Asymptotic formulas on flat surfaces

Alex Eskin* and Howard Masur†

1 Introduction.

This paper concerns asymptotic formulas for the number of closed trajectories and saddle connections on a flat surface. It is inspired by some recent work of W. Veech, especially the recent paper [Ve5]. In particular, the analogy noted in [Ve5] between the problem discussed in this paper and the quantitative version of the Oppenheim Conjecture [EMM] allowed us to use some of the ideas developed by G.A. Margulis in the context of the Oppenheim conjecture to simplify and improve the quadratic upper bound of [Ma2]. This, together with an ergodic theorem proved for this purpose by A. Nevo [Nevo], allowed us to find asymptotic quadratic growth rates for the number of closed trajectories on a generic surface. This sharpens results of [Ve5] where these growth rates were found in the mean.

Flat metrics and quadratic differentials. Given a Riemann surface structure on a closed surface of genus $g > 1$, recall that a holomorphic quadratic differential ϕ is a tensor of the form $\phi(z)dz^2$ in local coordinates with ϕ holomorphic. Away from the zeroes, a coordinate ζ can be chosen so that $\phi = d\zeta^2$, which determines a Euclidean metric $|d\zeta^2|$ in that chart. The change of coordinates away from the zeroes of ϕ are of the form $\zeta \to \pm\zeta + c$, which preserves the Euclidean metric. Consequently, quadratic differentials are sometimes referred to as translation surfaces or flat structures with parallel line fields. A general flat structure or equivalently a surface with zero Gaussian curvature would allow for rotational components in the

*Research partially supported by NSF grant DMS-9704845, the Sloan Foundation and the Packard Foundation

†Research partially supported by NSF grant 9803497

transition functions and therefore would not have a parallel line field. If one can always take the + sign in the change of coordinates, the quadratic differential is the square of an Abelian differential and the parallel line field is a vector field. We will henceforth adopt the notation S to refer to the structure of a quadratic differential. A local coordinate ζ can be chosen in the neighborhood of a zero of order $k \geq 1$ so that quadratic differential is given by $\zeta^k d\zeta^2$. The metric is $|\zeta^k||d\zeta^2|$. The total angle around the zero is $(k+2)\pi$, so we say that S has a cone angle $(k+2)\pi$ singularity. We will denote the set of singularities of S (i.e the zeroes of ϕ) by Z_S. Note that the curvature of the metric is 0, except for δ-function singularities at the points of Z_S.

The number of zeroes counting multiplicity is $4g-4$. Let α be a partition of $4g-4$ (i.e. a representation of $4g-4$ as a sum of positive integers), and let ϵ denote a sign (+ or −). We say a quadratic differential ϕ determines (α, ϵ) if the orders of its zeroes are the partition α, and ϵ is + if ϕ is the square of an Abelian differential; −, if ϕ is not. An obvious necessary, but not sufficient condition for ϕ to be a square of an Abelian differential is to have all zeroes with even order.

Let $QD(g, \alpha, \epsilon)$ denote the space of pairs (M, ϕ) where M is a Riemann surface of genus g, and ϕ is a holomorphic quadratic differential which determines (α, ϵ). This space is called a stratum; this term is justified by the fact that the space of all quadratic differentials on Riemann surfaces of genus g is stratified by the spaces $QD(g, \alpha, \pm)$, as α varies over the partitions of $4g-4$. These spaces have been studied by various authors, e.g. [Ma1], [MaSm], [Ve2], [Ve4]. The strata $QD(g, \alpha, \pm)$ have projection maps to the moduli space R_g of Riemann surfaces of genus g. We note that $QD(g, \alpha, \epsilon)$ are not always connected, but the number of connected components is always finite: see [KZ] for a classification. Let j be a label for a connected component, and let P denote the triple (α, ϵ, j). Thus $QD(P)$ refers to a connected component of a stratum.

One way to visualize a point $S \in QD(P) = QD(\alpha, \epsilon, j)$ where $\alpha = \{k_1, \dots k_n\}$ is as a finite union of polygons glued along parallel (or antiparallel) sides, such that each side is glued to exactly one other, and with special vertices $v_1, \dots v_n$ such that at v_i there is a conical singularity of excess angle πk_i.

There is an $SL(2, \mathbb{R})$ action on $QD(P)$. In each coordinate patch $\zeta \in \mathbb{R}^2$, and for $A \in SL(2, \mathbb{R})$, the action is the linear action $\zeta \to A\zeta$. The fact that the change of coordinates is given by $\zeta \to \pm\zeta + c$ says that this is well-defined. If one visualizes $S \in QD(P)$ as a union of polygons $R_1 \cup \cdots \cup R_n$, then $gS = gR_1 \cup \cdots \cup gR_n$. The orbits of this action (in Teichmuller space) are called Teichmuller disks, and have been studied extensively: see e.g. [Kra], [Ve3], [EG], [GuJ]. The $SL(2, \mathbb{R})$ action preserves an ergodic probability measure μ_0 on $QD(P)$ ([Ma1], [Ve1]). This measure is absolutely continuous with respect to the orbitfold structure and may be interpreted as the Liouville measure. One can also construct other invariant measures: those which come from certain invariant suborbitfolds of $QD(P)$. The authors are not aware of any other constructions.

Cylinders of closed geodesics and saddle connections. A saddle connection of S is a geodesic segment joining two zeroes of S which has no zeroes in its interior. We note that the metric on the surface has non-positive curvature. Every homotopy class contains a geodesic; however this geodesic typically consists of a union of several saddle connections. If we give the saddle connection an orientation then it determines a vector in $\mathbb{R}^2$ since in each coordinate chart it is geodesic with respect to the Euclidean metric. This vector can be thought of as the holonomy along the path given by the oriented saddle connection. The $SL(2,\mathbb{R})$ action on surfaces maps saddle connections to saddle connections. Given a (oriented) saddle connection β we can decompose the associated vector (which we also denote by β) as a sum of its horizontal part h_β and its vertical part v_β. The action on these components is the linear action

$$\begin{bmatrix} h_\beta \\ v_\beta \end{bmatrix} \mapsto A \cdot \begin{bmatrix} h_\beta \\ v_\beta \end{bmatrix}.$$

A closed geodesic that does not pass through any zeroes determines a cylinder of parallel freely homotopic closed geodesics of the same length. The cylinder is metric in that it induces an isometry from a right-angled flat Euclidean cylinder into the surface. Each boundary component of the cylinder consists of one or more parallel saddle connections. A cylinder of closed geodesics is called *regular* if both its boundaries consist of a single saddle connection.

Pointwise asymptotic formulas. One goal of this paper is to prove the following asymptotic formulas for the number of (oriented) closed geodesics and saddle connections on translation surfaces.

Theorem 1.1. *For a quadratic differential S and $T > 0$, let $N_1(S,T)$ denote the number of cylinders of closed geodesics on S such that the norm of the associated vector is at most T. For each P there exists a constant $c_P > 0$ such that for μ_0 almost all $S \in QD(P)$,*

$$N_1(S,T) \sim c_P T^2$$

Theorem 1.2. *For a quadratic differential S and $T > 0$ let $N_2(S,T)$ denote the number of saddle connections on S such that the norm of the associated vector is at most T. Then, for any P there exists a constant $s_P > 0$ such that for μ_0 almost all*

$S \in QD(P)$,

$$N_2(S,T) \sim s_P T^2$$

Versions of the above theorems where convergence a.e. replaced by convergence in L^1 are proved in [Ve5]. We also note some related results all of which will follow from our general result, Theorem 2.1:

Theorem 1.3. *For a quadratic differential S and $T > 0$ let $N_3(S,T)$ denote the number of irregular cylinders of closed geodesics on S such that the norm of the associated vector is at most T. Then, for any P, for μ_0 almost all $S \in QD(P)$,*

$$N_3(S,T) = o(T^2).$$

Note that in Theorem 1.3, μ_0 cannot be replaced by any ergodic $SL(2,\mathbb{R})$ invariant measure μ; for example on a Veech surface, one may have $N_3 \sim cT^2$ for some $c > 0$, see [Ve3].

Some results on the value of the constants c_P and s_P are known, see [EO], [EMZ]. In this paper, for two related problem, we prove the asymptotic formula and also write down the constants explicitly. The expression for these constants is much simpler then the expressions for the constants c_P and s_P in the known cases.

Let $Y_1 \to QD(P)$ the fiber bundle over $QD(P)$ where the fiber over each point in $QD(P)$ is the corresponding surface with flat structure S. Similarly, let $Y_2 \to QD(P)$ be the bundle whose fiber is $S \times S$. Thus, a point in Y_1 is a pair (S,p), where S is a surface with a flat structure, and p is a point in S. A point in Y_2 is a triple (S,p_1,p_2) where S is a surface with a flat structure, and p_1, p_2 are points in S. The fiber bundle Y_i inherits a natural $SL(2,\mathbb{R})$ invariant probability measure μ_i, which is the measure μ_0 on the base $QD(P)$ and is just Euclidean measure induced by the flat structure (product of flat structures) on the fiber.

Theorem 1.4.

(1) *Given a flat structure S and $p \in S \setminus Z_S$ let $N_4(S,p,T)$ be the number of segments of length at most T joining p to a zero of the quadratic differential. Then for μ_1 a.e. $(S,p) \in Y_1$,*

$$N_4(S,p,T) \sim \pi k T^2/2,$$

where $k\pi$ is the sum of the cone angle at the zeroes, Z_S.

(2) *For $(p_1,p_2) \in (S \setminus Z_S) \times (S \setminus Z_S)$ let $N_5(S,p_1,p_2,T)$ denote the number of segments joining p_1 and p_2 of length at most T. Then for μ_2 a.e (S,p_1,p_2) in Y_2,*

$$N_5(S,p_1,p_2,T) \sim \pi T^2.$$

The lack of connection to billiards. We note that there is a well known construction which associates a surface with a flat structure to each rational polygon (see [ZK], [Gut], [KMS]). In particular counting families of periodic trajectories of the billiard in the polygon is equivalent to counting cylinders of closed geodesics on the surface constructed from the polygon. However, the set of surfaces which are constructed from polygons has μ_0 measure 0 in the stratum (see however [EZ]). Thus, Theorem 1.1 cannot be used to obtain information about billiards in rational polygons. This is in contrast with the quadratic upper bound [Ma2] and lower bound [Ma3] theorems, and the asymptotic formulas for Veech surfaces [Ve3]. However, the techniques of this paper, together with Ratner's theorem [Rat] can also be used to obtain asymptotic formulas for billiards in some special polygons which do not give rise to Veech surfaces; see [EMaS]. We also note that we give another proof of the upper bound of [Ma2] in §5.

Spherical averages of K-finite functions. Let $r_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. Let $K = \{r_\theta \;:\; \theta \in [0, 2\pi]\}$ denote the standard $SO(2) \subset SL(2,\mathbb{R})$. Let $\mathcal{M}$ be a space on which $SL(2,\mathbb{R})$ acts. A function f on $\mathcal{M}$ is called K-finite if the linear span of the functions $\{f \circ r_\theta \;:\; \theta \in [0, 2\pi]\}$ is finite dimensional. (Thus f is K-finite if and only if there exists an integer m such that for almost all $x \in SO(2)\backslash\mathcal{M}$, $p(\theta) = f(r_\theta x)$ is a trigonometric polynomial of degree at most m.) The main ergodic tool we use is the following general ergodic theorem:

Theorem 1.5 (Nevo). *Let $a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$, and let K denote the standard $SO(2) \subset SL(2,\mathbb{R})$. Let $\mathcal{M}$ be a space on which $SL(2,\mathbb{R})$ acts and let μ be a ergodic $SL(2,\mathbb{R})$ invariant probability measure on $\mathcal{M}$. Suppose $\phi \in C_0^\infty(\mathbb{R})$ is any nonnegative function, $\epsilon > 0$ and suppose $f \in L^{1+\epsilon}(\mathcal{M}, \mu)$ is a K-finite function. Then, for μ-almost every x in $\mathcal{M}$,*

$$\lim_{\tau\to\infty} \int_{-\infty}^{\infty} \phi(\tau - t) \left(\frac{1}{2\pi} \int_0^{2\pi} f(a_t r_\theta x)\, d\theta \right) dt = \int_{-\infty}^{\infty} \phi(t)\, dt \int_{\mathcal{M}} f(y)\, d\mu(y).$$

Note in this theorem, ϕ is a smoothing function. Basically the assertion of the theorem is a smoothed version of the assertion

$$\lim_{t\to\infty} \frac{1}{2\pi} \int_0^{2\pi} f(a_t r_\theta x)\, d\theta = \int_{\mathcal{M}} f(y)\, d\mu(y).$$

Readers Guide. The paper is organized as follows: in §2 we formulate the general case of the counting problem considered in this paper. We then state our main theorem Theorem 2.1. The constants in that formula are given by the Siegel-Veech formula Theorem 2.2. As a consequence of these two theorems we prove Theorem 1.1, Theorem 1.2, and Theorem 1.3 in §2. The rest of the paper is devoted to the proof of Theorem 2.1. In §3, following [Ve5], we relate the counting problem to consideration of certain integrals over orbits of $SL(2, \mathbb{R})$. In §4 we use Nevo's Theorem 1.5 to obtain asymptotic results almost everywhere for the integrals in §3.

The results of §3-§4 are conditional on the assumptions (A), (B) and (C_μ) formulated in §2. For the key case of counting saddle connections, (A) is immediate, but (B) and (C_μ) require certain a priori upper bounds. These upper bounds are verified in §5, based on two results of independent interest, Theorem 5.1 and Theorem 5.2. We also show in §5 that Theorem 5.1 and Theorem 5.2 imply the quadratic upper bound result of [Ma2].

In §6-§8, which are more technical then the rest of the paper, we prove Theorem 5.1 and Theorem 5.2. In §9, we compute the constants in the asymptotic formulas of Theorem 1.4.

Acknowledgements. We would like to thank W. Veech, who besides inspiring this project also corrected an error in the initial version of this paper. We would also like to thank the referee for many valuable comments.

2 The counting problems

The space $\mathcal{M}$ and the measure μ. It is convenient to axiomatize the properties we will need. Let $\mathcal{M}$ denote a space on which $SL(2, \mathbb{R})$ acts. For most of the paper, $\mathcal{M}$ will denote $QD(P)$, but in §9, $\mathcal{M}$ will denote the fiber bundles Y_i. Let μ be a probability measure on $\mathcal{M}$ invariant under the $SL(2, \mathbb{R})$ action. Let $B(T, 0)$ denote the Euclidean ball of center $(0, 0)$ and radius T, and let $|\cdot|$ denote the cardinality of a finite set.

In the rest of the paper $V(S)$ will denote a subset of $\mathbb{R}^2 - \{(0,0)\}$ with multiplicity; that is, a set of vectors assigned positive weights. They typically will be integer weights, but in some applications, they need not be. Suppose to every point $S \in \mathcal{M}$ we associate such a set so that the following hold:

(A) V varies linearly under $SL(2, \mathbb{R})$ action. More precisely, for $g \in SL(2, \mathbb{R})$, $V(gS) = gV(S)$.

(B) For any $S \in \mathcal{M}$, there exists a constant $c = c(S) > 0$, such that for any $T > 1$ the cardinality of $V(S) \cap B(T, 0)$ is at most cT^2. The constant c can be chosen uniformly as S varies over any compact subset of $\mathcal{M}$.

(C_μ) For each $T > 0$ let $N_V(T,S)$ denote the cardinality of $V(S) \cap B(T,0)$. Then there exist some $R > 0$ and $\epsilon > 0$ such that as a function of $S \in \mathcal{M}$, $N_V(R,S) \in L^{1+\epsilon}(\mathcal{M},\mu)$.

We note that if $V(\cdot)$ denotes the holonomy vectors associated to oriented closed geodesics or to saddle connections, then assumption (A) is clearly satisfied. The fact that assumption (B) is satisfied is a theorem of H. Masur [Ma2] which we will reprove as Theorem 5.4 in §5. We will also verify that condition (C_μ) is also satisfied for every invariant probability measure μ. These statements together with the following theorem will allow us to prove our asymptotic formulas.

Theorem 2.1. *Let μ be an ergodic invariant probability measure on $\mathcal{M}$. Suppose $V(\cdot)$ satisfies assumptions* (A), (B) *and* (C_μ)*. Then there exists a constant $c_{V,\mu} = c(V,\mu,\mathcal{M})$ such that for μ-almost all $S \in \mathcal{M}$,*

$$\lim_{T\to\infty} \frac{1}{T^2} N_V(T,S) = \pi c_{V,\mu}$$

The constant $c_{V,\mu}$ is the constant in a "Siegel-Veech formula"; see Theorem 2.2 below.

We note that we do not know the value of $c_{V,\mu}$ in many cases. For some results, see [EO], [EMZ].

Remark. Conditions (A), (B) and (C_μ) are not sufficient to ensure that $c_{V,\mu} > 0$; indeed, Theorem 1.3 is an example when $c_{V,\mu} = 0$. See the statement of Theorem 2.2 for a necessary and sufficient condition to ensure that $c_{V,\mu} > 0$ (and thus that the asymptotic formula is quadratic).

2.1 The Siegel-Veech formula

In this subsection we recall some constructions and results from [Ve5]. For completeness, and since our assumptions differ somewhat from those of [Ve5], we reproduce the proofs below.

The functions f and $\hat{f}$. For $f \in C_0^\infty(\mathbb{R}^2)$, let $\hat{f} : QD(P) \to \mathbb{R}$ be defined by

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

The terms in the sum are counted with multiplicity. By assumption (B), the sum is finite, and also $\hat{f}$ is a bounded function on compact subsets of $QD(P)$

For technical reasons one needs f to be a smooth function, but if one takes f to be the characteristic function of a ball of radius T centered at the origin, then $\hat{f}(S)$ will count the number of elements of $V(S)$ which lie in a ball of radius T.

A generalization of the following theorem is proved in [Ve5]:

Theorem 2.2 (Siegel-Veech formula). *Suppose V satisfies assumptions (A), (B), and (C_μ). Let μ be any ergodic invariant probability measure on $\mathcal{M}$. Then there exists a constant $c_{V,\mu}$ such that for any $f \in C_0^\infty(\mathbb{R}^2)$,*

$$\int_{\mathcal{M}} \hat{f}(S)\, d\mu(S) = c_{V,\mu} \int_{\mathbb{R}^2} f(x,y)\, dxdy.$$

The constant $c_{V,\mu}$ is always non-negative; it is zero if and only if $\hat{f} = 0$ μ-almost everywhere.

Proof. We present the idea of the proof from [Ve5] for completeness. By assumption (C_μ) and since μ is a probability measure, $\hat{f} \in L^{1+\epsilon}(\mathcal{M},\mu) \subset L^1(\mathcal{M},\mu)$. Consider the linear functional

$$\Phi(f) = \int_{\mathcal{M}} \hat{f}(S)\, d\mu(S). \tag{1}$$

Since μ is invariant under the $SL(2,\mathbb{R})$ action on $\mathcal{M}$, Φ is invariant under the $SL(2,\mathbb{R})$ action on continuous functions on $\mathbb{R}^2$. Thus it follows from the general form of such functionals that for all $f \in C_0^\infty(\mathbb{R}^2)$,

$$\Phi(f) = af(0,0) + b\int_{\mathbb{R}^2} f. \tag{2}$$

where a and b are independent of f. Pick a function $\phi \in C_0^\infty(\mathbb{R}^2)$ so that $\phi(0,0) = 1$, and $\phi \leq 1$, and let $f_j(x,y) = \phi(jx,jy)f(x,y)$. Then, $f_j(0,0) = f(0,0)$, and $\int_{\mathbb{R}^2} f_j \to 0$ by dominated convergence. Note that for every fixed $S \in \mathcal{M}$, since $\hat{f}(S) < \infty$ and $(0,0) \notin V(S)$, $\lim_{j\to\infty} \hat{f}_j(S) = 0$. Hence, using (1) and the fact that $\hat{f} \in L^1$, we get (by dominated convergence) $\Phi(f_j) \to 0$. Substituting f_j instead of f into (2) and taking the limit $j \to \infty$, we get $a = 0$. Now set $c_{V,\mu} = b$. The final assertion is immediate. $\square$

Proof of Theorem 1.2 This is immediate from Theorem 2.1 and Theorem 2.2 since every surface has a saddle connection of bounded length so that for some f, $\hat{f}(S) \neq 0$ for all S. $\square$

Proof of Theorem 1.3 Now we need to show that $\hat{f}(S) = 0$ for μ_0 a.e. S. In other words we have to show that a generic surface does not have an irregular cylinder. We first consider the case of differentials that are squares of Abelian differentials. Note that all the periodic trajectories comprising a cylinder are all homotopic to the same curve on the surface; we call this curve the "core curve" of the cylinder. The holonomy along the core curve is the height of the cylinder, and is thus non-zero. Since the holonomy along any curve γ is equal to the integral of the Abelian differential along γ, this implies that the core curve of the cylinder is not homologous to 0. Thus the saddle connections on the boundary of an irregular cylinder are linearly independent in $H_1(S, Z_S)$ (the homology relative to the zeroes). Local coordinates on the stratum are given by the holonomy vectors of a basis for $H_1(S, Z_S)$. Thus the fact that these linearly independent holonomy vectors are real multiples of each other means that the set of flat structures defining an irregular cylinder near the given one is a positive codimension set, and hence has μ_0 measure 0. The result follows by taking the union over the at most countable set of homotopy classes.

If the stratum is not defined by Abelian differentials, the proof is the same if the cylinder is not homologous to 0. If it is, we pass to a double cover S', ramified over the zeroes. The quadratic differential lifts to an Abelian differential and the cylinder lifts to a pair of cylinders, each of which is not homologous to 0; although they are homologous to each other, and are interchanged by the involution on the double cover. Now local coordinates on the original space $QD(P)$ are given by the holonomy vectors of a basis for the part of the relative homology group $H_1(S', Z_{S'})$ which is odd with respect to the involution. Hence, again the set of surfaces with an irregular metric cylinder are a positive codimension subset. $\square$

Proof of Theorem 1.1 We again apply Theorem 2.1 and Theorem 2.2. We need to show that the constant c_P given in Theorem 2.2 is positive. Let S_0 be any surface with a metric cylinder. Such surfaces are easy to construct. For example, one can take a right angled flat Euclidean cylinder and make certain isometric gluings of the sides on the top and bottom to form a surface. Such a cylinder is not regular but it is immediate that every nearby surface also has a cylinder whose core curve is in the same homotopy class. Now by Theorem 1.3 the set of nearby surfaces such that the cylinder is irregular has positive codimension. Thus there must be a nearby surface with a regular cylinder and that surface has an open neighborhood of surfaces with a regular cylinder. Thus for an appropriate f, the function $\hat{f}$ is positive on an open set and we are done. $\square$

3 Counting and integration over $\mathcal{M}$

Let $a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$, and for $\theta \in [0, 2\pi]$, let $r_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$. Let $\mathbb{R}^2_+$ denote the upper half-plane.

Definition 3.1 (Generic Points). Suppose μ is an ergodic invariant probability measure on $\mathcal{M}$. A point $S \in \mathcal{M}$ is called *μ-generic* if for any $f \in C_0^\infty(\mathbb{R}^2_+)$ and all nonnegative $\phi \in C_0^\infty(\mathbb{R})$,

$$\lim_{\tau\to\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(\tau - t) \left(\int_0^{2\pi} \hat{f}(a_t r_\theta S)\, d\theta \right) dt = \int_{-\infty}^{\infty} \phi(t)\, dt \int_{\mathcal{M}} \hat{f}(S_1)\, d\mu(S_1). \tag{3}$$

Note that the integral on the left is smoothed version of $\frac{1}{2\pi}\int_0^{2\pi} \hat{f}(a_t r_\theta S)\, d\theta$. The proof of Theorem 2.1 breaks up naturally into two parts, which we state as Proposition 3.2 and Proposition 3.3 below.

Proposition 3.2. *Suppose μ is an ergodic invariant probability measure on $\mathcal{M}$, $V(\cdot)$ satisfies conditions* (A),(B) *and* (C_μ) *and S is μ-generic. Then,*

$$\lim_{T\to\infty} \frac{1}{T^2} N_V(T, S) = \pi c_{V,\mu},$$

where $c_{V,\mu}$ is as in Theorem 2.2.

Proposition 3.3. *Suppose μ is a ergodic invariant probability measure on $\mathcal{M}$, and $V(\cdot)$ satisfies conditions* (A) (B) *and* (C_μ). *Then, for μ-almost all $S \in \mathcal{M}$, S is μ-generic.*

Proposition 3.3 will be derived from Theorem 1.5. In this section, we will prove Proposition 3.2. We will also prove an elementary result, Proposition 3.5, which will be used in the proofs of the upper bounds (B).

The results in this section are essentially contained in §5 of [Ve5] (which were in turn inspired in part by [EMM] following an analogy noted in [Ve5]). In fact, our treatment, in this section and the next, is an almost verbatim translation of §3 of [EMM].

Lemma 3.4. *(cf. [EMM, Lemma 3.6]) For $f \in C_0^\infty(\mathbb{R}^2_+)$, let*

$$J_f(y) = \frac{1}{2\pi y} \int_{-\infty}^{\infty} f(x, y)\, dx$$

Then for every $\epsilon > 0$ there exists $T_0 > 0$ such that for every t with $e^t > T_0$, and every $v \in \mathbb{R}^2$ with $\|v\| > T_0$,

$$\left| \frac{e^{2t}}{2\pi} \int_0^{2\pi} f(a_t r_\theta v)\, d\theta - J_f(\|v\| e^{-t}) \right| < \epsilon. \tag{4}$$

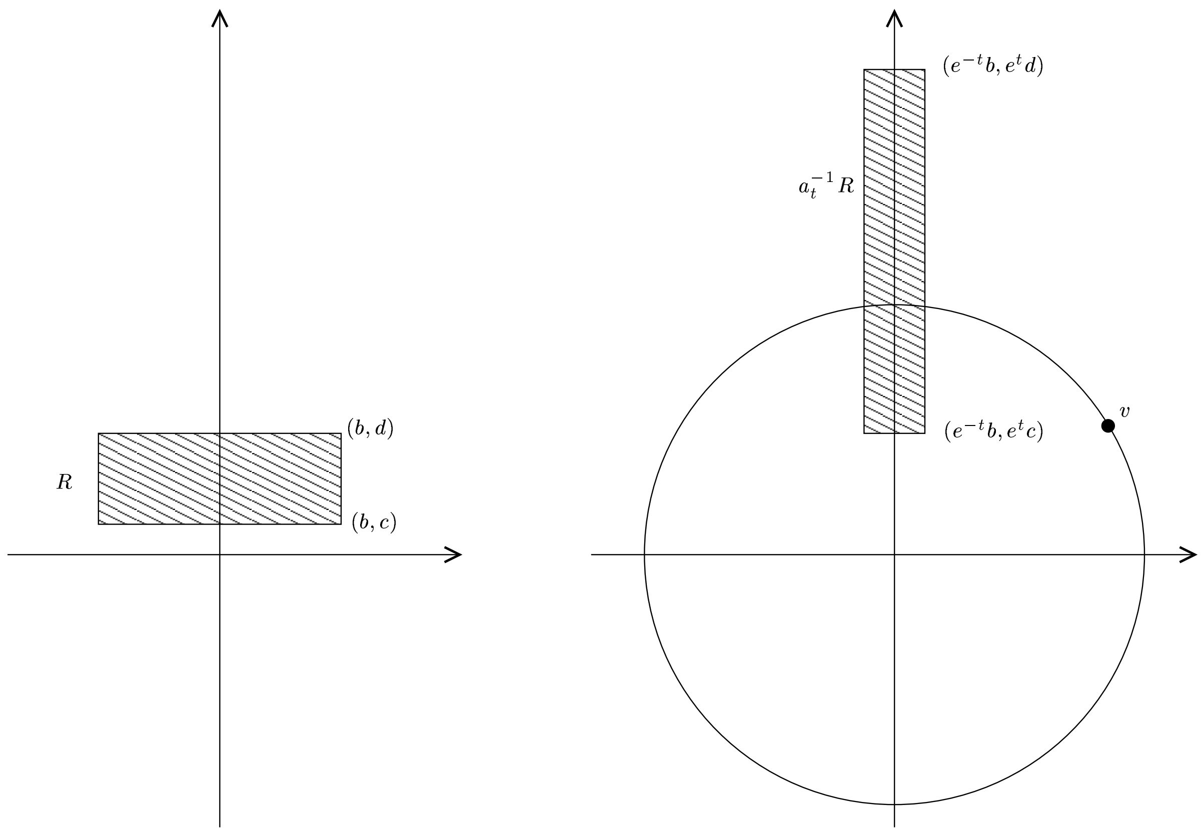

Figure 1. Lemma 3.4.

Heuristic argument for Lemma 3.4. (See figure 1) Suppose f is the characteristic function of the rectangle $R = [-b, b] \times [c, d]$ (where $b > 0$, and $0 < c < d$). Then, by definition, $J_f(\|v\|e^{-t}) = 2b/(2\pi\|v\|e^{-t})$ if $ce^t \leq \|v\| \leq de^t$ and is zero otherwise. Also, $a_t^{-1}R$ is the rectangle $[-be^{-t}, be^{-t}] \times [e^t c, e^t d]$. Since the integrand in (4) is zero unless $r_\theta v \in a_t^{-1} R$, $\frac{1}{2\pi}\int_0^{2\pi} f(a_t r_\theta v)\, d\theta$ is $\frac{1}{2\pi\|v\|}$ times the length of the intersection of the circle of radius $\|v\|$ with the rectangle $a_t^{-1}R$ (the factor of $1/(2\pi\|v\|)$ arises because the length of an arc of angle θ on a circle of radius $\|v\|$ is $\|v\|\theta$). Now, the intersection of the circle with $a_t^{-1}R$ is empty if $\|v\|$ is outside of $[ce^t, de^t]$ and is essentially the horizontal segment $[-be^{-t}, be^{-t}] \times \{\|v\|\}$ otherwise. Thus, $\frac{1}{2\pi}\int_0^{2\pi} f(a_t r_\theta v)\, d\theta \approx \frac{2}{2\pi}(be^{-t})/\|v\| = e^{-2t} J_f(\|v\|e^{-t})$.

Proof of Lemma 3.4. Clearly, the left hand side of (4) depends on v only via the norm $\|v\|$. Thus, we may assume that $v = (0, \|v\|)$. Then,

$$\frac{1}{2\pi}\int_0^{2\pi} f(a_t r_\theta v)\, d\theta = \frac{1}{2\pi}\int_0^{2\pi} f(e^t\|v\|\sin\theta, e^{-t}\|v\|\cos\theta)\, d\theta \tag{5}$$

Since t and $\|v\|$ are large, and f has compact support, the integrand is zero unless $\sin\theta$ is near 0. Hence θ must be near 0 or π. Since f is supported on the upper half plane $\mathbb{R}^2_+$, θ must be near 0. From the form of (5) it is clear that the integrand is 0 unless $\|v\|e^{-t}$ is bounded between c and c^{-1}, and unless $\theta = O(e^{-t}\|v\|^{-1}) = O(e^{-2t})$, where the $c > 1$ and the implied constants depend only on the support of f.

Let $\delta(\rho) = \sup\{|f(x, y_1) - f(x, y_2)| \text{ such that } |y_1 - y_2| < \rho\}$. By the uniform continuity of f, $\delta(\rho) \to 0$ as $\rho \to 0$.

We now make the change of variable $u = e^t\|v\|\sin\theta$ in (5); the integral becomes:

$$\begin{aligned} &\tfrac{e^{-t}}{2\pi\|v\|}\textstyle\int_{-\infty}^{\infty} f\left(u, e^{-t}\|v\|\sqrt{1-(e^{-t}u/\|v\|)^2}\right)(1-(e^{-t}u/\|v\|)^2)^{-1/2}\, du = \\ &\qquad \tfrac{e^{-t}}{2\pi\|v\|}\textstyle\int_{-\infty}^{\infty} f(u, e^{-t}\|v\|)\, du + E = e^{-2t}J_f(e^{-t}\|v\|) + E, \end{aligned} \tag{6}$$

where the error term E is bounded by:

$$|E| = O\left((e^{-t}/\|v\|)\delta((e^{-t}/\|v\|)^2)\right) + O\left((e^{-t}/\|v\|)\|f\|_\infty(e^{-t}/\|v\|)^2)\right) = O(e^{-2t}\delta(c^2e^{-4t})).$$

After multiplying both sides by e^{2t}, the lemma follows. □

As an illustration of the basic technique, we first prove the following elementary result:

Proposition 3.5. *Suppose $V(\cdot)$ satisfies condition* (A) *and the intersection of $V(\cdot)$ with any compact set is finite. Then, there exists an absolute constant $c < \infty$ such that for any $T > 0$ and any $S \in \mathcal{M}$,*

$$N_V(2T, S) - N_V(T, S) \le cT^2 \int_0^{2\pi} N_V(4, a_t r_\theta S)\, d\theta$$

where $t = \log T$.

Proof of Proposition 3.5. Let $\mathbb{R}_+$ denote the positive reals. Note that any function in $C_0^\infty(\mathbb{R}_+)$ can be written as J_f for some $f \in C_0^\infty(\mathbb{R}_+^2)$. Let χ denote the characteristic function of the interval $[1, 2]$, and let $\epsilon > 0$ be arbitrary. Choose $f \in C_0^\infty(\mathbb{R}_+^2)$ so that $J_f \geq (1 + \epsilon)\chi$, and the support of f is contained in $[-2.5, 2.5] \times [0, 2.5]$. Let T_0 be as in Lemma 3.4, and suppose $e^t > T_0$. By Lemma 3.4, if $e^t \leq \|v\| \leq 2e^t$, then

$$1 \leq J_f(e^{-t}\|v\|) - \epsilon \leq \frac{e^{2t}}{2\pi} \int_0^{2\pi} f(a_t r_\theta v)\, d\theta. \tag{7}$$

Let S be a point in $\mathcal{M}$. Summing the outermost inequality in (7) over those v in $V(S)$ that belong to the annulus $e^t \leq \|v\| \leq 2e^t$, we get

$$N_V(2e^t, S) - N_V(e^t, S) = \sum_{\substack{v \in V(S) \\ e^t \leq \|v\| \leq 2e^t}} 1 \leq \frac{e^{2t}}{2\pi} \sum_{\substack{v \in V(S) \\ e^t \leq \|v\| \leq 2e^t}} \int_0^{2\pi} f(a_t r_\theta v)\, d\theta \leq \tag{8}$$

$$\leq \frac{e^{2t}}{2\pi} \sum_{v \in V(S)} \int_0^{2\pi} f(a_t r_\theta v)\, d\theta = \frac{e^{2t}}{2\pi} \int_0^{2\pi} \hat{f}(a_t r_\theta S)\, d\theta \tag{9}$$

Set $c = \sup f$. The function f is bounded by c times the characteristic function of the rectangle $[-2.5, 2.5] \times [0.2.5]$. This rectangle is a subset of the disk $B(4, 0)$ of radius 4 centered at the origin, so f is majorized by cg where g is the characteristic function of the disk. Hence $\hat{f}(S)$ is majorized by $c\hat{g}(S)$. But $\hat{g}(S) = N_V(4, S)$ and so (8) becomes

$$N_V(2T, S) - N_V(T, S) < \frac{cT^2}{2\pi} \int_0^{2\pi} N_V(4, a_t r_\theta S)\, d\theta$$

where we set $T = e^t$. This gives the Lemma (for a new constant c). □

Proposition 3.5 will be used in §5 and a variation, Lemma 8.1 will be used in §8.

Proof of Proposition 3.2. Let χ be the characteristic function of the interval $[1, 2]$, and let $\epsilon > 0$ be arbitrary. Let c be as in condition (B) for the generic surface S. Choose $h_+ \in C_0^\infty(\mathbb{R}_+)$ $h_- \in C_0^\infty(\mathbb{R}_+)$ so that $h_- \leq \chi \leq h_+$ and $\int_0^\infty y(h_+(y) - h_-(y))\, dy \leq \epsilon/c$. Choose $f_\pm \in C_0^\infty(\mathbb{R}_+^2)$ so that $J_{f_\pm} = h_\pm$. Choose $\phi \in C_0^\infty(\mathbb{R})$ be non-negative with $\int_{-\infty}^\infty \phi = 1$, and with the support of ϕ so close to 0 that $J_{\phi * f_-} \leq \chi \leq J_{\phi * f_+}$, where we use the notation $(\phi * f)(v) = \int_{-\infty}^\infty \phi(t) f(a_{-t}v) dt$. Then, for any $v \in \mathbb{R}^2$ and any $\tau > 0$,

$$J_{\phi * f_-}(\|v\| e^{-\tau}) \leq \chi(\|v\| e^{-\tau}) \leq J_{\phi * f_+}(\|v\| e^{-\tau}).$$

Using Lemma 3.4, for sufficiently large τ we get

$$\frac{e^{2\tau}}{2\pi}\int_0^{2\pi}(\phi * f_-)(a_\tau r_\theta v)\,d\theta - \epsilon/c \leq \chi(\|v\|e^{-\tau}) \leq \frac{e^{2\tau}}{2\pi}\int_0^{2\pi}(\phi * f_+)(a_\tau r_\theta v)\,d\theta + \epsilon/c. \tag{10}$$

Summing over $v \in V(S)$ and noting that by condition (B), the number of terms with $\chi(\|v\|e^{-\tau}) \neq 0$ is at most $ce^{2\tau}$, we get

$$\begin{aligned} \frac{e^{2\tau}}{2\pi}\int_0^{2\pi}(\phi * \hat{f}_-)(a_\tau r_\theta S)\,d\theta - \epsilon e^{2\tau} \leq N_V(2e^\tau, S) - N_V(e^\tau, S) \leq \\ \leq \frac{e^{2\tau}}{2\pi}\int_0^{2\pi}(\phi * \hat{f}_+)(a_\tau r_\theta S)\,d\theta + \epsilon e^{2\tau} \end{aligned} \tag{11}$$

Since S is μ-generic and $\int_{-\infty}^{\infty}\phi = 1$,

$$\lim_{\tau\to\infty}\frac{1}{2\pi}\int_0^{2\pi}(\phi * \hat{f}_\pm)(a_\tau r_\theta S)\,d\theta = \int_{\mathcal{M}}\hat{f}_\pm\,d\mu = c_{V,\mu}\int_{\mathbb{R}^2} f_\pm, \tag{12}$$

where in the last identity we used the Siegel-Veech formula Theorem 2.2. But, by construction,

$$\int_{\mathbb{R}^2} f_\pm(x,y)\,dx\,dy = 2\pi\int_0^\infty yJ_{f_\pm}(y)\,dy = 2\pi(3/2 + E) \tag{13}$$

where $|E| < \epsilon/c$. Hence, after dividing (11) by $e^{2\tau}$, taking the limit as $\tau \to \infty$, and using (12) and (13) we get

$$\begin{aligned} 3\pi c_{V,\mu} - (2\pi c_{V,\mu}/c + 1)\epsilon &\leq \liminf_{\tau\to\infty} e^{-2\tau}(N_V(2e^\tau, S) - N_V(e^\tau, S)) \\ &\leq \limsup_{\tau\to\infty} e^{-2\tau}(N_V(2e^\tau, S) - N_V(e^\tau, S)) \\ &\leq 3\pi c_{V,\mu} + (2\pi c_{V,\mu}/c + 1)\epsilon. \end{aligned}$$

Since $\epsilon > 0$ is arbitrary, this implies

$$\lim_{T\to\infty}\frac{1}{T^2}\left(N_V(2T, S) - N_V(T, S)\right) = 3\pi c_{V,\mu}$$

where we let $T = e^\tau$. From this, it is easy to conclude that

$$\lim_{T\to\infty}\frac{1}{T^2}N_V(T, S) = \pi c_{V,\mu}$$

as required. $\square$

4 Approximation by K-finite functions, Proof of Proposition 3.3.

Lemma 4.1. *Let $\mathcal{C}$ be a compact set in $\mathbb{R}^2 - \{0\}$. There exists a countable family of (complex valued) functions $\mathcal{F} \in C_0^\infty(\mathbb{R}^2 - \{0\})$ with the following properties:*

(1) *For each $f \in \mathcal{F}$ there exists $m \in \mathbb{Z}$ such that for all $v \in \mathbb{R}^2$ and all $\theta \in [0, 2\pi]$, $f(r_\theta v) = e^{im\theta} f(v)$.*

(2) *For every $q \in C_0(\mathbb{R}^2 - \{0\})$ such that the support of q is contained in $\mathcal{C}$ and for every $\epsilon > 0$ there exist f_1 and f_2 finite linear combinations of functions in $\mathcal{F}$ such that the f_i are real-valued,*

$$f_1 \leq q \leq f_2$$

and

$$\int_{\mathbb{R}^2} (f_2 - f_1) < \epsilon \tag{14}$$

Also there exists a countable family Φ of functions in $C_0^\infty(\mathbb{R})$ such that for every $\phi \in C_0^\infty(\mathbb{R})$ there exist ϕ_1, ϕ_2 finite linear combinations of functions in Φ such that $\phi_1 \leq \phi \leq \phi_2$ and $\int_{\mathbb{R}} (\phi_2 - \phi_1) < \epsilon$.

Proof. We may fix a radially symmetric function $h \in C_0^\infty(\mathbb{R}^2 - \{0\})$ such that $h = 1$ on $\mathcal{C}$. Let $\mathcal{F} = \{h f_{m,n}(r, \theta) \ : \ m \in \mathbb{N}, n \in \mathbb{Z}\}$, where $f_{m,n}(r, \theta) = r^m e^{in\theta}$. Then $\mathcal{F}$ satisfies the conditions of the lemma by e.g. the Stone-Weierstrass theorem. For Φ one may use a similar construction. $\square$

Proof of Proposition 3.3. We note that if f is K-finite, then so is $\hat{f}$. By Theorem 1.5 there exists a set $E \subset \mathcal{M}$ with $\mu(E) = 1$ such that for any $S \in E$ any $f \in \mathcal{F}$, and any $\phi \in \Phi$, (3) holds. Let $\epsilon > 0$ be arbitrary, and let ϕ_1, ϕ_2, f_1, f_2 be as in Lemma 4.1. Then $\hat{f}_1 \leq \hat{f} \leq \hat{f}_2$, hence

$$\begin{aligned} \int_{-\infty}^{\infty} \phi_1(\tau - t) \left(\int_0^{2\pi} \hat{f}_1(a_t r_\theta S) \, d\theta \right) dt &\leq \int_{-\infty}^{\infty} \phi(\tau - t) \left(\int_0^{2\pi} \hat{f}(a_t r_\theta S) \, d\theta \right) dt \\ &\leq \int_{-\infty}^{\infty} \phi_2(\tau - t) \left(\int_0^{2\pi} \hat{f}_2(a_t r_\theta S) \, d\theta \right) dt \end{aligned} \tag{15}$$

Suppose $S \in E$. Since f_i and ϕ_i are finite linear combinations of functions in $\mathcal{F}$ and Φ,

$$\lim_{\tau\to\infty} \int_{-\infty}^{\infty} \phi_i(\tau - t) \left(\int_0^{2\pi} \hat{f}_i(a_t r_\theta S)\, d\theta \right) dt = \int_{\mathbb{R}} \phi_i \int_{\mathcal{M}} \hat{f}_i \, d\mu = c_{V,\mu} \int_{\mathbb{R}} \phi_i \int_{\mathbb{R}^2} f_i \quad (16)$$

where in the last equality we used the Siegel-Veech formula. Since $\epsilon > 0$ is arbitrary, (3) follows from (15), (16) and (14). □

5 Upper bounds

5.1 Statement of Results

Letting $V(\cdot)$ denote the set of vectors (with multiplicity) associated to saddle connections, then in this section we will prove that conditions (B) and (C_μ) hold for every μ. This allows us to apply Theorem 2.1. Conditions (B) and (C_μ) will follow from the following two theorems:

Theorem 5.1. *Suppose $V(\cdot)$ denotes the set of vectors associated to saddle connections. Then, for any stratum $QD(P)$ there exists $\kappa > 0$ such that*

(a) *For any $\delta > 0$ any $R < \kappa$ and any $S \in QD(P)$,*

$$|V(S) \cap B(R, 0)| \le C \left(\frac{R}{\ell(S)} \right)^{1+\delta}, \quad (17)$$

where $C = C(\delta, P)$ and $\ell(S)$ denotes the length of the shortest saddle connection in S.

(b) *For any $\delta > 0$ any $R < \infty$ and any $S \in QD(P)$, (17) holds with $C = C(\delta, R, P)$.*

Theorem 5.2. *For any $S \in QD(P)$ and any $\beta < 2$,*

$$\sup_{t>0} \int_0^{2\pi} \ell(a_t r_\theta S)^{-\beta}\, d\theta < \infty$$

and the upper bound can be chosen uniformly as S varies over compact subsets of $QD(P)$.

Corollary 5.3. *Let $K = \{r_\theta \ : \ 0 \leq \theta \leq 2\pi\}$, so that $K \cong SO(2)$. Let S any flat structure and let ν denote the K-invariant measure supported on the set KS. Then the set of measures $a_t\nu$ is compact (in the weak-* topology),*

The proofs of Theorem 5.1 and Theorem 5.2, which combines the ideas of [Ma2] with the ideas of G. Margulis, will be given in §6-§8. In the rest of this section we will show how Theorem 5.1 and Theorem 5.2 can be used to derive condition (B) for the saddle connections) and also to verify that condition (C_μ) holds for the saddle connections for every invariant measure μ.

We begin, however by showing how Corollary 5.3 follows from Theorem 5.2.

Proof of Corollary 5.3 For any $\epsilon > 0$, let $QD_\epsilon(P) = \{S \in QD(P) : \ell(S) < \epsilon\}$. Then $QD(P) \setminus QD_\epsilon(P)$ is compact. Taking $\beta = 1$ in Theorem 5.2, there is a constant M such that for all ϵ and t, $a_t\nu(QD_\epsilon(P)) \leq M\epsilon$. Since the set of probability measures on a compact space is weak-* compact, we are done. □

Theorem 5.4. *Suppose $V(S)$ denotes the set associated to saddle connections on S. Then, $V(\cdot)$ satisfies condition (B), i.e. for any $S \in QD(P)$, there exists a constant $c = c(S)$ such that for all $T > 0$,*

$$N_V(T, S) \leq cT^2$$

The constant c can be chosen uniformly as S varies over compact subsets of $QD(P)$.

This theorem was originally proved in [Ma2]. Of course Theorem 5.4 implies the analogous result for the number of cylinders of closed geodesics since the number of such cylinders on a surface is bounded trivially by the number of saddle connections.

Proof of Theorem 5.4. By Proposition 3.5, Theorem 5.1 and Theorem 5.2, there exist $R > 0$ and $\beta < 2$ such that we have

$$N_V(2T, S) - N_V(T, S) \leq cT^2 \int_0^{2\pi} N_V(R, a_t r_\theta S)\, d\theta \leq c'T^2 \int_0^{2\pi} \frac{d\theta}{\ell(a_t r_\theta S)^\beta} \leq c''T^2$$

Now the theorem follows from an easy geometric series argument. □

Lemma 5.5. *(cf. [Ve5, Corollary 2.8]) Let μ be any ergodic invariant probability measure on $QD(P)$. Then for any $1 \leq \beta < 2$, $\ell(\cdot)^{-1} \in L^\beta(QD(P), \mu)$.*

Proof. If $\mu = \mu_0$, it is possible to prove this directly. For general μ, we derive this by an indirect argument from Theorem 5.2 and the ergodic Theorem 1.5. In fact, we will only need the K-fixed version of Theorem 1.5, see [NS]. It suffices to prove that for any $\beta < 2$, $\ell^{-\beta} \in L^1(QD(P), \mu)$. Assume this is false.

Let χ be the characteristic function of $[1, \infty]$ and for $j \in \mathbb{N}$ let $h_j(S) = \ell(S)^{-\beta}\chi(j\ell(S))$. Then h_j are bounded with compact support, since the set of quadratic differentials whose shortest saddle is bounded below is compact. As $j \to \infty$, $h_j \uparrow \ell^{-\beta}$. Then by monotone convergence, $\int_{QD(P)} h_j(S)\, d\mu(S) \to \infty$.

Fix a positive $\phi \in C_0^\infty(\mathbb{R})$, and consider the integrals

$$I_j(\tau) = \int_{-\infty}^{\infty} \phi(\tau - t) \left(\int_0^{2\pi} h_j(a_t r_\theta S)\, d\theta \right) dt.$$

Since h_j is K-fixed (i.e. $h_j(r_\theta S) = h_j(S)$ for all $\theta \in [0, 2\pi]$ and all S), by Theorem 1.5, for almost all S and all j, $\lim_{\tau\to\infty} I_j(\tau) = \int_{QD(P)} h_j d\mu(S) \int_{-\infty}^{\infty} \phi$. Since for all j, $\ell^{-\beta} \geq h_j$ we get, for all j,

$$\liminf_{\tau\to\infty} \int_{-\infty}^{\infty} \phi(\tau - t) \left(\int_0^{2\pi} \ell(a_t r_\theta S)^{-\beta}\, d\theta \right) dt \geq \lim_{\tau\to\infty} I_j(\tau) = \int_{QD(P)} h_j d\mu(S) \int_{-\infty}^{\infty} \phi.$$

Taking $j \to \infty$ we get a contradiction since the right-hand side tends to infinity, and the left hand side remains bounded by Theorem 5.2. $\square$

Proof of condition (C_μ) for the saddle connections. This follows immediately from Theorem 5.1 and Lemma 5.5.

$\square$

6 Combinatorial Constructions

Saddle connections, p_0, complexes and domains. We say that two saddle connections are *disjoint* if the only common points (if any) are zeroes. A union of disjoint saddle connections Γ is called a system. By $|\Gamma|$ we mean the maximum length of any member of Γ. By a *complex* K we mean a closed subset of S, whose boundary, denoted ∂K, is a system of saddle connections. Here ∂K refers to the subset of K consisting of those points such that every neighborhood in S of the point intersects the complement of K. Thus if K consists of a single saddle connection then $\partial K = K$. Note that a complex need not be connected and is not assumed to be triangulated. However we will assume that if ∂K contains three saddle connections that bound a triangle, then the triangle is included in K.

We now show that there is a uniform bound p_0 to the number of disjoint saddle connections. Indeed, given any set of disjoint saddle connections, we may further add disjoint saddle connections so that all complimentary regions are discs and then further subdivide the discs into triangles. This gives a triangulation of the surface with edges which are saddle connections and whose vertices are the m zeroes. Let e the number of edges and t the number of triangles. Then Euler's formula gives $m - e + t = \chi$, the Euler characteristic of the surface. Each edge is contained in 2 triangles and each triangle has 3 edges; thus $3t = 2e$. This implies that the maximum number p_0 of disjoint saddle connections is at most $3(m - \chi)$. By a *domain* Y we will mean an open set in S with boundary a system Γ of saddle connections. A domain is not considered to be triangulated.

Area, complexity, and domain-complexity. The area of a complex K is denoted $||K||$. Note that if K is a complex, then $K \setminus \partial K$ is a (possibly empty) domain. The *complexity* of a complex is the number of saddle connections needed to triangulate the complex, including the boundary. The complexity of a domain is the number of saddle connections needed to triangulate it, not counting the number on the boundary.

Cross Products. Given a saddle connection β considered as a vector we may decompose β as a sum of its horizontal component β_h and its vertical component β_v. Given ω and β saddle connections we can form the length of the vector product

$$|\omega \times \beta| = |\beta_h \omega_v - \beta_v \omega_h|.$$

This number is invariant under the $SL(2, \mathbb{R})$ actions on flat structures and saddle connections. For example if S has area 1 and ω and β are two sides of a triangle, then since the triangle has area at most 1, and the length of the vector product measures the area of the corresponding parallelogram, we must have $|\omega \times \beta| \leq 2$.

Proposition 6.1 (Cross Products). *Suppose $\gamma_1, \gamma_2, \beta$ are saddle connections. Then*

(1) $|\gamma_1 \times \gamma_2| \leq \frac{|\gamma_1 \times \beta||\gamma_2|}{|\beta|} + \frac{|\gamma_2 \times \beta||\gamma_1|}{|\beta|}$. *(Hence if $|\gamma_1|$ and $|\gamma_2|$ are bounded in terms of $|\beta|$, then $|\gamma_1 \times \gamma_2|$ can be bounded in terms of $|\gamma_1 \times \beta|$ and $|\gamma_2 \times \beta|$).*

(2) *If $\gamma_1, \gamma_2, \beta$ are three sides of a triangle and γ is another saddle connection, then $|\beta \times \gamma| \leq \sum_{i=1}^{2} |\gamma_i \times \gamma|$.*

Proof. To prove (1), rotate the coordinate system so that β is vertical; $|\gamma_{i,h}| = \frac{|\gamma_i \times \beta|}{|\beta|}$. Since $|\gamma_{i,v}| \leq |\gamma_i|$, (1) follows. To prove (2), rotate so γ is vertical so $|\gamma_{i,h}| = \frac{|\gamma_i \times \gamma|}{|\gamma|}$. Then (2) follows from the fact that $|\beta_h| \leq |\gamma_{1,h}| + |\gamma_{2,h}|$ and $|\beta \times \gamma| = |\beta_h||\gamma|$. $\square$

In the next section we will need to combine or merge intersecting complexes to form bigger complexes. The main combinatorial lemma is the following.

Proposition 6.2 (Combining complexes). *Suppose K is a complex of complexity i with nonempty boundary and there is a saddle connection σ which is either disjoint from K or crosses ∂K. Then there is a complex K' containing K of complexity $j > i$ with the following properties:*

(i) $|\partial K'| \le |\partial K| + \max(|\partial K|, |\sigma|)$.

(ii) $||K'|| \le ||K|| + |\partial K| \max(|\sigma|, |\partial K|)$.

(iii) *For any $\omega' \subset K' \setminus K$ and any γ we have $|\omega' \times \gamma| \le \max_{\omega \in \partial K}(|\omega \times \gamma| + \max(|\omega \times \gamma|, |\sigma \times \gamma|))$.*

(iv) *If β is a saddle connection disjoint from both K and σ, then β is disjoint from K' or $\beta \subset K'$.*

Proof. Heuristically, to get K', we add σ to K and retriangulate the boundary, if necessary. More carefully, the construction can be done as follows:

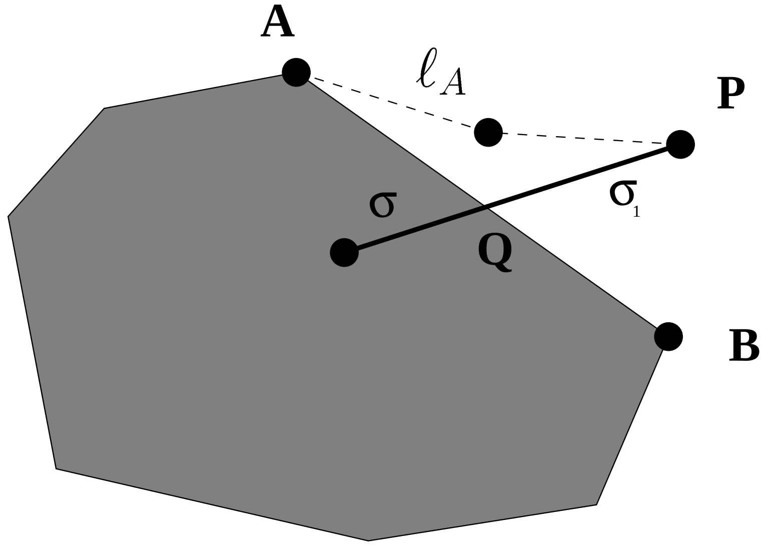

Figure 2.

If σ is disjoint from K, we may just add it to K and all the estimates hold. Thus we may assume that σ crosses ∂K. Let Q be the point where σ crosses ∂K on a segment ω with endpoints A, B. Suppose in following σ from Q into the exterior of K σ hits a singularity P in the exterior of K before again crossing ∂K or hits a singularity on ∂K. Let σ_1 be the subsegment of σ joining Q and P. (see figure 2) Consider the path $AQ * \sigma_1$ joining A to P, where $*$ denotes juxtaposition of paths. There is a unique geodesic l_A joining A to P in the homotopy class of that path. If not all saddle connections in l_A are already in K, we add those that are not in K giving us a larger complex K'. Since $|l_A| \leq |\omega| + |\sigma|$ and $||\Delta|| \leq |\sigma||\omega|$,(recall that $|\partial K|$ is not the length of ∂K, but rather the length of the longest saddle connection in ∂K), the length and area bounds (i) and (ii) are immediate. The second statement of Proposition 6.1 implies that for any segment ω' of l_A and any γ,

$$|\omega' \times \gamma| \leq |\sigma_1 \times \gamma| + |\omega \times \gamma| \leq |\sigma \times \gamma| + \max_{\omega \in \partial K} |\omega \times \gamma|$$

giving (iii). Since l_A, σ_1 and AQ bound a simply connected domain which contains no singularities in its interior, it is clear that if β is disjoint from K and σ, it is either one of the segments of l_A or disjoint from K' so we have (iv).

If every saddle connection of l_A is already a subset of K, then we apply the same analysis to l_B. Now l_A, l_B and ω bound a simply connected domain Δ. If Δ is a triangle, then it cannot be the case that the single segment l_B is also already in K, for then by assumption, Δ would already be in K and σ would not cross ω into the exterior of K. In this case can add l_B to form K'. If Δ is not a triangle and every segment of l_B is already in K, then we may add a saddle connection in Δ that goes from an endpoint of ω, say A to a singularity on l_B. The same estimates still hold.

Thus we are reduced to the case that before next hitting a singularity, σ crosses ∂K at Q' an interior point of a segment ω' with endpoints C, D. (see figure 3) It may be the case that $A = D$ or $B = C$ or both. Let σ_1 be the subsegment of σ joining Q to Q'. There are two cases. The first case is if σ_1 is homotopically nontrivial rel ∂K; in that case move σ_1 parallel to itself with one endpoint moving along ω toward one endpoint, say A, the other along ω', towards say D. We can choose the direction so that the lengths of the parallel segments do not increase. This can be done until for a first time there is a singularity P and a pair of segments, σ_2, one of which goes from ω to P, the other from ω' to P. (If P is either A or D, then σ_2 is a single segment.) The assumption that σ_1 is homotopically nontrivial guarantees that σ_2 does not lie in ∂K. It is clear that if β does not intersect σ, ω or ω', then it does not intersect σ_2. We proceed exactly as in the previous paragraph. We can use P to build a bigger complex whose boundary satisfies the same estimates on lengths and cross products as before.

The second possibility is that σ_1 is homotopically trivial rel ∂K. (see figure 4) In

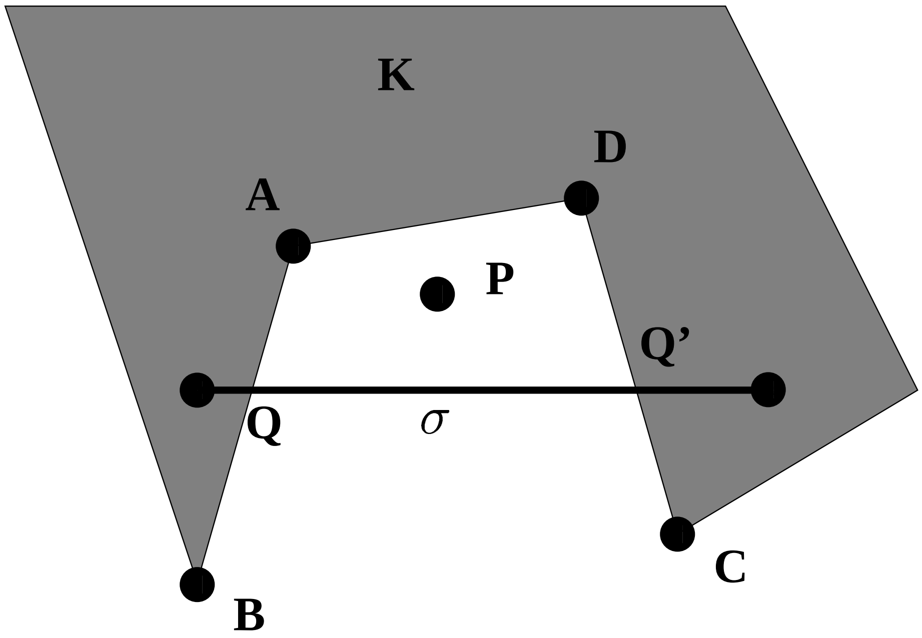

Figure 3.

that case σ_1 together with segments of ω and ω' and possibly other segments of ∂K bound a simply connected domain in the complement of K. This implies that there is a geodesic γ in the complement of K joining an endpoint of ω and an endpoint of ω' such that γ, ω, ω' and the same other segments of ∂K bound a simply connected domain in the complement of K. Then we may add the segments of γ to K. The above estimates hold. It is also the case that if β is disjoint from σ and ∂K, then either it is disjoint from γ or coincides with a segment of γ. $\square$

Corollary 6.3 (Complex generated by short saddle connections). *There are constants $c = c(p_0) < \infty$ and $c' = c'(p_0)$ such that for any $a > 0$, there is a complex Y containing all saddle connections of S of length at most a with $|\partial Y| \leq c'a$ and $||Y|| \leq ca^2$.*

Proof. We start with a complex consisting of a maximal set of disjoint saddle connections of length at most a. We fill in any triangles. Each has area at most a^2. If there are no additional segments of length at most a, we are done, with c the number of triangles. Suppose otherwise and suppose inductively we have built a complex Y_j of complexity j, with boundary length at most $c'_j a$ and area at most $c_j a^2$. If any

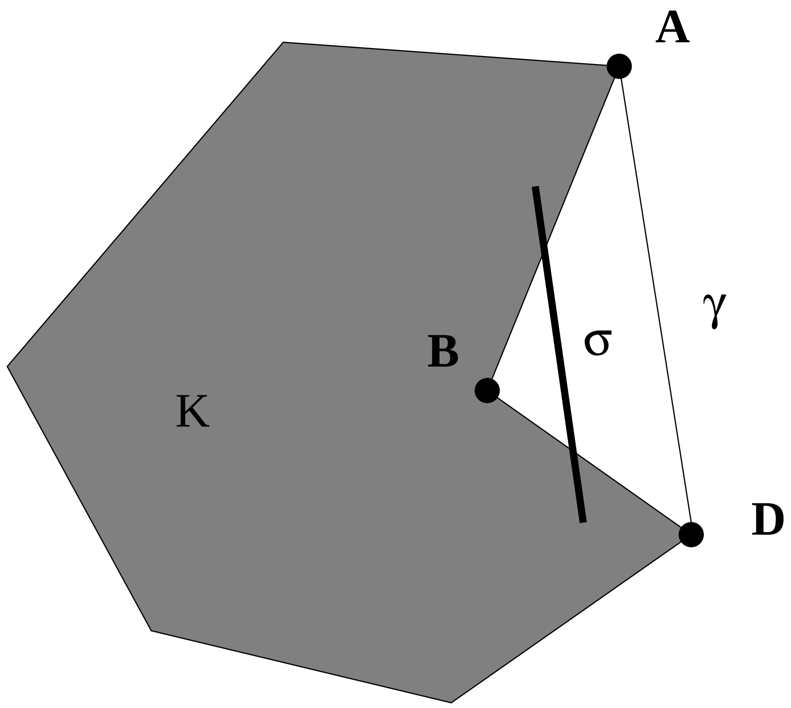

Figure 4.

saddle connection of length at most a crosses ∂Y_j by the above Proposition we can find a complex Y_{j+1} containing Y_j whose boundary has length has length at most $2c'_j a + a = c'_{j+1} a$ and area at most $c_j a^2 + c'_j a(a + c'_j a) = c_{j+1} a^2$. Since the maximum complexity of any complex is p_0, the induction ends after at most p_0 steps. □

Definition 6.4. For Y satisfying the conclusions of Corollary 6.3 we say Y is generated by saddle connections of length a.

The number ϵ. We now fix a number $\epsilon = \epsilon(p_0)$ so that $c\epsilon^2 < 1$ where c is as in Corollary 6.3. This implies that if $||Y|| = 1$, any complex $K \subset Y$ generated by saddle connections of length ϵ must in fact be a proper subset of Y.

Corollary 6.5 (Intersections of complexes).

(a) *Suppose Y is a complex and Ω is a set of saddle connections contained in Y. Then there is a complex $Z \subset Y$ containing every saddle connection in Ω.*

(b) *Suppose Y_1, Y_2 are two complexes. Then there exists a complex $Z \subset Y_1 \cap Y_2$ which contains every saddle connection in $Y_1 \cap Y_2$.*

Proof. To prove (a), as in the proof of Corollary 6.3 we build inductively Z out of the saddle connections in Ω. The only thing to check is that at each stage the resulting complex is a subset of Y. So suppose K is a complex which is a subset of Y and $\sigma \in \Omega$ crosses ∂K. Then either $\sigma \in \partial Y$ or every segment in ∂Y must be either contained in ∂K or disjoint from both σ and ∂K. Hence, by Proposition 6.2 (iv), ∂Y cannot cross the enlarged K'. This implies that $K' \subset Y$. This completes the inductive step in the proof of (a). To prove (b), apply (a) with Ω equal to the set of saddle connections in $Y_1 \cap Y_2$. $\square$

Definition 6.6 (Filled in Complex). Given a complex Y with boundary ∂Y and a complex K which does not contain ∂Y, we define the *filled in complex* K_f to be the set of points which are separated from ∂Y by K. Equivalently. K_f is the union of K with all the components of $Y \setminus K$ whose closure does not intersect ∂Y.

Definition 6.7 (Separate). A system Ω separates two disjoint saddle connection β and γ from each other if every arc from β to γ intersects Ω. We allow Ω to coincide with one of the saddle connections.

Definition 6.8 (Accessible, Weakly Accessible). Suppose Y is a domain, $\gamma \subset \partial Y$, and let $M > 1$ and $C > 0$. We say that a saddle connection $\beta \subset Y$ is *M accessible* from γ if for any system Ω that separates β from γ

(1) $|\Omega| > |\gamma|/M$.

(In particular, since we are allowing Ω to coincide with β, we must have $|\beta| > |\gamma|/M$).

We say β is (M, C) *weakly accessible* from γ if for any separating system Ω either (1) holds or for some saddle connection $\omega \subset \Omega$,

(2) $|\omega \times \gamma| > C||Y||$.

Remarks. We will use Definition 6.8 with $M \gg 1$ and $C \gg 1$. Note that always, if β is M-accessible from γ, then β is (M, C) weakly accessible from γ. If $||Y|| = 1$, $C \geq 2$, $|\gamma| \leq 1$ and (2) holds, then $|\omega| \geq 1$ and so (1) holds; this means that if $||Y|| = 1$, $|\gamma| \leq 1$, and $C \geq 2$, then β is M-accessible from γ if and only if β is (M, C) weakly-accessible from γ.

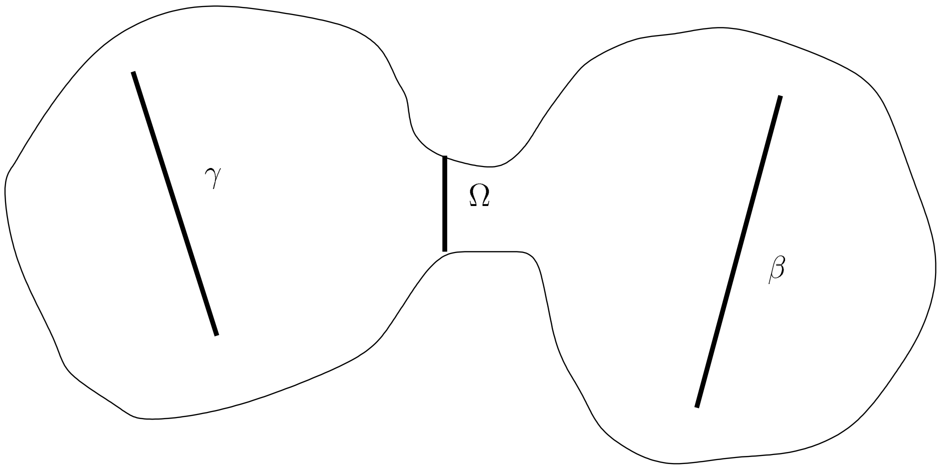

Figure 5. Ω is separating β from γ.

Proposition 6.9. *There exists a sequence of constants D_j such that if Y is a domain which can be triangulated by j saddle connections (not counting the saddle connections on the boundary of Y), $||Y|| = 1$, $\beta \subset Y$ is a saddle connection weakly (p_0, D_j) accessible from $\gamma \subset \partial Y$ then there is a chain of saddle connections $\beta = \beta_1, \beta_2, \dots, \beta_n = \gamma$ such that*

(1) *β_i is disjoint from β_{i+1}*

(2) *$|\beta_i \times \beta_{i+1}| \leq D_j$.*

(3) *β_i is weakly (p_0, D_j) accessible from β_{i+1}.*

The proof of Proposition 6.9 is based on the following:

Claim 6.10. *Suppose $C' > 0$. Let Y be a domain with $||Y|| = 1$, and let $\beta \subset Y$ and $\gamma \subset \partial Y$ be saddle connections.*

For $D > 2$ let $\mathcal{F}_D$ denote the set of saddle connections $\lambda \subset \bar{Y}$ such that $|\lambda \times \gamma| \leq D$ and λ and β are disjoint. Let $n(D)$ denote the maximal cardinality of a disjoint system contained in $\mathcal{F}_D$. Then at least one of the following is true:

(a) There exists a saddle connection $\sigma \subset Y$ with $|\sigma| \leq |\gamma|/p_0$ and $|\sigma \times \gamma| \leq D$.

(b) There exists a saddle connection $\lambda \in \mathcal{F}_D$ such that β is (p_0, C') weakly accessible from λ.

(c) *for* $D' = p_0 C' + (\frac{1}{p_0} + 2)D$ *we have* $n(D') > n(D)$.

(d) $|\beta \times \gamma| < p_0 C' + (\frac{1}{p_0} + 2)D$.

Proof of Claim 6.10. Suppose (a) and (b) do not hold. We note that $\mathcal{F}_D$ is not empty, since there exists a saddle connection τ which bounds a triangle with γ, so that $|\tau \times \gamma| \leq 2$.

Choose a maximal system Γ of disjoint saddle connections in $\mathcal{F}_D$, and let K be the complex that they form. Let Y_β be the connected component of $Y \setminus \partial K$ which contains β, and choose a saddle connection $\lambda \subset K \cap \partial Y_\beta$. Since $|\lambda \times \gamma| \leq D$ and (a) does not hold,

$$|\lambda| \geq |\gamma|/p_0. \tag{18}$$

Since (b) does not hold, β is not (p_0, M') weakly accessible from λ. This means that there is some system Ω separating β from λ such that neither (1) nor (2) holds. Choose any $\omega \in \Omega \setminus K$. Since (1) and (2) do not hold

$$|\lambda| \geq p_0 |\omega| \qquad \text{and} \qquad |\omega \times \lambda| \leq C'. \tag{19}$$

Now from Proposition 6.1 (i) applied to the saddle connections ω, γ, λ we get

$$|\omega \times \gamma| \leq |\omega \times \lambda| \frac{|\gamma|}{|\lambda|} + |\gamma \times \lambda| \frac{|\omega|}{|\lambda|} \leq p_0 C' + \frac{D}{p_0} \tag{20}$$

We now apply Proposition 6.2 to K and ω in place of σ to get a complex $K' \supset K$. Let ω_1 be any saddle connection in K' which is disjoint from K. Then it follows from Proposition 6.2 part (iii) and (20) that

$$|\omega_1 \times \gamma| \leq |\omega \times \gamma| + 2 \max_{\lambda \in K} |\lambda \times \gamma| \leq p_0 C' + \frac{D}{p_0} + D. \tag{21}$$

Now by Proposition 6.2 (iv), either $w_1 = \beta$ or w_1 is disjoint from β. In the former case, (d) holds in view of (21). In the latter case, $\Gamma \cup \omega_1 \in \mathcal{F}_{D'}$ where D' is the right-hand-side of (21), so (c) holds. $\square$

Proof of Proposition 6.9. The proof is by induction on the number $d(Y)$ of saddle connections needed to triangulate Y, not counting the saddle connections on the boundary of Y. (We note that $d(Y)$ is not exactly the complexity of the complex $\overline{Y}$ since we do not count boundary saddle connections in $d(Y)$). The proposition is clearly true if $d(Y) = 0$, so that Y is a triangle or a disjoint union of triangles. Now suppose the proposition has been proved for all domains Z with $d(Z) < j$, and Y is a domain with $d(Y) = j$. We assume $D_i, i < j$ have been chosen. Let $C' = D_{j-1}$, and for $D > 0$, let $f(D) = p_0 C' + (\frac{1}{p_0} + 2)D$. Let $D_j = f^{(p_0+1)}(D_{j-1})$ where $f^{(k)}$ denotes f iterated k times.

Suppose first that for some $D < D_j$, (a) of Claim 6.10 occurs. Consider the domain $Y \setminus \sigma$. We claim that if β is not (p_0, D_{j-1}) weakly accessible from γ on $Y \setminus \sigma$, it is not (p_0, D_j) weakly accessible from γ on Y, contrary to assumption. For adding σ to any system Ω that separates β from γ on $Y \setminus \sigma$ gives a system Ω' that separates β from γ on Y for which neither (1) nor (2) hold. Now, since $d(Y \setminus \sigma) = j - 1$, we can apply the induction hypothesis to $Y \setminus \sigma$ to get the required chain of saddle connections.

Hence we may assume that (a) of Claim 6.10 does not occur for any $D_{j-1} \leq D \leq D_j$. This implies that every saddle connection on Y is (p_0, D_j) weakly accessible from $\beta_n = \gamma$. We now apply Claim 6.10 repeatedly to

$$D_{j-1}, f(D_{j-1}), f^{(2)}(D_{j-1}), \dots, f^{(p_0+1)}(D_{j-1}) = D_j.$$

Note that (c) of Claim 6.10 can occur at most p_0 times. Thus, for some $D < D_j$, either (d) or (b) occurs. In the first case we set $n = 2$ and we have the chain β, γ. In the latter case we can set $\beta_{n-1} = \lambda$, and apply the induction hypothesis. $\square$

7 A system of integral inequalities

In this section, we prove Theorem 5.2, together with a variation, Proposition 7.2, which will be used in §8 in the proof of Theorem 5.1. The main idea of this section (i.e. the system of inequalities Lemma 7.5) is similar to ideas developed by G.A. Margulis for the "quantitative version of the Oppenheim conjecture" (cf. [EMM]).

For a domain Y, let

$$\ell(Y) = \min\{|\lambda| \ : \ \lambda \text{ is a saddle connection contained in } Y\}.$$

For γ a saddle connection on S, for some angle θ, γ determines a vector $[0, |\gamma|]$ on the surface $r_\theta S$. The angle θ is called the vertical angle for γ.

Given a domain Y of area 1 we may use the $SL(2, \mathbb{R})$ action to form the domains $a_t r_\theta Y$. Fix $\gamma \subset \partial Y$. Recall the constant p_0, which is the maximum number of disjoint saddle connections.

Definition 7.1 (Accessible saddle connection). For any θ and t, we say that a a saddle connection $\lambda \subset a_t r_\theta Y$ is *accessible* if $r_\theta^{-1} a_t^{-1} \lambda \subset Y$ is p_0-accessible from γ.

In this definition notice that the accessible condition refers to lengths on Y. Let

$$\ell_\gamma(a_t r_\theta Y) = \min \{|\lambda| \ : \ \lambda \text{ is an accessible saddle connection contained in } a_t r_\theta Y\}.$$

Proposition 7.2. *Let Y be a domain, $\gamma \subset \partial Y$ and $|\gamma| \leq 1$. By rotating let 0 be the vertical angle for γ. For each constant ρ and t, let $I(t) = I(t, \gamma, \rho)$ be the interval of angles $[\rho e^{-t}/|\gamma|, \rho e^{-t}/|\gamma|]$. Then for any δ, $0 < \delta < 1/2$ there exists ρ and C' (depending only on δ and ρ) such that for any t,*

$$\int_{I(t)} \frac{d\theta}{\ell_\gamma(a_t r_\theta X)^{1+\delta}} \leq C' \frac{e^{-(1-2\delta)t}}{|\gamma|^{1+2\delta}}.$$

Remarks. Proposition 7.2 differs from Theorem 5.2 as follows: In Proposition 7.2 we integrate only over the exponentially small sector $I(t)$ (instead of $[0, 2\pi]$) and consider only accessible saddle connections (instead of all saddle connections). Also Proposition 7.2 asserts that the integral is small (instead of bounded). However, the proofs of Proposition 7.2 and Theorem 5.2 are virtually identical. We will only give the proof of Proposition 7.2 in detail, and indicate how to modify it to obtain Theorem 5.2 at the end of this section. In §8, Proposition 7.2 will be used in the inductive step of the proof of Theorem 5.1.

7.1 Admissible complexes and the function α.

Definition 7.3 (Admissible Complex). Fix $\gamma \subset \partial Y$. For any θ, t, a complex $K' \subset a_t r_\theta Y$ with boundary $\partial K'$ and of complexity i is admissible on $a_t r_\theta Y$ if

(1) $||K'|| \leq 2^i \epsilon^2$ where ϵ is the constant defined following Corollary 6.3.

(2) for any system $\Omega \subset Y$ of saddle connections separating $r_\theta^{-1} a_t^{-1} K' \subset Y$ from $\gamma \subset \partial Y$, $|\Omega| > |\gamma|/p_0$.

In particular if K' is admissible then $|\partial(r_\theta^{-1} a_t^{-1} K')| \geq |\gamma|/p_0$.

The functions α_i. Fix $\delta > 0$. Let Y be a domain with $||Y|| = 1$. We define

$$\alpha_i(a_t r_\theta Y) = \max\left(\sup_{K_i} \frac{1}{|\partial K_i|^{1+\delta}}, 1\right)$$

where the sup is over all admissible subcomplexes K_i of $a_t r_\theta Y$ of complexity i.

The operator A_τ. For any $\tau > 0$ let A_τ denote the averaging operator which sends the function f defined on the $SL(2, R)$ orbit of a surface S_0 to

$$(A_\tau f)(S_0) = \frac{1}{2\pi} \int_0^{2\pi} f(a_\tau r_\theta S_0)\, d\theta,$$

We will be applying A_τ to functions f on $SL(2, \mathbb{R})$ which are invariant under the action of $SO(2)$ on the right, and can thus be considered as functions on the hyperbolic plane $SO(2)\backslash SL(2, R)$. For such f, $(A_\tau f)(z)$ is the average of f over the circle of radius τ centered at z.

7.2 A system of inequalities.

In this subsection we show that the functions α_i satisfy a certain system of integral inequalities (see Lemma 7.5 below). The idea is that if α_i is large, so that S_0 is near infinity in a stratum, then most points on the circle of radius τ centered at S_0 are further from infinity, which means that the average of the boundary length of K_i is bigger on the circle than at S_0. If K_i defines the supremum over the entire circle in the definition of α_i, then $A_\tau \alpha_i(S_0)$ is smaller than S_0. If for some point on the circle, some other K_i' of complexity i has a shorter boundary than K_i, then we "merge" K_i and K_i' as in §6 to form a complex K_j of complexity $j > i$, and bound α_i in terms of α_j.

Lemma 7.4. *Fix $0 < \delta < 1$. Then there exists a constant $c(\delta)$ such that for any $\tau > 0$ and any $v \in \mathbb{R}^2 - \{(0, 0)\}$,*

$$\frac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{||a_\tau r_\theta v||^{1+\delta}} \leq \frac{c(\delta) e^{-\tau(1-\delta)}}{||v||^{1+\delta}} \tag{22}$$

Proof. By rescaling and rotating may assume that $v = (0, 1)$. Then, the left-hand-side of (22) becomes:

$$\frac{1}{2\pi} \int_0^{2\pi} (e^{2\tau} \sin^2\theta + e^{-2\tau} \cos^2\theta)^{-(1+\delta)/2}\, d\theta$$

We decompose $[0, 2\pi] = R_1 \cup R_2$, where $R_1 = \{\theta \ : \ e^{2\tau} \sin^2\theta \leq e^{-2\tau} \cos^2\theta\}$ and R_2 is the set where the opposite inequality holds. Note that there exist absolute constants $0 < c_1 < c_2$ such that

$$c_1 e^{-2\tau} \leq |R_1| \leq c_2 e^{-2\tau} \tag{23}$$

On R_1, the integrand is bounded by a constant multiple of $e^{\tau(1+\delta)}$. Hence, in view of (23), the integral over R_1 is $O(e^{-\tau(1-\delta)})$ as required. Now the integral over R_2 is bounded by

$$e^{-\tau(1+\delta)} \int_{R_2} |\sin\theta|^{-(1+\delta)}\, d\theta = O(e^{-\tau(1+\delta)} \int_{R_2} |\theta|^{-(1+\delta)}\, d\theta = O(e^{-\tau(1-\delta)}),$$

where in the last estimate we used (23). $\square$

Lemma 7.5 (System of Inequalities). *For any $0 < \delta < 1$ and any $\tau > 0$ there exist constants $\omega = \omega(\tau, \delta)$ and $b = b(\tau, \delta)$ such that for any i and any $S_0 = a_t r_\phi Y$*

$$(A_\tau \alpha_i)(S_0) \leq c e^{-(1-\delta)\tau} \alpha_i(S_0) + b + \omega \sum_{j>i} \alpha_j(S_0).$$

where c is the constant of Lemma 7.4

Proof. We can assume that on S_0 the length of the boundary of the complex K_i that defines the supremum in α_i has length at most $e^{-2\tau}\epsilon$. Otherwise on the circle of radius τ centered at S_0, α_i is uniformly bounded, and thus we can absorb the left-hand-side in the constant b.

Suppose first K_i is the complex that defines α_i throughout the circle of radius τ. Then $A_\tau \alpha_i$ is the average of the minimum of

$$\frac{1}{||a_\tau r_\theta v||^{1+\delta}},$$

the minimum taken over $v \in \partial K_i$. Since the average of a minimum is smaller than the minimum of an average, by Lemma 7.4, we have the desired bound for α_i. Now suppose for some $\theta \in [0, 2\pi]$ some other complex $K_i^\theta \subset S_0$ with the same complexity defines $\alpha_i(a_\tau r_\theta S_0)$, i.e. $|\partial(a_\tau r_\theta K_i^\theta)| \leq |\partial(a_\tau r_\theta K_i)|$. Since lengths on $a_\tau r_\theta S_0$ differ from lengths on S_0 by a factor of at most e^τ,

$$|\partial K_i^\theta| \leq e^\tau |\partial(a_\tau r_\theta K_i^\theta)| \tag{24}$$

Here $|\partial K_i^\theta|$ refers to lengths on S_0. Assume first that K_i and K_i^θ have boundaries that cross. We take $\sigma \in \partial K_i^\theta \setminus \partial K_i$ and using Proposition 6.2 we can build a complex K_j^θ of complexity $j = j(\theta) > i$ containing K_i. We check that K_j^θ is admissible. On the contrary, if there were a system Ω separating $r_\phi^{-1} a_t^{-1} K_j^\theta$ from γ such that $|\Omega| \leq |\gamma|/p_0$, then Ω would separate $r_\phi^{-1} a_t^{-1} K_i$ from γ, contradicting that K_i is admissible. Thus

(2) in the definition of admissible holds. Since the boundary lengths of each of K_i and K_i^θ are at most ϵ and the area of K_i is at most $2^i\epsilon^2$, the area bound (1) follows from Proposition 6.2.

We also have

$$|\partial K_j^\theta| \leq |\partial K_i| + \max(|\partial K_i|, |\partial K_i^\theta|) \leq 2|\partial K_i^\theta|, \tag{25}$$

the first inequality by Proposition 6.2 and the second since K_i defines the supremum at S_0. Thus, in view of (25), for any θ

$$\alpha_i(a_\tau r_\theta S_0) = \frac{1}{|\partial(a_\tau r_\theta K_i^\theta)|^{1+\delta}} \leq \frac{\omega}{|\partial K_j^\theta|^{1+\delta}} \leq \omega\alpha_j(S_0), \tag{26}$$

where $\omega = (2e^\tau)^{1+\delta}$. Hence the desired estimate also holds in this case.

If the boundaries of K_i and K_i^θ do not cross, then we take $K_j^\theta = K_i \cup K_i^\theta$. Note that if Ω separates $r_\phi^{-1}a_t^{-1}K_j^\theta$ from γ, then Ω separates $r_\phi^{-1}a_t^{-1}K_i$ from γ, a contradiction. This shows that (2) in the definition of admissible holds. We have $\|K_j^\theta\| \leq \|K_i\| + \|K_i^\theta\| \leq 2^{i+1}\epsilon^2 \leq 2^j\epsilon^2$, so K_j^θ is admissible. Clearly, (25) and hence (26) hold in this case as well. This completes the proof of the lemma. $\square$

7.3 Averages over arcs.

In this subsection we will deduce Theorem 5.2 and Proposition 7.2 from Lemma 7.5.

We identify the right cosets $SO(2)\backslash SL(2,\mathbb{R})$ with the hyperbolic disk D^2; thus a function on $SL(2,\mathbb{R})$ such that $f(r_\theta g) = f(g)$ for all $\theta \in [0, 2\pi], g \in SL(2,\mathbb{R})$ may be considered to be a function on D^2. Let o denote the origin of D^2 (corresponding to the identity coset). For $\theta \in S^1$ and $t > 0$ the point $z_0 = o \cdot a_t r_\theta$ in the hyperbolic disk D^2 is on the circle of hyperbolic radius t centered at the origin and at angle θ. Now as ϕ varies over S^1, consider the points $o \cdot a_\tau r_\phi a_t r_\theta$. These points form the circle of radius τ centered at z_0. In polar coordinates we write

$$o \cdot a_\tau r_\phi a_t r_\theta = o \cdot a_s r_{\theta+\psi}$$

where $s = s(t,\tau,\phi)$ and $\psi = \psi(t,\tau,\phi)$. Thus $o \cdot a_\tau r_\phi a_t r_\theta$ is at hyperbolic distance $s(t,\tau,\phi)$ from the origin and is visible at the angle $\theta + \psi(t,\tau,\phi)$.

Recall the intervals $I(t,\gamma,\rho)$ defined in Proposition 7.2. For each $t > 0$, $\kappa > 0$ and $\theta \in I(t+\tau) \subset I(t)$ let $U_\theta(\kappa,t,\tau) \subset S^1$ denote the set of points $\phi \in S^1$ such that $s(t,\tau,\phi) \geq t+\tau-\kappa$ and $\theta + \psi(t,\tau,\phi) \in I(t+\tau)$, i.e. $o \cdot a_\tau r_\phi a_t r_\theta$ lies outside the circle of hyperbolic radius $t+\tau-\kappa$ and its radial projection to the circle of radius $t+\tau$ lies in $I(t+\tau)$.

Lemma 7.6. *There exists $c'' > 0$ so that for each κ, there is a τ and ρ such that for any $\theta \in I(t+\tau)$, $|U_\theta| \geq c''|S^1|$. The map $\Psi : U_\theta \to S^1$ sending ϕ to $\psi(t,\tau,\phi)$ is a diffeomorphism onto its image and*

$$2\pi c'' \leq |U_\theta| = \int_{U_\theta} d\phi = \int_{\Psi(U_\theta)} |\frac{d\phi}{d\psi}|\, d\psi \tag{27}$$

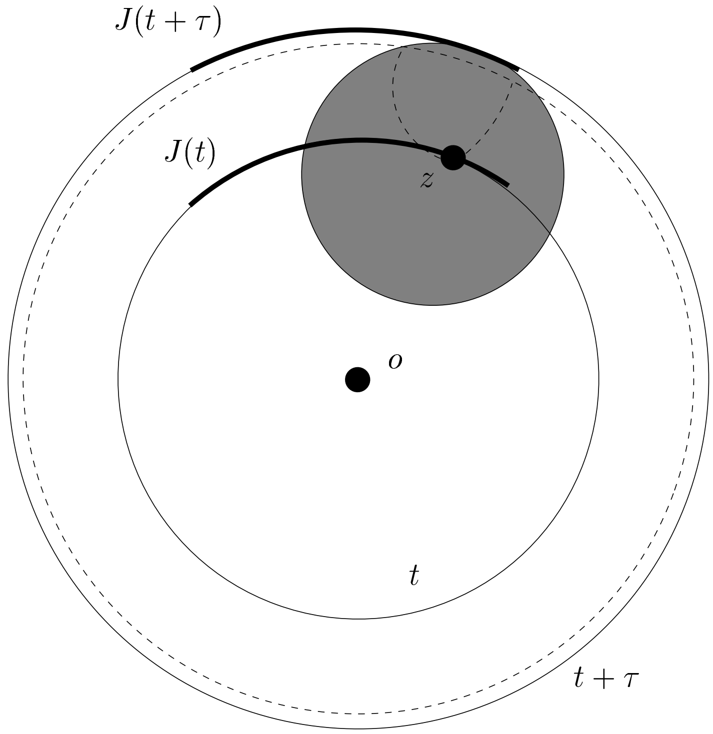

Figure 6. (Lemma 7.6). The dotted circle has center o and radius $t+\tau-\kappa$. The shaded circle has hyperbolic radius τ and is centered at a point $z \in J(t)$. The dotted lines starting at z are hyperbolic geodesics; $|U_\theta|$ is the angle between them. Hyperbolic geometry and the assumptions that τ is much larger then κ and that the diameter of $J(t+\tau) \geq 2\tau + 2\kappa$ imply that $|U_\theta|$ is bounded from below.

Proof. (See figure 6.) We use the notation $J(t) = \{o \cdot a_t r_\theta \quad : \quad \theta \in I(t)\}$, so $J(t) \subset D^2$ is the part of the circle of radius t around o which is seen from the origin at angles in $I(t)$. Given κ, for τ large enough, the assertion that $o \cdot a_\tau r_\phi a_t r_\theta$

lies outside the circle of hyperbolic radius $t+\tau-\kappa$ for most angles ϕ is just an exercise in hyperbolic geometry. Now choose $\rho = \rho(\kappa, |\gamma|)$ large enough so that for any t, $J(t+\tau)$ has hyperbolic diameter at least $2\tau + 2\kappa$ (it is enough to do this for $t=0$). By construction then, for any $\theta \in I(t+\tau)$ for at least $1/2$ of the angles ϕ, $\theta + \psi(t,\tau,\phi) \in I(t+\tau)$. The second assertion follows formally from the first. $\square$

Lemma 7.7. *Let f be any non-negative continuous function on $SL(2,\mathbb{R})$ invariant under left multiplication by $SO(2)$; (that is, a function on D^2) with the property that given any $\sigma > 1$ there exists $\kappa > 0$ such that for $g \in D^2$ of hyperbolic distance at most κ from o, and any $h \in SL(2,R)$ $\sigma^{-1} f(h) \le f(gh) \le \sigma f(h)$. Choose τ, ρ such that Lemma 7.6 holds. Then*

$$\int_{I(t+\tau)} f(a_{t+\tau} r_\theta)\, d\theta \le c' \sigma \int_{I(t)} (A_\tau f)(a_t r_\theta)\, d\theta$$

where c' does not depend on τ.

Proof. For each $\theta \in I(t+\tau)$,

$$\begin{aligned}(A_\tau f)(a_t r_\theta) &= \frac{1}{2\pi}\int_0^{2\pi} f(a_\tau r_\phi a_t r_\theta)\, d\phi \ge \frac{1}{2\pi}\int_{U_\theta} f(a_\tau r_\phi a_t r_\theta)\, d\phi \ge \\ &\ge \frac{\sigma^{-1}}{2\pi}\int_{\Psi(U_\theta)} f(a_{t+\tau} r_{\theta+\psi}) |\frac{d\phi}{d\psi}|\, d\psi\end{aligned} \tag{28}$$

where for the last estimate we used the definition of U_θ and the condition on f. Now integrating (28) over $\theta \in I(t+\tau)$ we get

$$\begin{aligned}\int_{I(t+\tau)} (A_\tau f)(a_t r_\theta)\, d\theta &\ge \frac{\sigma^{-1}}{2\pi}\int_{I(t+\tau)}\int_{\Psi(U_\theta)} f(a_{t+\tau} r_{\theta+\psi}) |\frac{d\phi}{d\psi}|\, d\psi d\theta \\ &= \frac{\sigma^{-1}}{2\pi}\int_{I(t+\tau)} f(a_{t+\tau} r_\xi) \left(\int_{-\Psi(U_\xi)} |\frac{d\phi}{d\psi}|\, d\psi\right) d\xi \\ &\ge c''\sigma^{-1}\int_{I(t+\tau)} f(a_{t+\tau} r_\xi)\, d\xi\end{aligned} \tag{29}$$

where in the second line we made the change of variable $\xi = \psi + \theta$ and for the last inequality we used (27) to estimate from below the inner integral. (To see the domain of the inner integral in (29) one can argue as follows: Take point on circle of radius t and on the real axis. Take the circle of radius τ centered at that point. Let p_-, p_+

be the two points where that circle intersects the circle of radius $t+\tau-\kappa$ centered at the origin. Denote the angles of p_- and p_+ as viewed from the origin by $-w(t)$ and $w(t)$ respectively. Then set $W=[-w(t),w(t)]$. The domain of the inner integral in (29) is $-\Psi(U_\xi)$ since before the change of variables the domain of integration was $\{(\theta,\psi)\ :\ \theta\in I(t+\tau) \text{ and } \psi\in W\cap(I(t+\tau)-\theta)\}$. After the change of variables the domain becomes $\{(\xi,\psi)\ :\ \xi\in I(t+\tau) \text{ and } \psi\in W\cap(-I(t+\tau)+\xi)\}$ which for the inner integral translates exactly to $-\Psi(U_\xi)$.)

Since $I(t)\supset I(t+\tau)$, (29) implies the estimate of the lemma. □

Proof of Proposition 7.2. For notational convenience we set $\alpha_i=1$ for $i>p_0$. the maximal complexity of a complex. Fix $\sigma>1$. Since α_i is invariant under left multiplication by rotations, and since elements of $SL(2,\mathbb{R})$ near the identity distort lengths of saddle connections by a multiplicative factor at most σ, we may choose κ so that α_i satisfies the hypotheses of Lemma 7.7. We may then choose τ sufficiently large so that applying Lemma 7.5 with Lemma 7.7 we get

$$\int_{I(t+\tau)}\alpha_i(a_{t+\tau}r_\theta Y)\,d\theta\le cc'\sigma e^{-(1-\delta)\tau}\int_{I(t)}\alpha_i(a_tr_\theta Y)\,d\theta+c'\sigma b|I(t)|+$$
$$+c'\sigma\omega\sum_{j=i+1}^{p_0}\int_{I(t)}\alpha_j(a_tr_\theta Y)\,d\theta \quad (30)$$

We want to prove that for all i, there is a constant $C_i(\rho,\tau,b)$ such that

$$\int_{I(t)}\alpha_i(a_tr_\theta Y)\,d\theta\le C_i\frac{e^{-t}e^{2\delta t}}{|\gamma|^{1+2\delta}} \quad (31)$$

This will be proved by reverse induction on i. Then, (31) is trivially true for $i>p_0$ (recall that $|I(t)|=2\rho e^{-t}/|\gamma|$ and $|\gamma|\le 1$). Now suppose (31) has already been proved for all $j>i$. It is clearly enough to prove (31) for i when $t=n\tau$. Let $a_n=\int_{I(n\tau)}\alpha_i(a_{n\tau}r_\theta Y)\,d\theta$. Then, the inequality (30) can be rewritten for a new constant $\omega'=\omega'(\omega,b,C_j,c',\sigma)$ as

$$a_{n+1}\le e^{-\tau}(cc'\sigma e^{\delta\tau})a_n+\frac{\omega'}{|\gamma|^{1+2\delta}}e^{-n\tau}e^{2\delta n\tau} \quad (32)$$

Iterating this equation, and taking out the common factor of $e^{-n\tau}$ we get for a new constant ω''

$$a_n\le e^{-n\tau}\left((cc'\sigma)^ne^{n\delta\tau}a_0+\frac{\omega''}{|\gamma|^{1+2\delta}}\sum_{k=1}^{n-1}(cc'\sigma e^{\delta\tau})^ke^{2(n-k)\delta\tau}\right)$$

Note that $a_0 \le (p_0/|\gamma|)^{1+\delta}$, since by definition an admissible complex has boundary length at least $|\gamma|/p_0$. If we choose τ large enough (e.g. so that $cc'\sigma e^{-\delta\tau} < 1/2$), the geometric series above is bounded, up to an absolute constant, by the term with $k = 1$. Thus, we get for all n,

$$a_n \le \frac{C_4 e^{-n\tau} e^{2n\delta\tau}}{|\gamma|^{1+2\delta}}$$

This implies (31). Now Proposition 7.2 follows from the fact that $\alpha_1(K) = 1/\ell_\gamma(K)^{1+\delta}$. $\square$

Proof of Theorem 5.2. The proof of this theorem is a simplified version of the proof of Proposition 7.2. The only differences are that one declares all complexes to be admissible, and one integrates over the entire circle $[0, 2\pi]$ instead of over $I(t)$. Then, since Lemma 7.5 holds on the entire hyperbolic plane, the proof goes through without modifications. $\square$

8 The Induction

In this section we prove Theorem 5.1. We will use the following version of Proposition 3.5:

Lemma 8.1. *Let Y be a domain, and for $\Theta > 0$ and $R_2 > R_1 > 0$ let $A(R_1, R_2, \Theta)$ denote the set of vectors in $\mathbb{R}^2$ of length between R_1 and R_2 which make an angle of at most Θ with the vertical. Let $V(\cdot)$ be a (set-valued) function on the orbit $SL(2,\mathbb{R})Y$ which satisfies condition (A) in §2. For $T > 0$ and $S \in SL(2,\mathbb{R})Y$ let $N_V(T, S)$ denote the cardinality of $V(S) \cap B(T, 0)$. Then, for any ϵ, $0 < \epsilon < 1$, there exists a constant $c = c(\epsilon) > 0$ such that for any $T > 1$, and any $\Theta > 1/T$,*

$$|A(\epsilon T/4, \epsilon T/2, \Theta) \cap V(Y)| \le cT^2 \int_{-2\Theta}^{2\Theta} N_V(\epsilon, a_t r_\theta Y)\, d\theta \tag{33}$$

where $t = \log T$.

Outline of proof. Let $f(x, y)$ be $16\pi/\epsilon^2$ times the characteristic function of the rectangle $[-\epsilon/4, \epsilon/4] \times [\epsilon/8, 3\epsilon/4]$. Then, it is easy to check (cf. the proof of Lemma 3.4) that for any $t > 0$, any $\Theta > e^{-t}$ and any $v \in A(\epsilon e^t/4, \epsilon e^t/2, \Theta)$,

$$e^{2t} \int_{-2\Theta}^{2\Theta} f(a_t r_\theta v)\, d\theta \ge 1.$$

The rest of the argument follows exactly the proof of Proposition 3.5. $\square$

The next theorem is the major step in the proof of Theorem 5.1. First recall the constants D_l given by Proposition 6.9

Theorem 8.2. *For any l, the following hold:*

(U_l) Let Y be a domain of complexity at most l and $||Y|| \leq 1$. Suppose $\gamma \in \partial Y$, and for $R > 0$, let $N_1(\gamma, R)$ denote the number of saddle connections β such that $|\beta| \leq R$, β is (p_0, D_l)-weakly accessible from γ and $|\beta \times \gamma| \leq D_l||Y||$. Then, for each $\delta > 0$ there is $C_1 = C_1(l, \delta, D_l)$ such that for each $R > 0$

$$N_1(\gamma, R) \leq C_1 \left(\frac{R}{|\gamma|}\right)^{1+\delta}$$

(V_l) Let Y and γ be as in (U_l). Let $N_2(\gamma, R)$ denote the number of saddle connections β such that $|\beta| < R$ and β is (p_0, D_l) weakly- accessible from γ. Then, for each $\delta > 0$ there is $C_2 = C_2(l, \delta)$ such that for each $R > 0$

$$N_2(\gamma, R) \leq C_2 \left(\frac{R}{|\gamma|}\right)^{1+\delta}$$

(W_l) Let Y be as in (U_l), and K a possibly empty complex that does not contain ∂Y. Let Y_K denote the complement of the filled in complex K_f. Let $N(Y, K, R)$ denote the number of saddle connections in Y of length at most R which intersect Y_K; that is to say, are not contained in K_f. Then, for each $\delta > 0$, there is $C_3 = C_3(l, \delta)$ such that for each $R > 0$

$$N(Y, K, R) \leq C_3 \left(\frac{R}{\ell(Y, K)}\right)^{1+\delta}$$

where $\ell(Y, K)$ denotes the length of the shortest saddle connection in Y that intersects Y_K (i.e. not contained in K_f).

Remark. We note that Theorem 5.1 (a) follows easily from (W_l). Indeed, let $\epsilon > 0$ be the constant defined following Corollary 6.3, and let $S \in QD(P)$ be arbitrary. Form the complex $Y \subset S$ containing all saddle connections of length at most ϵ. Y is a complex with boundary, and the shortest saddle connection on S is contained in Y. We take K to be the empty complex and apply (W_l). Then, $\ell(Y, K) = \ell(S)$, and thus (17) holds.

The proof of Theorem 8.2 is by induction on l. Clearly (U_1), (V_1) and (W_1) are all true since only 1 saddle connection can be added to the domains. The induction is done in the following steps:

(W_{l-1}) and Proposition 7.2 imply (U_l). The assumption on β implies that $|\beta|/|\gamma| \geq 1/p_0$; otherwise the curve β itself would separate itself from γ, contradicting the weak accessibility of β to γ, contrary to the assumption. By rescaling, we may assume that $\|Y\| = 1$. Suppose first $|\gamma| \leq 1$. This implies that the notions of p_0 accessible and (p_0, D_l) weakly accessible coincide.

By Corollary 6.3 we form a complex $K_0 \subset Y$ which generated by the saddle connections of length at most $|\gamma|/p_0 c'$; then $|\partial K_0| \leq |\gamma|/p_0$ so any saddle connection contained in K_0 is not p_0-accessible from γ. We replace K_0 by its "filled-in" version (see §6). It is still the case that any saddle connection contained in K_0 is not p_0 accessible from γ.

Now, for $S = a_t r_\theta Y \in SL(2,\mathbb{R})Y$, let $V(S)$ denote the set of saddle connections $\lambda \subset S$ such that $\lambda \not\subset a_t r_\theta K_0$. Then, $V(\cdot)$ satisfies the condition (A) of §2.

Let 0 correspond to the angle for which γ is vertical. Suppose a saddle connection $\beta \subset Y$ contributes to $N_1(\gamma, R) - N_1(\gamma, R/2)$. Then,

$$R/2 \leq |\beta| \leq R \tag{34}$$

and

$$|\beta \times \gamma| \leq D_l. \tag{35}$$

From (35) and the fact that γ is vertical, $|h_\beta| \leq D_l/|\gamma|$ (where as in §6, h_β denotes the horizontal component of β). Let ϕ denote the angle, smaller than $\pi/2$ between β and γ. Then from (34),

$$\sin\phi = \frac{|h_\beta|}{|\beta|} \leq \frac{2D_l}{|\gamma| R}.$$

Choose $\epsilon > 0$ to be as in §6 (following Corollary 6.3). Let $T = \frac{2R}{\epsilon}$ so that saddle connections satisfying (34) have length between $\epsilon T/4$ and $\epsilon T/2$. Also set $e^t = T$, and $\Theta = \sin^{-1}(\min(1, 2D_l/|\gamma|R))$. Then, $|\phi|$ is always bounded by Θ, hence

$$N_1(\gamma, R) - N_1(\gamma, R/2) \leq |A(\epsilon T/4, \epsilon T/2, \Theta) \cap V(Y)|. \tag{36}$$

Assume that $T \geq 1$. Then, by Lemma 8.1,

$$|A(\epsilon T/4, \epsilon T/2, \Theta) \cap V(Y)| \leq cT^2 \int_{-2\Theta}^{2\Theta} N_V(\epsilon, a_t r_\theta Y)\, d\theta. \tag{37}$$

Note that

$$\Theta \leq \frac{4D_l}{|\gamma|R} = \frac{8D_l e^{-t}}{\epsilon|\gamma|} \tag{38}$$

Now suppose $\theta \in [-2\Theta, 2\Theta]$. We will estimate $N_V(\epsilon, a_t r_\theta Y)$. By Corollary 6.3 we build the complex $L \subset Y$ of complexity of at most $l-1$ generated by the saddle connections of length at most ϵ on $a_t r_\theta Y$. Thus, all the saddle connections which contribute to $N_V(\epsilon, a_t r_\theta Y)$ are contained in L. It is our intention to count the number of saddle connections of length ϵ in L which are not contained in K_0, since the saddle connections in K_0 are not accessible. By Corollary 6.5 part (b) there exists a complex $K' \subset L \cap K_0$ which contains all the saddle connections in $L \cap K_0$. By construction, if a saddle connection in L is contained in $V(Y)$, it is not contained in the filled in K'_f, with respect to L. We may assume that K' does not contain ∂L (otherwise $L \subset K_0$ and there is nothing to count). Thus for some constant $C_3 = C_3(\delta, \epsilon)$,

$$\begin{aligned} N_V(\epsilon, a_t r_\theta Y) &\leq N(L, K'_f, \epsilon) \leq \\ &\leq \frac{C_3}{\ell(L, K'_f)^{1+\delta/2}} = \frac{C_3}{\ell(a_t r_\theta Y, a_t r_\theta K_0)^{1+\delta/2}} = \frac{C_3}{\ell_\gamma(a_t r_\theta Y)^{1+\delta/2}} \end{aligned} \tag{39}$$

where in the second estimate we have used (W_{l-1}) and the last equality uses the definition of ℓ_γ in §7.

Combining (36), (37) and (39) we obtain:

$$N_1(\gamma, R) - N_1(\gamma, R/2) \leq C_3 e^{2t} \int_{-2\Theta}^{2\Theta} \frac{d\theta}{\ell_\gamma(a_t r_\theta Y)^{1+\delta/2}}$$

Now set $\rho = \frac{16D_l}{\epsilon}$ so by (38) the domain of integration in the above integral is contained in the interval $[\frac{-\rho e^{-t}}{|\gamma|}, \frac{\rho e^t}{|\gamma|}]$. By Proposition 7.2, for some $C'' = C''(D_l, \delta, \epsilon)$, $C_1 = C_1(D_l, \delta, \epsilon)$,

$$N_1(\gamma, R) - N_1(\gamma, R/2) \leq C'' e^{2t} \frac{e^{-t(1-\delta)}}{|\gamma|^{1+\delta}} = C'' \left(\frac{e^t}{|\gamma|}\right)^{1+\delta} = C_1 \left(\frac{R}{|\gamma|}\right)^{1+\delta}.$$

Now, an easy geometric series argument shows that

$$N_1(\gamma, R) \leq C_1'(D_l, \delta, \epsilon) \left(\frac{R}{|\gamma|}\right)^{1+\delta}.$$

This completes the proof of (U_l) in the case that $|\gamma| \le 1$ and $T \ge 1$.

If $|\gamma| \le 1$ and $T \le 1$, then $R \le \epsilon$, and in view of the upper bound in (34),

$$N_1(\gamma, R) \le N_V(R, Y). \tag{40}$$

We do exactly the same except that we set L to be the complex generated by the saddle connections of length $R \le \epsilon$ on the original domain Y. Again, set $K' \subset L \cap K_0$ to be the complex containing all the saddle connections in $L \cap K_0$. Then, using (W_{l-1}) as in (39),

$$N_V(R, Y) \le N(L, K'_f, R) \le C_4 \left(\frac{R}{\ell(L, K'_f)} \right)^{1+\delta} = C_4 \left(\frac{R}{\ell(Y, K_0)} \right)^{1+\delta} \le C_5 \left(\frac{R}{|\gamma|} \right)^{1+\delta}, \tag{41}$$

where the constants C_4 and C_5 depend on l and δ, and the last estimate follows from Corollary 6.3 and the definition of K_0. Now (U_l) follows from (40) and (41). This completes the proof of (U_l) in the case $|\gamma| \le 1$.

If $|\gamma| \ge 1$ then we can rotate γ so it is vertical and then flow by a_t so that its length is 1. Since rotations are isometries, a_t contracts a vertical saddle connection the most and cross products are preserved, it is clear that after this action by an element of $SL(2, R)$ that any β which is (p_0, D_l)- weakly accessible before is still (p_0, D_l)- weakly accessible. Let $|\cdot|'$ denote lengths on the deformed surface. Then we have, for some C_6,

$$\frac{R}{|\gamma|} \le |\beta|' \le \frac{2R}{|\gamma|} + D_l \le C_6 \frac{R}{|\gamma|},$$

with the first inequality on the right following from $|\beta|' \ge \frac{|\beta|}{|\gamma|} \ge \frac{R}{|\gamma|}$, the middle inequality following from the condition on the cross product, and the rightmost inequality following from the fact that $R/|\gamma|$ is bounded below by $\frac{1}{p_0}$. This implies that the estimate for $|\gamma| = 1$ gives the general estimate.

(U_l) **and Proposition 6.9 imply** (V_l)**.** We say a saddle connection σ has size n if $e^{n-1} < |\sigma| \le e^n$. Up to an additive constant there are $\log R/|\gamma|$ sizes between $|\gamma|$ and R. For each saddle connection β which contributes to $N_2(\gamma, R)$, first note that β is (p_0, D_l) weakly accessible from γ. Thus we find the chain (see Proposition 6.9) $\beta = \beta_1, \ldots, \beta_n = \gamma$ where $|\beta_i \times \beta_{i+1}| \le D_l$, β_i is (p_0, D_l) weakly-accessible from β_{i+1}. This implies that $|\beta_i|/|\beta_{i+1}| \ge \frac{1}{p_0}$ and so up to an additive constant β_i has size between size($|\gamma|$) and size($|\beta|$). Now (U_l) implies that the number of β is bounded by a constant times a product

$$\left(\frac{|\beta_1|}{|\beta_2|}\right)^{1+\delta/2} \times \ldots \times \left(\frac{|\beta_{n-1}|}{|\beta_n|}\right)^{1+\delta/2} = C_3 \left(\frac{|\beta|}{|\gamma|}\right)^{1+\delta/2} \leq C_3 \left(\frac{R}{|\gamma|}\right)^{1+\delta/2},$$

where the product has to be summed over all the possible sizes of $|\beta_i|$. This introduces an extra factor of $\log(R/|\gamma|)^{p_0}$ into the above quantity, which is incorporated into the exponent $1+\delta$.

(V_l) **and** (W_{l-1}) **imply** (W_l)**.** Let ω be the shortest saddle connection intersecting Y_K; that is, not contained in the filled in K_f. Let $K' \subset Y$ denote the complex generated by all the saddle connections in Y of length smaller than $|\omega|/c'$ where c' is as in Corollary 6.3 so $K' \subset K_f$. Then $K'_f \subset K_f$ as well so it is sufficient to prove (W_l) for saddle connections intersecting $Y_{K'}$; that is, not contained in K'_f. Let γ the longest saddle connection in $\partial Y \setminus K$ so $|\gamma| \geq |\omega|$. Let L be the subcomplex of Y generated by all the saddle connections of length smaller then $|\gamma|/c'$, so $K' \subset L$ and therefore $K'_f \subset L_f$. Since, $\gamma \notin L_f$, L_f is a proper subcomplex of Y. Thus, its complexity is at most $l-1$.

We wish first to count the number of saddle connections in $L_f \setminus K'_f$, if this set is nonempty. By (W_{l-1}),

$$N(L_f, K'_f, R) \leq C_3(l,\delta) \left(\frac{R}{\ell(L_f, K'_f)}\right)^{1+\delta} \leq C_4(l,\delta) \left(\frac{R}{|\omega|}\right)^{1+\delta}$$

Now we wish to count the saddle connections that intersect $Y_{K'}$ that are not contained in L_f. Each such saddle connection is p_0 accessible from γ, and hence is (p_0, D_l) weakly accessible from γ. Hence, by (V_l) the total number $N(\gamma, R)$ of such saddle connections is at most

$$N(\gamma, R) \leq C_2(l,\delta) \left(\frac{R}{|\gamma|}\right)^{1+\delta} \leq C_2(l,\delta) \left(\frac{R}{|\omega|}\right)^{1+\delta}$$

This completes the proof of (W_l) and the inductive step. □

Proof of Theorem 5.1 (b). Let $V(\cdot)$ denote the set of vectors, counting multiplicity associated to saddle connections. Let $\epsilon = \epsilon(P) > 0$ be as in Theorem 5.1 (a) (and §6). Let $C(\delta)$ be as in Theorem 5.1 (a).

We note that for any $R > \epsilon$ there exists an integer $m = m(R)$ and m elements $g_1, \ldots, g_m$ of $SL(2,\mathbb{R})$ depending only on R such that $B(R,0) \subset \bigcup_{j=1}^m g_j B(\epsilon, 0)$. Hence, for any $S \in QD(g,P)$,

$$|V(S) \cap B(R,0)| \leq \sum_{j=1}^{m} |V(S) \cap g_j B(\epsilon,0)| = \sum_{j=1}^{m} |V(g_j^{-1}S) \cap B(\epsilon,0)|. \qquad (42)$$

By Theorem 5.1 (a), for each j,

$$|V(g_j^{-1}S) \cap B(\epsilon,0)| \leq C(\delta) \left(\frac{\epsilon}{\ell(g_j^{-1}S)} \right)^{1+\delta}. \qquad (43)$$

Since lengths on $g_j^{-1}S$ differ from lengths on S by a factor depending only on g_j (and hence only on R), (42) and (43) imply the theorem. □

Remark. An easy geometric series argument shows that Theorem 5.1 (a), and hence (B) and (C_μ) hold for cylinders of imprimitive closed geodesics. This implies that the pointwise asymptotic formula holds for cylinders of imprimitive closed geodesics on almost every surface.

We now relate the constants in the asymptotic formulas for primitive and imprimitive geodesics. Let $V(\cdot)$ denote the primitive closed geodesics, and let $W(\cdot)$ denote the imprimitive closed geodesics. For a a bounded function f on $\mathbb{R}^2$ and a surface S let

$$\hat{f}(S) = \sum_{v \in V(S)} f(v) \qquad \text{and} \qquad \tilde{f}(S) = \sum_{v \in W(S)} f(v)$$

We claim that for any ergodic invariant probability measure μ, the Siegel-Veech constants are related by $c_{V,\mu} = c_{W,\mu}/\zeta(2)$. Choose any $R > 0$, and let f denote the characteristic function of the ball of radius R. Let f_k denote the characteristic function of the annulus $\{v \ : \ R/(k+1) < \|v\| \leq R/k\}$. Then, for any $S \in QD(P)$,

$$\tilde{f}(S) = \sum_{k=1}^{\infty} k \hat{f}_k(S) \qquad (44)$$

From the Siegel-Veech formula Theorem 2.2 we have,

$$\int_{QD(P)} \tilde{f}(S)\, d\mu(S) = \pi c_{W,\mu} R^2, \quad \text{and} \quad \int_{QD(P)} \hat{f}_k(S)\, d\mu(S) = \pi c_{V,\mu} \left(\frac{R^2}{k^2} - \frac{R^2}{(k+1)^2} \right) \qquad (45)$$

Integrating (44) over $QD(P)$ with respect to μ and substituting (45), we get

$$\pi c_{W,\mu} R^2 = \pi c_{V,\mu} \sum_{k=1}^{\infty} k \left(\frac{R^2}{k^2} - \frac{R^2}{(k+1)^2} \right) = \pi c_{V,\mu} R^2 \sum_{k=1}^{\infty} \frac{1}{k^2} = \pi c_{V,\mu} R^2 \zeta(2).$$

Hence $c_{W,\mu} = \zeta(2) c_{V,\mu}$.

9 Proof of Theorem 1.4

In this section we use the notations Y_i, μ_i, $N_4(\cdot)$, $N_5(\cdot)$ defined in §1.

Theorem 9.1. *Each μ_i is ergodic under the $SL(2,\mathbb{R})$ action on Y_i.*

Proof. We follow the Hopf proof of ergodicity of the diagonal action a_t. Let $f \in C_0(Y_i)$ a continuous function with compact support. For each $x \in Y_i$ we form

$$f^*(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(a_t x) dt$$

and

$$f_*(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(a_{-t} x) dt,$$

the limits existing for μ_i a.e. $x \in Y_i$. Moreover $f^* = f_*$ a.e. as well. We need to show that $f^* = f_* = c$ a.e. for some constant c. Now for each x, f^* is constant along the vertical leaves (pairs of vertical leaves) of the flat structure (pair of flat structures) corresponding to x. If the vertical foliation of the flat structure is ergodic then f^* is constant a.e. on each (product) fiber. Since μ_i a.e. fibers have ergodic vertical foliations, this implies that f^* is constant on fibers. This reduces the proof to that of [Ma1]. $\square$

Lemma 9.2. *Let $\epsilon > 0$ be as in §6.*

(1) For any $\delta > 0$ there exists $C_5 = C_5(P, \delta)$ such that for any $S \in QD(P)$ and any $p \in S$, $N_4(S, P, p, \epsilon) \leq C_5/\ell(S)^{1+\delta}$.

(2) For any $\delta > 0$ there exists C_6 such that for any p, q, $N_5(S, P, p, q, \epsilon) \leq C_6/\ell(S)^{1+\delta}$.

Proof. Recall that Z_S denotes the set of singularities of S. To prove (1), let $Z'_S = Z_S \cup p$. Let S' be the same surface as S except that on S' we define a saddle connection to be a geodesic segment which begins and ends at a point of Z'_S, and does not contain a point of Z'_S in its interior.

By Corollary 6.3, there exists a complex $Y' \subset S'$ with boundary such that every saddle connection on S' of length at most ϵ is contained in Y'. Note that any saddle connection on S' of length shorter then $\ell(S)$ must begin or end at p. Let $b = \ell(S)/3$, and form the complex K' generated by saddle connections of length at most b. Now we claim K' consists of at most a single segment from p to a vertex v. Suppose on the contrary there are two segments joining p to vertices v_1, v_2 or a closed loop from p to itself with length at most b. In the first case there would be geodesic joining v_1 to v_2 of length at most $2b < \ell(S)$, contradicting that the shortest segment on S has length at most $\ell(S)$. In the second case, in the homotopy class of the closed loop would be a segment of length at most b, again a contradiction.

Since K' consists of at most single segment which is not a closed loop, it cannot contain the entire boundary of Y'. Thus by Theorem 8.2 (W_l),

$$N(Y', K', \epsilon) \leq \frac{C_3}{\ell(Y', K')^{1+\delta}} = \frac{C_5}{\ell(S)^{1+\delta}}.$$

The proof of (2) is similar. This time set $b = \ell(S)/4$. We add both p, q to the vertex set Z_S and again form Y' and K'. Now K' can consist of at most 3 segments; no two of which can be both be from p (resp. q) to vertices or from p to q. Nor there can there be a segment from p (resp. q) to itself. Thus the only possible closed loop must be homotopically trivial consisting of segments joining a vertex v to p and q and a segment joining p to q. This closed loop bounds a triangle Δ. Now K' can be the entire boundary of Y' only if $Y' = \Delta$, in which case the bound in the lemma is trivial. Otherwise we apply (W_l) again. □

Proof of Theorem 1.4. We first verify that Theorem 2.1 applies. Let $\mathcal{M}$ be one the fiber bundles Y_i. If $i = 1$, let $V(S, p) \subset \mathbb{R}^2$ denote the set of segments on S joining $p \in S$ to a vertex. If $i = 2$, let $V(S, p, q) \subset \mathbb{R}^2$ denote the set of segments on S joining $p \in S$ to $q \in S$. The condition (A) is clearly satisfied. Also (C_{μ_i}) is clearly satisfied in view of Lemma 9.2 and Lemma 5.5. Also, (B) is satisfied in view of Lemma 9.2 and Theorem 5.2 (cf. the proof of Theorem 5.4 in §5). Thus, to prove Theorem 1.4, it remains to evaluate the Siegel-Veech constants.

Fix a small ϵ and let $f \in C_0^\infty(R^2)$ be a positive function which is 1 on the disc of radius ϵ, is 0 outside the disc of radius 2ϵ and has integral $\pi\epsilon^2 + o(\epsilon^2)$. To prove (1), for $p \in S \setminus Z_S$, let $V = V(S, P, p)$ denote the set of segments joining p to a singularity.

By the Siegel-Veech formula

$$\int_{Y_1} \hat{f}\, d\mu_1 = c_{V,\mu_1} \int_{\mathbb{R}^2} f = c_{V,\mu_1} \pi \epsilon^2 + o(\epsilon^2). \tag{46}$$

We will show that

$$c_{V,\mu_1} = \frac{k}{2} \tag{47}$$

We will compute c_{V,μ_1} by computing the left hand side of (46), dividing both sides by ϵ and let $\epsilon \to 0$.

Let $QD_\epsilon(P)$ denote the set of flat structures S such that the shortest saddle connection on S has length at most 4ϵ. For each $S \notin QD_\epsilon(P)$ and $v \in P$ let $D_\epsilon(v)$ denote the ball of radius ϵ centered at v. For each such S,

$$\int_S \hat{f}(S,p)\, d\lambda_S(p) = \sum_{v \in P} D_\epsilon(v) + o(\epsilon^2) = k\pi\epsilon^2/2 + o(\epsilon^2). \tag{48}$$

where λ_S denotes the Euclidean measure on S.

Let $Y_{1,\epsilon}$ be the fiber space over $QD_\epsilon(P)$. It is enough to show that

$$\int_{Y_{1,\epsilon}} \hat{f}\, d\mu_1 = o(\epsilon^2) \tag{49}$$

for then adding the contribution of (48) integrated over $QD(P)$ to (49), substituting into (46), dividing both sides by ϵ and let $\epsilon \to 0$, we get (47).

Now by Lemma 9.2, for each $S \in QD_\epsilon(P)$,

$$\hat{f}(S) \leq \frac{C_5}{\ell(S)^{1+\delta}}.$$

Since $\hat{f}$ is supported on a set of measure $O(\epsilon^2)$ on S,

$$\int_S \hat{f}(S,p)\, d\lambda_S(p) = O\left(\frac{\epsilon^2}{\ell(S)^{1+\delta}}\right).$$

Hence, since by Lemma 5.5 the function $\frac{1}{\ell(S)^{1+\delta}} \in L^1(QD(P))$, and the measure of $QD_\epsilon(P) \to 0$ as $\epsilon \to 0$, (49) holds. This concludes the proof of (1).

For (2), let $V = V_{S,P,p_1,p_2}$ denote the set of segments joining p_1, p_2 and again apply the Siegel-Veech formula to f. This time we will show

$$c_{V,\mu_2} = 1$$

and then use Theorem 2.1.

Again we need to compute the left side of the Siegel-Veech formula, divide by ϵ and let $\epsilon \to 0$. Again we first restrict to fibers over $QD(P) \setminus QD_\epsilon(P)$. In each fiber $S \times S$, let W_ϵ be the set of pairs (p_1, p_2) where either p_1 or p_2 is in the 2ϵ neighborhood of a singularity. For $(p_1, p_2) \in S \times S \setminus W_\epsilon$, $\hat{f} = 1$ if p_1, p_2 belong to the same ϵ ball. Thus

$$\int_{S \times S \setminus W_\epsilon} \hat{f} = \pi\epsilon^2 + o(\epsilon^2).$$

If $(p_1, p_2) \in W_\epsilon$, say, p_1 lies in a 2ϵ neighborhood of a singularity, then p_2 must lie within a 4ϵ neighborhood of the singularity; otherwise $\hat{f} = 0$. The integral of $\hat{f}$ over the set of pairs (p_1, p_2) such that both p_i lie in a 4ϵ neighborhood of a singularity is $O(\epsilon^4) = o(\epsilon^2)$. Thus the integral of $\hat{f}$ on the fiber space over $QD(P) \setminus QD_\epsilon(P)$ is $\pi\epsilon^2 + o(\epsilon^2)$.

Now using the second part of Lemma 9.2, on each fiber S over $QD_\epsilon(P)$, $\hat{f}$ is $O(\frac{1}{\ell(S)^{1+\delta}})$ and since $\hat{f}$ is supported on a set of measure $O(\epsilon^2)$, the integral of $\hat{f}$ is $O(\frac{\epsilon^2}{\ell(S)^{1+\delta}})$ on each such fiber. Exactly as before this implies that the total contribution to $\int \hat{f}$ is $o(\epsilon^2)$ on the fibers over $QD_\epsilon(P)$. $\square$

References

[EG] C. Earle, F. Gardiner. Teichmuller disks and Veech's $\mathcal{F}$-structures. *Extremal Riemann surfaces* (San Francisco, CA, 1995), 165–189, *Contemp. Math.*, **201**, Amer. Math. Soc., Providence, RI, 1997.

[EMaS] A. Eskin, H. Masur, M. Schmoll *In preparation.*

[EMM] A. Eskin, G. Margulis, and S. Mozes. Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture. *Ann. of Math. (2)* **147**(1998), no. 1, 93–141

[EMZ] A. Eskin, H. Masur, A. Zorich. "The Siegel-Veech constants." *In preparation.*

[EO] A. Eskin, A. Okounkov. Asymptotics of numbers of branched coverings of a torus and volumes of moduli spaces of holomorphic differentials *Preprint.*

[EZ] A. Eskin, A. Zorich Billiards in right-angled polygons. *In preparation.*

[Gut] E. Gutkin. Billiards on almost integrable polyhedral surfaces. *Ergodic Th. Dyn Syst.* **4** (1984) pp. 569–584

[GuJ] E. Gutkin, C. Judge. Affine mappings of translation surfaces: geometry and arithmetic. *Preprint.*

[KMS] S. Kerckhoff, H. Masur and J. Smillie. Ergodicity of Billiard flows and quadratic differentials, *Annals of Math.* **124** (1986), 293–311.

[KZ] M. Kontsevich and A. Zorich. Connected components of the space of holomorphic differentials with prescribed singularities. *Preprint.*

[Kra] I. Kra, On the Nielsen-Thurston-Bers type of some self-maps of Riemann surfaces. *Acta Math.* **146** (1981), no. 3-4, 231–270.

[Ma1] H. Masur. Interval exchange transformations and measured foliations. *Annals of Math.* **115**(1982), 169–200.

[Ma2] H. Masur. The growth rate of trajectories of a quadratic differential. *Ergodic Th. and Dynam. Syst.* **10** (1990), 151–176.

[Ma3] H. Masur. Lower bounds for the number of saddle connections and closed trajectories of a quadratic quadratic differential. In *Holomorphic Functions and Moduli*, vol. 2 Springer-Verlag, Berlin, New York 1988

[MaSm] H. Masur, J. Smillie. Hausdorff dimension of sets of foliations. *Annals of Math.* **134**(1991), 455–543.

[NS] A. Nevo, E. Stein. Pointwise ergodic theorems for radial averages on simple Lie groups II. Duke Math. J **86** (1997) 239-259

[Nevo] A. Nevo. "Equidistribution in measure preserving actions of semisimple groups". *Preprint.*

[Rat] M. Ratner. On Raghunathan's measure conjecture. *Annals of Math.* **134**(1991), 545–607.

[Ve1] W. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.* (2) **115** (1982), no. 1, 201–242.

[Ve2] W. Veech. The Teichmuller geodesic flow. *Ann. of Math.* (2) **124** (1986), no. 3, 441–530.

[Ve3] Teichmuller curves in moduli space, Eisenstein series and an application to triangular billiards. *Invent. Math.* **97** (1989), no. 3, 553–583.

[Ve4] W. Veech. Moduli spaces of quadratic differentials. *J. Analyse. Math* **55** (1990), 117–171

[Ve5] W. Veech, *Siegel measures. Annals of Mathematics* **148** (1998), 895-944

[ZK] A.N. Zemlyakov, A.B. Katok. Topological transitivity of billiards in polygons. *Matem. Zametki* **18**(2) (1975) pp. 291–300. English translation in *Math. Notes* **18**(2) (1976) pp. 760–764.