrm{def}}{=} \tau(t_n, s).$$

By a matrix multiplication in N, and using Corollary 4.2 we see that these choices ensure that for any surface M with no horizontal saddle connections,

$$g_{\tau_n}\mathrm{Rel}_{t_n} M = \mathrm{Rel}_{t_n+s}\, g_{\tau_n} M. \tag{68}$$

Moreover $g_{\tau_n} \to \mathrm{Id}$ as $n \to \infty$. Then for n large enough, by (67),

$$\left|\int \varphi\, d\mathrm{Rel}_{t_n *}\mu - \int \varphi \circ g_{\tau_n}\, d\mathrm{Rel}_{t_n *}\mu\right| < \varepsilon.$$

On the other hand, using (68) and the fact that μ is supported on $\mathcal{L}_\infty$, we obtain

$$\int \varphi \circ g_{\tau_n}\, d\mathrm{Rel}_{t_n *}\mu = \int_{\mathcal{L}_\infty} \varphi(g_{\tau_n}\mathrm{Rel}_{t_n}(M))\, d\mu(M)$$
$$= \int_{\mathcal{L}_\infty} \varphi(\mathrm{Rel}_{t_n+s} g_{\tau_n} M)\, d\mu(M) = \int_{\mathcal{L}_\infty} \varphi(\mathrm{Rel}_{t_n+s} M) d\mu(M).$$

In the last line we used the fact that μ is invariant under g_τ for all τ. Putting these together we find that for sufficiently large n,

$$\left| \int \varphi\, d\mathrm{Rel}_{t_n *}\mu - \int \varphi\, d\left(\mathrm{Rel}_{t_n+s}\right)_* \mu \right| < \varepsilon,$$

and since ε was arbitrary,

$$\nu = \lim_{n\to\infty} \mathrm{Rel}_{t_n *}\mu = \lim_{n\to\infty} (\mathrm{Rel}_{s+t_n})_* \mu = \mathrm{Rel}_{s*}\nu.$$

□

Generalizing Theorem 1.4 we have:

Theorem 12.1. *In each of the cases (ii), (iii), (iv) of §8.4, let $T > 0$ and let Ψ_T be the map described in §8.4. Let μ_T be the pushforward of Haar measure under Ψ_T, as in Proposition 8.8. Then as $T \to \infty$, μ_T tends to the flat measure on $\mathcal{E}_D(1,1)$.*

The proof of Theorem 1.4 goes through almost verbatim. We leave the details to the reader.

Proof of Theorem 1.3. We apply Proposition 11.6 to

$$\mu_t \stackrel{\mathrm{def}}{=} g_{t*}\mu, \tag{69}$$

where $\int f\, d\mu = \frac{1}{p}\int_0^p f(u_s M) ds$ and $\{u_s M : s \in [0,p]\}$ is a closed horocycle of period p.

To verify that there is no escape of mass we use Theorem 10.3. Let ρ_0 be the length of the shortest horizontal saddle connection on M, let $t_0 \stackrel{\mathrm{def}}{=} -2\log\rho_0$ and let $t \geqslant t_0$. The choice of t_0 ensures that the shortest horizontal saddle connection on $g_t M$ has length at least 1. Since the orbit Ug_tM is periodic, the measure μ_t is identical to the measure obtained by averaging along any integer multiple of the period $e^t p$ of this orbit. The set of saddle connections for $g_t M$ whose length is shorter than 1 is finite and none of these is horizontal. Thus for each such saddle connection δ, $\mathrm{hol}(u_s g_t M, \delta)$ diverges as $s \to \infty$. Therefore if we take a sufficiently large multiple of the period $s_0 = ke^t p$, $k \in \mathbb{N}$, any saddle connection on M will have length greater than 1 either for the surface $g_t M$ or for the surface $u_{s_0} g_t M$. As a consequence, the hypothesis of Theorem 10.3 is satisfied with $\rho = 1$, and there is no escape of mass. This verifies hypothesis (a) of Proposition 11.6, and hypothesis (b) is obvious.

To verify hypothesis (c) we adapt the argument given in the proof of Theorem 11.1, retaining the same notation. Namely, in step 2 we verify (c) for beds with $\Xi(\mathcal{B}) \neq \varnothing$. We fix i and $\varepsilon > 0$, and note that the argument in the preceding paragraph, using Theorem 10.3, implies that there is $\delta > 0$ such that if t is large enough so that $g_t M$ has no horizontal saddle of length shorter than 1, then

$$\mu_t(C(\delta)) < \varepsilon. \tag{70}$$

Then we take t_0 small enough so that all surfaces in $g_{t_0}(K_i)$ have horizontal saddle connections shorter than δ, and set $\mathcal{U} \stackrel{\mathrm{def}}{=} g_{-t_0}(C(\delta))$. Then (70) and $\mu_{t+t_0} = g_{t_0 *}\mu_t$ imply that for all sufficiently large t, $\mu_t(\mathcal{U}) < \varepsilon$, as required.

Continuing with Step 3, it remains to verify (c) for beds with $\Xi(\mathcal{B}) = \varnothing$. We define F, L_i, K_i as in the proof of Theorem 11.1, recalling that surfaces in L_i have no horizontal saddle connections. Given ε and i, we define $\mathcal{U}$ via (64), and need to show that $\mu_t(\mathcal{U}) < \varepsilon$ for all sufficiently large t. For this it suffices to prove that

$$\left|\widehat{\mathcal{I}}\right| < \varepsilon\, e^t p, \quad \text{where } \widehat{\mathcal{I}} \stackrel{\mathrm{def}}{=} \{s \in [0, e^t p] : u_s g_t M \in \mathcal{U}\}. \tag{71}$$

Before proving (71), we claim that for all sufficiently large t, if $s_0 \in \widehat{\mathcal{I}}$ and $g_t u_{s_0} M = F(M', x, y)$ with $M' \in \mathcal{W} \subset \mathcal{L}$, then $y \neq 0$. Indeed, suppose otherwise, that is there are $t_n \to \infty$, $M'_n \in \mathcal{W}$, $|x_n| \leqslant i+1$ and $s_n \in [0, e^{t_n} p]$ such that

$$u_{s_n} g_{t_n} M = \mathrm{Rel}_{x_n}(M'_n).$$

Set $s'_n \stackrel{\mathrm{def}}{=} e^{-t_n} s_n$ so that $g_{t_n} u_{s'_n} M = \mathrm{Rel}_{x_n}(M'_n)$. which implies

$$u_{s'_n} M = \mathrm{Rel}_{x'_n} g_{-t_n} M'_n, \quad \text{where } x'_n \stackrel{\mathrm{def}}{=} e^{-t_n/2} x_n \to 0.$$

Our hypothesis that M does not belong to a closed G-orbit implies that $x_n \neq 0$. Since U commutes with $\mathrm{Rel}_{x'_n}$, it follows that

$$UM = \mathrm{Rel}_{x'_n} U g_{-t_n} M'_n, \tag{72}$$

and hence $U g_{-t_n} M'_n$ is a closed horocycle of period p on $\mathcal{L}$. There are only finitely many such closed horocycles on $\mathcal{L}$ and their union is a compact set disjoint from the closed orbit UM. Since $x'_n \to 0$, this contradicts (72), and proves the claim.

We now note that the argument given in the proof of Theorem 11.1 for proving (65) goes through, with the same notations, and proves (71). Indeed the only information we needed in the proof of (65) was that the number $y = y(s_0)$ considered in the proof was nonzero, which is exactly what we have shown in the preceding paragraph. □

Remark 12.2. *When D is a square, Theorem 1.3 can also be proved by exploiting the connection between $\mathcal{E}_D(1,1)$ and a homogeneous space* $\mathrm{SL}_2(\mathbb{R}) \ltimes \mathbb{R}^2/\Gamma$ *(see* [EMS]*) and using a theorem of Shah* [KSS, Thm. 3.7.6].

Generalizing Theorem 1.3 we have:

Theorem 12.3. *Let $\mathcal{O}$ be a minimal set for the U-action, and let μ be the U-invariant measure on $\mathcal{O}$. Suppose that $\mathcal{O}$ is not contained in a closed G-orbit. Then $g_{t*}\mu$ tends to the flat measure on $\mathcal{E}_D(1,1)$.*

Proof. If $\mathcal{O}$ is a closed U-orbit this follows from Theorem 1.3. According to Corollary 8.3, in the remaining case $\dim \mathcal{O} = 2$, the cylinder decomposition is of type (A), and the measure μ is invariant under UZ. Let ν be any limit point of $g_{t_n*}\mu$, for $t_n \to \infty$. By repeating the arguments given in the proof of Theorem 1.3 we find that ν is a probability measure, and gives zero mass to surfaces with horizontal saddle connections. Also clearly ν is UZ-invariant. Therefore, by Claim 1, ν is the flat measure on $\mathcal{E}_D(1,1)$. □

Proof of Theorem 1.5. We show the existence of $\nu = \lim_{t\to\pm\infty} g_{t*}\mu$ and describe this limit explicitly, for each of the measures μ in Theorem 9.1. We treat the cases $t \to +\infty$ and $t \to -\infty$ separately, beginning with the case $t \to +\infty$. If μ is of type 1, then the limit measure ν is either the flat measure on $\mathcal{E}_D(1,1)$ or a G-invariant measure on a closed G-orbit, by Theorem 1.3. If μ is of type 2, then Theorem 12.3 implies that ν is flat measure. If μ is of type 3, 4, or 6, then there is some $T \neq 0$ such that $\mu = \mu_T$ is the pushforward of a G-invariant measure under Ψ_T, for the map Ψ_T described in §8.4. Note that this map satisfies the following equivariance rule: for $M \in \mathcal{B}$, $g_t\Psi_T(M) = \Psi_{e^tT}(g_tM)$. This can be seen by examining the definition of Ψ_T in each case and using Proposition 3.4. Therefore we have $g_{t*}\mu = \mu_{e^tT}$, and by Theorem 12.1, the limit ν is the flat measure. If μ is of type 5 then $\mu = \mathrm{Rel}_{T*}\mu_0$, where μ_0 is a G-invariant measure on a closed G-orbit. If $T = 0$ then $\mu = \mu_0$ is G-invariant and in particular $\nu = \mu_0$. If $T \neq 0$ then the relation $g_t\mathrm{Rel}_T = \mathrm{Rel}_{e^tT}g_t$ (see Proposition 3.4) implies that $g_{t*}\mu = \mathrm{Rel}_{e^tT*}\mu_0$ and by Theorem 1.4, the limit measure ν is the flat measure. Finally in case 7 the measure μ is G-invariant and there is nothing to prove.

When $t \to -\infty$, for each of the measures of type 1, 2, 3, 4 or 6, almost every surface in the support of μ has at least one horizontal saddle connection of some fixed length. As $t \to -\infty$, the length of this saddle connection tends to zero and hence $g_{t*}\mu$ diverges in the space of measures on $\mathcal{H}(1,1)$. If μ is of type 5 then $\mu = \mathrm{Rel}_{T*}\mu_0$, where μ_0 is a G-invariant measure on a closed G-orbit. The commutation relation $g_t\mathrm{Rel}_T = \mathrm{Rel}_{e^tT}g_t$ implies that $g_{t*}\mu = \mathrm{Rel}_{e^tT*}\mu_0$ and since $t \to -\infty$, the measure tends to the G-invariant measure μ_0. Finally in case 7 the measure μ is G-invariant and there is nothing to prove. $\square$

We now collect some results which will be used in the proof of Theorem 1.6. Let $B(a,b) \stackrel{\mathrm{def}}{=} [-a,a] \times [-b,b]$, and for any $t > 0$ and $v \in \mathbb{R}^2$, set

$$E_{t,v} \stackrel{\mathrm{def}}{=} \{g_t r_\theta v : \theta \in [0, 2\pi]\}.$$

The following is an elementary fact about ellipses whose proof we omit.

Proposition 12.4. *Given $\varepsilon > 0$, suppose $r_1, r_2, \delta_1, \delta_2$ satisfy the inequalities*

$$r_1 < \varepsilon r_2, \quad \delta_1 < \varepsilon\delta_2. \tag{73}$$

Then for any $t > 0$, either $E_{t,v} \subset B(r_2, \delta_2)$, or

$$\frac{|\{\theta \in [0,2\pi] : g_t r_\theta v \in B(r_1,\delta_1)\}|}{|\{\theta \in [0,2\pi] : g_t r_\theta v \in B(r_2,\delta_2)\}|} < \varepsilon. \tag{74}$$

$\square$

We will also need the following analogue of Corollary 10.5.

Proposition 12.5. *Given positive ε, T, η, and a compact subset $L \subset \mathcal{L}_\infty$, there are positive t_0, δ, a neighborhood $\mathcal{W}$ of L and a compact set $\Omega \subset \mathcal{L}$ containing $\mathcal{W}$, such that the map* (51) *is well-defined, continuous and injective on $\Omega \times [-T,T] \times [-\delta,\delta]$, and for any $t \geqslant t_0$, for any interval $I \subset J \stackrel{\mathrm{def}}{=} \left[\frac{\pi}{2} - \eta, \frac{\pi}{2} + \eta\right] \cup \left[\frac{3\pi}{2} - \eta, \frac{3\pi}{2} + \eta\right]$, and any $M \in \mathcal{H}$, if there is $\theta_0 \in I$ such that $g_t r_{\theta_0} g_{-t} M \in \mathcal{W}$ then*

$$\frac{|\{\theta \in I : g_t r_\theta g_{-t} M \notin \Omega\}|}{|I|} < \varepsilon. \tag{75}$$

Proof. Using Proposition 3.3, we see that it suffices to prove an analogue of Proposition 10.4; namely, that for any positive η, ε, r and any compact $L \subset \mathcal{L}_\infty$, there is a neighborhood $\mathcal{W}$ of L, a compact $\Omega \subset \mathcal{H}$ containing $\mathcal{W}$, and t_0, such that surfaces in Ω have no horizontal saddle connections shorter than r, and for any $t \geqslant t_0$, any interval $I \subset J$ which contains θ_0 with $g_t r_{\theta_0} g_{-t} M \in \mathcal{W}$, the estimate (75) holds. This statement can be obtained from Proposition 10.4 as follows.

A matrix computation shows that

$$u_{e^t \tan\theta} = a(t,\theta) g_t r_\theta g_{-t}, \quad \text{where } a(t,\theta) \stackrel{\mathrm{def}}{=} \begin{pmatrix} (\cos\theta)^{-1} & 0 \\ -e^{-t}\sin\theta & \cos\theta \end{pmatrix}. \tag{76}$$

We write $a(t,\theta) = a_2(\theta)\, a_1(t,\theta)$, where

$$a_2(\theta) \stackrel{\text{def}}{=} \begin{pmatrix} (\cos\theta)^{-1} & 0 \\ 0 & \cos\theta \end{pmatrix}, \quad a_1(t,\theta) \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 0 \\ -e^t \tan\theta & 1 \end{pmatrix}.$$

Thus $a_1(t,\theta) \to \mathrm{Id}$ as $t \to \infty$, uniformly for $\theta \in J$, and $a_2(\theta)$ preserves horizontal saddle connections, changing their length by a factor of at most $C_1 \stackrel{\text{def}}{=} \max_{\theta\in J}(\cos\theta)^{-1}$. Also let C_2 be an upper bound on the derivative of the map $\theta \mapsto \tan\theta$ on the interval J. Let $L' = \bigcup_{\theta\in J} a_2(\theta)L$, which is also a compact subset of $\mathcal{L}_\infty$. We apply Proposition 10.4 with $C_1(r+1), \varepsilon/C_2, L'$ in place of r, ε, L, to obtain a neighborhood $\mathcal{W}'$ of L' and a compact set Ω' containing $\mathcal{W}'$, such that surfaces in Ω' have no horizontal saddle connections shorter than $C_1(r+1)$, and (55) holds. We define:

$$\Omega'' \stackrel{\text{def}}{=} \bigcup_{\theta\in J} a_2(\theta)^{-1}\Omega'.$$

Then Ω'' is a compact set, containing no surfaces with horizontal saddle connections of length less than $r+1$. Therefore for t_0 sufficiently large, surfaces in

$$\Omega \stackrel{\text{def}}{=} \bigcup_{t\geqslant t_0} a_1(t,\theta)^{-1}\Omega'' = \bigcup_{\theta\in J, t\geqslant t_0} a(t,\theta)^{-1}\Omega'$$

have no horizontal saddle connections of length less than r. We will make t_0 larger below. We also note that $a(t,\theta)^{-1}\Omega' \subset \Omega$ for every $t \geqslant t_0, \theta \in J$. Now we set

$$\mathcal{W} \stackrel{\text{def}}{=} \Omega \cap \bigcap_{\theta\in J, t\geqslant t_0} a(t,\theta)^{-1}\mathcal{W}',$$

so that if $g_t r_{\theta_0} g_{-t} M \in \mathcal{W}$ for some $\theta_0 \in J$ and $t \geqslant t_0$, then by (76), $u_{e^t \tan\theta_0} M \in \mathcal{W}'$. Since (55) holds for Ω' (with ε/C_2 instead of ε), we obtain that (75) holds for Ω. Finally we note that if t_0 is chosen sufficiently large, then $\mathcal{W}$ contains an open set containing L. ☐

Proposition 12.6. *Suppose $\mathcal{L}$ is a closed G-orbit in $\mathcal{H}(1,1)$ and $M \notin \mathcal{L}$. Then there are positive $\widetilde{\delta}, \widetilde{t}$ so that for all $t \geqslant \widetilde{t}$, if there is $\theta \in [0,2\pi]$ such that $g_t r_\theta M = M' + (x,y)$ with $M' \in \mathcal{L}$, then $e^{-t}x^2 + e^t y^2 \geqslant \widetilde{\delta}$.*

Proof. Assume by contradiction that there are $M_n \in \mathcal{L}$, $t_n \to \infty$, $\theta_n \in [0,2\pi]$ and x_n, y_n such that

$$e^{-t_n}x_n^2 + e^{t_n}y_n^2 \to 0 \tag{77}$$

and $g_{t_n} r_{\theta_n} M = M_n + (x_n, y_n)$. By Corollary 6.2, there are neighborhoods $\mathcal{U} \subset \mathcal{H}(1,1)$ containing M and $\mathcal{V} \subset \mathbb{R}^2$ containing 0, such that the map $(M', v) \mapsto M' + v$ is well-defined on $\mathcal{U}\times\mathcal{V}$. Set $v_n \stackrel{\text{def}}{=} r_{-\theta_n} g_{t_n}(x_n, y_n)$.

Then (77) implies that $v_n \to 0$ and hence for all sufficiently large n, the maps $M' \mapsto M' \mathbin{+\!\!\!\!\circ} \pm v_n$ are well-defined on $\mathcal{U}$. Set $\widetilde{M}_n \stackrel{\text{def}}{=} M \mathbin{+\!\!\!\!\circ} -v_n$, so that $\widetilde{M}_n \to M$. We have

$$\begin{aligned} M &= r_{-\theta_n} g_{-t_n} \left(M_n \mathbin{+\!\!\!\!\circ} (x_n, y_n) \right) = r_{-\theta_n} g_{-t_n} M_n \mathbin{+\!\!\!\!\circ} \left(r_{-\theta_n} g_{-t_n} (x_n, y_n) \right) \\ &= r_{-\theta_n} g_{-t_n} M_n \mathbin{+\!\!\!\!\circ} v_n, \end{aligned}$$

and this implies that $\widetilde{M}_n = r_{-\theta_n} g_{-t_n} M_n$. We have found a sequence in $\mathcal{L}$ converging to M, contrary to the assumption that $M \notin \mathcal{L}$. □

Theorem 12.7. *For any $M \in \mathcal{E}_D(1,1)$ which is not a lattice surface, let μ_t be the measure on $\mathcal{E}_D(1,1)$ defined by*

$$\int \varphi \, d\mu_t = \frac{1}{2\pi} \int_0^{2\pi} \varphi(g_t r_\theta M) \, d\theta, \text{ for all } \varphi \in C_c \left(\mathcal{E}_D(1,1) \right). \tag{78}$$

Then μ_t converges to the flat measure on $\mathcal{E}_D(1,1)$ as $t \to \infty$.

Proof. We will use Proposition 11.6. Hypothesis (a) was verified (for any translation surface M) in [EM]. To prove hypothesis (b), we note that μ_t is invariant under the conjugated group $\{g_t r_\theta g_{-t} : \theta \in [0, 2\pi]\}$. Conjugating we see that

$$g_t r_\theta g_{-t} = \begin{pmatrix} \cos\theta & -e^t \sin\theta \\ e^{-t} \sin\theta & \cos\theta \end{pmatrix}.$$

For every fixed s and t we set

$$\theta(s,t) \stackrel{\text{def}}{=} \arcsin\left(-s e^{-t} \right), \text{ so that } -e^t \sin\theta(s,t) = s.$$

Then $g_t r_{\theta(s,t)} g_{-t} \to u_s$ as $t \to \infty$. Therefore the limit measure ν is invariant under each u_s.

We now verify (c) for beds $\mathcal{B}$ with $\Xi(\mathcal{B}) \neq \varnothing$. Let $K_1, K_2, \ldots$ be the sequence filling out the bed $\mathcal{B}$, as in Proposition 11.5. Fix i and let $\varepsilon > 0$. Then any surface in K_i contains a horizontal saddle connection of length at most i. As before, let $C(\delta)$ be the open set of surfaces whose shortest saddle connection is shorter than δ. By [EM], there are positive t_1, δ such that for all $t \geqslant t_1$, $\mu_t(C(\delta)) < \varepsilon$. Let t_0 satisfy (60), so that $K_i \subset \mathcal{U} \stackrel{\text{def}}{=} g_{-t_0}(C(\delta))$. Since $\mu_{t+t_0} = g_{t_0 *}\mu_t$, for all $t > t_0 + t_1$ we have

$$\mu_t(\mathcal{U}) = \mu_{t_0 *}\mu_{t-t_0}(\mathcal{U}) = \mu_{t-t_0}(C(\delta)) < \varepsilon,$$

so (c) is satisfied.

It remains to verify (c) for beds with $\Xi(\mathcal{B}) = \varnothing$. Let $F, L_i, K_i = F(L_i \times [-i, i])$ be as in the proof of Theorem 11.1. Given i and $\varepsilon > 0$, choose

$$j > \frac{8(i+1)}{\varepsilon}. \tag{79}$$

Choose $\eta > 0$ small enough so that

$$|J| = 4\eta < \pi\varepsilon, \quad \text{where } J \stackrel{\text{def}}{=} \left[\frac{\pi}{2} - \eta, \frac{\pi}{2} + \eta\right] \cup \left[\frac{3\pi}{2} - \eta, \frac{3\pi}{2} + \eta\right]. \tag{80}$$

Let $\mathcal{W}, \Omega, t_0, \delta$ be as in Proposition 12.5, where we apply the proposition with η and with $j, L_i, \varepsilon/8$ instead of T, L, ε. Let $\widetilde{\delta}, \widetilde{t}$ be as in Proposition 12.6. Making δ smaller and t_0 larger, we can assume that $t_0 \geqslant \widetilde{t}$ and

$$j\delta < \widetilde{\delta}^2. \tag{81}$$

Now let

$$\delta' \in \left(0, \frac{\varepsilon\delta}{8}\right) \tag{82}$$

and set

$$\mathcal{U} \stackrel{\text{def}}{=} F\big(\mathcal{W} \times (-(i+1), i+1) \times (-\delta', \delta')\big). \tag{83}$$

Then $\mathcal{U}$ is a neighborhood of K_i, and we need to show that $\mu_t(\mathcal{U}) < \varepsilon$ for all $t \geqslant t_0$. That is, we need to verify

$$\left|\widehat{\mathcal{I}}\right| < 2\pi\varepsilon, \text{ where } \widehat{\mathcal{I}} \stackrel{\text{def}}{=} \{\theta \in [0, 2\pi] : g_t r_\theta M \in \mathcal{U}\}.$$

In light of (80) it suffices to show that $\left|\widehat{\mathcal{I}} \cap J\right| < \pi\varepsilon$. Whenever $\theta_0 \in \widehat{\mathcal{I}} \cap J$ there are $M' = M'(\theta_0) \in \mathcal{W}$ and $(x, y) = (x(\theta_0), y(\theta_0)) \in (-(i+1), i+1) \times (-\delta', \delta')$ such that $g_t r_{\theta_0} M = F(M', x, y)$. We define the following intervals:

$$\mathcal{I} = \mathcal{I}(\theta_0) \stackrel{\text{def}}{=} \{\theta \in [0, 2\pi] : g_t r_{\theta - \theta_0} g_{-t}(x, y) \in B(i+1, \delta')\}$$

$$\mathcal{J} = \mathcal{J}(\theta_0) \stackrel{\text{def}}{=} \{\theta \in [0, 2\pi] : g_t r_{\theta - \theta_0} g_{-t}(x, y) \in B(j, \delta)\}.$$

We think of $[0, 2\pi]$ as a circle by identifying 0 and 2π, and think of these subsets as arcs on the circle. In view of Proposition 12.6 and the choice of t_0, we have $e^{-t}x^2 + e^t y^2 \geqslant \widetilde{\delta}$. Using (81) and considering the two choices of θ which make $r_{\theta - \theta_0} g_{-t}(x, y)$ horizontal and vertical, we see that $\mathcal{J}$ does not coincide with the entire circle. Then (79), (82) and Proposition 12.4 imply that

$$\frac{|\mathcal{I}|}{|\mathcal{J}|} < \frac{\varepsilon}{8}. \tag{84}$$

We claim that if $\theta \in \mathcal{J} \smallsetminus \mathcal{I}$ then either $g_t r_{\theta - \theta_0} g_{-t} M' \notin \Omega$ or $g_t r_\theta M \notin \mathcal{U}$. Indeed, suppose $\theta \in \mathcal{J} \smallsetminus \mathcal{I}$ and $g_t r_{\theta - \theta_0} g_{-t} M' \in \Omega$. According to Proposition 3.4,

$$\begin{aligned} g_t r_\theta M &= g_t r_{\theta - \theta_0} g_{-t} g_t r_{\theta_0} M = g_t r_{\theta - \theta_0} g_{-t}(M' \mathbin{\text{⌖}} (x, y)) \\ &= g_t r_{\theta - \theta_0} g_{-t} M' \mathbin{\text{⌖}} g_t r_{\theta - \theta_0} g_{-t}(x, y). \end{aligned}$$

The injectivity of F on $\Omega \times [-j, j] \times [-\delta, \delta]$ implies that $g_t r_\theta M \notin \mathcal{U}$. This proves the claim.

The claim, combined with (74) and (75), implies that

$$\frac{\left|\widehat{\mathcal{I}} \cap \mathcal{J}\right|}{|\mathcal{J}|} \leqslant \frac{|\mathcal{I}|}{|\mathcal{J}|} + \frac{|\{\theta \in \mathcal{J} : g_t r_{\theta-\theta_0} g_{-t} M' \notin \Omega|}{|\mathcal{J}|} < \frac{\varepsilon}{4}.$$

We have covered $\widehat{\mathcal{I}} \cap J$ by the intervals $\left\{\mathcal{J}(\theta_0) + \theta_0 : \theta_0 \in \widehat{\mathcal{I}}\right\}$, and we pass to a subcover such that

$$\theta \in \widehat{\mathcal{I}} \cap J \implies 1 \leqslant \#\{\ell : \theta \in \mathcal{J}'_\ell\} \leqslant 2,$$

and obtain:

$$\left|\widehat{\mathcal{I}} \cap J\right| \leqslant \sum_\ell \left|\widehat{\mathcal{I}} \cap \mathcal{J}'_\ell\right| < \frac{\varepsilon}{4} \sum_\ell |\mathcal{J}'_\ell| < \pi\varepsilon.$$

□

Proof of Theorem 1.6. This follows immediately from Theorem 12.7 by an argument developed by Eskin and Masur [EM]. See [EMS, EMaMo, B2] for more details. □

References

[AEM] A. Avila, A. Eskin and M. Möller, *Symplectic and isometric* $\mathrm{SL}(2,\mathbb{R})$*-invariant subbundles of the Hodge bundle,* preprint (2012).

[B1] M. Bainbridge, *Euler characteristics of Teichmüller curves in genus two*, Geom. Topol. **11** (2007), 1887–2073.

[B2] M. Bainbridge, *Billiards in L-shaped tables with barriers,* Geom. Funct. Anal. **20** (2010), no. 2, 299–356.

[Bo] N. Bourbaki, **Éléments de Mathematique, Livre 6, Integration, Chaptre IX**, Hermann, Paris 1969.

[Boi] C. Boissy, *Labeled Rauzy classes and framed translation surfaces*, Ann. Inst. Fourier, to appear.

[BoSh] A. Borevich, I. Shafarevich, **Number theory**, Pure and Applied Mathematics, Vol. 20, Academic Press, New York-London, 1966

[BH] M. Bridson and A. Haefliger, **Metric spaces of non-positive curvature**, Grund. der Math. Wiss. **319** Springer-Verlag, Berlin, 1999.

[BS] T. Bridgeland and I. Smith, *Quadratic differentials as stability conditions*, arXiv:1302.7030 (2014).

[C] K. Calta, *Veech surfaces and complete periodicity in genus 2*, J. Amer. Math. Soc. **17** (2004) 871–908.

[CaSm] K. Calta and J. Smillie, *Algebraically periodic translation surfaces*, J. Mod. Dyn. Vol. 2, No. 2, 2008, 209–248., arXiv

[CW] K. Calta and C. Wortman, *On unipotent flows in* $\mathcal{H}(1,1)$, Erg. Th. Dyn. Sys. **30** (2010), no. 2, 379–398.

[D] S. G. Dani, *Invariant measures of horospherical flows on noncompact homogeneous spaces,* Invent. Math. **47** (1978), no. 2, 101–138.

[DS] S. G. Dani and J. Smillie, *Uniform distribution of horocycle orbits for Fuchsian groups*, Duke Math. J. **51** (1984), 185-194.

[E] W. Ebeling, **Functions of Several Complex Variables and their Singularities**, AMS (2007).

[EW] M. Einsiedler and T. Ward, **Ergodic theory with a view toward number theory**, Graduate texts in math. **259** (2011).

[EM] A. Eskin and H. Masur. Asymptotic formulas on flat surfaces. *Ergodic Theory Dynam. Systems*, 21(2):443–478, 2001.

[EMaMo] A. Eskin, J. Marklof and D. W. Morris, *Unipotent flows on the space of branched covers of Veech surfaces*, Erg. Th. Dyn. Sys. **26** (2006) 129–162.

[EMS] A. Eskin, H. Masur and M. Schmoll, *Billiards in rectangles with barriers.* Duke Math. J. **118** (2003) 427–463.

[EMZ] A. Eskin, H. Masur and A. Zorich, *Moduli spaces of abelian differentials: the principal boundary, counting problems, and the Siegel-Veech constants,* Publ. I.H.E.S. **97** (2003) 631–678.

[EMi] A. Eskin and M. Mirzakhani, *Invariant and stationary measures for the* $\mathrm{SL}(2,\mathbb{R})$ *action on moduli space,* preprint (2014).

[EMiMo] A. Eskin, M. Mirzakhani and A. Mohammadi, *Isolation theorems for* $\mathrm{SL}_2(\mathbb{R})$*-invariant submanifolds in moduli space,* (preprint) 2013.

[FaMa] B. Farb and D. Margalit, **A Primer on Mapping Class Groups**, Princeton Mathematical Series, 49. Princeton University Press (2012).

[F] S. Filip, *Splitting mixed Hodge structures over affine invariant manifolds,* preprint (2013).

[HKK] F. Haiden, L. Katzarkov and M. Konstevich, *Flat surfaces and stability structures*, arXiv:1409.8611 (2015).

[H] A. Hatcher, **Algebraic Topology**, Cambridge University Press (2002).

[HW] P. Hooper and B. Weiss, *The rel leaf and real-rel ray of the Arnoux-Yoccoz surface in genus 3,* preprint (2015).

[HPV] J. Hubbard, P. Papadopol and V. Veselov, *A compactification of Hénon maps in* $\mathbb{C}^2$ *as dynamical systems*, Acta Math. **184** (2000), 203–270.

[I] N. V. Ivanov, *Mapping class groups*, in **Handbook of Geometric Topology,** p. 523–633 (2002) North Holland.

[K] A. Kechris, **Classical Descriptive Set Theory**, Graduate Texts in Mathematics **156**, Springer (1995).

[KoZo] M. Kontsevich and A. Zorich, *Connected components of the moduli spaces of Abelian differentials with prescribed singularities,* Invent. Math. 153 (2003), no. 3, 631–678.

[KSS] D. Kleinbock, N. A. Shah and A. Starkov, *Dynamics of subgroup actions on homogeneous spaces of Lie groups and applications to number theory*, in **Handbook on Dynamical Systems, Volume 1A**, Elsevier Science, North Holland (2002).

[LM] E. Lindenstrauss and M. Mirzakhani, *Ergodic theory of the space of measured laminations,* Int. Math. Res. Not. IMRN **4** (2008).

[Ma] H. Masur, *Interval exchange transformations and measured foliations*, Ann. Math. (2) **115** (1982) 169–200.

[MaSm] H. Masur and J. Smillie, *Hausdorff dimension of sets of nonergodic measured foliations*, Ann. Math. **134** (1991) 455–543.

[MaTa] H. Masur and S. Tabachnikov, *Rational billiards and flat structures*, in **Handbook of dynamical systems, Enc. Math. Sci. Ser.** (2001).

[MYZ] C. Matheus, J.-Ch. Yoccoz and D. Zmiaikou, *Homology of origamis with symmetries*, preprint (2013), to appear in Ann. Inst. Fourier.

[McM1] C. T. McMullen, *Teichmüller geodesics of infinite complexity*, Acta Math. **191** (2003) 191–223.

[McM2] C. T. McMullen, *Billiards and Teichmüller curves on Hilbert modular surfaces,* J. Amer. Math. Soc. **16** (2003) 857–885.

[McM3] C. T. McMullen, *Dynamics of* $\mathrm{SL}_2(\mathbb{R})$ *over moduli space in genus two*, Ann. Math. (2) **165** (2007), no. 2, 397–456.

[McM4] C. T. McMullen, *Teichmller curves in genus two: torsion divisors and ratios of sines*, Invent. Math. **165** (2006), no. 3, 651–672.

[McM5] C. T. McMullen. *Teichmüller curves in genus two: Discriminant and spin,* Math. Ann., **333** 87–130 (2005).

[McM6] C. McMullen, *Prym varieties and Teichmüller curves,* Duke Math. J. **133** (2006), no. 3, 569–590.

[McM7] C. T. McMullen, *Foliations of Hilbert modular surfaces*, Amer. J. Math. **129** (2007) 183–215.

[McM8] C. T. McMullen, *Navigating moduli space with complex twists*, J. Eur. Math. Soc. (JEMS) **15** (2013) 1223–1243.

[MW1] Y. Minsky and B. Weiss, *Non-divergence of horocyclic flows on moduli spaces*, J. Reine Angew. Math. **552** (2002) 131–177.

[MW2] Y. Minsky and B. Weiss, *Cohomology classes represented by measured foliations, and Mahler's question for interval exchanges*, Annales scientifiques de l'ENS **47** (2014).

[R] M. Ratner, *Raghunathan's conjectures for* $\mathrm{SL}(2,\mathbb{R})$, Israel J. Math. **80** (1992), 1-31.

[Sch] M. Schmoll, *Spaces of elliptic differentials*, in **Algebraic and topological dynamics**, S. Kolyada, Yu. I. Manin and T. Ward eds., Cont. Math. **385** (2005) 303–320.

[SmWe1] J. Smillie and B. Weiss, *Minimal sets for flows on moduli space*, Isr. J. Math, **142** (2004) 249–260.

[SmWe2] J. Smillie and B. Weiss, *Finiteness results for flat surfaces: a survey and problem list*, in **Partially hyperbolic dynamics, laminations, and Teichmüller flow** (Proceedings of a conference, Fields Institute, Toronto Jan 2006), G. Forni (ed.)

[SmWe3] J. Smillie and B. Weiss, *Examples of horocycle invariant measures on the moduli space of translation surfaces.* In preparation.

[SmWe4] J. Smillie and B. Weiss, *Dynamics of the strong stable foliation on loci of quadratic differentials*. In preparation.

[Th] W. T. Thurston, **The geometry and topology of 3-manifolds**, lecture notes (1980), available at `http://library.msri.org/books/gt3m/`.

[Ve1] W. A. Veech, *Moduli spaces of quadratic differentials*, J. Analyse Math. **55** (1990) 117–171.

[Ve2] W. A. Veech, *Teichmüller curves in moduli space, Eisenstein series and an application to triangular billiards*, Invent. Math. **97** (1989), no. 3, 553–583.

[W1] A. Wright, *Cylinder deformations in orbit closures of translation surfaces*, preprint (2014).

[W2] A. Wright, *The field of definition of affine invariant submanifolds of the moduli space of abelian differentials*, Geom. Top. (2014).

[Y] J.-C. Yoccoz, *Interval exchange maps and translation surfaces*, in **Homogeneous flows, moduli spaces and arithmetic**, Clay Math. Proc. **10** (2007) 1–70.

[Zo] A. Zorich, *Flat surfaces*, in **Frontiers in number theory, physics and geometry,** P. Cartier, B. Julia, P. Moussa and P. Vanhove (eds), Springer (2006).

University of Indiana mabainbr@indiana.edu

University of Warwick j.smillie@warwick.ac.uk

Tel Aviv University barakw@post.tau.ac.il