Geodesics on flat surfaces

Anton Zorich*

Abstract. Various problems of geometry, topology and dynamical systems on surfaces as well as some questions concerning one-dimensional dynamical systems lead to the study of closed surfaces endowed with a flat metric with several cone-type singularities. In an important particular case, when the flat metric has trivial holonomy, the corresponding flat surfaces are naturally organized into families which appear to be isomorphic to moduli spaces of holomorphic one-forms.

One can obtain much information about the geometry and dynamics of an individual flat surface by studying both its orbit under the Teichmüller geodesic flow and under the linear group action on the corresponding moduli space. We apply this general principle to the study of generic geodesics and to counting of closed geodesics on a flat surface.

Mathematics Subject Classification (2000). Primary 57M50, 32G15; Secondary 37D40, 37D50, 30F30.

Keywords. Flat surface, Teichmüller geodesic flow, moduli space, asymptotic cycle, Lyapunov exponent, interval exchange transformation, renormalization.

Introduction: families of flat surfaces as moduli spaces of Abelian differentials

Consider a collection of vectors $\vec{v}_1, \dots, \vec{v}_n$ in $\mathbb{R}^2$ and construct from these vectors a broken line in a natural way: a j-th edge of the broken line is represented by the vector $\vec{v}_j$. Construct another broken line starting at the same point as the initial one by taking the same vectors in the order $\vec{v}_{\pi(1)}, \dots, \vec{v}_{\pi(n)}$, where π is some permutation of n elements. By construction the two broken lines share the same endpoints; suppose that they bound a polygon like in Figure 1. Identifying the pairs of sides corresponding to the same vectors $\vec{v}_j$, $j = 1, \dots, n$, by parallel translations we obtain a surface endowed with a flat metric. (This construction follows the one in [M1].) The flat metric is nonsingular outside of a finite number of cone-type singularities corresponding to the vertices of the polygon. By construction the flat metric has trivial holonomy: a parallel transport of a vector along a closed path does not change the direction (and length) of the vector. This implies, in particular, that all cone angles are integer multiples of 2π.

The polygon in our construction depends continuously on the vectors $\vec{v}_j$. This means that the combinatorial geometry of the resulting flat surface (its genus g, the

*The author is grateful to the MPI (Bonn) and to IHES for hospitality during the preparation of this paper.

Proceedings of the International Congress
of Mathematicians, Madrid, Spain, 2006

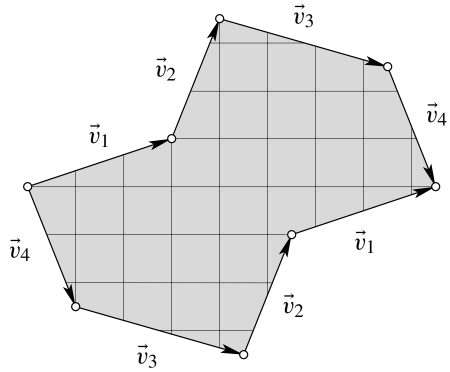

Figure 1. Identifying corresponding pairs of sides of this polygon by parallel translations we obtain a surface of genus two. It has single conical singularity with cone angle 6π; the flat metric has trivial holonomy.

number m and types of the resulting conical singularities) does not change under small deformations of the vectors $\vec{v}_j$. This allows to consider a flat surface as an element of a family of flat surfaces sharing common combinatorial geometry; here we do not distinguish isometric flat surfaces. As an example of such family one can consider a family of flat tori of area one, which can be identified with the space of lattices of area one:

$$^{\mathrm{SO}(2,\mathbb{R})}\backslash\ \mathrm{SL}(2,\mathbb{R})\ /_{\mathrm{SL}(2,\mathbb{Z})} \quad = \quad {}^{\mathbb{H}^2}/_{\mathrm{SL}(2,\mathbb{Z})}$$

The corresponding "modular surface" is not compact, see Figure 2. Flat tori representing points, which are close to the cusp, are almost degenerate: they have a very short closed geodesic. Similarly, families of flat surfaces of higher genera also form noncompact finite-dimensional orbifolds. The origin of their noncompactness is the same as for the tori: flat surfaces having short closed geodesics represent points which are close to the multidimensional "cusps".

We shall consider only those flat surfaces, which have trivial holonomy. Choosing a direction at some point of such flat surface we can transport it to any other point. It would be convenient to include the choice of direction in the definition of a flat structure. In particular, we want to distinguish the flat structure represented by the polygon in Figure 1 and the one represented by the same polygon rotated by some angle different from 2π.

Consider the natural coordinate z in the complex plane. In this coordinate the parallel translations which we use to identify the sides of the polygon in Figure 1 are represented as $z' = z + \text{const}$. Since this correspondence is holomorphic, it means that our flat surface S with punctured conical points inherits the complex structure. It is easy to check that the complex structure extends to the punctured points. Consider now a holomorphic 1-form dz in the complex plane. When we pass to the surface S the coordinate z is not globally defined anymore. However, since the changes of

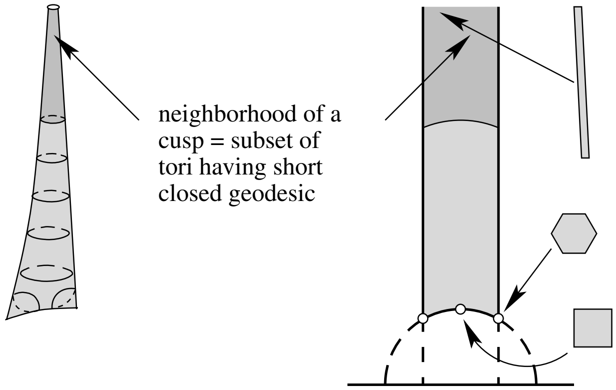

Figure 2. "Modular surface" $\mathbb{H}^2 / \mathrm{SL}(2, \mathbb{Z})$ representing the space of flat tori is a noncompact orbifold of finite volume.

local coordinates are defined as $z' = z + \text{const}$, we see that $dz = dz'$. Thus, the holomorphic 1-form dz on $\mathbb{C}$ defines a holomorphic 1-form ω on S which in local coordinates has the form $\omega = dz$. It is easy to check that the form ω has zeroes exactly at those points of S where the flat structure has conical singularities.

Reciprocally, one can show that a pair (Riemann surface, holomorphic 1-form) uniquely defines a flat structure of the type described above.

In an appropriate local coordinate w a holomorphic 1-form can be represented in a neighborhood of zero as $w^d\, dw$, where d is called the degree of zero. The form ω has a zero of degree d at a conical point with cone angle $2\pi(d+1)$. The sum of degrees $d_1 + \cdots + d_m$ of zeroes of a holomorphic 1-form on a Riemann surface of genus g equals $2g-2$. The moduli space $\mathcal{H}_g$ of pairs (complex structure, holomorphic 1-form) is a $\mathbb{C}^g$-vector bundle over the moduli space $\mathcal{M}_g$ of complex structures. The space $\mathcal{H}_g$ is naturally stratified by the strata $\mathcal{H}(d_1, \ldots, d_m)$ enumerated by unordered partitions of the number $2g-2$ in a collection of positive integers $2g-2 = d_1 + \cdots + d_m$. Any holomorphic 1-forms corresponding to a fixed stratum $\mathcal{H}(d_1, \ldots, d_m)$ has exactly m zeroes, and $d_1, \ldots, d_m$ are the degrees of zeroes. Note, that an individual stratum $\mathcal{H}(d_1, \ldots, d_m)$ in general does not form a fiber bundle over $\mathcal{M}_g$.

It is possible to show that if the permutation π which was used to construct a polygon in Figure 1 satisfy some explicit conditions, vectors $\vec{v}_1, \ldots, \vec{v}_n$ representing the sides of the polygon serve as coordinates in the corresponding family $\mathcal{H}(d_1, \ldots, d_m)$. Consider vectors $\vec{v}_j$ as complex numbers. Let $\vec{v}_j$ join vertices P_j and P_{j+1} of the polygon. Denote by ρ_j the resulting path on S joining the points $P_j, P_{j+1} \in S$. Our interpretation of $\vec{v}_j$ as of a complex number implies that

$$\int_{\rho_j} \omega = \int_{P_j}^{P_{j+1}} dz = v_j \in \mathbb{C}.$$

The path ρ_j represents a relative cycle: an element of the relative homology group

$H_1(S, \{P_1, \dots, P_m\};\, \mathbb{Z})$ of the surface S relative to the finite collection of conical points $\{P_1, \dots, P_m\}$. Relation above means that $\vec{v}_j$ represents a period of ω: an integral of ω over the relative cycle ρ_j. In other words, a small domain in $H^1(S, \{P_1, \dots, P_m\}; \mathbb{C})$ containing $[\omega]$ can be considered as a local coordinate chart in our family $\mathcal{H}(d_1, \dots, d_m)$ of flat surfaces.

We summarize the correspondence between geometric language of flat surfaces and the complex-analytic language of holomorphic 1-forms on a Riemann surface in the dictionary below.

Geometric language	Complex-analytic language
flat structure (including a choice of the vertical direction)	complex structure and a choice of a holomorphic 1-form ω
conical point with a cone angle $2\pi(d+1)$	zero of degree d of the holomorphic 1-form ω (in local coordinates $\omega = w^d\, dw$)
side $\vec{v}_j$ of a polygon	relative period $\int_{P_j}^{P_{j+1}} \omega = \int_{\vec{v}_j} \omega$ of the 1-form ω
family of flat surfaces sharing the same cone angles $2\pi(d_1+1), \dots, 2\pi(d_m+1)$	stratum $\mathcal{H}(d_1, \dots, d_m)$ in the moduli space of Abelian differentials
coordinates in the family: vectors $\vec{v}_i$ defining the polygon	coordinates in $\mathcal{H}(d_1, \dots, d_m)$: relative periods of ω in $H^1(S, \{P_1, \dots, P_m\}; \mathbb{C})$

Note that the vector space $H^1(S, \{P_1, \dots, P_m\};\, \mathbb{C})$ contains a natural integer lattice $H^1(S, \{P_1, \dots, P_m\};\, \mathbb{Z} \oplus \sqrt{-1}\,\mathbb{Z})$. Consider a linear volume element $d\nu$ in the vector space $H^1(S, \{P_1, \dots, P_m\};\, \mathbb{C})$ normalized in such a way that the volume of the fundamental domain in the "cubic" lattice

$$H^1(S, \{P_1, \dots, P_m\};\, \mathbb{Z} \oplus \sqrt{-1}\,\mathbb{Z}) \;\subset\; H^1(S, \{P_1, \dots, P_m\};\, \mathbb{C})$$

equals one. Consider now the real hypersurface $\mathcal{H}_1(d_1, \dots, d_m) \subset \mathcal{H}(d_1, \dots, d_m)$ defined by the equation $area(S) = 1$. The volume element $d\nu$ can be naturally restricted to the hypersurface defining the volume element $d\nu_1$ on $\mathcal{H}_1(d_1, \dots, d_m)$.

Theorem (H. Masur; W. A. Veech). *The total volume* $\mathrm{Vol}(\mathcal{H}_1(d_1,\dots,d_m))$ *of every stratum is finite.*

The values of these volumes were computed by A. Eskin and A. Okounkov [EO].

Consider a flat surface S and consider a polygonal pattern obtained by unwrapping S along some geodesic cuts. For example, one can assume that our flat surface S is glued from a polygon $\Pi \subset \mathbb{R}^2$ as on Figure 1. Consider a linear transformation $g \in \mathrm{GL}^+(2,\mathbb{R})$ of the plane $\mathbb{R}^2$. The sides of the new polygon $g\Pi$ are again arranged into pairs, where the sides in each pair are parallel and have equal length. Identifying the sides in each pair by a parallel translation we obtain a new flat surface gS which, actually, does not depend on the way in which S was unwrapped to a polygonal pattern Π. Thus, we get a continuous action of the group $\mathrm{GL}^+(2,\mathbb{R})$ on each stratum $\mathcal{H}(d_1,\dots,d_m)$.

Considering the subgroup $\mathrm{SL}(2,\mathbb{R})$ of area preserving linear transformations we get the action of $\mathrm{SL}(2,\mathbb{R})$ on the "unit hyperboloid" $\mathcal{H}_1(d_1,\dots,d_m)$. Considering the diagonal subgroup $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t}\end{pmatrix} \subset \mathrm{SL}(2,\mathbb{R})$ we get a continuous action of this one-parameter subgroup on each stratum $\mathcal{H}(d_1,\dots,d_m)$. This action induces a natural flow on the stratum which is called the *Teichmüller geodesic flow*.

Key Theorem (H. Masur; W. A. Veech). *The actions of the groups* $\mathrm{SL}(2,\mathbb{R})$ *and* $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t}\end{pmatrix}$ *preserve the measure* $d\nu_1$. *Both actions are ergodic with respect to this measure on each connected component of every stratum* $\mathcal{H}_1(d_1,\dots,d_m)$.

The following basic principle (which was first used in the pioneering works of H. Masur [M1] and of W. Veech [V1] to prove unique ergodicity of almost all interval exchange transformations) appeared to be surprisingly powerful in the study of flat surfaces. Suppose that we need some information about geometry or dynamics of an individual flat surface S. Consider the "point" S in the corresponding family of flat surfaces $\mathcal{H}(d_1,\dots,d_m)$. Denote by $\mathcal{N}(S) = \overline{\mathrm{GL}^+(2,\mathbb{R})\,S} \subset \mathcal{H}(d_1,\dots,d_m)$ the closure of the $\mathrm{GL}^+(2,\mathbb{R})$-orbit of S in $\mathcal{H}(d_1,\dots,d_m)$.

In numerous cases knowledge about the structure of $\mathcal{N}(S)$ gives a comprehensive information about geometry and dynamics of the initial flat surface S. Moreover, some delicate numerical characteristics of S can be expressed as averages of simpler characteristics over $\mathcal{N}(S)$. We apply this general philosophy to the study of geodesics on flat surfaces.

Actually, there is a hope that this philosophy extends much further. A closure of an orbit of an abstract dynamical system might have extremely complicated structure. According to the optimistic hopes, the closure $\mathcal{N}(S)$ of a $\mathrm{GL}^+(2,\mathbb{R})$-orbit of any flat surface S is a nice complex-analytic variety, and all such varieties might be classified. For genus two the latter statements were recently proved by C. McMullen (see [Mc1] and [Mc2]) and partly by K. Calta [Ca].

The following theorem supports the hope for some nice and simple description of orbit closures.

Theorem (M. Kontsevich). *Suppose that the closure in the stratum $\mathcal{H}(d_1, \dots, d_m)$ of a $\mathrm{GL}^+(2,\mathbb{R})$-orbit of some flat surface S is a complex-analytic subvariety. Then in cohomological coordinates $H^1(S, \{P_1, \dots, P_m\}; \mathbb{C})$ this subvariety is represented by an affine subspace.*

1. Geodesics winding up flat surfaces

In this section we study geodesics on a flat surface S going in generic directions. According to the theorem of S. Kerckhoff, H. Masur and J. Smillie [KeMS], for any flat surface S the directional flow in almost any direction is uniquely ergodic. This implies, in particular, that for such directions the geodesics wind around S in a relatively regular manner. Namely, it is possible to find a cycle $c \in H_1(S;\mathbb{R})$ such that a long piece of geodesic pretends to wind around S repeatedly following this asymptotic cycle c. Rigorously it can be described as follows. Having a geodesic segment $X \subset S$ and some point $x \in X$ we emit from x a geodesic transversal to X. From time to time the geodesic would intersect X. Denote the corresponding points as $x_1, x_2, \dots$. Closing up the corresponding pieces of the geodesic by joining the starting point x_0 and the point x_j of j-th return to X with a path going along X we get a sequence of closed paths defining the cycles $c_1, c_2, \dots$. These cycles represent longer and longer pieces of the geodesic. When the direction of the geodesic is uniquely ergodic, the limit

$$\lim_{N\to\infty} \frac{1}{N}\, c_N = c$$

exists and the corresponding asymptotic cycle $c \in H_1(S;\mathbb{R})$ does not depend on the starting point $x_0 \in X$. Changing the transverse interval X we get a collinear asymptotic cycle.

When S is a flat torus glued from a unit square, the asymptotic cycle c is a vector in $H_1(\mathbb{T}^2;\mathbb{R}) = \mathbb{R}^2$ and its slope is exactly the slope of our flat geodesic in standard coordinates. When S is a surface of higher genus the asymptotic cycle belongs to a $2g$-dimensional space $H_1(S;\mathbb{R}) = \mathbb{R}^{2g}$. Let us study how the cycles c_j deviate from the direction of the asymptotic cycle c. Choose a hyperplane W in $H_1(S,\mathbb{R})$ orthogonal (transversal) to the asymptotic cycle c and consider a parallel projection to this screen along c. Projections of the cycles c_N would not be necessarily bounded: directions of the cycles c_N tend to direction of the asymptotic cycle c provided the norms of the projections grow sublinearly with respect to N.

Let us observe how the projections are distributed in the screen W. A heuristic answer is given by Figure 3.

We see that the distribution of projections of the cycles c_N in the screen W is anisotropic: the projections accumulate along some line. This means that in the original space $\mathbb{R}^{2g}$ the vectors c_N deviate from the asymptotic direction L_1 spanned by c not arbitrarily but along some two-dimensional subspace L_2 containing L_1, see

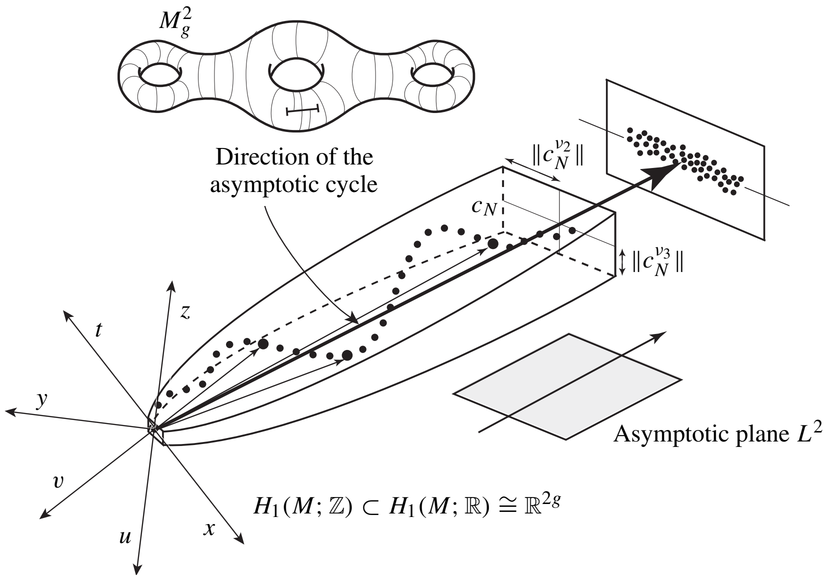

Figure 3. Deviation from the asymptotic direction exhibits anisotropic behavior: vectors deviate mainly along two-dimensional subspace, a bit more along three-dimensional subspace, etc. Their deviation from a Lagrangian g-dimensional subspace is already uniformly bounded.

Figure 3. Moreover, measuring the norms $\|\operatorname{proj}(c_N)\|$ of the projections we get

$$\limsup_{N\to\infty} \frac{\log\|\operatorname{proj}(c_N)\|}{\log N} = \nu_2 < 1.$$

Thus, the vector c_N is located approximately in the two-dimensional plane L_2, and the distance from its endpoint to the line L_1 in L_2 is at most of the order $\|c_N\|^{\nu_2}$, see Figure 3.

Consider now a new screen $W_2 \perp L_2$ orthogonal to the plane L_2. Now the screen W_2 has codimension two in $H_1(S,\mathbb{R}) \simeq \mathbb{R}^{2g}$. Taking the projections of c_N to W_2 along L_2 we eliminate the asymptotic directions L_1 and L_2 and we see how the vectors c_N deviate from L_2. On the screen W_2 we observe the same picture as in Figure 3: the projections are again located along a one-dimensional subspace.

Coming back to the ambient space $H_1(S,\mathbb{R}) \simeq \mathbb{R}^{2g}$, this means that in the first term of approximation all vectors c_N are aligned along the one-dimensional subspace L_1 spanned by the asymptotic cycle. In the second term of approximation, they can deviate from L_1, but the deviation occurs mostly in the two-dimensional subspace L_2, and has order $\|c_N\|^{\nu_2}$ where $\nu_2 < 1$. In the third term of approximation we see that the vectors c_N may deviate from the plane L_2, but the deviation occurs mostly in a three-dimensional space L_3 and has order $\|c_N\|^{\nu_3}$ where $\nu_3 < \nu_2$.

Going on we get further terms of approximation. However, getting to a subspace L_g which has half of the dimension of the ambient space we see that, in there is no more deviation from L_g: the distance from any c_N to L_g is uniformly bounded. Note that the intersection form endows the space $H_1(S,\mathbb{R}) \simeq \mathbb{R}^{2g}$ with a natural symplectic structure. It can be checked that the resulting g-dimensional subspace L_g is a Lagrangian subspace for this symplectic form.

A rigorous formulation of phenomena described heuristically in Figure 3 is given by the theorem below.

By convention we always consider a flat surface together with a choice of direction which is called the vertical direction, or, sometimes, "direction to the North". Using an appropriate homothethy we normalize the area of S to one, so that $S \in \mathcal{H}_1(d_1,\ldots,d_m)$.

We chose a point $x_0 \in S$ and a horizontal segment X passing through x_0; by $|X|$ we denote the length of X. The interval X is chosen in such way, that the interval exchange transformation induced by the vertical flow has the minimal possible number $n = 2g+m-1$ of subintervals under exchange. (Actually, almost any other choice of X would also work.) We consider a geodesic ray γ emitted from x_0 in the vertical direction. (If x_0 is a saddle point, there are several outgoing vertical geodesic rays; choose any of them.) Each time when γ intersects X we join the point x_N of intersection and the starting point x_0 along X producing a closed path. We denote the homology class of the corresponding loop by c_N.

Let ω be the holomorphic 1-form representing S; let g be genus of S. Choose some Euclidean metric in $H_1(S;\mathbb{R}) \simeq \mathbb{R}^{2g}$ which would allow to measure a distance from a vector to a subspace. Let by convention $\log(0) = -\infty$.

Theorem 1. *For almost any flat surface S in any stratum $\mathcal{H}_1(d_1,\ldots,d_m)$ there exists a flag of subspaces*

$$L_1 \subset L_2 \subset \cdots \subset L_g \subset H_1(S;\mathbb{R})$$

in the first homology group of the surface with the following properties.

Choose any starting point $x_0 \in X$ in the horizontal segment X. Consider the corresponding sequence $c_1, c_2, \ldots$ of cycles.

– The following limit exists:

$$|X| \lim_{N\to\infty} \frac{1}{N} c_N = c,$$

where the nonzero asymptotic cycle $c \in H_1(M_g^2;\mathbb{R})$ is Poincaré dual to the cohomology class of $\omega_0 = \mathrm{Re}[\omega]$, and the one-dimensional subspace $L_1 = \langle c\rangle_{\mathbb{R}}$ is spanned by c.

– For any $j = 1,\ldots,g-1$ one has

$$\limsup_{N\to\infty} \frac{\log \operatorname{dist}(c_N, L_j)}{\log N} = \nu_{j+1}$$

and

$$\operatorname{dist}(c_N, L_g) \le \text{const},$$

where the constant depends only on S and on the choice of the Euclidean structure in the homology space.

The numbers $2, 1 + \nu_2, \dots, 1 + \nu_g$ *are the top g Lyapunov exponents of the Teichmüller geodesic flow on the corresponding connected component of the stratum* $\mathcal{H}(d_1, \dots, d_m)$*; in particular, they do not depend on the individual generic flat surface S in the connected component.*

It should be stressed, that the theorem above was formulated in [Z3] as a conditional statement: under the conjecture that $\nu_g > 0$ there exist a Lagrangian subspace L_g such that the cycles are in a bounded distance from L_g; under the further conjecture that all the exponents ν_j, for $j = 2, \dots, g$, are distinct, there is a complete Lagrangian flag (i.e. the dimensions of the subspaces L_j, where $j = 1, 2, \dots, g$, rise each time by one). These two conjectures were later proved by G. Forni [Fo1] and by A. Avila and M. Viana [AvVi] correspondingly.

Currently there are no methods of calculation of individual Lyapunov exponents ν_j (though there is some experimental knowledge of their approximate values). Nevertheless, for any connected component of any stratum (and, more generally, for any $\mathrm{GL}^+(2; \mathbb{R})$-invariant suborbifold) it is possible to evaluate the *sum* of the Lyapunov exponents $\nu_1 + \dots + \nu_g$, where g is the genus. The formula for this sum was discovered by M. Kontsevich; morally, it is given in terms of characteristic numbers of some natural vector bundles over the strata $\mathcal{H}(d_1, \dots, d_m)$, see [K]. Another interpretation of this formula was found by G. Forni [Fo1]; see also a very nice formalization of these results in the survey of R. Krikorian [Kr]. For some special $\mathrm{GL}^+(2; \mathbb{R})$-invariant suborbifolds the corresponding vector bundles might have equivariant subbundles, which provides additional information on corresponding subcollections of the Lyapunov exponents, or even gives their explicit values in some cases, like in the case of Teichmüller curves considered in the paper of I. Bouw and M. Möller [BMö].

Theorem 1 illustrates a phenomenon of *deviation spectrum*. It was proved by G. Forni in [Fo1] that ergodic sums of smooth functions on an interval along trajectories of interval exchange transformations, and ergodic integrals of smooth functions on flat surfaces along trajectories of directional flows have deviation spectrum analogous to the one described in Theorem 1. L. Flaminio and G. Forni showed that the same phenomenon can be observed for other parabolic dynamical systems, for example, for the horocycle flow on compact surfaces of constant negative curvature [FlFo].

Idea of the proof: renormalization. The reason why the deviation of the cycles c_j from the asymptotic direction is governed by the Teichmüller geodesic flow is illustrated in Figure 4. In a sense, we follow the initial ideas of H. Masur [M1] and of W. Veech [V1].

Fix a horizontal segment X and emit a vertical trajectory from some point x in X. When the trajectory intersects X for the first time join the corresponding point $T(x)$ to the original point x along X to obtain a closed loop. Here $T: X \to X$

denotes the first return map to the transversal X induced by the vertical flow. Denote by $c(x)$ the corresponding cycle in $H_1(S;\mathbb{Z})$. Let the interval exchange transformation $T\colon X\to X$ decompose X into n subintervals $X_1\sqcup\cdots\sqcup X_n$. It is easy to see that the "first return cycle" $c(x)$ is piecewise constant: we have $c(x)=c(x')=: c(X_j)$ whenever x and x' belong to the same subinterval X_j, see Figure 4. It is easy to see that

$$c_N(x)=c(x)+c(T(x))+\cdots+c(T^{N-1}(x)).$$

The average of this sum with respect to the "time" N tends to the asymptotic cycle c. We need to study the deviation of this sum from the value $N\cdot c$. To do this consider a shorter subinterval X' as in Figure 4. Its length is chosen in such way, that the first return map of the vertical flow again induces an interval exchange transformation $T'\colon X'\to X'$ of n subintervals. New first return cycles $c'(X'_k)$ to the interval X' are expressed in terms of the initial first return cycles $c(X_j)$ by the linear relations below; the lengths $|X'_k|$ of subintervals of the new partition $X'=X'_1\sqcup\cdots\sqcup X'_m$ are expressed in terms of the lengths $|X_j|$ of subintervals of the initial partition by dual linear relations:

$$c'(X'_k)=\sum_{j=1}^{n}A_{jk}\cdot c(X_j),\qquad |X_j|=\sum_{k=1}^{n}A_{jk}\cdot|X'_k|,$$

where a nonnegative integer matrix A_{jk} is completely determined by the initial interval exchange transformation $T\colon X\to X$ and by the choice of $X'\subset X$.

To construct the cycle c_N representing a long piece of leaf of the vertical foliation we followed the trajectory $x, T(x),\dots,T^{N-1}(x)$ of the initial interval exchange transformation $T\colon X\to X$ and computed the corresponding ergodic sum. Passing to a shorter horizontal interval $X'\subset X$ we can follow the trajectory $x, T'(x),\dots,(T')^{N'-1}(x)$ of the new interval exchange transformation $T'\colon X'\to X'$ (provided $x\in X'$). Since the subinterval X' is shorter than X we cover the initial piece of trajectory of the vertical flow in a smaller number N' of steps. In other words, passing from T to T' we accelerate the time: it is easy to see that the trajectory $x, T'(x),\dots,(T')^{N'-1}(x)$ follows the trajectory $x, T(x),\dots,T^{N-1}(x)$ but jumps over several iterations of T at a time.

This approach would not be efficient if the new first return map $T'\colon X'\to X'$ would be more complicated than the initial one. But we know that passing from T to T' we stay within a family of interval exchange transformations of some fixed number n of subintervals, and, moreover, that the new "first return cycles" $c'(X'_k)$ and the lengths $|X'_k|$ of the new subintervals are expressed in terms of the initial ones by means of the $n\times n$-matrix A, which depends only on the choice of $X'\subset X$ and which can be easily computed.

Our strategy can be now formulated as follows. One can define an explicit algorithm (generalizing Euclidean algorithm) which canonically associates to an interval exchange transformation $T\colon X\to X$ some specific subinterval $X'\subset X$ and, hence,

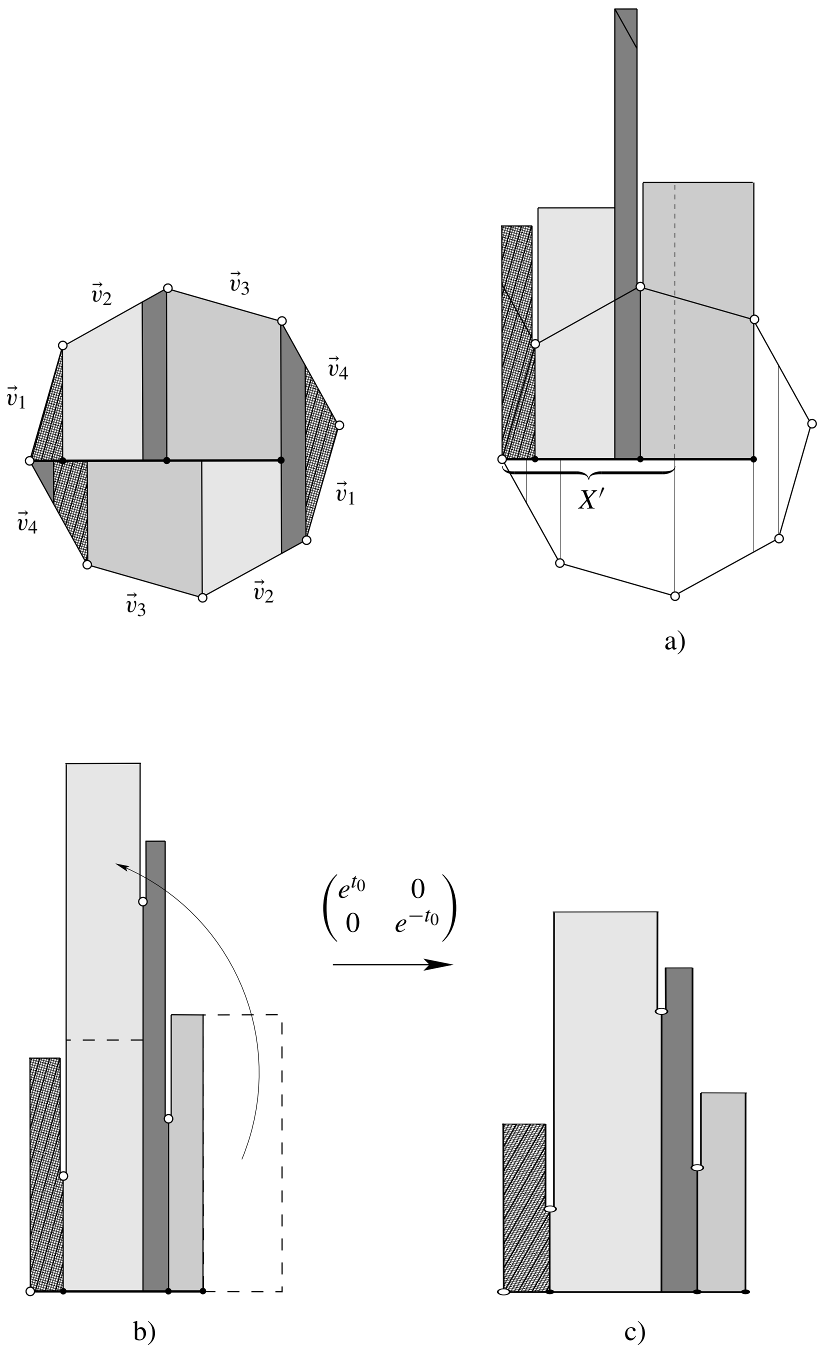

Figure 4. Idea of renormalization. a) Unwrap the flat surface into "zippered rectangles". b) Shorten the base of the corresponding zippered rectangles. c) Expand the resulting tall and narrow zippered rectangle horizontally and contract it vertically by same factor e^{t_0}.

a new interval exchange transformation $T'\colon X' \to X'$. Similarly to the Euclidean algorithm our algorithm is invariant under proportional rescaling of X and X', so, when we find it convenient, we can always rescale the length of the interval to one. This algorithm can be considered as a map $\mathcal{T}$ from the space of all interval exchange transformations of a given number n of subintervals to itself. Applying recursively this algorithm we construct a sequence of subintervals $X = X^{(0)} \supset X^{(1)} \supset X^{(2)} \supset \cdots$ and a sequence of matrices $A = A(T^{(0)}), A(T^{(1)}), \dots$ describing transitions form interval exchange transformation $T^{(r)}\colon X^{(r)} \to X^{(r)}$ to interval exchange transformation $T^{(r+1)}\colon X^{(r+1)} \to X^{(r+1)}$. Taking a product $A^{(s)} = A(T^{(0)}) \cdot A(T^{(1)}) \dots A(T^{(s-1)})$ we can immediately express the "first return cycles" to a microscopic subinterval $X^{(s)}$ in terms of the initial "first return cycles" to X. Considering now the matrices A as the values of a matrix-valued function on the space of interval exchange transformations, we realize that we study the products of matrices A along the orbits $T^{(0)}, T^{(1)}, \dots, T^{(s-1)}$ of the map on the space of interval exchange transformations. When the map is ergodic with respect to a finite measure, the properties of these products are described by the Oseledets theorem, and the cycles c_N have a deviation spectrum governed by the Lyapunov exponents of the cocycle A on the space of interval exchange transformations.

Note that the first return cycle to the subinterval $X^{(s)}$ (which is very short) represents the cycle c_N corresponding to a very long trajectory $x, T(x), \dots, T^{N-1}(x)$ of the initial interval exchange transformation. In other words, our renormalization procedure $\mathcal{T}$ plays a role of a time acceleration machine: morally, instead of getting the cycle c_N by following a trajectory $x, T(x), \dots, T^{N-1}(x)$ of the initial interval exchange transformation for the exponential time $N \sim \exp(\mathrm{const} \cdot s)$ we obtain the cycle c_N applying only s steps of the renormalization map $\mathcal{T}$ on the space of interval exchange transformations.

It remains to establish the relation between the cocycle A over the map $\mathcal{T}$ and the Teichmüller geodesic flow. Conceptually, this relation was elaborated in the original paper of W. Veech [V1].

First let us discuss how can one "almost canonically" (that is up to a finite ambiguity) choose a zippered rectangles representation of a flat surface. Note that Figure 4 suggests the way which allows to obtain infinitely many zippered rectangles representations of the same flat surface: we chop an appropriate rectangle on the right, put it atop the corresponding rectangle and then repeat the procedure recursively. This resembles the situation with a representation of a flat torus by a parallelogram: a point of the fundamental domain in Figure 2 provides a canonical representative though any point of the corresponding $\mathrm{SL}(2, \mathbb{Z})$-orbit represents the same flat torus. A "canonical" zippered rectangles decomposition of a flat surface also belongs to some fundamental domain. Following W. Veech one can define the fundamental domain in terms of some specific choice of a "canonical" horizontal interval X. Namely, let us position the left endpoint of X at a conical singularity. Let us choose the length of X in such way that the interval exchange transformation $T\colon X \to X$ induced by the first return of the vertical flow to X has minimal possible number $n = 2g + m - 1$

of subintervals under exchange. Among all such horizontal segments X choose the shortest one, which length is greater than or equal to one. This construction is applicable to almost all flat surfaces; the finite ambiguity corresponds to the finite freedom in the choice of the conical singularity and in the choice of the horizontal ray adjacent to it.

Since the interval X defines a decomposition of (almost any) flat surface into "zippered rectangles" (see Figure 4) we can pass from the space of flat surfaces to the space of zippered rectangles (which can be considered as a finite ramified covering over the space of flat surfaces). Teichmüller geodesic flow lifts naturally to the space of zippered rectangles. It acts on zippered rectangles by expansion in horizontal direction and contraction in vertical direction; i.e. the zippered rectangles are modified by linear transformations $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$. However, as soon as the Teichmüller geodesic flow brings us out of the fundamental domain, we have to modify the zippered rectangles decomposition to the "canonical one" corresponding to the fundamental domain. (Compare to Figure 2 where the Teichmüller geodesic flow corresponds to the standard geodesic flow in the hyperbolic metric on the upper half-plane.) The corresponding modification of zippered rectangles (chop an appropriate rectangle on the right, put it atop the corresponding rectangle; repeat the procedure several times, if necessary) is illustrated in Figure 4.

Now everything is ready to establish the relation between the Teichmüller geodesic flow and the map $\mathcal{T}$ on the space of interval exchange transformations.

Consider some codimension one subspace Υ in the space of zippered rectangles transversal to the Teichmüller geodesic flow. Say, Υ might be defined by the requirement that the base X of the zippered rectangles decomposition has length one, $|X| = 1$. This is the choice in the original paper of W. Veech [V1]; under this choice Υ represents part of the boundary of the fundamental domain in the space of zippered rectangles. Teichmüller geodesic flow defines the first return map $\mathcal{S}\colon \Upsilon \to \Upsilon$ to the section Υ. The map $\mathcal{S}$ can be described as follows. Take a flat surface of unit area decomposed into zippered rectangles Z with the base X of length one. Apply expansion in horizontal direction and contraction in vertical direction. For some $t_0(Z)$ the deformed zippered rectangles can be rearranged as in Figure 4 to get back to the base of length one; the result is the image of the map $\mathcal{S}$. Actually, we can first apply the rearrangement as in Figure 4 to the initial zippered rectangles Z and then apply the transformation $\begin{pmatrix} e^{t_0} & 0 \\ 0 & e^{-t_0} \end{pmatrix}$ – the two operations commute. This gives, in particular, an explicit formula for $t_0(Z)$. Namely let $|X_n|$ be the width of the rightmost rectangle and let $|X_k|$ be the width of the rectangle, which top horizontal side is glued to the rightmost position at the base X. (For the upper zippered rectangle decomposition in Figure 4 we have $n = 4$ and $k = 2$.) Then

$$t_0 = -\log\big(1 - \min(|X_n|, |X_k|)\big).$$

Recall that a decomposition of a flat surface into zippered rectangles naturally defines an interval exchange transformation – the first return map of the vertical flow to the base X of zippered rectangles. Hence, the map $\mathcal{S}$ of the subspace Υ of zippered rectangles defines an induced map on the space of interval exchange transformations. It remains to note that this induced map is exactly the map $\mathcal{T}$. In other words, the map $\mathcal{S}\colon \Upsilon \to \Upsilon$ induced by the first return of the Teichmüller geodesic flow to the subspace Υ of zippered rectangles is the suspension of the map $\mathcal{T}$ on the space of interval exchange transformations.

We complete with a remark concerning the choice of a section. The natural section Υ chosen in the original paper of W. Veech [V1] is in a sense too large: the corresponding invariant measure (induced from the measure on the space of flat surfaces) is infinite. Choosing an appropriate subset $\Upsilon' \subset \Upsilon$ one can get finite invariant measure. Moreover, the subset Υ' can be chosen in such way that the corresponding first return map $\mathcal{S}'\colon \Upsilon' \to \Upsilon'$ of the Teichmüller geodesic flow is a suspension of some natural map $\mathcal{G}$ on the space of interval exchange transformations, see [Z1]. According to the results of H. Masur [M1] and W. Veech [V1] the Teichmüller geodesic flow is ergodic which implies ergodicity of the maps $\mathcal{S}'$ and $\mathcal{G}$. To apply Oseledets theorem one should, actually, consider the induced cocycle B over this new map $\mathcal{G}$ instead of the cocycle A over the map $\mathcal{T}$ described above.

2. Closed geodesics on flat surfaces

Consider a flat surface S; we always assume that the flat metric on S has trivial holonomy, and that the surface S has finite number of cone-type singularities. By convention a flat surface is endowed with a choice of direction, refereed to as a "vertical direction", or as a "direction to the North". Since the flat metric has trivial holonomy, this direction can be transported in a unique way to any point of the surface.

A geodesic segment joining two conical singularities and having no conical points in its interior is called *saddle connection*. The case when boundaries of a saddle connection coincide is not excluded: a saddle connection might join a conical point to itself. In this section we study saddle connections and closed regular geodesics on a generic flat surface S of genus $g \geq 2$. In particular, we count them and we explain the following curious phenomenon: saddle connections and closed regular geodesics often appear in pairs, triples, etc of parallel saddle connections (correspondingly closed regular geodesics) of the same direction and length. When all saddle connections (closed regular geodesics) in such configuration are short the corresponding flat surface is almost degenerate; it is located close to the boundary of the moduli space. A description of possible configurations of parallel saddle connections (closed geodesics) gives us a description of the multidimensional "cusps" of the strata.

The results of this section are based on the joint work with A. Eskin and H. Masur [EMZ] and on their work [MZ]. A series of beautiful results developing the counting problems considered here were recently obtained by Ya. Vorobets [Vo].

Counting closed geodesics and saddle connections. Closed geodesics on flat surfaces of higher genera have some similarities with ones on the torus. Suppose that we have a regular closed geodesic passing through a point $x_0 \in S$. Emitting a geodesic from a nearby point x in the same direction we obtain a parallel closed geodesic of the same length as the initial one. Thus, closed geodesics appear in families of parallel closed geodesics. However, in the torus case every such family fills the entire torus while a family of parallel regular closed geodesics on a flat surfaces of higher genus fills only part of the surface. Namely, it fills a flat cylinder having a conical singularity on each of its boundaries. Typically, a maximal cylinder of closed regular geodesics is bounded by a pair of closed saddle connections. Reciprocally, any saddle connection joining a conical point P to itself and coming back to P at the angle π bounds a cylinder filled with closed regular geodesics.

A geodesic representative of a homotopy class of a curve on a flat surface is realized in general by a broken line of geodesic segments with vertices at conical points. By convention we consider only closed *regular* geodesics (which by definition do not pass through conical points) or saddle connections (which by definition do not have conical points in its interior). Everywhere in this section we normalize the area of a flat surface to one.

Let $N_{\mathrm{sc}}(S, L)$ be the number of saddle connections of length at most L on a flat surface S. Let $N_{\mathrm{cg}}(S, L)$ be the number of maximal cylinders filled with closed regular geodesics of length at most L on S. It was proved by H. Masur that for any flat surface S both counting functions $N(S, L)$ grow quadratically in L. Namely, there exist constants $0 < \mathrm{const}_1(S) < \mathrm{const}_2(S) < \infty$ such that

$$\mathrm{const}_1(S) \le N(S, L)/L^2 \le \mathrm{const}_2(S)$$

for L sufficiently large. Recently Ya. Vorobets has obtained uniform estimates for the constants $\mathrm{const}_1(S)$ and $\mathrm{const}_2(S)$ which depend only on the genus of S, see [Vo]. Passing from *all* flat surfaces to *almost all* surfaces in a given connected component of a given stratum one gets a much more precise result, see [EM]:

Theorem (A. Eskin and H. Masur). *For almost all flat surfaces S in any stratum $\mathcal{H}(d_1, \dots, d_m)$ the counting functions $N_{\mathrm{sc}}(S, L)$ and $N_{\mathrm{cg}}(S, L)$ have exact quadratic asymptotics*

$$\lim_{L\to\infty} \frac{N_{\mathrm{sc}}(S, L)}{\pi L^2} = c_{\mathrm{sc}}(S), \quad \lim_{L\to\infty} \frac{N_{\mathrm{cg}}(S, L)}{\pi L^2} = c_{\mathrm{cg}}(S).$$

Moreover, the Siegel–Veech constants $c_{\mathrm{sc}}(S)$ (correspondingly $c_{\mathrm{cg}}(S)$) coincide for almost all flat surfaces S in each connected component $\mathcal{H}_1^{\mathrm{comp}}(d_1, \dots, d_m)$ of the stratum.

Phenomenon of higher multiplicities. Note that the direction to the North is well-defined even at a conical point of a flat surface, moreover, at a conical point P_1 with a cone angle $2\pi k$ we have k different directions to the North! Consider some saddle connection $\gamma_1 = [P_1 P_2]$ with an endpoint at P_1. Memorize its direction, say, let it be the North–West direction. Let us launch a geodesic from the same starting point P_1 in one of the remaining $k-1$ North–West directions. Let us study how big is the chance to hit P_2 ones again, and how big is the chance to hit it after passing the same distance as before. We do not exclude the case $P_1 = P_2$. Intuitively it is clear that the answer to the first question is: "the chances are low" and to the second one is "the chances are even lower". This makes the following theorem (see [EMZ]) somehow counterintuitive:

Theorem 2 (A. Eskin, H. Masur, A. Zorich). *For almost any flat surface S in any stratum and for any pair P_1, P_2 of conical singularities on S the function $N_2(S, L)$ counting the number of pairs of parallel saddle connections of the same length joining P_1 to P_2 has exact quadratic asymptotics*

$$\lim_{L\to\infty} \frac{N_2(S, L)}{\pi L^2} = c_2 > 0,$$

where the Siegel–Veech constant c_2 depends only on the connected component of the stratum and on the cone angles at P_1 and P_2.

For almost all flat surfaces S in any stratum one cannot find neither a single pair of parallel saddle connections on S of different length, nor a single pair of parallel saddle connections joining different pairs of singularities.

Analogous statements (with some reservations for specific connected components of certain strata) can be formulated for arrangements of 3, 4, ... parallel saddle connections. The situation with closed regular geodesics is similar: they might appear (also with some exceptions for specific connected components of certain strata) in families of $2, 3, \dots, g-1$ distinct maximal cylinders filled with parallel closed regular geodesics of equal length. A general formula for the Siegel–Veech constant in the corresponding quadratic asymptotics is presented at the end of this section, while here we want to discuss the numerical values of Siegel–Veech constants in a simple concrete example. We consider the principal strata $\mathcal{H}(1, \dots, 1)$ in small genera. Let $N_{k\_\mathrm{cyl}}(S, L)$ be the corresponding counting function, where k is the number of distinct maximal cylinders filled with parallel closed regular geodesics of equal length bounded by L. Let

$$c_{k\_\mathrm{cyl}} = \lim_{L\to\infty} \frac{N_{k\_\mathrm{cyl}}(S, L)}{\pi L^2}.$$

The table below (extracted from [EMZ]) presents the values of c_{k_cyl} for $g = 1, \dots, 4$. Note that for a generic flat surface S of genus g a configuration of $k \geq g$ cylinders is not realizable, so we do not fill the corresponding entry.

k	$g=1$	$g=2$	$g=3$	$g=4$
1	$\frac{1}{2}\cdot\frac{1}{\zeta(2)}\approx 0.304$	$\frac{5}{2}\cdot\frac{1}{\zeta(2)}\approx 1.52$	$\frac{36}{7}\cdot\frac{1}{\zeta(2)}\approx 3.13$	$\frac{3150}{377}\cdot\frac{1}{\zeta(2)}\approx 5.08$
2	–	–	$\frac{3}{14}\cdot\frac{1}{\zeta(2)}\approx 0.13$	$\frac{90}{377}\cdot\frac{1}{\zeta(2)}\approx 0.145$
3	–	–	–	$\frac{5}{754}\cdot\frac{1}{\zeta(2)}\approx 0.00403$

Comparing these values we see, that our intuition was not quite misleading. Morally, in genus $g = 4$ a closed regular geodesic belongs to a one-cylinder family with "probability" 97.1%, to a two-cylinder family with "probability" 2.8% and to a three-cylinder family with "probability" only 0.1% (where "probabilities" are calculated proportionally to the Siegel–Veech constants $5.08 : 0.145 : 0.00403$).

Rigid configurations of saddle connections and "cusps" of the strata. A saddle connection or a regular closed geodesic on a flat surface S persists under small deformations of S inside the corresponding stratum. It might happen that any deformation of a given flat surface which shortens some specific saddle connection necessarily shortens some other saddle connections. We say that a collection $\{\gamma_1, \dots, \gamma_n\}$ of saddle connections is *rigid* if any sufficiently small deformation of the flat surface inside the stratum preserves the proportions $|\gamma_1| : |\gamma_2| : \dots : |\gamma_n|$ of the lengths of all saddle connections in the collection. It was shown in [EMZ] that all saddle connections in any rigid collection are *homologous*. Since their directions and lengths can be expressed in terms of integrals of the holomorphic 1-form ω along corresponding paths, this implies that homologous saddle connections $\gamma_1, \dots, \gamma_n$ are parallel and have equal length and either all of them join the same pair of distinct singular points, or all γ_i are closed loops.

This implies that when saddle connections in a rigid collection are contracted by a continuous deformation, the limiting flat surface generically decomposes into several connected components represented by nondegenerate flat surfaces $S'_1, \dots, S'_k$, see Figure 5, where k might vary from one to the genus of the initial surface. Let the initial surface S belong to a stratum $\mathcal{H}(d_1, \dots, d_m)$. Denote the set with multiplicities $\{d_1, \dots, d_m\}$ by β. Let $\mathcal{H}(\beta'_j)$ be the stratum ambient for S'_j. The stratum $\mathcal{H}(\beta') = \mathcal{H}(\beta'_1) \sqcup \dots \sqcup \mathcal{H}(\beta'_k)$ of disconnected flat surfaces $S'_1 \sqcup \dots \sqcup S'_k$ is referred to as a *principal boundary* stratum of the stratum $\mathcal{H}(\beta)$. For any connected component of any stratum $\mathcal{H}(\beta)$ the paper [EMZ] describes all principal boundary strata; their

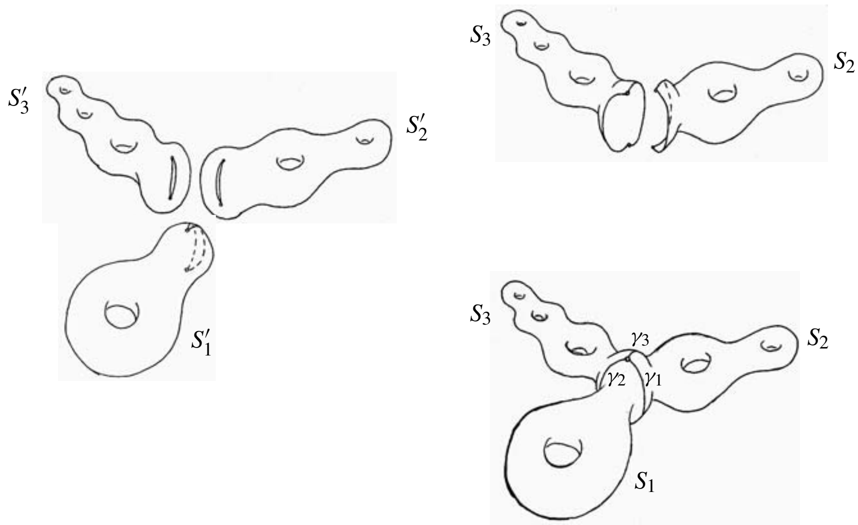

Figure 5. Multiple homologous saddle connections, topological picture (after [EMZ]).

union is called the principal boundary of the corresponding connected component of $\mathcal{H}(\beta)$.

The paper [EMZ] also presents the inverse construction. Consider any flat surface $S'_1 \sqcup \cdots \sqcup S'_k \in \mathcal{H}(\beta')$ in the principal boundary of $\mathcal{H}(\beta)$; consider a sufficiently small value of a complex parameter $\varepsilon \in \mathbb{C}$. One can reconstruct the flat surface $S \in \mathcal{H}(\beta)$ endowed with a collection of homologous saddle connections $\gamma_1, \ldots, \gamma_n$ such that $\int_{\gamma_i} \omega = \varepsilon$, and such that degeneration of S contracting the saddle connections γ_i in the collection gives the surface $S'_1 \sqcup \cdots \sqcup S'_k$. This inverse construction involves several surgeries of the flat structure. Having a disconnected flat surface $S'_1 \sqcup \cdots \sqcup S'_k$ one applies an appropriate surgery to each S'_j producing a surface S_j with boundary. The surgery depends on the parameter ε: the boundary of each S_j is composed from two geodesic segments of lengths $|\varepsilon|$; moreover, the boundary components of S_j and S_{j+1} are compatible, which allows to glue the compound surface S from the collection of surfaces with boundary, see Figure 5 as an example.

A collection $\gamma = \{\gamma_1, \ldots, \gamma_n\}$ of homologous saddle connections determines the following data on combinatorial geometry of the decomposition $S \setminus \gamma$: the number of components, their boundary structure, the singularity data for each component, the cyclic order in which the components are glued to each other. These data are referred to as a *configuration* of homologous saddle connections. A configuration $\mathcal{C}$ uniquely determines the corresponding boundary stratum $\mathcal{H}(\beta'_{\mathcal{C}})$; it does not depend on the collection γ of homologous saddle connections representing the configuration $\mathcal{C}$.

The constructions above explain how configurations $\mathcal{C}$ of homologous saddle connections on flat surfaces $S \in \mathcal{H}(\beta)$ determine the "cusps" of the stratum $\mathcal{H}(\beta)$. Consider a subset $\mathcal{H}_1^{\varepsilon}(\beta) \subset \mathcal{H}(\beta)$ of surfaces of area one having a saddle connection shorter than ε. Up to a subset $\mathcal{H}_1^{\varepsilon,\text{thin}}(\beta)$ of negligibly small measure the set $\mathcal{H}_1^{\varepsilon,\text{thick}}(\beta) = \mathcal{H}_1^{\varepsilon}(\beta) \setminus \mathcal{H}_1^{\varepsilon,\text{thin}}(\beta)$ might be represented as a disjoint union

$$\mathcal{H}_1^{\varepsilon,\text{thick}}(\beta) \approx \bigsqcup_{\mathcal{C}} \mathcal{H}_1^{\varepsilon}(\mathcal{C})$$

of neighborhoods $\mathcal{H}_1^{\varepsilon}(\mathcal{C})$ of the corresponding "cusps" $\mathcal{C}$. Here $\mathcal{C}$ runs over a finite set of configurations admissible for the given stratum $\mathcal{H}_1(\beta)$; this set is explicitly described in [EMZ].

When a configuration $\mathcal{C}$ is composed from homologous saddle connections joining *distinct* zeroes, the neighborhood $\mathcal{H}_1^{\varepsilon}(\mathcal{C})$ of the cusp $\mathcal{C}$ has the structure of a fiber bundle over the corresponding boundary stratum $\mathcal{H}(\beta'_{\mathcal{C}})$ (up to a difference in a set of a negligibly small measure). A fiber of this bundle is represented by a finite cover over the Euclidean disc of radius ε ramified at the center of the disc. Moreover, the canonical measure in $\mathcal{H}_1^{\varepsilon}(\mathcal{C})$ decomposes into a product measure of the canonical measure in the boundary stratum $\mathcal{H}(\beta'_{\mathcal{C}})$ and the Euclidean measure in the fiber (see [EMZ]), so

$$\operatorname{Vol}\left(\mathcal{H}_1^{\varepsilon}(\mathcal{C})\right) = (\text{combinatorial factor}) \cdot \pi\varepsilon^2 \cdot \prod_{j=1}^{k} \operatorname{Vol} \mathcal{H}_1(\beta'_j) \; + \; o(\varepsilon^2). \qquad (1)$$

Remark. We warn the reader that the correspondence between compactification of the moduli space of Abelian differentials and the Deligne–Mumford compactification of the underlying moduli space of curves is not straightforward. In particular, the desingularized stable curve corresponding to the limiting flat surface generically *is not* represented as a union of Riemann surfaces corresponding to $S'_1, \dots, S'_k$ – the stable curve might contain more components.

Evaluation of the Siegel–Veech constants. Consider a flat surface S. To every closed regular geodesic γ on S we can associate a vector $\vec{v}(\gamma)$ in $\mathbb{R}^2$ having the length and the direction of γ. In other words, $\vec{v} = \int_\gamma \omega$, where we consider a complex number as a vector in $\mathbb{R}^2 \simeq \mathbb{C}$. Applying this construction to all closed regular geodesic on S we construct a discrete set $V(S) \subset \mathbb{R}^2$. Consider the following operator $f \mapsto \hat{f}$ from functions with compact support on $\mathbb{R}^2$ to functions on a connected component $\mathcal{H}_1^{\text{comp}}(\beta)$ of the stratum $\mathcal{H}_1(\beta) = \mathcal{H}_1(d_1, \dots, d_m)$:

$$\hat{f}(S) := \sum_{\vec{v} \in V(S)} f(\vec{v}).$$

Function $\hat{f}(S)$ generalizes the counting function $N_{\text{cg}}(S, L)$ introduced in the beginning of this section. Namely, when $f = \chi_L$ is the characteristic function χ_L of the disc of radius L with the center at the origin of $\mathbb{R}^2$, the function $\hat{\chi}_L(S)$ counts the number of regular closed geodesics of length at most L on a flat surface S.

Theorem (W. Veech). *For any function $f : \mathbb{R}^2 \to \mathbb{R}$ with compact support the following equality is valid:*

$$\frac{1}{\operatorname{Vol} \mathcal{H}_1^{\mathrm{comp}}(\beta)} \int_{\mathcal{H}_1^{\mathrm{comp}}(\beta)} \hat{f}(S)\, d\nu_1 = C \int_{\mathbb{R}^2} f(x, y)\, dx\, dy, \tag{2}$$

where the constant C does not depend on the function f.

Note that this is an exact equality. In particular, choosing the characteristic function χ_L of a disc of radius L as a function f we see that for any positive L the *average* number of closed regular geodesics not longer than L on flat surfaces $S \in \mathcal{H}_1^{\mathrm{comp}}(\beta)$ is exactly $C \cdot \pi L^2$, where the Siegel–Veech constant C does not depend on L, but only on the connected component $\mathcal{H}_1^{\mathrm{comp}}(\beta)$.

The theorem of Eskin and Masur cited above tells that for large values of L one gets approximate equality $\hat{\chi}_L(S) \approx c_{\mathrm{cg}} \cdot \pi L^2$ "pointwisely" for almost all individual flat surfaces $S \in \mathcal{H}_1^{\mathrm{comp}}(d_1, \dots, d_m)$. It is proved in [EM] that the corresponding Siegel–Veech constant c_{cg} coincides with the constant C in equation (2) above.

Actually, the same technique can be applied to count separately pairs, triples, or any other specific configurations $\mathcal{C}$ of homologous saddle connections. Every time when we find a collection of homologous saddle connections $\gamma_1, \dots, \gamma_n$ representing the chosen configuration $\mathcal{C}$ we construct a vector $\vec{v} = \int_{\gamma_i} \omega$. Since all $\gamma_1, \dots, \gamma_n$ are homologous, we can take any of them as γ_i. Taking all possible collections of homologous saddle connections on S representing the fixed configuration $\mathcal{C}$ we construct new discrete set $V_{\mathcal{C}}(S) \subset \mathbb{R}^2$ and new functional $f \mapsto \hat{f}_{\mathcal{C}}$. Theorem of Eskin and Masur and theorem of Veech [V4] presented above are valid for $\hat{f}_{\mathcal{C}}$. The corresponding Siegel–Veech constant $c(\mathcal{C})$ responsible for the quadratic growth rate $N_{\mathcal{C}}(S, L) \sim c(\mathcal{C}) \cdot \pi L^2$ of the number of collections of homologous saddle connections of the type $\mathcal{C}$ on an individual generic flat surface S coincides with the constant $C(\mathcal{C})$ in the expression analogous to (2).

Formula (2) can be applied to $\hat{\chi}_L$ for any value of L. In particular, instead of taking large L we can choose a very small $L = \varepsilon \ll 1$. The corresponding function $\hat{\chi}_\varepsilon(S)$ counts how many collections of parallel ε-short saddle connections (closed geodesics) of the type $\mathcal{C}$ we can find on a flat surface $S \in \mathcal{H}_1^{\mathrm{comp}}(\beta)$. For the flat surfaces S outside of the subset $\mathcal{H}_1^{\varepsilon}(\mathcal{C}) \subset \mathcal{H}_1^{\mathrm{comp}}(\beta)$ there are no such saddle connections (closed geodesics), so $\hat{\chi}_\varepsilon(S) = 0$. For surfaces S from the subset $\mathcal{H}_1^{\varepsilon,\mathrm{thick}}(\mathcal{C})$ there is exactly one collection like this, $\hat{\chi}_\varepsilon(S) = 1$. Finally, for the surfaces from the remaining (very small) subset $\mathcal{H}_1^{\varepsilon,\mathrm{thin}}(\mathcal{C}) = \mathcal{H}_1^{\varepsilon}(\mathcal{C}) \setminus \mathcal{H}_1^{\varepsilon,\mathrm{thick}}(\mathcal{C})$ one has $\hat{\chi}_\varepsilon(S) \geq 1$. Eskin and Masur have proved in [EM] that though $\hat{\chi}_\varepsilon(S)$ might be large on $\mathcal{H}_1^{\varepsilon,\mathrm{thin}}$ the measure of this subset is so small (see [MS]) that

$$\int_{\mathcal{H}_1^{\varepsilon,\mathrm{thin}}(\mathcal{C})} \hat{\chi}_\varepsilon(S)\, d\nu_1 = o(\varepsilon^2)$$

and hence

$$\int_{\mathcal{H}_1^{\mathrm{comp}}(\beta)} \hat{\chi}_\varepsilon(S)\, d\nu_1 = \mathrm{Vol}\, \mathcal{H}_1^{\varepsilon,\mathrm{thick}}(\mathcal{C}) \;+\; o(\varepsilon^2).$$

This latter volume is almost the same as the volume $\mathrm{Vol}\,\mathcal{H}_1^{\varepsilon}(\mathcal{C})$ of the neighborhood of the cusp $\mathcal{C}$ evaluated in equation (1) above, namely, $\mathrm{Vol}\,\mathcal{H}_1^{\varepsilon,\mathrm{thick}}(\mathcal{C}) = \mathrm{Vol}\,\mathcal{H}_1^{\varepsilon}(\mathcal{C}) + o(\varepsilon^2)$ (see [MS]). Taking into consideration that

$$\int_{\mathbb{R}^2} \chi_\varepsilon(x, y)\, dx\, dy = \pi\varepsilon^2$$

and applying the Siegel–Veech formula (2) to χ_ε we finally get

$$\frac{\mathrm{Vol}\,\mathcal{H}_1^{\varepsilon}(\mathcal{C})}{\mathrm{Vol}\,\mathcal{H}_1^{\mathrm{comp}}(d_1, \dots, d_m)} + o(\varepsilon^2) = c(\mathcal{C}) \cdot \pi\varepsilon^2$$

which implies the following formula for the Siegel–Veech constant $c(\mathcal{C})$:

$$\begin{aligned} c(\mathcal{C}) &= \lim_{\varepsilon\to 0} \frac{1}{\pi\varepsilon^2} \frac{\mathrm{Vol}(\text{“}\varepsilon\text{-neighborhood of the cusp } \mathcal{C}\text{ ”})}{\mathrm{Vol}\,\mathcal{H}_1^{\mathrm{comp}}(\beta)} \\ &= (\text{explicit combinatorial factor}) \cdot \frac{\prod_{j=1}^{k} \mathrm{Vol}\,\mathcal{H}_1(\beta'_k)}{\mathrm{Vol}\,\mathcal{H}_1^{\mathrm{comp}}(\beta)}. \end{aligned}$$

Sums of the Lyapunov exponents $\nu_1 + \cdots + \nu_g$ discussed in Section 1 are closely related to the Siegel–Veech constants.

3. Ergodic components of the Teichmüller flow

According to the theorems of H. Masur [M1] and of W. Veech [V1] Teichmüller geodesic flow is ergodic on every connected component of every stratum of flat surfaces. Thus, the Lyapunov exponents $1 + \nu_j$ of the Teichmüller geodesic flow responsible for the deviation spectrum of generic geodesics on a flat surface (see Section 1), or Siegel–Veech constants responsible for counting of closed geodesics on a flat surface (see Section 2) are specific for each connected component of each stratum. The fact that the strata $\mathcal{H}_1(d_1, \dots, d_m)$ are not necessarily connected was observed by W. Veech.

In order to formulate the classification theorem for connected components of the strata $\mathcal{H}(d_1, \dots, d_m)$ we need to describe the classifying invariants. There are two of them: spin structure and hyperellipticity. Both notions are applicable only to part of the strata: flat surfaces from the strata $\mathcal{H}(2d_1, \dots, 2d_m)$ have even or odd spin structure. The strata $\mathcal{H}(2g-2)$ and $\mathcal{H}(g-1, g-1)$ have a special hyperelliptic connected component.

The results of this section are based on the joint work with M. Kontsevich [KZ].

Spin structure. Consider a flat surface S from a stratum $\mathcal{H}(2d_1, \dots, 2d_m)$. Let $\rho\colon S^1 \to S$ be a smooth closed path on S; here S^1 is a standard circle. Note that at any point of the surfaces S we know where is the "direction to the North". Hence, at any point $\rho(t) = x \in S$ we can apply a compass and measure the direction of the tangent vector $\dot{x}$. Moving along our path $\rho(t)$ we make the tangent vector turn in the compass. Thus we get a map $G(\rho)\colon S^1 \to S^1$ from the parameter circle to the circumference of the compass. This map is called the *Gauss map*. We define the *index* $\operatorname{ind}(\rho)$ of the path ρ as a degree of the corresponding Gauss map (or, in other words as the algebraic number of turns of the tangent vector around the compass) taken modulo 2.

$$\operatorname{ind}(\rho) = \deg G(\rho) \mod 2.$$

It is easy to see that $\operatorname{ind}(\rho)$ does not depend on parameterization. Moreover, it does not change under small deformations of the path. Deforming the path more drastically we may change its position with respect to conical singularities of the flat metric. Say, the initial path might go on the left of P_k and its deformation might pass on the right of P_k. This deformation changes the $\deg G(\rho)$. However, if the cone angle at P_k is of the type $2\pi(2d_k + 1)$, then $\deg G(\rho) \mod 2$ does not change! This observation explains why $\operatorname{ind}(\rho)$ is well-defined for a free homotopy class $[\rho]$ when $S \in \mathcal{H}(2d_1, \dots, 2d_m)$ (and hence, when all cone angles are odd multiples of 2π).

Consider a collection of closed smooth paths $a_1, b_1, \dots, a_g, b_g$ representing a symplectic basis of homology $H_1(S, \mathbb{Z}/2\mathbb{Z})$. We define the *parity of the spin-structure* of a flat surface $S \in \mathcal{H}(2d_1, \dots, 2d_m)$ as

$$\phi(S) = \sum_{i=1}^{g} (\operatorname{ind}(a_i) + 1)\,(\operatorname{ind}(b_i) + 1) \mod 2.$$

Lemma. *The value $\phi(S)$ does not depend on symplectic basis of cycles $\{a_i, b_i\}$. It does not change under continuous deformations of S in $\mathcal{H}(2d_1, \dots, 2d_m)$.*

The lemma above shows that the parity of the spin structure is an invariant of connected components of strata of those Abelian differentials (equivalently, flat surfaces) which have zeroes of even degrees (equivalently, conical points with cone angles which are odd multiples of 2π).

Hyperellipticity. A flat surface S may have a symmetry; one specific family of such flat surfaces, which are "more symmetric than others" is of a special interest for us. Recall that there is a one-to-one correspondence between flat surfaces and pairs (Riemann surface M, holomorphic 1-form ω). When the corresponding Riemann surface is hyperelliptic the hyperelliptic involution $\tau\colon M \to M$ acts on any holomorphic 1-form ω as $\tau^*\omega = -\omega$.

We say that a flat surface S is a hyperelliptic flat surface if there is an isometry $\tau\colon S \to S$ such that τ is an involution, $\tau \circ \tau = \mathrm{id}$, and the quotient surface S/τ

is a topological sphere. In flat coordinates differential of such involution obviously satisfies $D\tau = -\operatorname{Id}$.

In a general stratum $\mathcal{H}(d_1, \dots, d_m)$ hyperelliptic flat surfaces form a small subspace of nontrivial codimension. However, there are two special strata, namely, $\mathcal{H}(2g-2)$ and $\mathcal{H}(g-1, g-1)$, for which hyperelliptic surfaces form entire hyperelliptic connected components $\mathcal{H}^{\text{hyp}}(2g-2)$ and $\mathcal{H}^{\text{hyp}}(g-1, g-1)$ correspondingly.

Remark. Note that in the stratum $\mathcal{H}(g-1, g-1)$ there are hyperelliptic flat surfaces of two different types. A hyperelliptic involution $\tau S \to S$ may fix the conical points or might interchange them. It is not difficult to show that for flat surfaces from the connected component $\mathcal{H}^{\text{hyp}}(g-1, g-1)$ the hyperelliptic involution interchanges the conical singularities.

The remaining family of those hyperelliptic flat surfaces in $\mathcal{H}(g-1, g-1)$, for which the hyperelliptic involution keeps the saddle points fixed, forms a subspace of nontrivial codimension in the complement $\mathcal{H}(g-1, g-1) \setminus \mathcal{H}^{\text{hyp}}(g-1, g-1)$. Thus, the hyperelliptic connected component $\mathcal{H}^{\text{hyp}}(g-1, g-1)$ does not coincide with the space of all hyperelliptic flat surfaces.

Classification theorem for Abelian differentials. Now, having introduced the classifying invariants we can present the classification of connected components of strata of flat surfaces (equivalently, of strata of Abelian differentials).

Theorem 3 (M. Kontsevich and A. Zorich). *All connected components of any stratum of flat surfaces of genus $g \geq 4$ are described by the following list:*

- *The stratum $\mathcal{H}(2g-2)$ has three connected components: the hyperelliptic one, $\mathcal{H}^{\text{hyp}}(2g-2)$, and two nonhyperelliptic components: $\mathcal{H}^{\text{even}}(2g-2)$ and $\mathcal{H}^{\text{odd}}(2g-2)$ corresponding to even and odd spin structures.*

- *The stratum $\mathcal{H}(2d, 2d)$, $d \geq 2$ has three connected components: the hyperelliptic one, $\mathcal{H}^{\text{hyp}}(2d, 2d)$, and two nonhyperelliptic components: $\mathcal{H}^{\text{even}}(2d, 2d)$ and $\mathcal{H}^{\text{odd}}(2d, 2d)$.*

- *All the other strata of the form $\mathcal{H}(2d_1, \dots, 2d_m)$ have two connected components: $\mathcal{H}^{\text{even}}(2d_1, \dots, 2d_m)$ and $\mathcal{H}^{\text{odd}}(2d_1, \dots, 2d_n)$, corresponding to even and odd spin structures.*

- *The stratum $\mathcal{H}(2d-1, 2d-1)$, $d \geq 2$, has two connected components; one of them: $\mathcal{H}^{\text{hyp}}(2d-1, 2d-1)$ is hyperelliptic; the other $\mathcal{H}^{\text{nonhyp}}(2d-1, 2d-1)$ is not.*

All the other strata of flat surfaces of genera $g \geq 4$ are nonempty and connected.

In the case of small genera $1 \leq g \leq 3$ some components are missing in comparison with the general case.

Theorem 3′. *The moduli space of flat surfaces of genus $g = 2$ contains two strata: $\mathcal{H}(1, 1)$ and $\mathcal{H}(2)$. Each of them is connected and coincides with its hyperelliptic component.*

Each of the strata $\mathcal{H}(2, 2)$, $\mathcal{H}(4)$ of the moduli space of flat surfaces of genus $g = 3$ has two connected components: the hyperelliptic one, and one having odd spin structure. The other strata are connected for genus $g = 3$.

Since there is a one-to-one correspondence between connected components of the strata and extended Rauzy classes, the classification theorem above classifies also the extended Rauzy classes.

Connected components of the strata $\mathcal{Q}(d_1, \dots, d_m)$ of meromorphic quadratic differentials with at most simple poles are classified in the paper of E. Lanneau [L].

Bibliographical notes. As a much more serious accessible introduction to Teichmüller dynamics I can recommend a collection of surveys of A. Eskin [E], G. Forni [Fo2], P. Hubert and T. Schmidt [HSc] and H. Masur [M2], organized as a chapter of the Handbook of Dynamical Systems. I also recommend recent surveys of H. Masur and S. Tabachnikov [MT] and of J. Smillie [S] especially in the aspects related to billiards in polygons. The part concerning renormalization and interval exchange transformations is presented in the survey of J.-C. Yoccoz [Y]. The ideas presented in the current paper are illustrated in more detailed way in the survey [Z4].

Acknowledgements. A considerable part of the results presented in this survey is obtained in collaboration. I use this opportunity to thank A. Eskin, M. Kontsevich and H. Masur for the pleasure to work with them. I am grateful to M.-C. Vergne for her help with the preparation of the pictures.

References

[AEZ] Athreya, J., Eskin, A., Zorich, A., Rectangular billiards and volumes of spaces of quadratic differentials. In preparation.

[AvVi] Avila, A., Viana, M., Simplicity of Lyapunov spectra: proof of the Zorich–Kontsevich conjecture. Eprint math.DS/0508508, 2005, 36 pp.

[BMö] Bouw, I., and Möller, M., Teichmüller curves, triangle groups, and Lyapunov exponents. Eprint math.AG/0511738, 2005, 30 pp.

[Ca] Calta, K., Veech surfaces and complete periodicity in genus two. *J. Amer. Math. Soc.* **17** (2004), 871–908.

[E] Eskin, A., Counting problems in moduli space. In *Handbook of Dynamical Systems, Vol. 1B* (ed. by B. Hasselblatt and A. Katok), Elsevier Science B.V., Amsterdam 2006, 581–595.

[EM] Eskin, A., Masur, H., Asymptotic formulas on flat surfaces. *Ergodic Theory Dynam. Systems* **21** (2) (2001), 443–478.

[EMZ] Eskin, A., Masur, H., Zorich, A., Moduli spaces of Abelian differentials: the principal boundary, counting problems, and the Siegel–Veech constants. *Publ. Math. Inst. Hautes Études Sci.* **97** (1) (2003), 61–179.

[EO] Eskin, A., Okounkov, A., Asymptotics of number of branched coverings of a torus and volumes of moduli spaces of holomorphic differentials. *Invent. Math.* **145** (1) (2001), 59–104.

[FlFo] Flaminio, L., Forni, G., Invariant distributions and time averages for horocycle flows. *Duke Math. J.* **119** (3) (2003), 465–526.

[Fo1] Forni, G., Deviation of ergodic averages for area-preserving flows on surfaces of higher genus. *Ann. of Math.* (2) **155** (1) (2002), 1–103.

[Fo2] Forni, G., On the Lyapunov exponents of the Kontsevich–Zorich cocycle. In *Handbook of Dynamical Systems, Vol. 1B* (ed. by B. Hasselblatt and A. Katok), Elsevier Science B.V., Amsterdam 2006, 549–580.

[HSc] Hubert, P., and Schmidt, T., An introduction to Veech surfaces. In *Handbook of Dynamical Systems, Vol. 1B* (ed. by B. Hasselblatt and A. Katok), Elsevier Science B.V., Amsterdam 2006, 501–526.

[KeMS] Kerckhoff, S., Masur, H., and Smillie, J., Ergodicity of billiard flows and quadratic differentials. *Ann. of Math.* (2) **124** (1986), 293–311.

[Kr] Krikorian, P., Déviation de moyennes ergodiques, flots de Teichmüller et cocycle de Kontsevich–Zorich (d'après Forni, Kontsevich, Zorich). Séminaire Bourbaki 927, 2003; *Astérisque* **299** (2005), 59–93.

[K] Kontsevich, M., Lyapunov exponents and Hodge theory. In *The mathematical beauty of physics* (Saclay, 1996, in Honor of C. Itzykson), Adv. Ser. Math. Phys. 24, World Sci. Publishing, River Edge, NJ, 1997, 318–332.

[KZ] Kontsevich, M., Zorich, A., Connected components of the moduli spaces of Abelian differentials. *Invent. Math.* **153** (3) (2003), 631–678.

[L] Lanneau, E., Connected components of the moduli spaces of quadratic differentials. Eprint math.GT/0506136, 2005, 41pp.

[M1] Masur, H., Interval exchange transformations and measured foliations. *Ann. of Math.* (2) **115** (1982), 169–200.

[M2] Masur, H., Ergodic Theory of Translation Surfaces. In *Handbook of Dynamical Systems, Vol. 1B* (ed. by B. Hasselblatt and A. Katok), Elsevier Science B.V., Amsterdam 2006, 527–547.

[MS] Masur, H., Smillie, J., Hausdorff dimension of sets of nonergodic foliations. *Ann. of Math.* (2) **134** (1991), 455–543.

[MT] Masur, H., and Tabachnikov, S., Rational Billiards and Flat Structures. In *Handbook of Dynamical Systems, Vol. 1A* (ed. by B. Hasselblatt and A. Katok), Elsevier Science B.V., Amsterdam 2002, 1015–1089.

[MZ] Masur, H., and Zorich, A., Multiple Saddle Connections on Flat Surfaces and Principal Boundary of the Moduli Spaces of Quadratic Differentials. Eprint math.GT/0402197, 2004, 73pp.

[Mc1] McMullen, C., Billiards and Teichmüller curves on Hilbert modular surfaces. *J. Amer. Math. Soc.* **16** (4) (2003), 857–885.

[Mc2] McMullen, C., Dynamics of $\mathrm{SL}_2(\mathbb{R})$ over moduli space in genus two. *Ann. of Math.*, to appear.

[Ra] Rauzy, G., Echanges d'intervalles et transformations induites. *Acta Arith.* **34** (1979), 315–328.

[S] Smillie, J., The dynamics of billiard flows in rational polygons. In *Dynamical systems, ergodic theory and applications* (ed. by Ya. G. Sinai), Encyclopaedia Math. Sci. 100, Math. Physics 1, Springer-Verlag, Berlin 2000, 360–382.

[V1] Veech, W. A., Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.* (2) **115** (1982), 201–242.

[V2] Veech, W. A., The Teichmüller geodesic flow. *Ann. of Math.* (2) **124** (1986), 441–530.

[V3] Veech, W. A., Flat surfaces. *Amer. J. Math.* **115** (1993), 589–689.

[V4] Veech, W. A., Siegel measures. *Ann. of Math.* (2) **148** (1998), 895–944.

[Vo] Vorobets, Ya., Periodic geodesics on generic translation surfaces. In *Algebraic and Topological Dynamics* (ed. by S. Kolyada, Yu. I. Manin and T. Ward), Contemp. Math. 385, Amer. Math. Soc., Providence, RI, 2005, 205–258.

[Y] Yoccoz, J.-C., Continuous fraction algorithms for interval exchange maps: an introduction. In *Frontiers in Number Theory, Physics and Geometry* (Proceedings of Les Houches winter school 2003), Volume 1, Springer-Verlag, Berlin 2006, 403–437.

[Z1] Zorich, A., Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents. *Ann. Inst. Fourier* (*Grenoble*) **46** (2) (1996), 325–370.

[Z2] Zorich, A., Deviation for interval exchange transformations. *Ergodic Theory Dynam. Systems* **17** (1997), 1477–1499.

[Z3] Zorich, A., How do the leaves of a closed 1-form wind around a surface? In *Pseudoperiodic Topology* (ed. by V. I. Arnold, M. Kontsevich, A. Zorich), Amer. Math. Soc. Trans. Ser. 2 197, Amer. Math. Soc., Providence, RI, 1999, 135–178.

[Z4] Zorich, A., Flat surfaces. In *Frontiers in Number Theory, Physics and Geometry* (Proceedings of Les Houches winter school 2003), Volume 1, Springer-Verlag, Berlin 2006, 439–585.

IRMAR, Université Rennes 1, Campus de Beaulieu, 35042 Rennes, France
E-mail: Anton.Zorich@univ-rennes1.fr