Commentarii Mathematici Helvetici

Billiards on rational-angled triangles

Richard Kenyon* and John Smillie[†]

1. Introduction

Let P be a polygon in the plane. A point moves along a *billiard trajectory* in P if it moves with constant speed along a straight line in the interior of P and if it reflects off edges so that the speed is unchanged and the angle of incidence is equal to the angle of reflection. If a billiard trajectory hits a vertex then we do not define the path further. We are interested in statistical properties of trajectories. When are they periodic? If they are not periodic to what extent do they "fill up" the polygon?

Billiard trajectories are projections of orbits of a flow on a four dimensional phase space which we can think of as the tangent bundle of P. We can make this flow continuous away from the corners by identifying the appropriate inward and outward pointing vectors over the edges of P. We leave the flow undefined over the vertices of P. The billiard flow can be thought of as a geodesic flow and in particular as a Hamiltonian flow (a general reference for Hamiltonian flows is [Ad]). Hamiltonian flows can be quite simple or quite complex; a first step in analyzing the complexity of such a flow is to look for "integrals of motion". These are functions on the phase space which are invariant under the flow. For billiard flows the length of a tangent vector provides one such integral of motion. We have a second integral of motion in the special case in which the polygon P is rational i.e. each of its corner angles is a rational multiple of 2π. The existence of these two integrals of motion mean that the billiard flow is effectively taking place on

surfaces inside the phase space. Let M_θ be one of these surfaces.

In a Hamiltonian flow every integral of motion gives rise to a vector field. In our case the two integrals of motion give rise to a pair of non-zero vector fields on the surfaces M_θ away from the singular points. The two integrals of motion are in involution, which means that the vector fields commute. In particular this means that there is a canonical local coordinate system on M_θ with respect to which the flow is linear (away from the singularities). If the billiard flow were a smooth flow then the existence of two integrals of motion in involution would give us an *integrable* flow. In this case we could assert that the invariant surfaces are tori and that the flows on these tori are linear. However the presence of singularities for the billiard flow gives rise to singularities of the surfaces M_θ, and in fact these surfaces need not be tori–they can have any positive genus. Berry and Richens [BR] use the suggestive term *pseudo-integrable* for the billiard flow on a rational polygon. Even though the flows on these surfaces are locally linear the dynamics can be quite complicated and hard to describe explicitly, unlike the case of the torus. This suggests the idea that there is some class of rational billiard flows like the integrable flows in the smooth case whose special properties make the dynamics easy to describe. We review three such properties.

When M_θ is a torus with punctures then the flow is integrable in the strongest sense. Rational polygons with this property are called simply *integrable*. The integrable polygons are precisely the polygons which tile the plane by reflections. The list of integrable polygons is short. It consists of exactly the set of rectangles and three triangles: the equilateral triangle, the $30°$, $60°$, $90°$ triangle and the $45°$, $45°$, $90°$ triangle. In the integrable case the dynamics can be described with great precision. The following dichotomy property holds: in any direction either all bi-infinite orbits are closed or all bi-infinite orbits are uniformly distributed in the appropriate invariant surface. One can describe precisely the directions in which all orbits are closed and give the asymptotics of the counting function of closed orbits of a given length.

Gutkin (see [G] and [GJ2]) introduced a condition less restrictive than integrability which still allowed a precise description of the dynamics as in the integrable case. A billiard table is *almost integrable* if the surface M can be realized as a branched cover of the torus by an affine branching map so that the images of the singular points have rational coordinates. The class of polygons with this property can be described precisely. They are polygons which can be tiled by integrable polygons. Though this class of polygons is much larger than the integrable polygons it does not contain any "surprising" new examples. For a triangle, in particular, the condition of almost integrability is equivalent to the condition of classical integrability.

The situation changed with the paper by Veech [Ve1]. Veech made a twofold contribution. He defined a class of rational polygons, the *lattice polygons* (see definition in section 3), which contains the almost integrable polygons yet which continue to admit an explicit description of the dynamics. Secondly, Veech con-

structed a family of truly surprising examples which satisfy this lattice property: the regular n-gon (for $n \geq 5$) and many rational triangles. Unlike the case of the integrable and the almost integrable polygons it seems to be quite a subtle problem to determine when a given polygon satisfies the lattice condition. The determination of which acute and right triangles possess the lattice property is the subject of this paper.

The lattice polygons are those polygons for which the surface M_θ has a large affine automorphism group (see section 3). In this paper we introduce techniques to analyze affine automorphism groups and apply these techniques to the surfaces, M_θ, associated to acute and right rational triangles. In particular we find all acute lattice triangles whose angles are rational with denominator less than 10,000.

Theorem 1. *Let T be an acute non-isosceles rational triangle with angles α, β and γ which can be written as $p_1\pi/q$, $p_2\pi/q$ and $p_3\pi/q$ with $q \leq 10,000$. Then T is a lattice polygon if and only if (α, β, γ) is one of the following exceptional cases:*

$$(a)\left(\frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\right), \quad (b)\left(\frac{\pi}{5}, \frac{\pi}{3}, \frac{7\pi}{15}\right), \text{ or } (c)\left(\frac{2\pi}{9}, \frac{\pi}{3}, \frac{4\pi}{9}\right).$$

The examples (a) and (b) are mentioned in [Vo2]. Veech mentions the example (a) in [Ve1]. The remaining example (c) seems to be new. We discuss it in section 5.3.

For isosceles and right triangles we have the following result.

Theorem 2. *An acute isosceles triangle is a lattice polygon if and only if the apex angle is of the form π/n, $n \geq 3$. A right triangle is a lattice polygon if and only if the smallest angle is of the form π/n for $n \geq 4$.*

The hypothesis of Theorem 1, that q is less than or equal to 10,000, is satisfied by approximately 100,000,000,000 triangles. We conjecture that this hypothesis on q is unnecessary. That is to say we conjecture that the three triangles described in Theorem 1 are the only acute non-isosceles lattice triangles (see section 5.1).

The sufficiency of the hypotheses of Theorem 2 is not a new result. The case of right triangles is covered by Vorobets in [Vo2, sect. 4]. The isosceles case can be reduced to the right triangle case by an unfolding construction. See [Vo2, sect. 5] or [GJ2]. All of these examples are closely related to Veech's original example [Ve1, Ve2, EG]. We will prove the necessity of the hypotheses of Theorem 2 in section 6.

In addition to the work cited above we mention that Ward [W] and Vorobets find lattice examples and non-examples among obtuse triangles for which two angles have the form π/n.

In sections 2 and 3 we develop some well known background material in a manner adapted to the cutting and pasting constructions which we will use later on. Section 4 introduces a new scissors congruence invariant for surfaces with singular

flat structures (and having trivial rotational holonomy), which links the invariants of Sah-Arnoux-Fathi [Ax] and an invariant introduced by Kenyon [Ke]. In the case that the surface has a decomposition into parallel annuli of commensurable moduli (which, by a result of Veech, must be the case for lattice polygons), the invariant takes a particularly simple form.

In Section 5.1 we compute this invariant for the surfaces associated to acute rational triangles. Searching for those whose invariant has the desired form leads to a number-theoretic test for triangles to have the lattice property. This number theoretic test is easily implemented on a computer and it proves that all but 25 of the rational triangles with $q \leq 10000$ do not have the lattice property. Using standard techniques 22 of the 25 exceptional cases are shown not to have the lattice property. Two of the remaining three were known to have the lattice property and the third is shown to have the lattice property in section 5.3. Section 6 contains the proof of Theorem 2.

The appendix provides some results of a more elementary nature which are not required for the rest of the paper but nonetheless provide perspective on the main result. This section has overlap with the earlier independent work of Gutkin and Judge [GJ1] and [GJ2]. Our work and theirs have different, albeit related, perspectives. For example, we study translation surfaces with non-elementary automorphism groups while they investigate surfaces with lattice automorphism groups.

2. Background

2.1. Construction of the unit tangent bundle and frame bundle of a polygon

As explained in the introduction, the billiard flow takes place on the tangent bundle to P and it leaves invariant a family of surfaces, M_θ. Following [ZK], [BR] and [G] we will review the structure of the invariant surfaces in the tangent bundle $T(P)$. As is noted in [ZK], for certain values of θ the invariant surface M_θ is a surface with boundary. We will introduce a closely related object, the frame bundle $F(P)$, which has certain technical advantages.

A *polygon* is a region in $\mathbf{R}^2$ with piecewise linear boundary and connected interior. It need not be convex or even simply connected. Let $P \subset \mathbf{R}^2$ be a polygon in the plane. Let $e_1, \ldots, e_m$ denote the edges of P. Let $\rho_i \in O(2)$ denote the linear part of the reflection in the edge e_i. When a billiard trajectory, moving

in the direction v, hits an edge e_i the outgoing trajectory has direction $\rho_i(v)$. It is also possible and sometimes convenient to define a billiard flow on the space of frames. A frame is an orthonormal pair of vectors (v_1, v_2) based at a point p. A frame is translated to the frame based at $p + tv_1$ for t sufficiently small. When a trajectory hits an edge e_i the frame (v_1, v_2) is replaced by the frame $(\rho_i v_1, \rho_i v_2)$. It is convenient to identify a frame (v_1, v_2) with the matrix α whose columns are v_1 and v_2. Thus $\alpha \in O(2)$ is the linear map which takes (e_1, e_2) to (v_1, v_2). The effect of a reflection in edge e_i is to take a frame α to the frame $\rho_i \alpha$.

We will construct the tangent bundle, $T(P)$, of P. Let S denote the unit circle. We define an equivalence relation $\sim$ on $P \times S$ by setting $(p, v) \sim (p, \rho_i(v))$ when $p \in e_i$. Let $\approx$ denote the transitive closure of $\sim$ (this only affects the pairs (p, v) where p is a vertex of P and hence contained in two edges). Let $T(P) = P \times S / \approx$. In a similar way we now define the frame bundle, $F(P)$, of P (as we will see, there are certain technical advantages in considering the frame bundle instead of the tangent bundle). We define an equivalence relation on $P \times O(2)$ by setting $(p, \alpha) \sim (p, \rho_i \alpha)$ when $p \in e_i$ and letting $\approx$ be the transitive closure. We define $F(P)$ as a the quotient $P \times O(2) / \approx$. Both $T(P)$ and $F(P)$ admit billiard flows which are continuous away from the vertices. There is a semiconjugacy from the flow on the frame bundle $F(P)$ to the flow on the tangent bundle $T(P)$ which maps a frame (v_1, v_2) at a point p to the tangent vector v_1 at p. This map is generically two-to-one but it is one-to-one at vectors (p, v) where v is tangent to the edge containing p.

The frame bundle has a natural foliation or decomposition into surfaces (which we will denote by M_θ as in the case of the tangent bundle) arising from parallel translation. Each leaf of this foliation is invariant under the billiard flow. We can describe this foliation by describing the leaf topology on $F(P)$. The leaf topology is a new topology on the set $F(P)$ for which the path components are leaves. Let $F(P)_\ell$ denote the set $F(P)$ with this new topology which we now describe. Let $O(2)^\delta$ denote $O(2)$ with the discrete topology. Then $F(P)_\ell$ will be the result of taking the product topology on $P \times O(2)^\delta$ and then the quotient topology on $P \times O(2)^\delta / \approx$. In order to describe the path components of $F(P)_\ell$ it is convenient to introduce the group $\Gamma_0 \subset O(2)$ which is the group generated by the reflections $\rho_1, \ldots, \rho_m$. The polygon P is *rational* if Γ_0 is finite. If the boundary of P is connected this agrees with the previous definition. In this case Γ_0 is a dihedral group of order $2n$ for some $n \geq 2$.

Two points (p, α) and (p, β) are on the same leaf if $\alpha = \gamma \circ \beta$ for some $\gamma \in \Gamma_0$. Thus the space of path components (leaves) can be identified with the space of right Γ_0 cosets in $O(2)$. Let Γ_0^+ be the subgroup of orientation preserving elements of Γ_0. This group is a rotation subgroup of order n. The space of right cosets $\Gamma_0 \backslash O(2)$ is canonically identified with $\Gamma_0^+ \backslash SO(2)$ which we can realize as the interval $0 \leq \theta < 2\pi/n$ on the unit circle. Let us denote the corresponding leaf by M_θ.

There is a natural right action of $O(2)$ on $P \times O(2)$ coming from the right action

of $O(2)$ on itself. This corresponds to rotating or reflecting a frame relative to the coordinates determined by the frame. Since right multiplication commutes with left multiplication, the right action on $P \times O(2)$ preserves the equivalence relation $\approx$ and hence descends to an action of $O(2)$ on $F(P)$. This right action takes leaves of the foliation to leaves of the foliation and acts transitively on the space of leaves. We conclude that any two leaves are isomorphic and that any leaf can be obtained from any other simply by rotating or reflecting the parallel frame with which it comes equipped. (The tangent bundle has an analogous foliation. The generic leaf of the tangent bundle foliation is isomorphic to any leaf of the frame bundle foliation, but the tangent bundle foliation also contains singular leaves.)

Since any leaf can be obtained from any other we focus on the leaf M_0 which consists of pairs (p, γ) for $\gamma \in \Gamma_0$. The leaf M_0 has a natural framed translation structure. To show the dependence of M_0 on P we write it as $\tilde{P}$. Thus $\tilde{P} = P \times \Gamma_0 / \approx$. The group Γ_0 acts on $\tilde{P}$ by automorphisms which preserve the decomposition of $\tilde{P}$ into polygons. The surface $\tilde{P}$ was originally described in [ZK] and [BR].

In section 5 we will focus on the case of rational triangles. We will introduce some notation here and describe the group Γ_0. Let T be a rational triangle with angles α, β and γ. We can write these angles as $\alpha = r\pi/n$, $\beta = s\pi/n$ and $\gamma = t\pi/n$ where r, s and t are natural numbers and $n = r + s + t$. We further assume that the largest common divisor of r, s and t is 1. The triangle T is determined by the triple (r, s, t) and we will sometimes use this triple to denote T. The group Γ_0 is generated by reflections in edges and the group Γ_0^+ consists of words containing an even number of reflections. In particular Γ_0^+ is generated by pairs of reflections. When two edges e and e' meet at a vertex with angle θ the composition of the reflection through e and the reflection through e' is a rotation by 2θ. It follows that Γ_0^+ contains the rotations through angles 2α, 2β and 2γ. These rotations are contained in the subgroup of rotations generated by $2\pi/n$. The condition on the greatest common divisor implies that Γ_0^+ is equal to the group generated by $2\pi/n$. We conclude that Γ_0^+ has order n and Γ_0 has order $2n$.

2.2. Local cone structures on surfaces

The surfaces $\tilde{P}$ are obtained by gluing together polygons. Surfaces (with or without boundary) obtained this way are sometimes called "generalized polygons" or surfaces with "cone structures" [G] [Th] [Bo] [Ve1]. Such a surface without a boundary is sometimes called a flat surface with cone type singularities. We give a description here of this class of surfaces that will be useful for our purposes, focusing on their metric space properties. Our main interest is a subclass of these surfaces, the surfaces with translation structures which we define subsequently.

A *geodesic* in a metric space is a path $\rho : I \to X$ so that $d(\rho(x), \rho(y)) = |x - y|$. A metric space is a geodesic space if any two points can be connected by a geodesic. A general reference for metric spaces is [BH].

Let X be a metric space with metric d'. Following [ABN] we define the *cone* CX on X to be the quotient space $CX = X \times \mathbf{R}^+ / \sim$ where $(x, 0) \sim (y, 0)$ for all $x, y \in X$. We define a metric on CX as follows: If $s > 0$ and $t > 0$ then

$$d((x, s), (y, t)) = \begin{cases} \sqrt{s^2 + t^2 - 2st \cos(d'(x, y))}, & \text{if } d'(x, y) \leq \pi; \\ s + t, & \text{if } d'(x, y) > \pi. \end{cases}$$

If s or t is equal to zero (say for example that $t = 0$) we set

$$d((x, s), (y, 0)) = s.$$

The *cone point* in CX is the point corresponding to $(x, 0)$. The cone on a geodesic space is again a geodesic space.

Let S_ℓ be the circle of circumference ℓ with its natural length metric. Then CS_ℓ is homeomorphic to $\mathbf{R}^2$. When $\ell = 2\pi$ then CS_ℓ is isometric to $\mathbf{R}^2$. Let I_ℓ be the interval of length ℓ. The metric space CI_ℓ is homeomorphic to the upper half plane. It is isometric to the upper half plane when $\ell = \pi$. Let M be a geodesic space homeomorphic to a surface with boundary. We say that M has a *local cone structure* if every point p in M has a neighborhood which is isometric to a neighborhood of the cone point in CI_ℓ or CS_ℓ. We define the *cone angle* at p to be ℓ. We say that $p \in \text{int} M$ is a *regular* point if the cone angle is 2π. If p is a boundary point of P we say that p is a *regular boundary point* if the cone angle at p is π. The *singular points* are the points which are not regular. Let Σ denote the set of singular points. The set of singular points is discrete so that when M is compact Σ is finite. The set $M - \{\partial M \cup \Sigma\}$ is a smooth manifold. We can choose an atlas of charts for $M - \{\partial M \cup \Sigma\}$ each of which is an isometry to an open subset of $\mathbf{R}^2$.

Let $\gamma : I \to M$ be a geodesic in M. If $\gamma(t_0)$ is a non-singular point then γ is locally a straight line. We define a *segment* or *saddle connection* in M to be a geodesic connecting singularities which is non-singular in its interior.

One example of a surface (with boundary) with a natural cone structure is a polygon. The polygon $P \subset \mathbf{R}^2$ has a natural metric which gives it a local cone structure. This metric agrees with the metric inherited from $\mathbf{R}^2$ when P is convex but will otherwise not be the same. The geodesics in this metric are piecewise linear paths.

A second example is provided by the surfaces $\tilde{P}$ when P is a rational polygon. The appropriate metric on these surfaces will be described in the next section. The singularities of $\tilde{P}$ arise from vertices of P. It need not be the case that every vertex in P gives rise to a singularity in $\tilde{P}$. If P has a vertex with angle π/n for some $n \geq 2$ then the corresponding points in $\tilde{P}$ will have cone angle 2π and hence be non-singular. If all angles of a polygon have the form π/n then the polygon is integrable (see the introduction). The case of triangles with two angles of the form π/n has been studied by Veech, Ward and Vorobets. Our example (c) of Theorem 1 seems to be the first lattice example with only one angle of the form π/n.

If M has a local cone structure and Σ is the set of singular points of M then at each point of $M - \Sigma$ we have a coordinate chart with values in $\mathbf{R}^2$ which is a local isometry. These charts can be used to define a natural notion of *parallel translation* along a path in $P - \Sigma$. Given a path γ, from p to q, the parallel transport is a linear isometry from T_p, the tangent space at p, to T_q, the tangent space at q. Fix a non-singular interior point p in M and identify T_p with $\mathbf{R}^2$. A loop γ in $P - \Sigma$ based at p determines an element of $SO(2, \mathbf{R})$. This element depends only on the homotopy class of γ and defines a homomorphism from $\pi_1(M - \Sigma) \to SO(2, \mathbf{R})$ which we will refer to as the *rotational holonomy*.

We say M has a *translation structure* if the image of the rotational holonomy homomorphism of M is trivial. In this case there is a well defined notion of parallel translation of vectors which is independent of the path. Local cone structures on surfaces without boundary for which the image of the rotational holonomy is contained in group $\{Id, -Id\}$ arise in the study of quadratic differentials and pseudo-Anosov homeomorphisms (see [EG]). In [KMS] they are referred to as flat structures on surfaces. Veech refers to them as F-structures.

The result of parallel translation around loop centered at an interior singular point with cone angle α is a rotation by α. Thus when M has a translation structure all cone angles are multiples of 2π. The converse implication is not true however (see [Tr]). Thus the class of flat surfaces whose cone angles are multiples of 2π is strictly larger than the class of translation surfaces.

A *framed translation structure* on M is a translation structure with a choice of frame at some base point on the surface. Thus we have a natural notion of local $x-$ and $y-$ coordinates at every nonsingular point of M.

2.3. Gluing polygons and decomposing surfaces

We will give a formal description of the method of constructing surfaces with local cone structures by gluing together polygons. The construction of the surface $\tilde{P}$ from copies of P in section 2.1 provides an example but we will have other situations in which it is useful to build surfaces by gluing. In particular the precise definition of which gluings are allowed plays an important role in the construction of our geometric invariant in section 4.

Let $P_1, \dots, P_n$ be polygons in $\mathbf{R}^2$. Let $I_1, \dots, I_m$ be line segments in $\mathbf{R}^2$. Let $\phi_i : \coprod I_j \to \coprod P_k$ for $i = 1, 2$ be maps which are isometries on each component. Assume that the image $\phi_i(I_j)$ is contained in the boundary of P_k. Since ϕ_i is an isometry when restricted to I_j it follows that $\phi_i(I_j)$ is contained in an edge of P_k. We do not require that $\phi_i(I_j)$ be equal to the entire edge. We assume that interiors of segments are mapped disjointly, that is to say that if $\phi_i(\mathrm{int} I_j) \cap \phi_{i'}(\mathrm{int} I_{j'}) \neq \emptyset$ then $i = i'$ and $j = j'$. Say that $p \sim q$ if $\phi_1(x) = p$ and $\phi_2(x) = q$ for some x. Let $\approx$ be the equivalence relation generated by $\sim$. Define P to be the quotient space $\coprod P_k / \approx$.

We will next describe a metric on this quotient and mention some of its basic

properties. A metric space X with metric d is a *length space* if the distance between any pair of points is the infimum of lengths of paths between them. Whenever length spaces are glued together via isometries there is a natural metric on the resulting space which makes it a length space (see [BH]). In our case this metric is simple to describe. We have a notion of a piecewise linear curve in P. A curve $\rho : I \to P$ is piecewise linear if it is the concatenation of curves each of which is the image of a linear segment in a single polygon P_j. We have a metric on P so that the distance between points is the infimum of the lengths of piecewise linear paths between them.

Proposition 3. *Any compact surface with local cone structure can be obtained by gluing together polygons and any surface constructed by gluing finitely many polygons has a local cone structure.*

Proof. The fact that a surface with a local cone structure can be obtained by gluing together triangles is proved in [Bo]. The fact that gluing polygons yields a space with local cone structure is straightforward. The metric on this space is described above and makes it a length space. The fact that it is a geodesic space follows from the Hopf-Rinow theorem [BH] which proves that any compact length space is a geodesic space. $\qquad\square$

If the surface M has a translation structure then we can be more precise about the gluing functions.

Proposition 4. *If the maps $\phi_1 \circ \phi_2^{-1}$ in the above gluing definition are restrictions of translations of $\mathbf{R}^2$ then M admits a translation structure. Conversely if M admits a translation structure we can construct M from polygons in $\mathbf{R}^2$ so that the ϕ_i are translations.*

When the surface is defined as in this proposition then there is a natural framing of the translation structure, arising from the standard frame in $\mathbf{R}^2$.

The surfaces $\tilde{P}$ have translation structures. Indeed we can view these surfaces as the smallest "covers" of P which have translation structures.

A *decomposition* of a surface into polygons is an isometry from M to $\coprod P_k / \approx$ for some collection of polygons and maps. We say that the decomposition is *cellular* if the identifications ϕ_j map intervals onto entire edges. Cellular decompositions are also called "edge to edge" decompositions. A cellular decomposition of a surface M gives M a CW-structure with polygonal cells. The singularities of the translation structure of M will be vertices of this CW-structure but there may be other vertices of the decomposition which are not singularities of the translation structure. Any surface with a local cone structure has a cellular decomposition. If the surface has a translation structure then we can find a cellular decomposition as in Proposition 4.

2.4. Parallel forms and vector fields

Let us now assume that we are dealing with a compact surface M with a translation structure but without a boundary. We do not require that $M = \tilde{P}$ for some polygon P though that is our primary example.

Let p_0 be a point in $M - \Sigma$ which will serve as a base point. Let T_0 denote the tangent space at p_0. Let T denote the space of parallel vector fields on $M - \Sigma$. Let $\tau_* : T_0 \to T$ denote the map that assigns to a vector based at p_0 the parallel vector field on $M - \Sigma$ obtained by parallel translation. The spaces T_0 and T are canonically identified by τ. When M arises from a polygon P the parallel vector fields generate the "billiard flows" in different directions.

Let T_0^* denote the dual space to T_0. For each cotangent vector in T_0^* we can use parallel translation to extend it to a one-form in $M - \Sigma$. Let T^* denote the space of parallel one forms and let $\tau^* : T_0^* \to T^*$ denote the canonical map. A nonzero two-form on T_0 gives rise to a parallel volume form which gives a smooth measure λ on M. We will scale this two form so that the total λ-measure of M is one. Every parallel flow on M is defined λ-almost everywhere and every such flow is measure preserving.

Let $\gamma : [a, b] \to M$ be a continuous curve of finite length and assume that $\gamma((a, b)) \subset M - \Sigma$ and $\gamma|(a, b)$ is smooth. The *translational holonomy* of γ is the vector $h(\gamma) \in T_0$ defined by

$$h(\gamma) = \int_a^b \frac{d\gamma}{dt} dt$$

where we use the parallel translation to view $\frac{d\gamma}{dt}$ as an element of T_p. If there is no danger of confusing the translational holonomy with the rotational holonomy we will drop the word "translational".

Define a smooth chain c to be a formal sum of curves $c = \gamma_1 + \ldots + \gamma_n$ where the γ_j are curves as described above. We define the holonomy of a chain c to be $h(c) = \sum h(\gamma_i)$.

For any $w \in T_0^*$ the holonomy of c satisfies:

$$\langle h(c), w \rangle = \langle c, \tau^*(w) \rangle.$$

where the bracket on the left denotes the pairing of vectors and covectors and the bracket on the right denotes the evaluation of a chain on a one-form. If c is a smooth cycle contained in $M - \Sigma$ which is homologous to zero in $M - \Sigma$ then, since parallel forms are closed, it follows that $\langle c, \tau^*(w) \rangle = 0$ for any w. This implies that $h(c) = 0$. In fact the same conclusion holds more generally.

Lemma 5. *If c is a smooth cycle in M which is homologous to zero in M then* $h(c) = 0$.

Proof. Let p be an interior singular point with cone angle α. For ϵ sufficiently small the set of points at distance ϵ from p is a closed curve $\gamma_{p,\epsilon}$ of length $\alpha\epsilon$. The

homology class represented by $\gamma_{p,\epsilon}$ in $M - \Sigma$ is independent of ϵ and we denote it by $[\gamma_p]$.

It is clear that

$$||h(\gamma)|| \leq length(\gamma)$$

so

$$||h(\gamma_{p,\epsilon})|| \leq length(\gamma_{p,\epsilon}) = \alpha\epsilon.$$

Since $h(\gamma_{p,\epsilon})$ depends only on the homology class of $\gamma_{p,\epsilon}$ for $\epsilon > 0$ it is independent of ϵ. We conclude that $h(\gamma_{p,\epsilon}) = 0$.

For any cycle c in M there is a homologous cycle c_ϵ which is contained in $M - \Sigma$ and agrees with c outside of a set of length less than ϵ. It follows that $||h(c_\epsilon) - h(c)|| \leq \epsilon$. Now if c is homologous to zero in M then c_ϵ is homologous to a sum of classes $[\gamma_{p_j}]$ so

$$h(c) = \lim_{\epsilon \to 0} h(c_\epsilon) = \lim_{\epsilon \to 0} \sum h([\gamma_{p_j}]) = 0$$

$\square$

If c is a cycle let $[c]$ denote its cohomology class in $H_1(M, \Sigma; \mathbf{Z})$. It follows from the previous lemma that the holonomy defines a homomorphism $h : H_1(M, \Sigma; \mathbf{Z}) \to \mathbf{R}^2$ as well as $h : H_1(M; \mathbf{Z}) \to \mathbf{R}^2$. It also follows from the lemma that the parallel one forms, which a priori represent elements of $H^1(M - \Sigma; \mathbf{R})$ in fact represent elements of $H^1(M; \mathbf{R})$. Let $A^* \subset H^1(M; \mathbf{R})$ denote the subspace of classes represented by parallel one forms.

Just as parallel one-forms define cohomology classes, parallel vector fields define homology classes. Parallel vector fields give rise to certain asymptotic homology classes (cf [Sc]) as we now explain. Let $v \in T$ be a parallel vector field. Let f_t be the (partially defined) flow generated by v. Let p be a point so that the flow through p is defined for all time (λ-almost every point has this property). Given $R > 0$ the curve $t \mapsto f_t(p)$ for $t \in [0, R]$ can be closed up by adding a segment of bounded length. Let $[v, p, R]$ denote the homology class represented by this closed curve. Define

$$[v, p] = \lim_{R \to \infty} \frac{1}{R}[v, p, R].$$

The limit exists almost everywhere and is independent of the choices made. Define

$$[v] = \int [v, p] d\lambda(p).$$

If the flow f_t is ergodic then the class $[v, p]$ is the same for λ-almost every p. In this case $[v]$ is equal to this common value. Let $\tau_* : T \to H_1(M, \Sigma; \mathbf{R})$ denote the map that assigns a vector v to its asymptotic homology class $[v]$. Let $A_* \subset H_1(M, \Sigma; \mathbf{R})$ or $A_* \subset H_1(M; \mathbf{R})$ (as the case may be) denote the image of τ_*. We will call A_* the space of *parallel homology classes*.

Though the definition of parallel homology class is rather involved, the pairing between parallel homology classes and one forms has a simple expression. If θ is a 1-form defined on $M - \Sigma$ with bounded coefficient functions then

$$\langle [\tau_*(v)], [\theta] \rangle = \int \theta_p(v) d\lambda(p).$$

Here $\theta_p(v)$ denotes the value of the form θ evaluated on the vector v based at the point p. The brackets on the left denote the Kronecker pairing of homology and cohomology. The formula is proved by interchanging the order of taking limits and integrating. The formula gives two corollaries. We see that the value of the right hand side of the equation depends only on the cohomology class of θ. We also see that the value of the left hand side of the equation is linear in v.

When we specialize the previous formula to the case where θ is a parallel one-form we get:

$$\langle \tau_*(v), \tau^*(w) \rangle = \langle v, w \rangle.$$

Lemma 6. *The map $\tau_* \circ h$ is a projection from $H_1(M, \Sigma; \mathbf{R})$ to itself. The kernel is the space of cycles with trivial holonomy and the image is A_*, the space of asymptotic cycles.*

Proof. Let c be a cycle. To show that $\tau_* \circ h$ is a projection we need to show that $\tau_* \circ h \circ \tau_* \circ h(c) = \tau_* \circ h(c)$. Let $v = h(c)$. It suffices to show that $h \circ \tau_*(v) = v$. This will follow if we can show that

$$\langle h \circ \tau_*(v), w \rangle = \langle v, w \rangle$$

for any $w \in T^*$. But

$$\langle h \circ \tau_*(v), w \rangle = \langle \tau_*(v), \tau^*(w) \rangle = \langle v, w \rangle.$$

$\square$

We conclude that the kernel of the holonomy is a complementary subspace to the space of asymptotic homology classes.

3. Affine automorphism groups

Now let M be a surface without boundary with a translation structure. An *affine homeomorphism* $f : M \to M'$ is a homeomorphism which takes singularities of M to singularities of M' and is locally affine in the nonsingular part of M. Affine homeomorphisms take geodesics to geodesics but need not preserve length.

The image of a parallel vector field under an affine homeomorphism is again a parallel vector field. Thus an affine automorphism f gives rise to a linear transformation from T to itself which we denote by Df.

An affine homeomorphism from a surface to itself is called an *affine automorphism*. Fix a surface M with a framed translation structure and let Γ denote the group of affine automorphisms. The function which takes an affine automorphism f to its derivative Df gives a homomorphism from Γ into $GL(2, \mathbf{R})$. Since the area of M is preserved Df has determinant ± 1. Let $SL^{\pm}(2, \mathbf{R})$ denote the subgroup of $GL(2, \mathbf{R})$ for which the determinant is ± 1. The image $D(\Gamma) \subset SL^{\pm}(2, \mathbf{R})$ of Γ is a discrete subgroup and the kernel of the homomorphism is finite. These affine automorphism groups arise in different contexts. They occur as automorphism groups of Teichmüller disks [EG]. The question of what groups can occur as automorphism groups of translation structures was raised by Thurston in [Kr]. Theorem 28 and Corollary 29 give conditions which these groups must satisfy.

Example. If the surface $\tilde{P}$ arises from the Zemlyakov-Katok construction applied to a polygon P then the dihedral group Γ_0 of section 2.1 acts as a group of affine automorphisms. Let $\iota : \Gamma_0 \to \Gamma$ denote this homomorphism. Then ι is injective and $D(\iota(\gamma)) = \gamma$ where we view Γ_0 as a subgroup of $O(2)$.

A polygon P has the *lattice property* ([Ve1]) if $D(\Gamma)$ is a lattice in $SL^{\pm}(2, \mathbf{R})$, that is, the quotient $SL^{\pm}(2, \mathbf{R})/D(\Gamma)$ has finite volume with respect to Haar measure. (This property is called the *Veech property* in [GJ1].) Equivalently, the image of Γ is a lattice in $PGL(2, \mathbf{R})$.

There is a standard classification of an element of $SL^{\pm}(2, \mathbf{R})$ into three types: elliptic, parabolic and hyperbolic. There is a corresponding classification of an affine automorphism f into three types based on the type of Df.

We will recall the geometric description of the corresponding affine automorphisms. An automorphism is *elliptic* exactly when it has finite order.

A non-trivial affine automorphism f of M is said to be *parabolic* if the linear map Df is parabolic. This is equivalent to saying that Df has a unique invariant direction. A *cylinder* in M is a subset isometric to the metric product $S^1 \times I$. We call the length of the I interval the height, width or length of the cylinder depending on the situation. We call the length of the S^1 factor the circumference of the cylinder. We assume that our cylinders are *maximal* meaning they are not contained in larger cylinders in M. This is equivalent to assuming that each boundary contains a singularity. The *modulus* of a cylinder is the ratio of its length to its circumference. Veech ([Ve1] see also [Vo2]) proves the following two results.

Proposition 7. *If a surface M has a parabolic affine automorphism f, then there is a decomposition of M into cylinders parallel to the fixed direction of Df. Furthermore the moduli of the cylinders are commensurable (have rational ratios).*

Proposition 8. *If M has a decomposition into vertical cylinders with commensurable moduli $m_1, \dots, m_k$ then $D(\Gamma)$ contains*

$$\begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix}$$

where α is the least common multiple of the moduli.

See Figure 7 for an example of a decomposition of a surface into commensurable cylinders.

Hyperbolic automorphisms are examples of pseudo-Anosov diffeomorphisms. These play a central role in the topological classification of surface homeomorphisms. In the hyperbolic case Df has two real eigenvalues λ_+ and λ_- with $|\lambda_+| > 1 > |\lambda_-|$. Let v_+ and v_- be eigenvectors of Df corresponding λ_+ and λ_- and let v^+ and v^- be dual basis vectors. Then the parallel forms $\tau(v^+)$ and $\tau(v^-)$ are taken to multiples of themselves. It follows that λ_+ and λ_- are eigenvectors of the induced automorphism of $H^1(M; \mathbf{R})$. In particular they are eigenvalues of a monic integral polynomial [Fa] or [Fri].

We recall here some elementary facts about discrete subgroups of $SL^{\pm}(2, \mathbf{R})$ (see [Fo]). There are three mutually exclusive possibilities:

(1) They can be finite. In this case they contain only elliptic elements. Conversely any discrete group which contains only elliptic elements is finite.

(2) They can contain a subgroup of finite index consisting of parabolic elements. In this case the index of the subgroup is at most two and the subgroup of parabolic elements is cyclic.

(3) They can contain a hyperbolic element.

Lattice subgroups are in the third category. They are cocompact if and only if they do not contain parabolic elements. The groups $D(\Gamma)$ are never cocompact:

Proposition 9. [Ve1] *If M satisfies the lattice property then for any saddle connection γ with holonomy v there is a parabolic affine automorphism for which the derivative fixes v.*

If $M = \tilde{P}$ arises from a polygon P by the Zemlyakov-Katok construction then the affine automorphism group of M contains the dihedral group Γ_0. Such a dihedral subgroup cannot be a subgroup of a group of type 2 above. In particular in this case if $D(\Gamma)$ contains a parabolic element then $D(\Gamma)$ also contains a hyperbolic element.

4. A scissors congruence invariant

We say generalized polygons P and P' are *scissors congruent* if we can write P as a union of polygons $P_1 \cup \ldots \cup P_n$ (as described in §2.3) and we can write P' as a union of polygons $P'_1 \cup \ldots \cup P'_n$, such that for each j, P_j is congruent to P'_j. An *additive scissors congruence invariant* is a homomorphism ϕ from the group of formal $\mathbf{Z}$-linear combinations of polygons to an abelian group, which has the property that if P is scissors congruent to P' then $\phi(P) = \phi(P')$. An additive scissors congruence invariant gives us an invariant for surfaces with framed translation structures by the following procedure: we decompose the surface a union of planar polygons and

sum the invariants of the polygons.

Unfortunately there is essentially only one scissors congruence invariant for polygons and that is area: P is scissors congruent to P' if and only if $\text{area}(P) = \text{area}(P')$. We get a finer notion of scissors congruence if we consider generalized polygons with translation structures P and P' so that P'_j be a *translate* of P_j for each j. There is an additional invariant for this finer notion, the Hadwiger invariant [Sa1]. Unfortunately this does not yield an interesting invariant for closed surfaces with translation structures because if a closed surface can be written as $P_1 \cup \ldots \cup P_n$ then the sum of the Hadwiger invariants is zero.

We get a still finer equivalence relation if we only permit cellular decompositions (as defined in section 2.3). We say two polygons P and P' are *equidecomposable* if P and P' have cellular decompositions $P = P_1 \cup \cdots \cup P_k$ and $P' = P'_1 \cup \cdots \cup P'_k$, where for each i, P'_i is a translate of P_i.

In this section we produce an additive invariant for the relation of equidecomposability and show that it gives a non-trivial invariant for framed translation structures on surfaces.

The boundary of a polygon $P \subset \mathbf{R}^2$ comes equipped with a natural geometric decomposition into edges and vertices where the vertices are the corners. This natural decomposition gives a natural a CW-structure on P where the vertices are 0-cells, the edges are 1-cells and P itself is a 2-cell. It will be useful to discuss CW-structures on polygons other than the natural geometric one. In particular the invariant that we will define depends on the choice of a CW structure on the boundary of P. In particular we will permit the addition of vertices or 0-cells for the CW structure which lie in the interior of edges and hence are not geometric vertices.

To be precise let us define the notion of a *polygon with cellular structure* to be a polygon, P, in the plane together with a specified set of points $v_1, \ldots, v_n$ on ∂P which includes the set of vertices of P. If the danger of confusion is small we will call a polygon with cellular structure simply a polygon. We say that a polygon P with vertices $v_1, \ldots, v_n$ is a translation of P' with vertices $v'_1, \ldots, v'_n$ if there is a translation which takes P to P' and takes the vertices of P to the vertices of P'.

We now define an equidecomposability invariant for polygons taking values in the rational vector space $\mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2$. Let P be a polygon with vertices $v_1, \ldots, v_n$ in counterclockwise order around the boundary of P. Define $J(P)$ to be

$$J(P) = v_1 \wedge v_2 + v_2 \wedge v_3 + \ldots + v_n \wedge v_1.$$

Lemma 10. *J is invariant under translations.*

Proof. Assume that $P' = P + v$ and $v'_j = v_j + v$. Then

$$J(P') = \left(\sum_{j=1}^{n-1} v'_j \wedge v'_{j+1} \right) + v'_n \wedge v'_1$$

$$= \Big(\sum_{j=1}^{n-1} (v_j + v) \wedge (v_{j+1} + v) \Big) + (v_n + v) \wedge (v_1 + v)$$

$$= \Big(\sum_{j=1}^{n-1} v_j \wedge v_{j+1} + v \wedge v_{j+1} + v_j \wedge v \Big) + v_n \wedge v_1 + v \wedge v_1 + v_n \wedge v$$

$$= J(P) + \sum_{j=1}^{n-1} v_j \wedge v + \sum_{j=1}^{n-1} v \wedge v_j$$

$$= J(P).$$

$\square$

It is useful to compare $\mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2$ to $\mathbf{R}^2 \wedge_{\mathbf{R}} \mathbf{R}^2$. Note that if v is a non-zero vector in $\mathbf{R}^2$ and λ is an irrational real number then $v \wedge_{\mathbf{Q}} \lambda v$ is not zero. As we will see adding an extraneous vertex to P changes the invariant by terms of this form.

Let π denote the natural map from $\mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2$ to $\mathbf{R}$ which takes $v \wedge w$ to $det(v, w)$. This map factors through $\mathbf{R}^2 \wedge_{\mathbf{R}} \mathbf{R}^2$. Note that $\pi(J(P)) = 2 \cdot area(P)$, which shows that we can recover the scissors congruence invariant from J.

For $j = 2, \dots, n$ let e_j be the vector $e_j = v_j - v_{j-1}$. Let $e_1 = v_1 - v_n$. We have the following expression for $J(P)$ in terms of these edge coordinates.

Lemma 11. $J(P) = \sum_{1 \le k < j \le n-2} e_k \wedge e_j$.

Proof. Since $J(P)$ is invariant under translation we may translate P so that $v_n = 0$. We get $v_1 = e_1$, $v_2 = e_1 + e_2$ and in general $v_j = e_1 + \dots + e_j$. Now

$$J(P) = \sum_{j=1}^{n-2} v_j \wedge v_{j+1}$$

$$= \sum_{j=1}^{n-2} \Big(\sum_{k=1}^{j} e_k \Big) \wedge \Big(\sum_{k=1}^{j+1} e_k \Big)$$

$$= \sum_{j=1}^{n-2} \Big(\sum_{k=1}^{j} e_k \Big) \wedge e_{j+1}$$

$$= \sum_{1 \le k \le j \le n-2} e_k \wedge e_j.$$

$\square$

Example. If T is a triangle with edge vectors e_1, e_2 and e_3 then

$$J(T) = e_1 \wedge e_2. \tag{1}$$

Example. If R is a rectangle with edge vectors $e_1, \ldots, e_4$ with $e_3 = -e_1$ and $e_4 = -e_2$ then

$$
\begin{aligned}
J(R) &= e_1 \wedge e_2 + e_1 \wedge e_3 + e_2 \wedge e_3 \\
&= e_1 \wedge e_2 + e_1 \wedge (-e_1) + e_2 \wedge (-e_1) \\
&= 2e_1 \wedge e_2.
\end{aligned}
$$

Lemma 12. *If a planar polygon P has a cellular decomposition $P = P_1 \cup \cdots \cup P_n$ then $J(P) = \sum_j J(P_j)$.*

Proof. By definition $J(P_j)$ is a sum over oriented edges e_k^j of terms $v_1 \wedge v_2$ where v_1 and v_2 are the endpoints of e_k^j. Now $\sum_j J(P_j)$ is again a sum over edges. Each internal edge occurs twice in this sum, once with each orientation. Thus these terms cancel. We are left with the boundary edges. Each of these occurs once and the orientation inherited from P_j is the same as the orientation inherited from P. $\square$

This proof does *not* imply that $J = 0$ for any surface without boundary: such a surface has a nontrivial translational holonomy, so one cannot assign unambiguously a coordinate v_j to each vertex in the subdivision.

Lemma 13. *Let P and Q be disjoint polygons in $\mathbf{R}^2$ that share an edge e. Let v be a point on e. Let P' and Q' be the result of adding v as a vertex to P and Q respectively. Then $J(P) + J(Q) = J(P') + J(Q')$.*

Proof. Let v_1 and v_2 be the endpoints of e with the order inherited from P. Adding the vertex v changes $J(P)$ by $v_1 \wedge v + v \wedge v_2 - v_1 \wedge v_2$. As an edge of Q, e inherits the opposite orientation so $J(Q)$ changes by $v_2 \wedge v + v \wedge v_1 + v_1 \wedge v_2$. Thus the change in $J(P)$ and the change in $J(Q)$ cancel. $\square$

Theorem 14. *Let M be a surface with framed translation structure. Say that we have two cellular decompositions of M into planar polygons, $M = P_1 \cup \ldots \cup P_n$ and $M = Q_1 \cup \ldots \cup Q_m$. Then $\sum_{j=1}^{n} J(P_j) = \sum_{j=1}^{m} J(Q_j)$.*

Proof. Let R_{ij} denote the non-empty intersections $P_i \cap Q_j \in M$. If R_{ij} has an edge in common with the boundary of M then we count any vertex on this edge coming from M as a vertex of R_{ij}. Otherwise the vertices of R_{ij} are just the boundary singularities (its corners). Even though each P_i is a union of polygons $\bigcup_j R_{ij}$, the R_{ij} do not constitute a cellular decomposition because the union of polygons R_{ij} may have additional vertices. We can remedy this by adding vertices to P_i. Let P_i' denote P_i where we have added the vertices (of all the R_{ij}) that lie on the

boundary of P_i. None of the added vertices lie on the boundary of M. Similarly define Q'_j. The union of polygons R_{ij} that lie in P_i gives a cellular decomposition of P'_i. By Lemma 12 we have $J(P'_i) = \sum_j J(R_{ij})$. On the other hand Lemma 13 gives $\sum_i J(P_i) = \sum_i J(P'_i)$ and $\sum_j J(Q_j) = \sum J(Q'_j)$. So

$$\sum_i J(P_i) = \sum_i J(P'_i) = \sum_{ij} J(R_{ij}) = \sum_j J(Q'_j) = \sum_j J(Q_j).$$

□

Definition. *Let M be a surface with framed translation structure with a cellular decomposition into planar polygons $P_1 \cup \ldots \cup P_n$. We define $J(M)$ to be $\sum_j J(P_j)$.*

The previous theorem shows that $J(M)$ does not depend on the decomposition of M into polygons. It does however depend on the choice of frame v in the translation structure on M. For example if we identify $\mathbf{R}^2$ with $\mathbf{C}$ in the usual way then, if $v' = e^{i\theta}v$ is another frame, then $J(M_v) = (e^{i\theta} \wedge e^{i\theta})J(M_{v'})$. Note in particular that a rotation by an angle of π leaves the invariant unchanged. More generally, we have the following.

Proposition 15. *Let $f : M \to M'$ be an affine diffeomorphism with differential $Df : \mathbf{R}^2 \to \mathbf{R}^2$ then $J(M') = \varepsilon \cdot (Df \wedge Df)J(M)$, where $\varepsilon = \pm 1$ according to whether or not f preserves orientation.*

Corollary 16. *If f is an affine automorphism of M then $J(M) = \varepsilon \cdot (Df \wedge Df)J(M)$.*

4.1. Cohomological interpretation of J

In this section will give two cohomological interpretations of the invariant $J(M)$ under the assumption that M has no boundary. We note here that one of the properties of J that makes it useful for computation is that it is defined and additive for translation surfaces with and without boundary.

Let $E\mathbf{R}^2$ denote the simplicial complex whose n-simplices correspond to finite $(n+1)$-tuples of elements of $\mathbf{R}^2$, so that the inclusion of simplices corresponds to inclusion of sets. We denote such an n-simplex in $E\mathbf{R}^2$ by $(v_0, \ldots, v_n)$.

The space $E\mathbf{R}^2$ is contractible and there is a natural simplicial action of $\mathbf{R}^2$ given by

$$v(v_0, \ldots, v_n) = (v + v_0, \ldots, v + v_n).$$

Let $B\mathbf{R}^2$ denote the quotient space of $E\mathbf{R}^2$ by the action of $\mathbf{R}^2$. The n-simplices of $B\mathbf{R}^2$ are often written with the following bar notation:

$$[e_1|e_2|\ldots|e_n],$$

which represents the equivalence class of the simplex with vertices

$$0, e_1, e_1 + e_2, \ldots, e_1 + \cdots + e_n.$$

The space $B\mathbf{R}^2$ is a classifying space for the additive group $\mathbf{R}^2$ with the discrete topology, that is to say that homotopy classes of maps from X to $B\mathbf{R}^2$ are in one-to-one correspondence with homomorphisms from $\pi_1(X) \to \mathbf{R}^2$. The simplicial cochains on $B\mathbf{R}^2$ correspond to cochains in the standard resolution of $\mathbf{R}^2$ so there is a canonical identification of the space (co-)homology of $B\mathbf{R}^2$ with the group (co-)homology of $\mathbf{R}^2$.

Let M be a surface with a framed translation structure. We assume for the rest of this section that M has no boundary. A polygonal decomposition of M into triangles induces a map $c : M \to B\mathbf{R}^2$ as follows. Let $\tau_1, \ldots, \tau_n$ denote the triangles in the triangulation. The framed translation structure on M gives each triangle τ_j the structure of a triangle in $\mathbf{R}^2$ (well defined up to translation). Thus this triangle corresponds to a simplex in $B\mathbf{R}^2$ and we define c on τ_j to be the affine map from τ_j to the corresponding simplex. The induced map $c_* : \pi_1(M) \to \mathbf{R}^2$ is just the holonomy homomorphism. In particular the homotopy class of c is independent of the particular triangulation chosen.

The elements of the group $H^1(M; \mathbf{R}^2)$ correspond to homomorphisms from $\pi_1(M) \to \mathbf{R}^2$. Let θ be the element corresponding to the holonomy homomorphism. We can define a cup product map from $H^1(M; \mathbf{R}^2) \otimes H^1(M; \mathbf{R}^2) \to H^2(M; \mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2)$ corresponding to the coefficient pairing $\mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2$. Let $\theta \cup \theta$ denote the corresponding element of $H^2(M, \mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2)$.

Proposition 17. *We have $J(M) = \langle \theta \cup \theta, [M] \rangle$.*

Proof. The cohomology group $H^1(\mathbf{R}^2; \mathbf{R}^2)$ is isomorphic to the group of group homomorphisms from $\mathbf{R}^2$ to itself. Such a homomorphism is automatically $\mathbf{Q}$ linear but need not be $\mathbf{R}$ linear. Let ι be the element of group cohomology corresponding to the identity map from $\mathbf{R}^2$ to itself. The map c induces a map c^* from $H^1(\mathbf{R}^2; \mathbf{R}^2)$ to $H^1(M; \mathbf{R}^2)$, and $c^*(\iota) = \theta$. We have

$$
\begin{aligned}
\langle \theta \cup \theta, [M] \rangle &= \langle c^*(\iota) \cup c^*(\iota), [M] \rangle \\
&= \langle c^*(\iota \cup \iota), [M] \rangle \\
&= \langle \iota \cup \iota, c_*[M] \rangle,
\end{aligned}
$$

where $c_* : H_2(M; \mathbf{Z}) \to H_2(\mathbf{R}^2, \mathbf{Z})$. We can describe the cup product in the cohomology of $\mathbf{R}^2$ on the cochain level using the Alexander-Whitney diagonal approximation [p. 110, Br]. Let $\tau = (v_0, v_1, v_2)$ represent the class of a simplex in $B\mathbf{R}^2$. Then

$$
\begin{aligned}
(\iota \cup \iota)(\tau) &= (\iota \cup \iota)(v_0, v_1, v_2) \\
&= \iota(v_0, v_1) \wedge \iota(v_1, v_2) \\
&= (v_1 - v_0) \wedge (v_2 - v_1) \\
&= J(\tau).
\end{aligned}
$$

Recall that $\tau_1, \ldots, \tau_m$ is a triangulation of M. Now

$$
\begin{aligned}
\langle \iota \cup \iota, c_*[M] \rangle &= \langle \iota \cup \iota, \sum_j \tau_j \rangle \\
&= \sum_j J(\tau_j) \\
&= J(M).
\end{aligned}
\tag{2}
$$

$\square$

The Pontryagin product [Br] gives a map from $H_1(\mathbf{R}^2; \mathbf{Z}) \wedge H_1(\mathbf{R}^2; \mathbf{Z}) \mapsto H_2(\mathbf{R}^2; \mathbf{Z})$. This map is an isomorphism. If we identify $H_1(\mathbf{R}^2; \mathbf{Z})$ with $\mathbf{R}^2$ then we can think of the domain as $\mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2$.

Proposition 18. *With the previous identification of $H_2(\mathbf{R}^2; \mathbf{Z})$ with $\mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2$ we have $J(M) = 2c_*([M])$.*

Proof. Let $p : H_1(\mathbf{R}^2) \wedge H_1(\mathbf{R}^2) \to H_2(\mathbf{R}^2) = \mathbf{R}^2 \wedge \mathbf{R}^2$ denote the Pontryagin product. The proposition asserts that $p(J(M)) = 2c_*(M)$. It is useful to introduce a second method of identifying $H_2(\mathbf{R}; \mathbf{Z})$ with $\mathbf{R}^2 \wedge_{\mathbf{Q}} \mathbf{R}^2$. Let $q : H_2(\mathbf{R}^2, \mathbf{R}^2) \to \mathbf{R}^2 \wedge \mathbf{R}^2$ denote the map $q(\alpha) = \langle \iota \cup \iota, \alpha \rangle$. According to the previous proposition $q(c_*([M])) = J(M)$. We have an explicit formula for the Pontryagin product in terms of the bar notation [Br] which allows us to compare these two identifications by computing the composition $q \circ p$:

$$
\begin{aligned}
q \circ p(\alpha \wedge \beta) &= \langle \iota \cup \iota, [\alpha] \cdot [\beta] \rangle \\
&= \langle \iota \cup \iota, [\alpha|\beta] - [\beta|\alpha] \rangle \\
&= \langle \iota \cup \iota, [\alpha|\beta] \rangle - \langle \iota \cup \iota, [\beta|\alpha] \rangle \\
&= \langle \iota \cup \iota, (0, \alpha, \alpha + \beta) \rangle - \langle \iota \cup \iota, (0, \beta, \beta + \alpha) \rangle \\
&= \iota(0, \alpha) \wedge \iota(\alpha, \alpha + \beta) - \iota(0, \beta) \wedge \iota(\beta, \beta + \alpha) \\
&= \alpha \wedge \beta - \beta \wedge \alpha \\
&= 2\alpha \wedge \beta.
\end{aligned}
$$

Since the composition $q \circ p$ is an isomorphism it follows that both p and q are isomorphisms. We want to show that $p(J(M)) = 2c_*(M)$ or equivalently that $qp(J(M)) = 2q(c_*(M))$. Since $qp(J(M)) = 2J(M)$ we need to show that $q(c_*[M]) = J(M)$. But this is proved in formula (2) above. $\square$

4.2. Projections, the Sah-Arnoux-Fathi invariant and the Kenyon invariant

We define two linear projections from $\mathbf{R}^2 \wedge \mathbf{R}^2$ to $\mathbf{R} \wedge \mathbf{R}$. On basis elements they are defined by

$$\pi_{xx}\left(\begin{bmatrix} a \\ b \end{bmatrix} \wedge \begin{bmatrix} c \\ d \end{bmatrix}\right) = a \wedge c$$

$$\pi_{yy}\left(\begin{bmatrix} a \\ b \end{bmatrix} \wedge \begin{bmatrix} c \\ d \end{bmatrix}\right) = b \wedge d$$

We also define a linear projection from $\mathbf{R}^2 \wedge \mathbf{R}^2$ to $\mathbf{R} \otimes \mathbf{R}$. On basis elements it is defined by

$$\pi_{xy}\left(\begin{bmatrix} a \\ b \end{bmatrix} \wedge \begin{bmatrix} c \\ d \end{bmatrix}\right) = a \otimes d - c \otimes b.$$

Let us denote $\pi_{xx}(J)$, $\pi_{xy}(J)$ and $\pi_{yy}(J)$ by J_{xx}, J_{xy} and J_{yy}. These "projections" determine J and we will give a geometric interpretation of each of them.

A *rectilinear subdivision* of M is a decomposition of M into rectangles whose sides are either horizontal or vertical. We do not require that the edges of the rectangles meet each other in complete edges. Kenyon [Ke] has defined a rectilinear subdivision invariant $K(M)$ of M as follows. If M is decomposed into rectangles $R_1, \ldots, R_n$ where the height of R_j is h_j and the width is w_j then

$$K(M) = \sum_{j=1}^{n} h_j \otimes w_j$$

Proposition 19. *We have* $2K(M) = J_{xy}(M)$.

Proof. A rectilinear subdivision of M into rectangles $R_1, \ldots, R_n$ can be converted into a cellular decomposition by adding vertices to the edges. Let $R'_1, \ldots, R'_n$ denote the results of adding vertices to the boundaries of $R_1, \ldots, R_n$. Adding vertices to horizontal or vertical edges changes J by terms in the kernel of π_{xy}. Thus

$$\pi_{xy}(J(M)) = \sum_{j=1}^{n} \pi_{xy} J(R'_j) = \sum_{j=1}^{n} \pi_{xy} J(R_j) = 2 \sum_{j=1}^{n} h_j \otimes w_j \qquad \square$$

An *interval exchange transformation* is a map f from an interval to itself which is a piecewise isometry: let I and I' be two copies of the interval $[a, b]$. Let $a = v_0 < v_1 < \ldots < v_n = b$ be a partition of I and let $a = w_0 < \ldots < w_n = b$ be a partition of I'. Let $I_j = [v_{j-1}, v_j]$ for $j = 1, \ldots, n$ and let $I'_j = [w_{j-1}, w_j]$ for $j = 1, \ldots, n$. Let $\sigma : \{1 \ldots n\} \to \{1 \ldots n\}$ be a permutation, and assume that the length ℓ_j of I_j is equal to the length of $I'_{\sigma(j)}$. We define $f : [a, b] \to [a, b]$

(outside of the set $\{v_0, v_1, \ldots, v_n\}$) by setting $f(x) = x + \tau_j$ for $x \in \mathrm{int}I_j$, where $\tau_j = w_{\sigma(j)} - v_j$. Then f is an interval exchange. The Sah-Arnoux-Fathi, (SAF), invariant of an interval exchange is developed in [Ax]. The SAF invariant of f in $\mathbf{R} \wedge_{\mathbf{Q}} \mathbf{R}$ is defined to be

$$SAF(f) = \sum_{j=1}^{n} \ell_j \wedge \tau_j.$$

Given a surface with a transversally oriented measured foliation and an interval I transverse to the foliation, one can define a first return map $f_I : I \to I$. This map is an interval exchange. If I meets every leaf (and we will assume for simplicity that this is the case), the invariant $SAF(f_I)$ is independent of the transversal chosen. (See [Ax] for a proof.)

Proposition 20. J_{xx} is the SAF invariant of the vertical foliation of M. Similarly J_{yy} is the SAF invariant for the horizontal foliation of M.

Proof. The interval I gives us a decomposition of M into rectangles $R_1, \ldots, R_n$, where the bottom edge of R_j is attached to I at the interval I_j and the top edge of R_j is attached to I at the interval $I'_{\sigma(j)}$. These rectangles are maximal in the sense that each has a singularity or an endpoint of I on each vertical boundary. To make this a cellular decomposition, we must add vertices to each R_j along its vertical edges (at singularities of M and endpoints of I), as well as along I, to make the two decompositions of I compatible. The only contribution to J_{xx} comes from adding vertices along I.

To compute this contribution, think of I as a single segment, to which we have added vertices for the endpoints of I_j coming from rectangles "above" I and vertices from endpoints of the I'_j coming from rectangles "below" I.

The contribution of the endpoints of the I_j is

$$v_0 \wedge v_1 + v_1 \wedge v_2 + \ldots + v_{n-1} \wedge v_n + v_n \wedge v_0,$$

where v_0 is the left endpoint of I, $I_j = [v_{j-1}, v_j]$ and v_n is the right endpoint of I. Similarly the contribution for the endpoints of the I'_j is

$$-w_0 \wedge w_1 - w_1 \wedge w_2 - \ldots - w_{n-1} \wedge w_n - w_n \wedge w_0$$

where $w_0 = v_0$, $w_n = v_n$ and $I'_j = [w_{j-1}, w_j]$. The sum of these contributions gives J_{xx}: using the notation preceding the theorem, we have

$$
\begin{aligned}
J_{xx} = {}& v_0 \wedge v_1 + v_1 \wedge v_2 + \ldots + v_{n-1} \wedge v_n + v_n \wedge v_0 \\
& - w_0 \wedge w_1 - w_1 \wedge w_2 - \ldots - w_{n-1} \wedge w_n - w_n \wedge w_0 \\
= {}& v_0 \wedge (v_1 - v_0) + v_1 \wedge (v_2 - v_1) + \ldots + v_{n-1} \wedge (v_n - v_{n-1}) + v_n \wedge v_0 \\
& - w_0 \wedge (w_1 - w_0) - \ldots - w_{n-1} \wedge (w_n - w_{n-1}) - w_n \wedge w_0
\end{aligned}
$$

$$= v_0 \wedge \ell_1 + v_1 \wedge \ell_2 + \ldots + v_{n-1} \wedge \ell_n$$
$$\quad - w_0 \wedge \ell_{\sigma^{-1}(1)} - w_1 \wedge \ell_{\sigma^{-1}(2)} - \ldots - w_{n-1} \wedge \ell_{\sigma^{-1}(n)}$$
$$= v_0 \wedge \ell_1 + v_1 \wedge \ell_2 + \ldots + v_{n-1} \wedge \ell_n$$
$$\quad - w_{\sigma(0)} \wedge \ell_1 - w_{\sigma(1)} \wedge \ell_2 - \ldots + w_{\sigma(n-1)} \wedge \ell_n$$
$$= (v_0 - w_{\sigma(0)}) \wedge \ell_1 + (v_1 - w_{\sigma(1)}) \wedge \ell_2 + \ldots + (v_{n-1} - w_{\sigma(n-1)}) \wedge \ell_n$$
$$= -\tau_1 \wedge \ell_1 - \tau_2 \wedge \ell_2 - \ldots - \tau_n \wedge \ell_n$$
$$= SAF(f).$$

$\square$

In the next section we will use the invariant J to detect parabolic elements in the affine automorphism group. Recall from Proposition 9 that the existence of a parabolic automorphism gives a particular decomposition of M into cylinders.

Now from Propositions 20 and 19 we have:

Theorem 21. *If M has a decomposition into parallel cylinders in the direction y then $J_{xx} = 0$. If in addition the cylinders all have moduli λ then J_{xy} can be written $J_{xy} = \sum_j r_j \otimes \lambda r_j$*

Note that $J_{xy} \in \mathbf{R} \otimes \mathbf{R}$, and so can be considered as a bilinear form on $\mathrm{Hom}(\mathbf{R}, \mathbf{Q})$. If the λ in the above theorem were rational, then we could write $J_{xy} = \lambda \sum_j r_j \otimes r_j$. Since $\lambda > 0$ this would mean that J_{xy}, considered as a bilinear form on $\mathrm{Hom}(\mathbf{R}, \mathbf{Q})$, would be *symmetric* and *positive semidefinite*.

If λ is not rational, then if we change the translation structure by dividing the y-coordinate by λ, then the new invariant would be $J'_{xy} = \sum_j r_j \otimes r_j$ and would hence be symmetric and positive semidefinite.

5. Triangular tables

In this section we explicitly compute the J-invariant for the surface $\tilde{T}$ where T is a rational triangle. In case T is acute, we find a closed billiard trajectory and compute the modulus of the maximal cylinder containing it. We then use J to determine when the surface can be decomposed into cylinders of moduli commensurable to this one.

Let T be a triangular billiard table in the plane with angles α, β, and γ. As in section 2.1 we write $\alpha = r\pi/n$, $\beta = s\pi/n$ and $\gamma = t\pi/n$ where r, s and t are natural numbers with greatest common divisor one and $n = r + s + t$.

For the sake of our calculations it is convenient to identify $\mathbf{R}^2$ with $\mathbf{C}$ and to identify the rotation group with $\{|z| = 1\}$. Let us take the vertices of T to be 0, 1 and v where $\mathrm{Im}(v) > 0$ and T has angle α at 0, angle β at 1, and angle γ at

v. Let $a = e^{2\alpha i}$, $b = e^{2\beta i}$, and $c = e^{2\gamma i}$. Let $w = e^{2\pi i/n}$. We observe that a, b and c are powers of w and that $abc = 1$. The group Γ_0 is generated by w and complex conjugation $z \mapsto \bar{z}$. Let K denote the cyclotomic field generated by w, and $k = \mathbf{Q}[\cos\frac{2\pi}{n}]$ be its maximal real subfield. As we will show below in formulas (4) and (3), the holonomy of segments takes values in $K \subset \mathbf{C}$. If $n \equiv 0 \bmod 4$, the x- and y-coordinates of elements in K are in the real subfield k. If $n \not\equiv 0 \bmod 4$, the x-coordinates are in k and, after scaling by $1/\sin(2\pi/n)$, the y-coordinates are also in k. Therefore, after this change of coordinates, the holonomy takes values in k^2; in particular k is the holonomy field defined in section 7.

Let $\tilde{T}$ be the surface with framed translation structure associated to T. We wish to calculate $J(\tilde{T})$. The surface $\tilde{T}$ comes equipped with a decomposition into triangles each of which has a canonical representation as a triangle in $\mathbf{R}^2$ (up to translation). These triangles have the form γT for $\gamma \in \Gamma$. We use this to calculate J:

$$J(\tilde{T}) = \sum_{\gamma \in \Gamma} J(\gamma T) = \sum_{k=0}^{n-1} J(w^k T) + \sum_{k=0}^{n-1} J(\overline{w^k T}) = \sum_{k=0}^{n-1} J(w^k T) + \sum_{k=0}^{n-1} J(w^k \overline{T}).$$

The triangle $w^k T$ has vertices 0, w^k and $w^k v$ listed in counterclockwise order. The triangle $w^k \overline{T}$ has vertices 0, $w^k \bar{v}$ and w^k listed in counterclockwise order. Using formula (1) we have

$$J(w^k T) = w^k \wedge (w^k v - w^k) = w^k \wedge w^k v$$
$$J(w^k \overline{T}) = w^k \bar{v} \wedge (w^k - w^k \bar{v}) = w^k \bar{v} \wedge w^k = -w^k \wedge w^k \bar{v}.$$

Thus

$$J(\tilde{T}) = \sum_{k=0}^{n-1} w^k \wedge w^k (v - \bar{v}).$$

Lemma 22. *If λ is in the cyclotomic field K then*

$$\sum_{k=0}^{n-1} w^k \otimes \lambda w^k = \sum_{k=0}^{n-1} w^k \bar{\lambda} \otimes w^k.$$

Similarly $\sum_{k=0}^{n-1} w^k \otimes \lambda w^{-k} = \sum_{k=0}^{n-1} w^k \lambda \otimes w^{-k}$. The same holds when the tensor product is replaced with the wedge product.

Proof. If we take $\lambda = w^\ell$ and use the fact that $w^n = 1$ then

$$\sum_{k=0}^{n-1} w^k \otimes \lambda w^k = \sum_{k=0}^{n-1} w^k \otimes w^{k+\ell}$$

$$= \sum_{k=0}^{n-1} w^{k-\ell} \otimes w^k = \sum_{k=0}^{n-1} w^k w^{-\ell} \otimes w^k = \sum_{k=0}^{n-1} w^k \bar{\lambda} \otimes w^k.$$

The first formula now follows for any rational linear combination of powers of w. The other cases are similar. $\qquad\square$
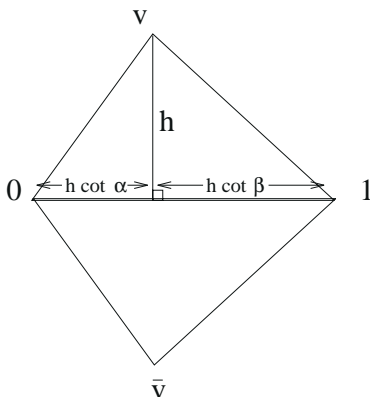

Figure 1.

Let h be the height of the triangle T as in Figure 1. It follows from Figure 1 that $v - \bar{v} = 2ih$ and $h \cot \alpha + h \cot \beta = 1$. Thus

$$v - \bar{v} = \frac{2i}{\cot \alpha + \cot \beta}.$$

Now

$$\cot \alpha = i \left(\frac{e^{\alpha i} + e^{-\alpha i}}{e^{\alpha i} - e^{-\alpha i}} \right) = i \left(\frac{e^{2\alpha i} + 1}{e^{2\alpha i} - 1} \right) = i \left(\frac{a+1}{a-1} \right)$$

and similarly

$$\cot \beta = i \left(\frac{b+1}{b-1} \right).$$

This gives

$$v - \bar{v} = \frac{(a-1)(b-1)}{ab - 1}. \tag{3}$$

Note also that the x-coordinate of v is

$$h \cot \alpha = \frac{\cot \alpha}{\cot \alpha + \cot \beta} = \frac{\frac{a+1}{a-1}}{\frac{a+1}{a-1} + \frac{b+1}{b-1}}. \tag{4}$$

In particular $v = h\cot\alpha + ih = h\cot\alpha + (v - \overline{v})/2$ is in the field K. Therefore the holonomy image of all segments of $\tilde{T}$ is in K.

Now assume the triangle T is acute. Let α denote the largest angle. The first step is to find a cylinder in $\tilde{T}$ explicitly. According to Proposition 7 if T is a lattice polygon then $\tilde{T}$ will decompose into a union of cylinders of moduli commensurable to this one. We use the invariant J and Theorem 21 (and the comments after that theorem) to test for this possibility.

In order to begin we need an explicitly given cylinder. A cylinder in $\tilde{T}$ corresponds to a closed billiard trajectory in T. Since T is acute we have a natural candidate, the Fagnano trajectory. This trajectory is the inscribed triangle whose vertices are the feet of the altitudes of T, see [Ta] and Figure 2. There is a family of parallel trajectories of twice the length of this triangular trajectory; these trajectories fill out a cylinder in $\tilde{T}$. The boundary of this cylinder is a trajectory which hits the vertex with largest angle. (See Figure 3.)
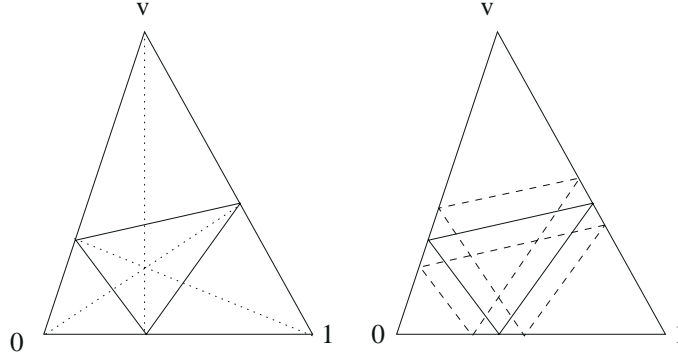

Figure 2.
The Fagnano trajectory and a nearby parallel trajectory.

In that figure the point z_2 has coordinate $1 - b^{-1}$ and the point z_1 has coordinate

$$1 - b^{-1} + a^{-1}b^{-1} - a^{-1} = (1 - b^{-1})(1 - a^{-1}).$$

With respect to the above coordinates for T, the cylinder containing this path lies parallel to the direction z_1 and the modulus of the cylinder is

$$m = \frac{\text{height}}{\text{circumference}} = \operatorname{Im}\left(\frac{-z_2}{z_1}\right) = -\operatorname{Im}\left(\frac{1}{1 - a^{-1}}\right) = -\frac{1}{2i}\left(\frac{a+1}{a-1}\right).$$

Let L denote the linear transformation $L(z) = z/z_1$. Let $\tilde{T}'$ be $L(\tilde{T})$, that is, $\tilde{T}'$ is the result of (rescaling and) rotating $\tilde{T}$ so that the cylinder is vertical (the
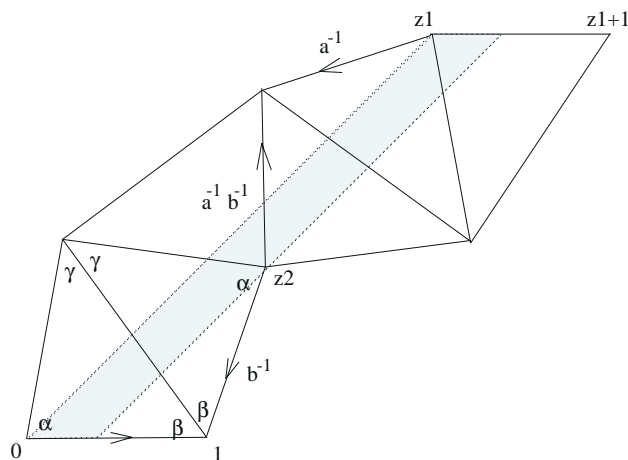
Figure 3.
Cylinder containing the Fagnano trajectory.

circumference is horizontal). We have

$$J(\tilde{T}') = \sum_{k=0}^{n-1} L(w^k) \wedge L(w^k(v - \bar{v}))$$

$$= \sum_{k=0}^{n-1} w^k z_1^{-1} \wedge w^k(v - \bar{v})z_1^{-1}$$

$$= \sum_{k=0}^{n-1} w^k \wedge w^k \lambda$$

where $\lambda = \frac{v - \bar{v}}{z_1 \bar{z}_1}$ is pure imaginary.

Now

$$J_{xy}(\tilde{T}') = \sum_{k=0}^{n-1} \mathrm{Re}(w^k) \otimes \mathrm{Im}(w^k \lambda) - \sum_{k=0}^{n-1} \mathrm{Re}(w^k \lambda) \otimes \mathrm{Im}(w^k)$$

$$= \sum_{k=0}^{n-1} \frac{1}{2}(w^k + w^{-k}) \otimes \frac{1}{2i}(\lambda w^k - \bar{\lambda} w^{-k})$$

$$\quad - \sum_{k=0}^{n-1} \frac{1}{2}(\lambda w^k + \bar{\lambda} w^{-k}) \otimes \frac{1}{2i}(w^k - w^{-k})$$

$$= \frac{1}{4}\Big(\sum_{k=0}^{n-1} w^k \otimes \frac{w^k \lambda}{i} + w^k \otimes \frac{w^{-k} \lambda}{i} + w^{-k} \otimes \frac{w^k \lambda}{i} + w^{-k} \otimes \frac{w^{-k} \lambda}{i}$$

$$-\lambda w^k \otimes \frac{w^k}{i} + \lambda w^k \otimes \frac{w^{-k}}{i} + \lambda w^{-k} \otimes \frac{w^k}{i} - \lambda w^{-k} \otimes \frac{w^{-k}}{i}\Big)$$

$$= \sum_{k=0}^{n-1} w^k \otimes \frac{w^k \lambda}{i} + w^k \otimes \frac{w^{-k}\lambda}{i}.$$

If we change the framed translation structure by dividing the y-coordinate by m we create a cylinder with modulus 1. Let $\tilde{T}''$ denote the resulting surface. We can obtain $J_{xy}(\tilde{T}'')$ from $J_{xy}(\tilde{T}')$ by dividing the second coordinate by m:

$$J_{xy}(\tilde{T}'') = \sum_{k=0}^{n-1} w^k \otimes \frac{w^k \lambda}{mi} + w^k \otimes \frac{w^{-k}\lambda}{mi}.$$

By the comments after Theorem 21, we need to determine when $J_{xy}(\tilde{T}'')$ is symmetric and positive semidefinite. In fact, it is already symmetric, as the reader may show.

To test for positive semidefiniteness, note first that $J_{xy}(\tilde{T}'') \in K \otimes K$ (in fact, despite appearances, $J_{xy}(\tilde{T}'') \in k \otimes k$). If σ, σ' are embeddings of K into $\mathbf{C}$ then the map $f_{\sigma,\sigma'}: K \otimes K \to \mathbf{C}$ given by

$$f_{\sigma,\sigma'}\left(\sum u_j \otimes v_j\right) = \sum \sigma(u_j)\sigma'(v_j)$$

is well-defined. If $\sigma' = \bar{\sigma}$ then $f_{\sigma,\bar{\sigma}}$ applied to a non-zero symmetric, positive semidefinite form is positive:

$$f_{\sigma,\bar{\sigma}}\left(\sum u_j \otimes u_j\right) = \sum \sigma(u_j)\overline{\sigma(u_j)} > 0.$$

In particular to test $J_{xy}(\tilde{T}'')$ for positive definiteness, we can check that $f_{\sigma,\bar{\sigma}}J_{xy}(\tilde{T}'') > 0$ for every embedding σ of K.

An embedding of K is determined by the image of w, which must be a Galois conjugate of w, i.e. of the form w^r for some r relatively prime to n.

Proposition 23. *Let $\sigma(w^k) = w^{\ell k}$ and let $\sigma'(w^k) = w^{\ell' k}$. Then*

$$f_{\sigma,\sigma'}(J_{xy}(\tilde{T}'')) = \begin{cases} n\sigma'(\frac{\lambda}{mi}), & \text{if } \ell = \ell' \text{ or } \ell = -\ell'; \\ 0, & \text{otherwise.} \end{cases}$$

Proof.

$$f_{\sigma,\sigma'}(J_{xy}(\tilde{T}'')) = \sum_{k=0}^{n-1} \sigma(w^k)\sigma'(w^k\frac{\lambda}{mi}) + \sigma(w^k)\sigma'(w^{-k}\frac{\lambda}{mi})$$

$$= \sum_{k=0}^{n-1} \sigma'(\frac{\lambda}{mi})w^{k(\ell+\ell')} + \sigma'(\frac{\lambda}{mi})w^{k(\ell-\ell')}.$$

Now if $X = \sum_{k=0}^{n-1} w^{pk}$ then changing the order of summation gives $X = w^p X$. Thus $\sum_{k=0}^{n-1} w^{pk} = 0$ unless $p \equiv 0 \bmod n$ in which case $\sum_{k=0}^{n-1} w^{pk} = n$. $\square$

Note that $\sigma(w^k) = w^{k\ell}$ and $\sigma'(w^k) = w^{-k\ell}$ correspond to the *same* embeddings of k into **R**.

We calculate

$$
\begin{aligned}
\frac{\lambda}{mi} &= \frac{v - \bar{v}}{miz_1\bar{z}_1} \\
&= \frac{(a-1)(b-1)}{ab-1} \left(\frac{1}{\frac{-(a+1)}{2(a-1)}} \right) \frac{1}{(1-b^{-1})(1-a^{-1})(1-b)(1-a)} \\
&= \frac{2}{(a+1)(1-b)(1-c)}
\end{aligned}
$$

where we used $abc = 1$.

This gives the following theorem.

Theorem 24. *Let T be an acute rational triangle with angles $\alpha \geq \beta \geq \gamma$. Let n be the least common denominator of the rational numbers $\frac{\alpha}{\pi}, \frac{\beta}{\pi}, \frac{\gamma}{\pi}$. Then T has the lattice property only if for all k prime to n we have*

$$
\cos(k\beta + k\gamma) \sin(k\beta) sin(k\gamma) < 0. \tag{5}
$$

One can show, although we will not use this fact, that this condition is sufficient for the form $J_{xy}(\tilde{T}'')$ to be positive semidefinite.

Proof. Recall that $abc = 1$. We have

$$
\begin{aligned}
(1+a)(1-b)(1-c) &= (1 + e^{2\alpha i})(1 - e^{2\beta i})(1 - e^{2\gamma i}) \\
&= e^{(\alpha+\beta+\gamma)i}(e^{-\alpha i} + e^{\alpha i})(e^{-\beta i} - e^{\beta i})(e^{-\gamma i} - e^{\gamma i}) \\
&= (-1)(e^{-\alpha i} + e^{\alpha i})(e^{\beta i} - e^{-\beta i})(e^{\gamma i} - e^{-\gamma i}) \\
&= 8\cos(\alpha)\sin(\beta)\sin(\gamma) \\
&= -8\cos(\beta+\gamma)\sin(\beta)\sin(\gamma)
\end{aligned}
$$

Therefore $\sigma(\frac{\lambda}{mi}) > 0$ if and only if $\sigma(\cos(\beta+\gamma)\sin(\beta)\sin(\gamma)) < 0$.

The group of Galois automorphisms of $\mathbf{Q}[w]$ is the group U_n of integers modulo n relatively prime to n, under multiplication: the automorphism corresponding to $k \in U_n$ sends $\sin(\beta)$ to $\sin(k\beta)$, $\sin(\gamma)$ to $\sin(k\gamma)$ and $\cos(\beta+\gamma)$ to $\cos(k\beta + k\gamma)$. $\square$
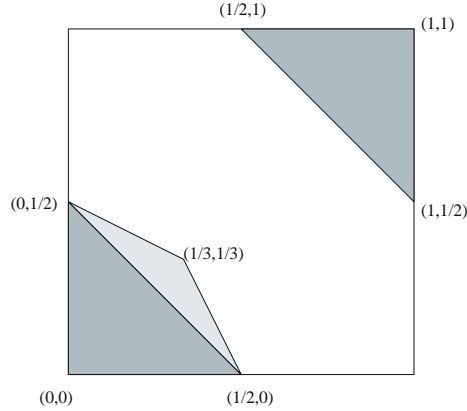
Figure 4.
The lightly shaded region is the set of pairs corresponding to acute triangles. The darkly shaded region is the set where $\cos(\beta + \gamma)\sin(\beta)\sin(\gamma)) > 0$.

5.1. Verifying the condition of Theorem 24

The set of pairs $\frac{1}{2\pi}(2\beta, 2\gamma)$ for which $\cos(\beta+\gamma)\sin(\beta)\sin(\gamma) < 0$ is the complement of the darkly shaded region in Figure 4: choosing $2\beta, 2\gamma \in [0, 2\pi)$ we must have $\pi < 2\beta + 2\gamma < 3\pi$.

Now n is the least common denominator of the two rational numbers $\beta/\pi, \gamma/\pi$, so $(2\beta, 2\gamma) = (\frac{2\pi s}{n}, \frac{2\pi t}{n})$ for positive integers s, t.

In other words the positivity condition of Theorem 24 is equivalent to: for all $p \in U_n$, $(\frac{ps}{n}, \frac{pt}{n})$ mod 1 is in the unshaded region in Figure 4.

Figure 5 shows the set of pairs $(\frac{ps}{n}, \frac{pt}{n})$ as p runs over U_n, first in the case $s = 23$, $t = 17$, $n = 78$ (example 25 in the table below) and secondly in the case $s = 23$, $t = 17$, $n = 78$. The lighter shaded triangle in that figure consists of pairs (s, t) for which the triangle $(\frac{\pi n - s - t}{n}, \frac{\pi s}{n}, \frac{\pi t}{n})$ is acute.

The condition of Theorem 24 can easily be checked on a computer. For denominators n running from 1 to 10000 we found exactly 25 non-isosceles acute triangles which satisfied the condition of Theorem 24. They are listed in the table below.

Based on this computer evidence we make the following conjecture.

Conjecture. *Only 25 non-isosceles acute triangles satisfy the condition of Theorem 24.*

Because of the results of the following subsection this conjecture implies that only the 3 non-isosceles acute triangles listed in Theorem 1 have the lattice property.

We restate this conjecture in strictly number theoretic terms taking into ac-
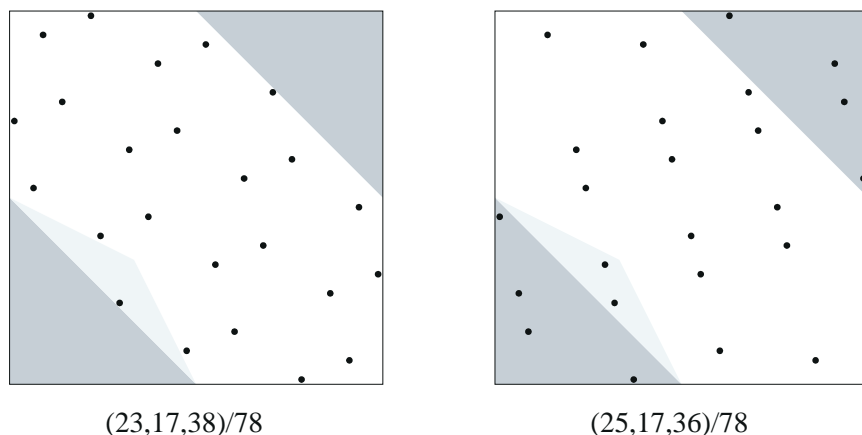
(23,17,38)/78   (25,17,36)/78

Figure 5.
Two examples.

count the results of our Theorem 1.

Restatement. For all $n \geq 10000$, and for all $s \neq t \in U_n$ for which $s + t \neq 0$, $s + 2t \neq 0$ and $2s + t \neq 0$ there exists $p \in U_n$ such that either $(ps \bmod n) + (pt \bmod n) < n/2$ or $(ps \bmod n) + (pt \bmod n) > 3n/2$. That is, for some $p \in U_n$ the pair $\left(\frac{ps}{n}, \frac{pt}{n}\right) \bmod 1$ is in the darker shaded region of Figure 4.

Note that the case $s + t = n$ occurs only if $\beta + \gamma = \pi$, an impossibility. The cases $2s + t = n$ and $s + 2t = n$ occur only when $2\beta + \gamma = \pi$ or $\beta + 2\gamma = \pi$, that is, when the triangle is isosceles.

5.2. 25 special cases

In the previous table the triangles in cases 1, 2 and 4 do have the lattice property. These correspond to the three cases mentioned in Theorem 1. Cases 1 and 2 have been considered in [Vo2]. We will consider case 4 in section 5.2. In this section we deal with the 22 remaining cases.

In these 22 cases, we construct directly two non-commensurable maximal cylinders in a singular direction (perpendicular to an edge). In the table, the column "side" gives the left and right angles of the side perpendicular to the cylinders we found; the entries 'x_1' and 'x_2' give the center x-coordinate of the cylinders, assuming the left endpoint of the edge "side" is at the origin and its right endpoint is at $2\sin\theta$, where θ is the opposite angle (note that this is a different convention for the side length than used previously). Thus these examples do not have the Veech property. The data for the ratio of the cylinders is in the table below.

By way of example, in example number 13 with angles $(14, 11, 5)\frac{\pi}{30}$ we found

case	n	r	s	t	side	$x1$	$x2$
1	9	4	3	2			
2	12	5	4	3			
3	14	6	5	3	6,5	.20	.82
4	15	7	5	3			
5	18	7	6	5	7,6	.296	1.149
6	20	9	8	3	9,8	.148	.603
7	20	9	7	4	7,4	.267	.591
8	21	10	6	5	10,6	.291	1.302
9	21	10	9	2	10,9	.073	.585
10	24	11	8	5	11,5	.288	.642
11	24	11	7	6	11,6	.092	625
12	30	14	13	3	14,3	.187	1.267
13	30	14	11	5	14,5	.180	.585
14	30	14	9	7	14,9	.085	1.004
15	30	12	11	7	12,11	.282	.665
16	30	13	9	8	13,9	.168	.665
17	42	17	13	12	17,13	.0096	.0812
18	42	19	12	11	19,12	.117	1.046
19	42	19	17	6	19,6	.360	.655
20	42	20	13	9	20,13	.160	.263
21	42	20	17	5	20,5	.316	.088
22	60	23	20	17	23,20	.110	.214
23	60	29	19	12	29,12	.494	.101
24	60	29	20	11	29,20	.045	.793
25	78	38	23	17	38,23	.032	.505

the cylinders of Figure 6 perpendicular to the edge whose angles are $14, 5$. The first cylinder is centered at $x = .180$ and the second at $x_2 = .585$, where the x coordinate runs from 0 to $2\sin 11\pi/30 = 1.827$. To compute their moduli, assume rather that the length of the shortest edge of the triangle is 1; then the circumference of the first cylinder is (letting $w = e^{\pi i/30}$)

$$2Im(w^{14} + w^6 + w^8 + w^{16} + w^{18} + w^{26} + w^{28} + w^{20} + w^{18} + w^{10} + w^8)$$

and its height (the width in the figure) is

$$Re(w^6 + w^8 + w^{16} + w^{18} + w^{26}).$$

Similarly the modulus of the second cylinder is computed to be

$$-\frac{2Im(w^{14} + w^6 + w^8 + w^{16} + w^{18} + w^{26} + w^{24} + w^{16})}{Re(w^8 + w^{16} + w^{18} + w^{26})}.$$
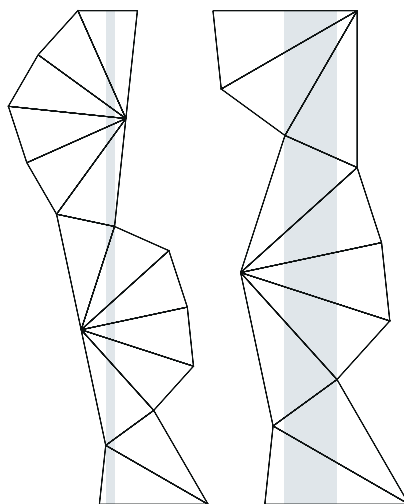
Figure 6.
Two cylinders in the $5, 11, 14$-example.

To show that these cylinders are non-commensurable, we computed the ratio of the moduli (in reduced form in $\mathbf{Q}[w]$) to be

$$\text{ratio } = 1 - w^2 - w^4 + w^8 + w^{10} - w^{14}.$$

Since w has algebraic degree $\phi(60) = 16$, this ratio is irrational (if the ratio were rational, w would satisfy a rational polynomial of degree 14). In fact the algebraic degree over $\mathbf{Q}$ of the ratio can be shown to be 4.

The table below lists the degree of w, gives the ratio as an element of $\mathbf{Q}[w]$, and gives the degree of the ratio. In each case the ratio is a polynomial in w of degree less than the degree of w, so the ratio is irrational.

5.3. The 2,3,4-example

In this subsection we consider the triangle which is example (c) of Theorem 1.

Theorem 25. *The triangle with angles $2\pi/9$, $\pi/3$ and $4\pi/9$ has the lattice property.*

Proof. Let T denote the triangle and let $\tilde{T}$ denote the corresponding surface. Let Γ denote the affine automorphism group of $\tilde{T}$. To prove the lattice property it suffices to show that the image of the group $D(\Gamma)$ in $PGL(2, \mathbf{R})$ contains a lattice. We use the action of $PGL(2, \mathbf{R})$ on the upper half-plane. We will show that the image of Γ contains the group generated by reflections in the sides of the hyperbolic

#	deg [w]	ratio=r	deg [r]
3	12	$\frac{1}{7}(6 + 3w^4 - w^6 + w^8 - 3w^{10})$	3
5	12	$\frac{1}{3}(w^2 + w^4 - 2w^8 + w^{10})$	3
6	16	$7 - 6w^2 + w^6 + 4w^8 - 3w^{10} - 4w^{12} + 5w^{14}$	4
7	16	$5 + 3w^8 - 3w^{12}$	2
8	12	$9 + 2w - 4w^3 - 2w^4 + 4w^5 + 4w^6 - 2w^7 - 6w^8 - 2w^9 + 6w^{11}$	6
9	12	$7 + 11w + 12w^2 + 7w^3 + 2w^4 + 2w^5 + 6w^6 + 5w^7 - w^8 - 7w^9$ $- 8w^{10} - 4w^{11}$	6
10	16	$\frac{1}{3}(-6 + 2w^2 + 8w^4 + 2w^6 - 7w^{10} - 4w^{12} + 5w^{14})$	4
11	16	$\frac{1}{5}(3 - w^2 + 2w^4 - w^6 - w^{12} + w^{14})$	4
12	16	$w^2 + w^4 + w^6 + w^8 - w^{12} - w^{14}$	4
13	16	$1 - w^2 - w^4 + w^8 + w^{10} - w^{14}$	4
14	16	$2 + w^2 - w^6 - w^8 - w^{10} + 2w^{14}$	4
15	16	$18 + 6w^2 - 7w^6 - 8w^8 - 7w^{10} + w^{12} + 14w^{14}$	4
16	16	$3 + 3w^4 + 3w^6 - 3w^{14}$	2
17	24	$\frac{1}{211}(-98 + 84w^2 - 339w^4 - 710w^6 + 485w^8 + 2051w^{10} + 1279w^{12} - 1195w^{14}$ $- 2474w^{16} - 856w^{18} + 1050w^{20} + 855w^{22})$	6
18	24	$-2 - 3w^2 + 2w^4 + 3w^6 + 2w^8 - 4w^{12} + w^{14} + 5w^{16} - w^{18} + 2w^{20} - 6w^{22}$	6
19	24	$2 + w^2 + w^4 - w^6 - w^8 + w^{10} + w^{12} - w^{16} - w^{18} + w^{22}$	6
20	24	$44 + 130w^2 + 30w^4 - 178w^6 - 140w^8 + 106w^{10} + 168w^{12} - 38w^{14}$ $- 206w^{16} - 68w^{18} + 123w^{20} + 93w^{22}$	6
21	24	$5 + 4w^2 + 4w^4 + w^6 + 2w^{10} + 3w^{12} + w^{14} - 2w^{16} - 3w^{18} - 2w^{20}$	6
22	32	$\frac{1}{552626999}(-492890382 + 390818622w^2 + 816824625w^4 + 760386855w^6 +$ $930005115w^8 + 830687855w^{10} + 573428027w^{12} + 577841758w^{14} +$ $103670449w^{16} + 208273525w^{18} - 356577088w^{20} - 390818622w^{22} -$ $460247537w^{24} - 577841758w^{26} - 216850939w^{28} - 324071379w^{30})$	8
23	32	$15 + 8w^2 + 16w^4 + 11w^6 + 15w^8 + 17w^{10} + 7w^{12} + 6w^{14} + 3w^{18} - 8w^{20}$ $- 8w^{22} - 8w^{24} - 6w^{26} + w^{28} - 6w^{30}$	8
24	32	$\frac{1}{59}(32 - 47w^2 - 31w^4 - 29w^6 - 17w^8 + 9w^{10} + 6w^{12} + 48w^{14} + 15w^{16}$ $+ 30w^{18} + 23w^{20} + 47w^{22} + 8w^{24} - 48w^{26} - 29w^{28} - 43w^{30})$	8
25	48	$\frac{1}{311}(266 - 37w^2 - 246w^4 + 24w^6 + 96w^8 - 111w^{10} + 159w^{12} + 69w^{14}$ $- 291w^{16} + 6w^{18} + 114w^{20} - 66w^{22} + 53w^{24} - 90w^{26} - 143w^{28} + 156w^{30}$ $+ 146w^{32} + 18w^{34} + 67w^{36} - 153w^{38} - 96w^{40} + 134w^{42} + 66w^{44} - 80w^{46})$	12

triangle with a vertex at i and ideal vertices at ∞ and $\tan(4\pi/9)$. This triangle has angles 0, 0 and $\pi/9$ and the group generated by reflections in its edges is a lattice (see Figure 8).

The group Γ contains the dihedral group Γ_0 and the image of this group is

generated by

$$z \mapsto \frac{\cos \frac{\pi}{9} z - \sin \frac{\pi}{9}}{\sin \frac{\pi}{9} z + \cos \frac{\pi}{9}}$$

(rotation by $2\pi/9$), and

$$z \mapsto -\overline{z},$$

(reflection in the horizontal direction). This subgroup contains the reflections in two edges of the triangle (see Figure 8).

The group $D(\Gamma)$ also contains a parabolic element corresponding to a decomposition into cylinders. A decomposition of the surface of the $(2, 3, 4)$-triangle is shown in Figure 7. In that figure, parallel sides (which are labeled with their corresponding angle with respect to the horizontal, in multiples of $\pi/9$) are identified by translation.
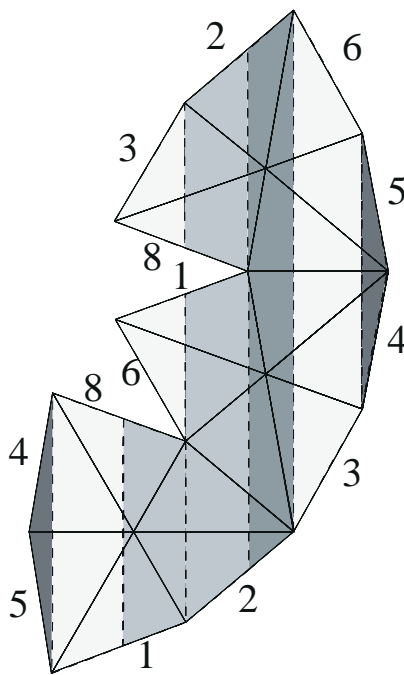

Figure 7.
Decomposition of the $2, 3, 4$-example surface into 4 commensurable cylinders (4 shades of gray).

The widths of the cylinders in the decomposition, from lightest to darkest, are

$$\cos \frac{3\pi}{9}, -\cos \frac{3\pi}{9} - \cos \frac{8\pi}{9}, \cos \frac{2\pi}{9} + \cos \frac{3\pi}{9} + \cos \frac{8\pi}{9}, \cos \frac{4\pi}{9},$$

and their heights are respectively

$$2\sin\frac{3\pi}{9} + 4\sin\frac{4\pi}{9}, 2\sin\frac{6\pi}{9} + 2\sin\frac{\pi}{9} + 4\sin\frac{2\pi}{9}, 2\sin\frac{3\pi}{9} + 2\sin\frac{4\pi}{9}, 2\sin\frac{4\pi}{9}.$$

One may easily verify that each of the cylinders has modulus

$$M = 2\cot 4\pi/9 = 2\tan\pi/18.$$

As a consequence the group $D(\Gamma)$ contains the parabolic map

$$z \mapsto z + 2\tan\frac{4\pi}{9}$$

fixing the vertical direction (see Proposition 8). The composition of the inverse of this parabolic element with the map $z \mapsto -\bar{z}$ gives the reflection in the third side of the triangle.   □
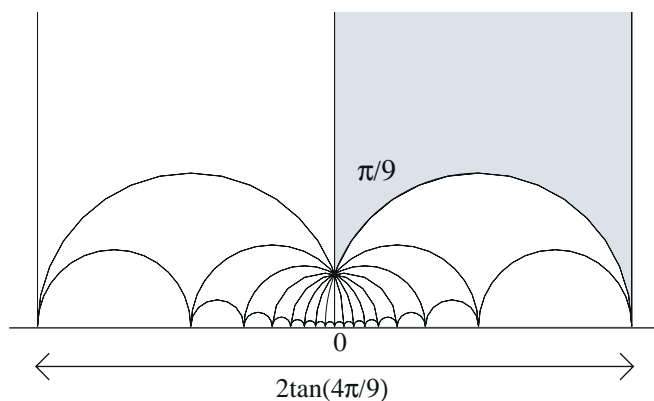

Figure 8.
Fundamental domain (shaded), in the upper-half plane, of the Veech group of the $(2, 3, 4)$-example. The center point is located at i. The points shown on the axis are $\tan(k\pi/18)$ where the integer k ranges from -8 to 8.

6. Right-angled and isosceles triangles

In this section we will prove the necessity of the hypotheses of Theorem 2. The sufficiency of the hypotheses has already been discussed following the statement of the Theorem. Theorem 2 deals with right and isosceles triangles. We begin

with the case of right triangles. Specifically we will show that a rational right triangle for which the smallest angle is not of the form π/n does not have the lattice property.

Proof of Theorem 2, right-triangle case. Let T be a rational right triangle. We will find a pair of maximal cylinders in the surface $\tilde{T}$ with incommensurable moduli. The fact that the automorphism group of $\tilde{T}$ is not a lattice will then follow from Propositions 7 and 9.

Let α, β be the two acute angles, with $\alpha \geq \beta$. There is a maximal cylinder perpendicular to the hypotenuse of the triangle as shown in Figure 9. Assuming the shorter leg has length $1/2$, the circumference of this cylinder is $2\sin(\alpha)$, and the height is $\cos(\alpha)$. This cylinder therefore has modulus
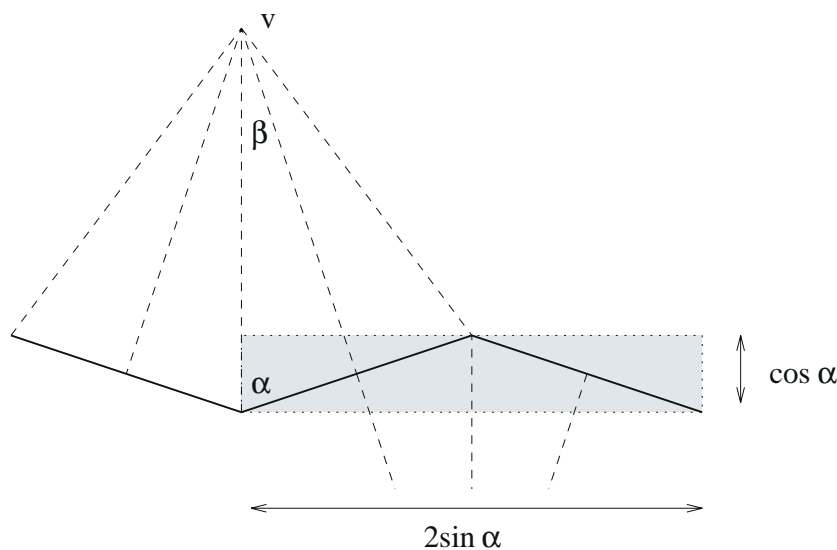
$$M = \frac{1}{2}\cot(\alpha) = \frac{1}{2}\tan(\beta).$$

Figure 9.
Cylinder perpendicular to the hypotenuse.

If β is not of the form π/n, then the surface has a singularity at the point v of Figure 9. Let k' be the integer for which $k'\beta \leq \pi/2 \leq (k'+1)\beta$. If k' is odd, say $k' = 2k - 1$, then there is a maximal cylinder which passes just below this singularity (see Figure 11): its circumference is $4A\sin((2k-1)\beta)$ and its height is $2A\cos((2k-1)\beta)$ (where A is the length of the longer leg of the triangle), and so

the modulus is

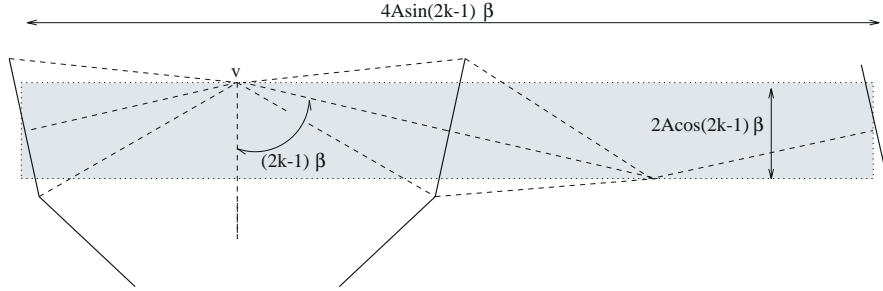$$M' = \frac{2A\cos((2k-1)\beta)}{4A\sin((2k-1)\beta)} = 2\cot((2k-1)\beta).$$

Figure 10.
The cylinder below the singularity.

If k' is even, say $k' = 2k$, then there is a cylinder of circumference $4A\sin((2k+1)\beta)$ and height $-2A\cos((2k+1)\beta)$ which passes just above the singularity (Figure 11). The modulus of this cylinder is

$$M' = -\frac{2A\cos((2k+1)\beta)}{4A\sin((2k+1)\beta)} = -2\cot((2k+1)\beta).$$

In the first case $(2k-1)\beta \in (0, \pi/2)$, and in the second $(2k+1)\beta \in (\pi/2, \pi)$.
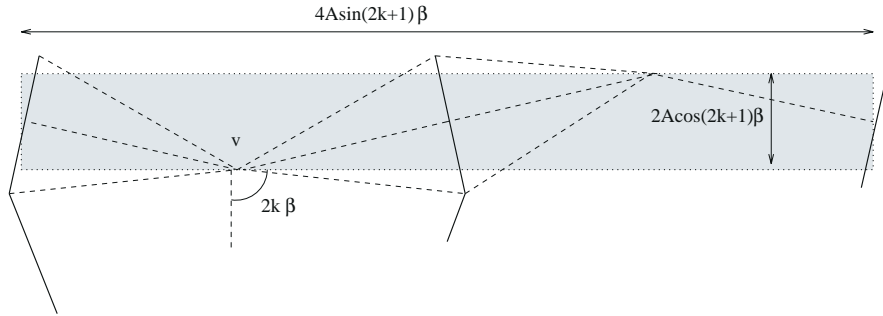

Figure 11.
The cylinder above the singularity.

In order for the cylinders of Figure 9 and (the appropriate one of) Figures 10 or 11 to be commensurable, we must have

$$M/M' = \pm\tan((2k\pm1)\beta)\tan(\beta) \in \mathbf{Q}.$$

We claim that this is only possible if $M/M' = \pm 1$. For, if $\alpha \in \mathbf{Q}$ is such that

$$\tan(\beta) = \alpha \cot((2k \pm 1)\beta) = \alpha \tan(\pi/2 - (2k \pm 1)\beta),$$

then (letting $\| \cdot \|$ denote the algebraic norm, i.e. the product of the Galois conjugates)

$$\| \tan(\beta) \| = \alpha^d \| \tan(\pi/2 - (2k \pm 1)\beta) \|,$$

where d is the algebraic degree of $\tan(\pi/2 - (2k \pm 1)\beta)$ (and hence the algebraic degree of $\tan(\beta)$). By Lemma 26 below, $\| \tan(\beta) \|$ and $\| \tan(\pi/2 - (2k \pm 1)\beta) \|$ are each (up to sign) either 1, prime or the reciprocal of a prime. Thus if $\alpha \neq \pm 1$ then $d = 1$ or $d = 2$.

Since

$$i \tan(p\pi/q) = \frac{e^{2\pi i p/q} - 1}{e^{2\pi i p/q} + 1}, \tag{6}$$

we have $\mathbf{Q}[i, \tan(p/\pi q)] = \mathbf{Q}[i, e^{2\pi i p/q}]$. So $\tan(\frac{p\pi}{q})$ has degree 1 only when $e^{2\pi i p/q} \in \mathbf{Q}[i]$, that is, only when $q = 2$ or $q = 4$. Similarly, $\tan(\frac{p\pi}{q})$ has degree 2 only when $e^{2\pi i p/q}$ is in a quadratic extension of $\mathbf{Q}[i]$. The degree of the cyclotomic extension of order q is $\phi(q)$ where ϕ is the Euler ϕ-function. The only values of q for which $\phi(q) = 2$ or $\phi(q) = 4$ are $q = 3, 6, 8, 12$. However since $\beta = \frac{p\pi}{q} < \pi/4$ and is not of the form π/n none of these exceptions occur.

So we conclude from Lemma 26 that $\tan(\beta) = \tan(\frac{\pi}{2} - (2k - 1)\beta)$ if k' is odd and $\tan(\beta) = -\cot((2k + 1)\beta)$ if k' is even. The first possibility implies $\beta = \frac{\pi}{2} - (2k - 1)\beta$ and so $\beta = \frac{\pi}{4k}$, which is not allowed. For the second possibility we must have $\tan(\frac{\pi}{2} - \beta) = -\tan((2k + 1)\beta) = \tan(\pi - (2k + 1)\beta)$, or $\beta = \frac{\pi}{4k}$ again. Again this is not allowed.

In conclusion, the cylinders are never commensurable. So these triangles do not have the lattice property. $\square$

If β has the form π/n then the cylinders constructed here are not maximal. Nevertheless if we assume that $\beta = \pi/n$ with n *odd* then it is easy to check that the previous argument for the incommensurability of the moduli still applies. We will use this fact below.

Lemma 26. *If m and n are relatively prime, then*

$$\| \tan(m\pi/n) \| = \begin{cases} \pm p & \text{if } n = p^\ell \text{ and } p \text{ is an odd prime} \\ \pm 1/p & \text{if } n = 2p^\ell \text{ and } p \text{ is an odd prime} \\ \pm 1 & \text{otherwise.} \end{cases}$$

Proof. The norm of an element is the constant term of its minimal polynomial (times $(-1)^{degree}$). The minimal polynomial of $e^{2\pi i m/n}$ is $\Phi_n(z)$, the n-th cyclotomic polynomial. The minimal polynomial for $e^{2\pi i m/n} \pm 1$ is $\Phi_n(z \mp 1)$. According

to (6) we have $\|\tan(p\pi/q)\| = \Phi_s(1)/\Phi_s(-1)$. If $n = p^\ell$ for some prime p, then

$$\Phi_n(z) = \frac{z^{p^\ell} - 1}{z^{p^{\ell-1}} - 1} = 1 + z^{p^{\ell-1}} + z^{2p^{\ell-1}} + \ldots + z^{(p-1)p^{\ell-1}}$$

and so $\Phi_n(1) = p$ and $\Phi_n(-1) = 1$, unless $p = 2$, in which case $\Phi_n(1) = 2 = \Phi_n(-1)$.

We have $\prod_{d|n,d\neq 1} \Phi_d(z) = 1 + z + \ldots + z^{n-1}$, and so

$$\prod_{d|n,\ d\neq 1} \Phi_d(1) = n.$$

If $n = 2^{\alpha_1} 3^{\alpha_2} \cdots p_k^{\alpha_k}$ is the prime factorization of n then dividing out by terms corresponding to the the prime power factors $2^1, 2^2, \ldots, 2^{\alpha_1}, 3^1, \ldots, 3^{\alpha_2}, \ldots, p_k^{\alpha_k}$ leaves 1 on the right hand side. So each of the remaining factors d must satisfy $\Phi_d(1) = \pm 1$. A similar argument for $\Phi_d(-1)$ using $\prod_{d|n,d>2} \Phi_d(z) = 1 + z^2 + \ldots + z^{n-2}$ when n is even yields the result. $\qquad\blacksquare$

Proof of Theorem 2, isoceles case. Let T' be an isosceles triangle with apex angle θ where θ is a rational multiple of π which is not of the form π/n. Let T be the right triangle obtained by folding T' in half. Thus the apex angle of T is $\beta = \theta/2$. Let $\tilde{T}'$ and $\tilde{T}$ be the surfaces associated with T' and T. There is a natural projection from T' to T.

Case 1. Assume that θ does not have the form $2\pi/n$ with n odd. In this case β does not have the form π/n. The annuli in T shown in figure 9 and figures 10 or 11 (as the case may be) are maximal and incommensurable in $\tilde{T}$. They lift to maximal and incommensurable annuli in $\tilde{T}'$.

Case 2. Assume that $\theta = 2\pi/n$ with n odd. In this case $\tilde{T}'$ is a branched double cover of $\tilde{T}$ branched at the points in $\tilde{T}$ corresponding to the apex angle of T (one of these points is marked v in figures 10 and 11). These points are non-singular points of $\tilde{T}$. In this case the annuli in figures 10 and 11 are not maximal in $\tilde{T}$ (because v is not a singular point in $\tilde{T}$) but their lifts are maximal in $\tilde{T}'$ (because v does correspond to a singular point in $\tilde{T}'$). The incommensurability of the moduli of the cylinder from figure 9 and the cylinder from figure 10 or 11 follows from the comment before Lemma 26. $\qquad\blacksquare$

7. Appendix: Consequences of having a non-elementary automorphism group

The criteria developed in section 5 (to show that polygons do not have the lattice property) are rather subtle. The results of this section were motivated by the

desire to find elementary criteria which could be used to show that a particular polygon did not give rise to a surface with the lattice property. In fact the criteria that we develop here test for the weaker property of having an automorphism group of 'type 3' as in section 3. On the other hand we will see that the translation surfaces for all rational triangles satisfy the hypotheses in the criteria of this appendix so that these criteria are not useful in distinguishing differences in the affine automorphisms groups of rational triangles. In particular the main results of this paper in section 5 do not depend on the results of this appendix.

The existence of an affine pseudo-Anosov automorphism of a translation structure or, what is the same, a hyperbolic element of $D(\Gamma)$ has consequences for the algebraic structure of the holonomy image. We describe these consequences in this appendix. We begin with an algebraic Lemma.

A polynomial $P(X)$ of even degree $2g$ is said to be *reciprocal* if $P(X) = X^{2g}P(1/X)$. This implies that each root λ occurs with the same multiplicity as λ^{-1}.

Lemma 27. *If λ is the root of a reciprocal polynomial $P(X) \in \mathbf{Z}[X]$ of degree $2g$ then $\lambda + \lambda^{-1}$ is the root of an integer polynomial of degree at most g.*

Proof. By induction on g we will prove $P(X)X^{-g} = r(X + 1/X)$ for some integer polynomial r. This is clearly true if $g = 1$, for then $P(X)X^{-1} = aX + b + a/X = a(X + 1/X) + b$. For $g \geq 1$, the terms of degree g and $-g$ in $P(X)X^{-g}$ have equal coefficients a_1; thus $P(X)X^{-g} - a_1(X + 1/X)^g$ is either zero or of the form $Q(X)X^{-g'}$ for some reciprocal integer polynomial Q of even degree $2g' < 2g$. $\square$

Let $\Lambda = h(H_1(M, \mathbf{Z}))$, that is, Λ is the subgroup of $\mathbf{R}^2$ generated by the integer holonomy. Then Λ is a free $\mathbf{Z}$-module; let $\mathrm{rk}(\Lambda)$ denote its rank.

Let $e_1, e_2 \in \Lambda$ be non-parallel vectors in $\mathbf{R}^2$. Define k to be the smallest subfield of $\mathbf{R}$ such that every element of Λ can be written $ae_1 + be_2$, with $a, b \in k$. In other words, k is the smallest field such that $\Lambda \otimes_k k \cong k^2$. We call k the *holonomy field*.

For a different approach to these notions see [GJ2]. Chris Judge communicated to us that he and Gutkin have independently obtained the Corollary 29 and the first part of Theorem 28 below.

Theorem 28. *If the affine automorphism group of M contains an orientation-preserving pseudo-Anosov element f with largest eigenvalue λ then $k = \mathbf{Q}[\lambda + \lambda^{-1}]$ and $[k : \mathbf{Q}] = \frac{1}{2}\mathrm{rk}(\Lambda)$.*

In particular since the holonomy field depends only on the translation structure we see that for every orientation-preserving affine pseudo-Anosov automorphism of M, the field $\mathbf{Q}[\lambda + \lambda^{-1}]$ is the same. In case f reverses orientation then $k = \mathbf{Q}[\lambda - \lambda^{-1}]$ but the rest of the statement is the same. The proof in this case is similar and left to the reader.

Proof. Let $k' = \mathbf{Q}[\lambda + \lambda^{-1}]$. We first show that $k' \subset k$. Since $Df \in SL(2, \mathbf{R})$ and Df maps $\Lambda \otimes_k k \cong k^2$ to itself, after choosing a basis in Λ, the map Df is in $SL(2, k)$. But trace $(Df) = \lambda + \lambda^{-1}$. As a consequence $\lambda + \lambda^{-1} \in k$ and so $k' \subset k$.

To show that $k \subset k'$ we will show that Λ is contained in a 2-dimensional k'-vector subspace of $\mathbf{R}^2$. Let f_* denote the action of f on $H_1(M, k')$. The characteristic polynomial of f_* factors over k' as

$$P(X) = (X^2 - (\lambda + \lambda^{-1})X + 1)Q(X),$$

where λ and λ^{-1} are not roots of $Q(X)$ (Fried [Fri] shows that λ and λ^{-1} occur with multiplicity 1). Furthermore f_* preserves $\mathrm{Ker}(h)$. By the primary decomposition theorem, $H_1(M, k')$ splits into invariant factors $W' \oplus V$, where W' is the k'-linear subspace on which f_* has eigenvalues λ, λ^{-1}. But from Lemma 6 we have another decomposition $H_1(M; k') = W \oplus W^\perp$, where $W^\perp = \mathrm{Ker}(h) \cap H_1(M; k')$ and $W = A_* \cap H_1(M; k')$. Now we must have $W' = W$ since this is the only subspace on which f_* has eigenvalues $\lambda^{\pm 1}$. Since $\mathrm{Ker}(h)$ is f_*-invariant, we must have $W^\perp = V$ as well.

The primary decomposition theorem implies that the projection map (along V) from $H_1(M, k')$ to $W' = W$ is k'-linear. Therefore Λ, the image of the projection of $H^1(M, \mathbf{Z}) \subset H_1(M, k')$ is contained in the 2-dimensional k'-vector space W. We conclude that $k \subset k'$.

Now write $P(X) = (X^2 - (\lambda + \lambda^{-1})X + 1)Q'(X)R(X)$ where $R(X)$ is the largest *rational* factor. The rational subspace W_R corresponding to $R(X)$ is in the kernel of h and so $\mathrm{rk}\Lambda = 2 + \deg Q'$. Now $(X^2 - (\lambda + \lambda^{-1})X + 1)Q'(X)$ is either the minimal polynomial for λ or the product of the minimal polynomials for λ and λ^{-1}, in case they are not algebraic conjugates. By Lemma 27 the algebraic degree of $\lambda + \lambda^{-1}$ is equal to $\frac{1}{2}(2 + \deg Q') = \frac{1}{2}\mathrm{rk}\Lambda$. $\qquad\square$

Corollary 29. *If the affine automorphism group of M contains a pseudo-Anosov element then the group $D(\Gamma)$ contains a subgroup of finite index which is contained in $SL(2, \mathcal{O})$ where $\mathcal{O}$ is the ring of integers in k.*

Proof. Both $\mathcal{O}^2 = \mathcal{O} \times \mathcal{O}$ and Λ are lattices in k^2 of maximal rank. Thus their intersection is a lattice of finite index n in Λ. There are finitely many subgroups of index n in Λ and $D(\Gamma)$ permutes them. Thus there is a subgroup $D(\Gamma')$ of finite index that fixes $\mathcal{O}^2 \cap \Lambda$. Now $\mathcal{O}^2 \cap \Lambda$ is a subgroup of $\mathcal{O}^2$ of finite index m and there is a finite set of lattices that contain $\mathcal{O}^2 \cap \Lambda$ as a subgroup of index m. So passing to a further finite index subgroup $D(\Gamma'')$ we may assume that $\mathcal{O}^2$ is stabilized. But the stabilizer of $\mathcal{O}^2$ is just $SL(2, \mathcal{O})$. $\qquad\square$

Recall that a *segment* is a curve connecting singularities of M. Let Λ_0 be the subgroup of $\mathbf{R}^2$ generated by the holonomy images of segments, that is, Λ_0 is the image of $H_1(M, \Sigma; \mathbf{Z})$ under the holonomy mapping. (Recall that Λ is the image of $H_1(M, \mathbf{Z})$ under holonomy.)

Theorem 30. *If the affine automorphism group of M contains a pseudo-Anosov element then Λ has finite index in Λ_0.*

Proof. It suffices to show that $H_1(M, \mathbf{Q})$ and $H_1(M, \Sigma; \mathbf{Q})$ have the same image under holonomy.

Let f be an affine pseudo-Anosov automorphism and let $\lambda^{\pm 1}$ be the eigenvalues of Df. The map f induces automorphisms of each of the following modules:

$$H_1(M; \mathbf{Q}) \xrightarrow{\iota} H_1(M, \Sigma; \mathbf{Q}) \xrightarrow{\partial} H_0(\Sigma; \mathbf{Q})$$
$$\downarrow h$$
$$\mathbf{R}^2$$

Because the top row is exact, the eigenvalues of f acting on $H_1(M, \Sigma, \mathbf{Q})$ are the union of the eigenvalues of f acting on $H_1(M, \mathbf{Q})$ and the eigenvalues of f acting on $H_0(\Sigma, \mathbf{Q})$. Since Σ is a finite set, f is a permutation matrix on $H_0(\Sigma, \mathbf{Q})$, and so the eigenvalues on this space are roots of unity. As a consequence the eigenvalues λ, λ^{-1} occur with multiplicity exactly 1 on $H_1(M, \Sigma; \mathbf{Q})$.

As in the proof of Theorem 28, $x \in H_1(M, \Sigma; \mathbf{Q})$ is in $\mathrm{Ker}(h)$ if and only if x is in an f-invariant $\mathbf{Q}$-subspace such that the eigenvalues of f on that subspace do not contain λ. $\square$

Theorems 28 and 30 are related to a result of Franks and Rykken ([FR]): if λ is quadratic, then (M, f) is a branched cover of an Anosov mapping of a torus. For then $\lambda + \lambda^{-1} \in \mathbf{Z}$ and so $\Lambda = \mathbf{Z}^2$ after change of coordinates (and homothety). In particular this implies that M is almost integrable. Compare this result with Theorem 4 in [GJ1] and Theorem 5.4 in [GJ2].

References

[ABN] A. D. Alexandrov, V. N. Berestovski and I. G. Nikolaev, Russian Math. Surveys **41** (3) (1986), 1-54.

[Ad] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, 1978.

[Ax] P. Arnoux, Thèse de 3e cycle, Université de Reims, 1981.

[Bo] B. Bowditch, Singular Euclidean structures on surfaces *J. London Math Soc.* **44** (1991), 553-565.

[BH] M. Bridson and A. Haefliger, Metric spaces of non-positive curvature, in preparation.

[Br] K. Brown, *Cohomology of Groups*, Springer-Verlag 1982.

[BR] M. V. Berry and P. J. Richens, Pseudointegrable systems in classical and quantum mechanics, *Physica* D **2**(3) (1981), 495-512.

[EG] C. J. Earle, F. P. Gardiner, Teichmuller disks and Veech's $\mathcal{F}$-structures. *Extremal Riemann surfaces* 165-189 Contemp. Math., 201 AMS Providence, RI, 1997.

[Fa] A. Fathi, Some compact invariant sets for hyperbolic linear automorphisms of tori, *Erg. Th. & Dyn. Sys.* **8** (1988), 191-204.

[Fo] L. Ford, *Automorphic Functions*, Chelsea Publishing Company, New York 1929 (reprinted 1972).

[FR] J. Franks and E. Rykken, Pseudo-Anosov homeomorphisms with quadratic expansion, (to appear Proc. of the A.M.S.).

[Fri] D. Fried, Growth rate of surface homeomorphisms and flow equivalence. *Erg. Th. & Dyn. Syst.* **5**(4) (1985), 539–563.

[G] E. Gutkin, Billiards in almost integrable polyhedral surfaces. *Erg. Th. & Dyn. Syst.* **4** (1984), 569-584.

[GJ1] E. Gutkin and C. Judge, The geometry and arithmetic of translation surfaces with applications to polygonal billiards, *Math. Res. Lett.* **3** (1996), 391-403.

[GJ2] E. Gutkin and C. Judge, Affine mappings of translation surfaces: geometry and arithmetic, preprint (1997).

[Ke] R. Kenyon, Tilings and discrete Dirichlet problems. *Isr. J. Math* **105** (1998), 61-84.

[Kr] R. Kirby, Problems in low dimensional manifold theory, in *Proceed. of Symp. in Pure Math.* **32** (1978), 273-312.

[L] G. Levitt, Propriètés homologiques des feuilletages des surfaces, *C.R. Acad. Sci Paris* Sér. I Math. **293** (1981), 597-600.

[MS] H. Masur, J. Smillie. Hausdorff dimension of sets of nonergodic measured foliations *Annals of Math.* **134** (1991), 455-543.

[Sa1] C. H. Sah, Hilbert's third problem: scissors congruence. Research Notes in Mathematics, 33. Pitman (Advanced Publishing Program), Boston, Mass.-London, 1979. vi+188 pp.

[Sa2] C. H. Sah, Scissors congruence of the interval, preprint.

[Sc] S. Schwartzman, Asymptotic cycles. *Ann. of Math.* **66**(2) (1957), 270–284.

[Ta] S. Tabachnikoff, Billiards. Panoramas et Synthèses **1** Soc. Math. France (1995).

[Th] W.P. Thurston, Shapes of polyhedra, preprint.

[Tr] M. Troyanov, Les surfaces euclidiennes à singulariéts coniques. (French) [Euclidean surfaces with conical singularities] *Enseign. Math.* **32**(2) (1986), no. 1-2, 79–94.

[Ve1] W. A. Veech, Teichmuller curves in modular space, Eisenstein series, and an application to triangular billiards, *Inventiones Math.* **97** (1989), 553-583.

[Ve2] W. A. Veech, The billiard in a regular polygon, *Geom. Func. Anal.* **2** (1992), 341-379.

[Vo1] Ya. B. Vorobets, Plane structures and billiards in rational polyhedra. (Russian) *Uspekhi Mat. Nauk* **51**(1) (1996), 145–146; translation in Russian Math. Surveys **51**(1) (1996), 177–178.

[Vo2] Ya. B. Vorobets, Plane structures and billiards in rational polygons: the Veech alternative. (Russian) *Uspekhi Mat. Nauk* **51**(5) (1996), 3–42. translation in Russian Math. Surveys **51**(5) (1996), 779-817.

[W] C. Ward, Fuchsian groups and polygonal billiards, Thesis, Rice University 1996.

[ZK] A. Zemlyakov, A. Katok. Topological transitivity of billiards in polygons, *Math. Notes of the USSR Acad. Sci.* **18**(2) (1975), 291-300. (English translation in Math. Notes **18**(2) (1976), 760-764.)

Richard Kenyon   John Smillie
CNRS UMR8628   Mathematics Dept.
Bât 425   Cornell University
Université Paris-Sud   Ithaca, NY 14853
91405 Orsay   USA
France