Billiards in rectangles with barriers

Alex Eskin[*], Howard Masur[†] and Martin Schmoll

May 30, 2002

1 Introduction

Let $P = P(p/q, \alpha)$ denote the polygon whose boundary is the union of the boundary of the unit square $\partial([0,1] \times [0,1])$ and the interval $\{p/q\} \times [0, \alpha]$ (see Figure 1). We consider "billiard paths"; straight lines in the polygon, that when hitting a side reflect with equal angles and that do not pass through vertices. If a trajectory closes up, then it determines a parallel family of closed trajectories of the same length. The boundary of such a parallel family consists of a pair of trajectories from the point $(p/q, \alpha)$ to itself. Let v_1, v_2 denote unit vectors in the directions where the trajectories leave $(p/q, \alpha)$ and let w_1, w_2 denote unit vectors in the directions where the trajectories return to $(p/q, \alpha)$. Because of the fact that the polygon has right angles, the set $\{\pm v_1, \pm v_2, \pm w_1, \pm w_2\}$ is stable under the action of the Klein group (i.e the group generated by reflections in the horizontal and vertical axes). For each T let $N_c(p/q, \alpha, T)$ denote the number of parallel families of closed trajectories of Euclidean length at most T. (In this formulation we allow trajectories to repeat themselves. We also distinguish between closed trajectories traveling in opposite directions). We also consider "generalized diagonals", i.e. billiard trajectories that join the point $(p/q, \alpha)$ to itself which are not on the boundary of a parallel family; These can occur singly or in pairs, in the sense that there are two that are of the same length and the set $\{\pm v_1, \pm w_1, \pm v_2, \pm w_2\}$ of incoming and outgoing directions of the two trajectories is stable under the Klein group. Let $N_s^{(1)}(p/q, \alpha, T)$ denote

*Research partially supported by the Packard Foundation
†Research partially supported by NSF grant 9803497

the number of such generalized diagonals of multiplicity 1 of length at most T and $N_s^{(2)}(p/q,\alpha)$ the number of multiplicity 2.

Theorem 1.1. *There exist constants $s_1(q), s_2(q)$ and $c(q)$ depending only on q and independent of p and α, such that whenever α is irrational,*

$$N_s^{(i)}(p/q,\alpha,T) \sim \frac{1}{8} s_i(q)\pi T^2, i=1,2$$

and

$$N_c(p/q,\alpha,T) \sim \frac{1}{8} c(q)\pi T^2$$

as $T \to \infty$.

In the classical integrable case of billiards in the unit square, the asymptotics are well-known, since a parallel family of length at most T corresponds to an integer lattice point $(2m,2n)$ with $4(m^2+n^2) \leq T^2$. The asymptotics are $\frac{1}{8}\pi T^2$. The equilateral triangle, the isoceles right triangle and the $\pi/3$ right triangle are also integrable and thus have easily identified asymptotics. Veech [20] found a surprising class of nonintegrable of billiards for which one can find precise asymptotics for the number of cylinders and saddle connections. One such set of examples are billiards in right triangles with angle $\pi/n, n \geq 5$. These billiards have the property that a certain affine automorphism group of the billiard is a lattice in $SL(2,\mathbb{R})$. Veech showed that whenever that is the case, the counting problems have exact quadratic asymptotics. Another proof is given in [11]. The billiards $P(p/q,\alpha)$ are Veech precisely when α is rational. Thus our theorem provides, for α irrational, a new class of billiards with precise asymptotics.

Theorem 1.2. *We have $c(2)=9/2$, $s_1(2)=0$, $s_2(2)=2$. For $q \geq 3$,*

$$c(q) = \frac{10q-11}{2q-2}$$

$$s_1(q) = \frac{27(q-2)}{8(q-1)}$$

$$s_2(q) = \frac{5q+6}{8(q-1)}.$$

We note that in the case where $q=2$, both Theorem 1.1 and Theorem 1.2 were previously observed in [10] using a more elementary argument.

Actually in this paper we prove a different set of formulae for the constants $c(q)$, $s_1(q)$ and $s_2(q)$. See Proposition 4.14, Proposition 4.13 and Proposition 4.15. The algebraic proofs which show that the expressions given in those propositions are in fact equal to the constants in Theorem 1.2 will appear in [5].

By a standard construction ([13], [9], [14]) a rational billiard gives rise to a holomorphic 1 form on a compact Riemann surface, or equivalently the structure of a flat surface. This construction has proven useful in many contexts. We describe this construction in the case of the billiard P.

We take four copies of $P = P(p/q, \alpha)$ which are images of P under reflection in the two coordinate axes and reflection in the origin and glue them as in Figure 1. Then we make the following identifications on the union of the 4 copies: we identify the top to the bottom, the left side to the right side. The interval $p/q \times [0, \alpha]$ gives rise to two vertical double lines. We glue the left side of the right line to the right side of the left line, and the right side of the right line to the left side of the left line. We get a surface $S = S(p/q, \alpha)$ of area 4 which is locally flat except at the two points located at the endpoints of the vertical lines. At these points, the total angle around the point is 4π.

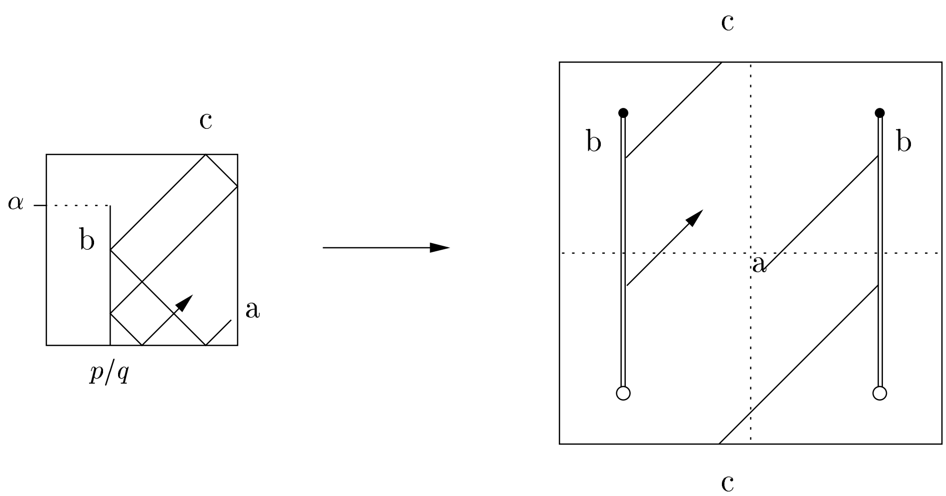

Figure 1. On the invariant surfaces $S(p/q, \alpha)$ of the wall billiard $P(p/q, \alpha)$ the trajectories are straight lines.

A billiard trajectory λ on $P(p/q, \alpha)$ gives rise to a straight line in a direction θ on $S = S(p/q, \alpha)$; if the line is in the upper left hand square of S then the trajectory λ is moving in the direction θ on $P = P(p/q, \alpha)$; if it is in the upper right hand square it is moving in the direction $\pi - \theta$; if it is in the lower right square it is in the direction $\pi + \theta$ and if it is in the lower left square it is $2\pi - \theta$. This analysis shows that the directions θ and $\pi - \theta$ determine the same closed orbit on the table. Thus the asymptotic constants for the surface are exactly twice that of the table. (see Figure 1).

A saddle connection on S that returns to the zero from which it left, returns at an angle of π and hence is on the boundary of a parallel family of closed geodesics. Thus a generalized diagonal on P that is not the boundary of a parallel family corresponds to a saddle connection joining distinct zeroes. In particular,

$$N_s^{(i)}(p/q, \alpha, T) = \frac{1}{2} N_s^{(i)}(S(p/q, \alpha), T),$$

and

$$N_c(p/q, \alpha, T) = \frac{1}{2} N_c(S(p/q, \alpha), T),$$

where $N_s^{(i)}(S, T)$ (resp. $N_c(S, T)$) denotes the number of saddle connections of multiplicity $i = 1, 2$ (resp. cylinders of closed geodesics which do not pass through singular points) on S of length at most T. (By a saddle connection of multiplicity 2 we mean a pair of saddle connections of the same length and direction, connecting the same zeroes).

We can generalize the notion of billiards to holomorphic 1 forms on compact Riemann surface. Given a pair (M, ω), where M is a compact Riemann surface and ω is a holomorphic 1 form, and a point $p \in M$ such that $\omega(p) \neq 0$, there exists a local coordinate z near p such that $\omega = dz$. Such a local coordinate is unique up to $z \to z + c$. Thus $|dz|^2$ is a flat metric on M; this metric develops conical singularities at the zeroes of ω. An easy calculation shows that the total angle at a zero of order k is $2\pi(k + 1)$.

Furthermore using such a coordinate near every nonsingular point determines a structure of a "translation surface" on M, namely an atlas of coordinate charts which cover the surface away from the singularities, such that the transition functions are translations $z \to z + c$. Hence we may visualize $S = (M, \omega)$ as a union of polygons contained in $\mathbb{R}^2$ glued along parallel sides, such that each side is glued to exactly one other, and the angle excess at each vertex is an integer multiple of 2π. The surface $S(p/q, \alpha)$ determines a holomorphic 1 form ω on a compact Riemann surface M with two simple zeroes at the points corresponding to the two points located at the end of the slit.

The study of the action of $SL(2, \mathbb{R})$ on the space of holomorphic 1 forms or translation surfaces has proved to be important in the subject. The action is as follows. Suppose $A \in SL(2,\mathbb{R})$ and a surface S is realized as a union of polygons Δ, glued isometrically along parallel sides. Then A acts linearly on each Δ and $A \cdot S$ is the union of the $A(\Delta)$ glued along the corresponding parallel sides.

On any surface $S = (M, \omega)$ we have the notion of geodesics which are straight lines away from the singularities in any coordinate system. We also define a *saddle connection* to be a straight line segment which begins and ends at a singularity. An oriented saddle connection γ defines a vector $v(\gamma)$ called its *holonomy* defined to be $v(\gamma) = \int_\gamma \omega$. A closed geodesic that does not meet a singularity determines a cylinder of closed geodesics that are parallel of the same length. If γ is a saddle connection or closed geodesic of S, with holonomy v, then γ is a saddle connection or closed geodesic of $A \cdot S$ and the holonomy of γ on $A \cdot S$ is given by $A(v)$.

Now let β be a partition of $2g-2$ (i.e. a representation of $2g-2$ as a sum of positive integers). Let $\mathcal{H}(\beta)$ denote the space of pairs (M, ω) where M is a Riemann surface of genus g, and ω is a holomorphic differential such that the orders of its zeroes is given by β. This space is called a stratum; this term is justified by the fact that the space of all Abelian differentials on Riemann surfaces of genus g is stratified by the spaces $\mathcal{H}(\beta)$, as β varies over the partitions of $2g-2$. In particular, the stratum corresponding to the partition $(1, \ldots, 1)$ is called the principal stratum: it corresponds to holomorphic differentials with simple zeroes.

In genus 2, there are 2 strata, namely the principal stratum $\mathcal{H}(1,1)$, consisting of holomorphic differentials with 2 simple zeroes, and $\mathcal{H}(2)$, consisting of holomorphic differentials with one double zero. Our construction shows that the surface $S(p/q, \alpha)$ belongs to $\mathcal{H}(1,1)$.

For a general stratum $\mathcal{H}(\beta)$, the subset $\mathcal{H}_1(\beta)$ of area 1 surfaces in $\mathcal{H}(\beta)$ carries a canonical finite $SL(2,\mathbb{R})$ invariant measure ν_1 defined as follows. Let $(M, \omega) \in \mathcal{H}(\beta)$ and let P the set of zeroes of ω. One chooses a basis for the relative homology group $H_1(M, P, \mathbb{Z})$. (To avoid confusion, we will henceforth refer to the group $H_1(M, \mathbb{Z})$ as the *absolute* homology). The holonomy of ω along the relative homology basis gives local coordinates for $\mathcal{H}(\beta)$. One first defines the measure ν on $\mathcal{H}(\beta)$ as the pull-back of Lebesgue measure on $\mathbb{R}^n$ by these local coordinates. One checks easily that the measure is well-defined, independent of the choice of holonomy basis and $SL(2,\mathbb{R})$ invariant. Then for $E \subset \mathcal{H}_1(\beta)$ define the cone $C(E)$ over E to be the set $(M, r\omega)$ where $(M, \omega) \in E$ and $0 \leq r \leq 1$. We define

$$\nu_1(E) = n\nu(C(E)).$$

For each surface $S \in \mathcal{H}(1,1)$ we can introduce the quantities $N_s^{(i)}(S,T)$ and $N_c(S,T)$ for the number of saddle connections and cylinders of length at most T.

It follows from [3], that there exist constants $s_1(1,1), s_2(1,1)$ and $c(1,1)$ such that for ν_1 almost all surfaces $S \in \mathcal{H}(1,1)$

$$N_s^{(i)}(S,T) \sim s_i(1,1)\pi T^2$$

and

$$N_c(S,T) \sim c(1,1)\pi T^2$$

as $T \to \infty$. These constants have been evaluated in [6]; we have $s_1(1,1) = 27/8, s_2(1,1) = 5/8$ and $c(1,1) = 5$. We emphasize that the constants $s_i(q)$ and $c(q)$ of Theorem 1.1 for the billiard $P(p/q,\alpha)$ differ from the "generic" constants $s_i(1,1)$ and $c(1,1)$. However, we have the following:

Theorem 1.3.

$$\lim_{q\to\infty} s_1(q) = \frac{27}{8} = s_1(1,1)$$

$$\lim_{q\to\infty} s_2(q) = \frac{5}{8} = s_2(1,1)$$

$$\lim_{q\to\infty} c(q) = 5 = c(1,1).$$

The above is consistent with the behaviour of unipotent flows on homogeneous spaces [7]. In the context of flat surfaces, it suggests that as $q \to \infty$, the q fold torus covers that are not Veech surfaces are approximating generic surfaces in $\mathcal{H}(1,1)$.

Remark: The proof of Theorem 1.1 will be an application of Ratner's measure classification theorem [17] and methods developed in [3]. The results about the evaluation of the constant (i.e. Theorem 1.2 or Proposition 4.14, Proposition 4.13 and Proposition 4.15) require also certain results about connectedness of the space of primitive torus covers (Theorem 4.4 below). Even though Theorem 1.3 follows immediately from Theorem 1.2, in order to make the paper self contained we derive it in Appendix A from the alternative expressions Propositions 4.13-4.15.

2 Torus Covers

As we have seen, the surface $S(p/q),\alpha)$ is given by a pair (M,ω) where M is a compact Riemann surface and ω is a holomorphic 1 form on M. The essential feature which we use is that $S = (M,\omega)$ is a *q-fold torus cover* i.e. there exists a holomorphic map $\pi : M \to \mathfrak{t}$, where $\mathfrak{t}$ is a torus, and such that ω is the pullback of the standard differential dz on $\mathfrak{t}$. The map π is a q-fold covering map, with branching at the zeroes of ω. Namely, we can divide the surface $S(p/q, \alpha)$ into q vertical strips of width $2/q$ and height 2 and map each to a torus of the same width and height.

Now let $S \in \mathcal{H}(1,1)$ be a surface which is a q fold cover. Since a torus covers itself, the torus covered by $S = (M, \omega)$ is not unique. We say that a torus cover $\pi' : M \to \mathfrak{t}'$ factors through $\pi : M \to \mathfrak{t}$ if there is an unbranched cover $\pi'' : \mathfrak{t} \to \mathfrak{t}'$ such that $\pi' = \pi'' \circ \pi$. Let z_1, z_2 denote the zeroes of ω. We say that the covering $\pi : M \to \mathfrak{t}$ is *unstable* if there exists a covering $\pi' : M \to \mathfrak{t}'$ which factors through π such that $\pi'(z_1) = \pi'(z_2)$. It is easy to see that $S(p/q, \alpha)$ is a stable q-fold torus cover if and only if α is irrational.

Let $\mathcal{M}_q(1,1)) \subset \mathcal{H}(1,1)$ denote the surfaces in $\mathcal{H}(1,1)$ which have area 1 and are q-fold torus covers.

Lemma 2.1. *The set $\mathcal{M}_q(1,1)$ is a closed subset of $\mathcal{H}(1,1)$ which is invariant under the action of $SL(2,\mathbb{R})$ on $\mathcal{H}(1,1)$.*

Proof. Suppose ω_n is a sequence in $\mathcal{M}_q(1,1)$ such that there is a sequence of tori $\mathfrak{t}_n$ so that ω_n is the pull-back of dz under a covering map $\pi : M \to \mathfrak{t}_n$. Suppose ω_n converges to ω_0. The sequence $\mathfrak{t}_n$ cannot converge to the cusp; otherwise the q fold cover would contain a curve whose length approaches 0. Thus passing to a subsequence we can assume $\mathfrak{t}_n$ converges to $\mathfrak{t}_0$. The holomorphic maps $\pi : M \to \mathfrak{t}_n$ must converge to a holomorphic map which is a q fold covering of ω_0 over $\mathfrak{t}_0$. Hence $\mathcal{M}_q(1,1)$ is closed. Since $SL(2,\mathbb{R})$ acts linearly on tori as well as on flat structures, it is clear that the $SL(2,\mathbb{R})$ orbit of a torus cover again consists of torus covers. $\square$

Let $\mathcal{T}^2$ denote the moduli space of tori with 2 marked points; we always assume that the marked points are distinguishable and one of them is at the origin. Then,

$$\mathcal{T}^2 \cong SL(2,\mathbb{R}) \ltimes \mathbb{R}^2 / SL(2,\mathbb{Z}) \ltimes \mathbb{Z}^2 \equiv G/\Gamma.$$

Lemma 2.2. *There is a covering map $p : \mathcal{M}_q(1,1) \to \mathcal{T}^2$. The covering map commutes with the $SL(2,\mathbb{R})$ action. (The action on $\mathcal{T}^2 = G/\Gamma$ is by left multiplication which makes sense since $SL(2,\mathbb{R}) \subset G$).*

Proof. The map p just sends $S = (M, \omega) \in \mathcal{M}_q^{(1,1)}$ to $(\pi(M), \pi(z_1), \pi(z_2))$, where z_1, z_2 are the zeroes of ω. $\square$

Let μ denote the Haar measure on $\mathcal{T}^2$; then $\tilde{\mu} = p^{-1}(\mu)$ is a smooth invariant measure for the $SL(2,\mathbb{R})$ action on $\mathcal{M}_q(1,1)$.

As in [3], let $r_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$ and $a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$.

The main step is

Lemma 2.3. *Let $K = \{r_\theta \ : \ 0 \leq \theta \leq 2\pi\}$, so that $K \cong SO(2)$. Let $S \subset \mathcal{M}_q(1,1)$ be a stable cover, and let $\mathcal{M}(S)$ denote the connected component of $\mathcal{M}_q(1,1)$ containing S. Let ν denote the K-invariant measure supported on the set KS. Then in the weak-* topology on $\mathcal{M}(S)$,*

$$\lim_{t\to\infty} a_t\nu = \frac{1}{\tilde{\mu}(\mathcal{M}(S))}\tilde{\mu}.$$

Proof. The Moore ergodicity theorem says that the action of $SL(2,\mathbb{R})$ on $\mathcal{T}^2$ is ergodic. Since $\mathcal{M}(S)$ is connected, the action on $\mathcal{M}(S)$ is ergodic.

By [3, Corollary 5.3], for any sequence t_i, there exists a subsequence, again denoted t_i such that the measures $a_{t_i}\nu$ converge weakly to a probability measure. Hence it is enough to show that whenever $a_{t_i}\nu$ converges to a measure ν_∞, we have $\nu_\infty = \lambda\tilde{\mu}$, where $\lambda = 1/\tilde{\mu}(\mathcal{M}(S))$.

Consider the orbit of $SL(2,\mathbb{R})$ through $p(S) \in \mathcal{T}^2 \equiv G/\Gamma$. By Ratner's theorem [17] the orbit closure is itself the orbit of a connected Lie group F such that $SL(2,\mathbb{R}) \subset F \subset G$. It is easy to check that the only connected subgroups F of G containing $SL(2,\mathbb{R})$ are $SL(2,\mathbb{R})$ itself and G. We show that the stability assumption forces $F = G$, i.e. that the $SL(2,\mathbb{R})$ orbit through $p(S)$ is dense in G/Γ. Suppose on the contrary that the $SL(2,\mathbb{R})$ orbit is closed. Intersecting with the fiber over the standard torus, the $SL(2,\mathbb{Z})$ orbit of the point $\pi(z_2)$ is closed. But this means that $\pi(z_2)$ is rational, and so $p(S)$ is not stable.

Now since p commutes with the $SL(2,\mathbb{R})$ action, it follows from the assumption that $a_{t_i}\nu \to \nu_\infty$ that $a_{t_i}p(\nu) \to p(\nu_\infty)$; the latter convergence takes place on $\mathcal{T}^2$. Now by [19, Corollary 1.2] (which uses Ratner's theorem), we have $p(\nu_\infty) = c\mu$, where $c \in \mathbb{R}$ and μ is the Haar measure on G/Γ. Then ν_∞ is a measure on $\mathcal{M}(S)$ which is absolutely continuous with respect to $\tilde{\mu} = p^{-1}(\mu)$. Also ν_∞ is $SL(2,\mathbb{R})$ invariant. Then by the Radon-Nykodim theorem and the ergodicity of the action, $\nu_\infty = \lambda\tilde{\mu}$, where $\lambda \in \mathbb{R}$. Since ν_∞ is a probability measure, $\lambda = 1/\tilde{\mu}(\mathcal{M}(S))$. $\square$

For any continuous compactly supported function $f : \mathbb{R}^2 \to \mathbb{R}$, let $\hat{f} : \mathcal{H}(1,1) \to \mathbb{R}$ denote the Siegel-Veech transform

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

where $V(S) \subset \mathbb{R}^2$ denotes either the vectors associated with saddle connections of multiplicity i, or (cylinders of) periodic geodesics on S which do not pass through singularities (see [3] for details on the notation). If $V(\cdot)$ denotes the vectors associated with saddle connections of multiplicity i, then $N_s^{(i)}(S,T) = |V(S) \cap B(T)|$, where $B(T)$

is the ball of radius T centered at the origin. If $V(\cdot)$ denotes the vectors associated with cylinders of periodic geodesics, then $N_c(S,T) = |V(S) \cap B(T)|$.

Theorem 2.4. *Let $S \in \mathcal{M}_q(1,1)$ denotes a stable q-fold torus cover . Then there exists a constant $\kappa(S)$ such that as $T \to \infty$,*

$$|V(S) \cap B(T)| \sim \pi\kappa T^2$$

where $B(T)$ is the ball of radius T (cf. [3, Theorem 2.1]). The constant κ depends only on the connected component $\mathcal{M}(S)$ of $\mathcal{M}_q(1,1)$ containing S. In fact, $\kappa(S)$ is given by the following Siegel-Veech formula: for any continous compactly supported $f : \mathbb{R}^2 \to \mathbb{R}$,

$$\frac{1}{\tilde{\mu}(\mathcal{M}(S))} \int_{\mathcal{M}(S)} \hat{f} \, d\tilde{\mu} = \kappa(S) \int_{\mathbb{R}^2} f.$$

We have already observed that the surface $S(p/q, \alpha)$ is a stable q-fold torus cover for α irrational, (and after rescaling by a factor of $1/2$ will have area 1). Now Theorem 1.1 will follow from Theorem 2.4, and Theorem 4.4 which says that the $S(p/q, \alpha)$ all belong to the same connected component.

Proof of Theorem 2.4 . The proof relies on results and methods of [3]. First, in view of Lemma 2.2, $\mathcal{M}(S)$ is a finite cover of $\mathcal{T}^2$. We again denote the covering map by p. Also we continue to denote the restriction of $\tilde{\mu}$ to $\mathcal{M}(S)$ by $\tilde{\mu}$. The main step is to show

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

Now equation (1) does not immediately follow from Lemma 2.3 because the function $\hat{f}$ is not bounded (it is continuous on an open dense set of full measure).

We proceed as follows. As in [3], let $\ell(S)$ denote the length of the shortest saddle connection on S. For $\epsilon > 0$, let h_ϵ be a smooth function such that $h_\epsilon(S) = 1$ if $\ell(S) > \epsilon$ and $h_\epsilon(S) = 0$ if $\ell(S) < \epsilon/2$. Then for any $\eta > \delta > 0$,

$$\hat{f}(S)(1 - h_\epsilon(S)) \le \frac{C(\delta)}{\ell(S)^{1+\delta}}(1 - h_\epsilon(S)) \le \epsilon^{\eta-\delta} \frac{C(\delta)}{\ell(S)^{1+\eta}}, \tag{2}$$

where the left inequality is by [3, Theorem 5.1] and the right is by the fact that $1 - h_\epsilon$ is supported on the set where $\ell(\cdot) < \epsilon$. Let A_t denote the averaging functional $A_t(\phi) = \frac{1}{2\pi} \int_0^{2\pi} \phi(a_t r_\theta S) d\theta$. Then, by (2),

$$A_t(\hat{f} h_\epsilon) \le A_t(\hat{f}) = A_t(\hat{f} h_\epsilon) + A_t(\hat{f}(1 - h_\epsilon)) \le A_t(\hat{f} h_\epsilon) + \epsilon^{\eta-\delta} C(\delta) A_t(\ell(\cdot)^{-1-\eta}).$$

By [3, Theorem 5.2] $A_t(\ell(\cdot)^{-1-\eta})$ is uniformly bounded by a constant $C_1(\eta)$ as $t \to \infty$. By Lemma 2.3 and since $\hat{f}h_\epsilon$ is continuous and compactly supported, $A_t(\hat{f}h_\epsilon)$ converges to $\lambda \int_{\mathcal{M}(S)} \hat{f}h_\epsilon d\tilde{\mu}$, where $\lambda = \frac{1}{\tilde{\mu}(\mathcal{M}(S))}$. Hence

$$\lambda \int_{\mathcal{M}(S)} \hat{f}h_\epsilon d\tilde{\mu} \leq \liminf_{t\to\infty} A_t(\hat{f}) \leq \limsup_{t\to\infty} A_t(\hat{f}) \leq \lambda \int_{\mathcal{M}(S)} \hat{f}h_\epsilon d\tilde{\mu} + C(\delta)C_1(\eta)\epsilon^{\eta-\delta}.$$

Letting $\epsilon \to 0$ we get

$$\lim_{t\to\infty} \frac{1}{2\pi} \int_0^{2\pi} \hat{f}(a_t r_\theta S) d\theta = \lambda \int_{\mathcal{M}(S)} \hat{f} d\tilde{\mu}$$

as required to prove equation (1).

Now equation (1) implies the following smoothed version. For all S, $f \in C_0^\infty(\mathbb{R}_+^2)$ and non-negative $\phi \in C_0^\infty(\mathbb{R})$,

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

Now we apply Proposition 3.2 of [3]. This Proposition says that for any S satisfying the above smoothed version, Theorem 2.4 holds. (The Proposition also requires that certain growth conditions be satisfied, but these were proved in Theorem 5.4, Theorem 5.1 and Lemma 5.5 of [3]). $\square$

Remark. The methods of this section can be applied almost without modification in a the more general context of branched covers over a torus which are branched at $k > 2$ points. We do not do so here because the results are more cumbersome to state.

3 Cylinder decompositions

In this section we introduce coordinates on the spaces $\mathcal{M}_d(1,1)$ and $\mathcal{M}_d(2)$. These coordinates will be used throughout the paper.

3.1 The space $\mathcal{M}_d(1,1)$.

Recall that $S = (M, \omega) \in \mathcal{M}_d(1,1)$ if and only if S has area 1 and there exists a d-fold holomorphic covering map $\pi : M \to \mathfrak{t}$, where $\mathfrak{t}$ is a torus, and ω is the pullback of the differential dz on $\mathfrak{t}$. The covering π is branched of order 2 at the zeroes z_1 and z_2 of ω. Note that if (M, ω) has area 1, $(\mathfrak{t}, dz)$ has area $1/d$.

Let $\mathcal{T}^2$ denote the moduli space of tori with two marked points. Then we have the covering map $p : \mathcal{M}_d(1,1) \to \mathcal{T}^2$, sending (M, ω) to $(\pi(M), \pi(z_1), \pi(z_2))$, (see Lemma 2.2). We can assume $\pi(z_1)$ is the origin.

We may write $\mathfrak{t} = \mathbb{R}^2/\Delta$, where $\Delta \subset \mathbb{R}^2$ is a lattice. Let (v_1, v_2) be a reduced basis for the lattice; i.e. (v_1, v_2) is a basis for Δ such that $(v_1, v_2) \in \mathcal{F}$, where the fundamental domain $\mathcal{F}$ is given by

$$\mathcal{F} = \{(v_1, v_2) \in \mathbb{R}^2 \times \mathbb{R}^2 \ : \ \|v_1\| \leq \|v_2\|, \ |\langle v_1, v_2 \rangle| \leq \tfrac{1}{2}\langle v_1, v_1 \rangle, \ |v_1 \times v_2| = 1/d\}.$$

We always assume that $\pi(z_1)$ is the origin in $\mathbb{R}^2/\Delta$. We may write $\pi(z_2) = \delta_1 v_1 + \delta_2 v_2$, where $0 \leq \delta_1 < 1$, and $0 \leq \delta_2 < 1$. Hence the tuple $((v_1, v_2), \delta_1, \delta_2) \in \mathcal{F} \times [0,1) \times [0,1)$ can be used as a coordinate system on $\mathcal{T}^2$.

Let $(\mathcal{T}')^2$ denote the subset of $\mathcal{T}^2$ with $\delta_2 \neq 0$, and let $\mathcal{M}'_d(1,1)$ denote the surfaces in $\mathcal{M}_d(1,1)$ which project under p to $(\mathcal{T}')^2$. It is clear that $\mathcal{M}'_d(1,1)$ is an open and dense subset of $\mathcal{M}_d(1,1)$. Our coordinate system will only be valid on $\mathcal{M}'_d(1,1)$.

We define the direction of v_1 to be *horizontal* and the direction of v_2 to be *vertical*. In using these terms we are not assuming that the directions are perpendicular. However, all figures are drawn in the case where the horizontal and vertical directions are perpendicular. Recall that the Abelian differential ω is the pullback $\pi^* dz$ of the canonical differential dz on $\mathfrak{t}$. Consequently any closed direction on the base surface is a closed direction on M. In particular, the horizontal (and vertical) trajectories through any point of M are closed.

We now assume $(M, w) \in \mathcal{M}'_d(1,1)$. By taking the preimages (under π) of the two horizontal closed geodesics through the origin and $\pi(z_2)$ one obtains four saddle connections on (M, ω), with two connecting each z_i with itself. Because the horizontal direction on the base torus is closed, any horizontal trajectory away from the four saddle connections, is a closed geodesic on M. The closed geodesics occur in parallel families of the same length. Metrically they are right Euclidean cylinders whose boundary consists of saddle connections. Note that by construction, the length of these geodesics has to be an integer multiple of $\|v_1\|$. We refer to these lengths as the *widths* of the cylinders.

Lemma 3.1. *Suppose $(M, \omega) \in \mathcal{M}'_d(1,1)$. Then there are exactly three cylinders of closed trajectories in the horizontal direction.*

Proof. Each of the two singularities has a saddle connection returning to itself at angle π. Each of these saddles is the boundary of a cylinder. This accounts for 2 cylinders. Each boundary component of a third cylinder consists of the pair of these saddle connections returning to the singularity. □

Note that the width of this wider cylinder is the sum of the widths of the narrower ones. Let $w_1\|v_1\|, w_2\|v_1\|$ (where $w_i \in \mathbb{Z}$) be the widths of the small cylinders C_1 and C_2, and let $w_3\|v_1\| = (w_1 + w_2)\|v_1\|$ be the width of the wide cylinder C_3.

We note that the three cylinders may be assembled in two different ways. Either the two narrow cylinders are glued above z_1 (and wide cylinder is glued below z_1), in the sense that as we leave z_1 in the positive vertical direction we enter one of the narrow cylinders. The second possibility is that the wide cylinder is glued above z_1, and the narrow cylinders are glued below z_1. Let σ be a parameter which is $+1$ in the first case, and -1 in the second.

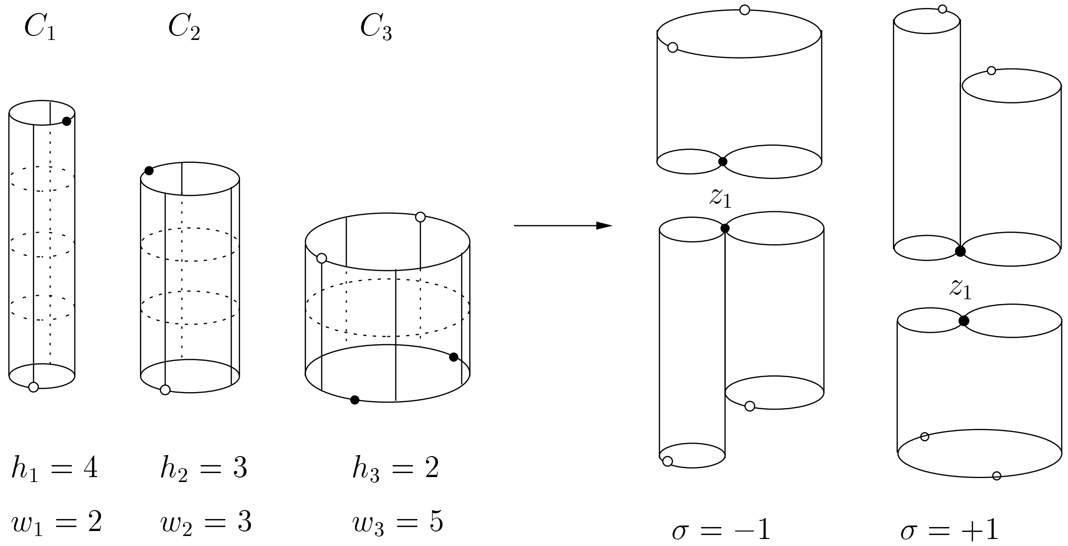

Figure 2. How to glue together three cylinders to obtain surfaces of genus 2.

Note that each narrow cylinder has a zero at both the top side and the bottom side. For each narrow cylinder, let z' denote the intersection with the top side of the vertical trajectory passing through the zero at the bottom side. The *twist* of a narrow cylinder is defined to be the (clockwise) horizontal distance along the top of the cylinder from z' to the zero at the top side. To define the twist for the wide cylinder C_3, choose the zero x on the bottom so that as we enter C_3 from that zero, the saddle connection on the boundary of C_1 lies to the left, and choose the zero x' on the top so that as we enter C_3, the saddle connection on the boundary of C_1 lies to the right. Again let z' denote the intersection with the top end of the vertical trajectory passing through x. The *twist* is defined to be the clockwise horizontal distance along the top of the cylinder from z' to x'. With this definition, if the twist is 0, then the two boundary components of C_1 lie directly opposite each other on C_3.

By construction, for $i = 1, 2$, the twist of the C_i can be written as $(t_i + \sigma\delta_1)\|v_1\|$, where $t_i \in \mathbb{Z}$, $0 \leq t_i < w_i$. The twist of the wide cylinder can be written as $(t_3 - \sigma\delta_1)\|v_1\|$, where $t_3 \in \mathbb{Z}$, $0 \leq t_3 < w_1 + w_2$.

We define the *height* of a cylinder to be the length of a vertical (i.e parallel to v_2) trajectory going from the bottom edge to the top edge. (This does not represent distance perpendicular to the horizontal). Similarly, the heights of the cylinder C_i are of the form $h_i\|v_2\|$ where the h_i satisfy $h_i \in \mathbb{Z} + \sigma\delta_2$ for $i = 1, 2$, and $h_3 \in \mathbb{Z} - \sigma\delta_2$. Clearly we must also have $h_i > 0$. We prefer to use the coordinates $s_1 = h_1 + h_3$, and $s_2 = h_2 + h_3$ and h_3 instead of (h_1, h_2, h_3). Then $1 \leq s_1 \in \mathbb{Z}$, $1 \leq s_2 \in \mathbb{Z}$, and $0 < h_3 < \min(s_1, s_2)$. (The upper bound on h_3 is determined by the conditions $h_i = s_i - h_3 > 0$, $i = 1, 2$).

Since the area of $(M, \omega) = 1$, the area of the base torus is given by $|v_1 \times v_2| = 1/d$. The area of C_i is $|w_i v_1 \times h_i v_2| = h_i w_i |v_1 \times v_2| = h_i w_i / d$, hence we have $w_1 h_1 + w_2 h_2 + w_3 h_3 = d$. Using $w_3 = w_1 + w_2$ and rewriting we get:

$$w_1 s_1 + w_2 s_2 = d \tag{3}$$

We also note that interchanging the coordinates (w_1, s_1, t_1) and (w_2, s_2, t_2) interchanges C_1 and C_2 and produces the same surface in $\mathcal{M}_d(1, 1)$. Summarizing this discussion we get the following:

Proposition 3.2. *Let Ω denote the set of $\{(v_1, v_2), \delta_1, \delta_2, \sigma, w_1, w_2, s_1, s_2, h_3, t_1, t_2, t_3\}$ where $(v_1, v_2) \in \mathcal{F}$, $0 \leq \delta_1 < 1$, $0 < \delta_2 < 1$, $\sigma \in \{+1, -1\}$, w_1, w_2, s_1 and s_2 are positive integers satisfying (3), h_3 is an element of $\mathbb{Z} - \sigma\delta_2$ satisfying $0 < h_3 < \min(s_1, s_2)$, and t_i are integers mod w_i, $i = 1, 2, 3$ (where $w_3 = w_1 + w_2$). Let $\tau : \Omega \to \Omega$ denote the map which exchanges (w_1, s_1, t_1) with (w_2, s_2, t_2) and fixes the other coordinates. Then there is a one to one map Ψ between Ω/τ and the open dense subset $\mathcal{M}'_d(1, 1)$ of $\mathcal{M}_d(1, 1)$. The surface $\Psi((v_1, v_2), \delta_1, \delta_2, \sigma, w_1, w_2, s_1, s_2, h_3, t_1, t_2, t_3)$ corresponding to a point of Ω has the following properties:*

- *It covers the torus $\mathbb{R}^2/\Delta$ where Δ is spanned by (v_1, v_2).*
- *The point z_1 projects to the origin; the point z_2 projects to $\delta_1 v_1 + \delta_2 v_2$.*
- *The surface is the union of three cylinders C_1, C_2 and C_3. For $i = 1, 2$, C_i has width $w_i\|v_1\|$, height $(s_i - h_3)\|v_2\|$ and twist $(t_i + \sigma\delta_1)\|v_1\|$. The cylinder C_3 has width $w_3 = w_1 + w_2$, height h_3 and twist $(t_3 - \sigma\delta_1)\|v_1\|$.*

- *Interchanging C_1 and C_2 produces the same surface in $\mathcal{M}_d(1,1)$.*

Corollary 3.3. *Let $N_d(1,1)$ denote the number of covers of a fixed torus of degree d with 2 fixed simple branch points. Then,*

$$N_d(1,1) = \sum_{s_1 w_1 + s_2 w_2 = d} w_1 w_2 (w_1 + w_2) \min(s_1, s_2) + \sum_{2sw=d} (2w^2 s). \tag{4}$$

Proof. We perform the sum over the t_i, h_3 and σ. In the first there is a factor of 2 from the 2 choices of σ which cancels the factor of $1/2$ from the action of τ. However, this does not count correctly the fixed points of τ, since we should count them with a factor of $1/2$. To compensate, we add the second term.

For similar results in a more general setting see [2], [1], [4], and also [23]. We note that here we do not weight each cover by the inverse of its automorphism group. □

Recall the definition of the measure ν_1 from the Introduction. We have the following: (see [4], [23])

Proposition 3.4. *For any stratum $\mathcal{H}(\beta)$,*

$$\nu_1(\mathcal{H}_1(\beta)) = n \lim_{D\to\infty} \frac{\sum_{d=1}^{D} N_d(\beta)}{D^{n/2}},$$

where $n = \dim_{\mathbb{R}} \mathcal{H}(\beta)$, and $N_d(\beta)$ is the degree of the covering map $p : \mathcal{M}_d(\beta) \to \mathcal{T}^{|\beta|}$ (i.e. the number of degree d covers of a fixed torus with ramification given by β).

Combining Corollary 3.3 with Proposition 3.4, we obtain

Lemma 3.5.

$$\nu_1(\mathcal{H}_1(1,1)) = \frac{\pi^4}{135}.$$

This calculation was first done, using the method presented in this paper by A. Zorich. (see [23]).

Proof of Lemma 3.5: See Appendix A. □

3.2 The Space $\mathcal{M}_d(2)$.

Now we have a covering map $p : \mathcal{M}_d(2) \to \mathcal{T}^1$, where $\mathcal{T}^1$ is the space of tori with one marked point, which we always place at the origin. As in §3.1, we may write the base torus as $\mathbb{R}^2/\Delta$, where the lattice Δ is spanned by $(v_1, v_2) \in \mathcal{F}$. Also as in §3.1, we consider the inverse image of the horizontal trajectory through the marked point at the origin; this breaks up the surface into cylinders.

Either there are one or two cylinders. We first consider the 2 cylinder case. There are a pair of homotopic saddle connections joining the zero to itself, bounding a cylinder C_1 of width $w_1\|v_1\|$. There is another saddle connection returning to the zero with angle 3π with length $w\|v_1\|$. Each side of it together with one of the previous saddles forms the boundary of a cylinder C_2 which then has width $w_2\|v_1\| = (w_1 + w)\|v_1\|$.

The height of C_i is $h_i\|v_2\|$ where $0 < h_i \in \mathbb{Z}$. Also h_1, h_2 satisfy

$$h_1 w_1 + h_2 w_2 = d. \tag{5}$$

The twists $t_i \in \mathbb{Z}$ satisfy $0 \leq t_i < w_i$. Hence

Proposition 3.6. *Let $\Omega_2 = \{((v_1, v_2), w_1, w_2, h_1, h_2, t_1, t_2)\}$ where $(v_1, v_2) \in \mathcal{F}$, w_i and h_i are positive integers satisfying (5), $w_1 < w_2$, and t_i are integers satisfying $0 < t_i < w_i$. Then there is a bijection Ψ_2 between the elements of Ω_2 and the surfaces of $\mathcal{M}_d(2)$ which have 2 cylinders. The surface $\Psi_2((v_1, v_2), w_1, w_2, h_1, h_2, t_1, t_2)$ covers the torus $\mathbb{R}^2/\Delta$ where Δ is the lattice spanned by v_1 and v_2. For $i = 1, 2$, the cylinder C_i has width $w_i\|v_1\|$, height $h_i\|v_2\|$ and twist $t_i\|v_1\|$.*

Now we consider the covers with one cylinder. There are three closed curves on each boundary component of the cylinder with lengths $l_1\|v_1\|, l_2\|v_1\|, l_3\|v_1\|$ such that $l_1 + l_2 + l_3 = w$ and $hw = d$, where h is the height. Cyclically permuting the l_i does not change the surface. Hence,

Proposition 3.7. *Let $\Omega_1 = \{((v_1, v_2), l_1, l_2, l_3, h)\}$ where $(v_1, v_2) \in \mathcal{F}$, l_1, l_2, l_3, h are positive integers satisfying $(l_1 + l_2 + l_3)h = d$. Let $\tau : \Omega_1 \to \Omega_1$ denote the map which cyclically permutes the l_i. Then there is a bijection Ψ_1 between Ω_1/τ and the points of $\mathcal{M}_d(2)$ with one cylinder. The surface $\Psi_1((v_1, v_2), l_1, l_2, l_3, h)$ covers the torus $\mathbb{R}^2/\Delta$ where Δ is the lattice spanned by v_1 and v_2. The single cylinder has height $h\|v_2\|$ and width $(l_1 + l_2 + l_3)\|v_1\|$.*

Corollary 3.8. *Let $N_d(2)$ denote the number of covers of a fixed torus with one double branch point. Then,*

$$N_d(2) = \sum_{\substack{h_1 w_1 + h_2 w_2 = d \\ w_1 < w_2}} w_1 w_2 + \frac{1}{3} \sum_{h|d} \sum_{l_1+l_2+l_3=\frac{d}{h}} \frac{d}{h} + \frac{2}{3} \sum_{3l=\frac{d}{h}} \frac{d}{h}. \tag{6}$$

The first term comes from the 2-cylinder surfaces, the second from the 1-cylinder surfaces with no symmetry, and the third from the 1-cylinder surfaces with $l_1 = l_2 = l_3$ (which have an extra $\mathbb{Z}_3$ symmetry).

Lemma 3.9.

$$\nu_1(\mathcal{H}_1(2)) = \frac{\pi^4}{120}.$$

Proof. This is done by computing the asymptotics of (6) as $d \to \infty$ and using Proposition 3.4. This again was done by A. Zorich, see [23]. The details of the calculation are in Appendix A. $\square$

3.3 The constants $s_i(1,1)$ and $c(1,1)$.

The results of [6] say that

$$s_1(1,1) = 3\frac{\nu_1(\mathcal{H}_1(2))}{\nu_1(\mathcal{H}_1(1,1))} = 3\frac{\pi^4/120}{\pi^4/135} = \frac{27}{8},$$

and

$$s_2(1,1) = \frac{1}{24}\frac{\nu_1(\mathcal{H}_1(\emptyset))\nu_1(\mathcal{H}_1(\emptyset))}{\nu_1(\mathcal{H}_1(1,1))} = \frac{5}{8}.$$

The results of [6] also say that

$$c(1,1) = 2\zeta(2)\frac{1}{3}\frac{\nu_1(\mathcal{H}_1(\emptyset))}{\nu_1(\mathcal{H}_1(1,1))} = 2\zeta(2)\frac{1}{3}\frac{\pi^2/3}{\pi^4/135} = 5.$$

(The numbers in [6] refer to cylinders of primitive geodesics; we need here to multiply by $\zeta(2)$ to count cylinders of imprimitive geodesics, and then again by 2 since when we count imprimitive geodesics, we count the cylinder in both directions).

4 Primitive Covers

We say that a cover $\pi : M \to \mathfrak{t}$ is *primitive* if it does not factor through any other torus cover (i.e. π has minimal degree under all possible torus coverings). This is to be contrasted with stable, introduced earlier, which says that there is no cover which factors through π mapping a pair of branch points to a single one. Clearly to compute the asymptotics of $N_c(S,T)$ and $N_s^{(i)}(S,T)$ we may assume that $\pi : M \to \mathfrak{t}$ is a primitive cover.

We recall the following general fact:

Lemma 4.1. *A cover of degree d of a surface of area d in $\mathcal{H}(1,1)$ over the standard torus is primitive if and only if the absolute homology generates the lattice $\mathbb{Z} \oplus \mathbb{Z}$.*

Lemma 4.2. *Using the coordinates of Proposition 3.2, the surface $\Psi((v_1, v_2), \delta_1, \delta_2, \sigma, w_1, w_2, s_1, s_2, h_3, t_1, t_2, t_3)$ is a primitive cover of the torus $\mathbb{R}^2/(\mathbb{Z}v_1 + \mathbb{Z}v_2)$ if and only if:*

$$(s_1, s_2) = 1$$

and

$$(s_1(t_2 + t_3) - s_2(t_1 + t_3), w_1, w_2) = 1.$$

Proof. Without loss of generality, we may assume that $v_1 = (1,0)$, $v_2 = (0,1)$. Let C_1, C_2, C_3 be as in Proposition 3.2. Recall the distinguished zeroes x, x' on the bottom and top of C_3 used to define the twist t_3. We may take a basis for the absolute homology as follows:

- a_1 is the core horizontal curve of C_1.
- a_2 is the core horizontal curve of C_2.
- b_1 is the curve transverse to the horizontal foliation which starts at the zero on the bottom of C_1, crosses C_1 to the zero on the top and then crosses C_3 from x on the bottom to x' on the top.
- b_2 is the curve transverse to the horizontal foliation which starts at the zero on the bottom of C_2, crosses C_2 to the zero on the top and then crosses C_3 from x on the bottom to x' on the top.

This absolute homology basis determines holonomy vectors:

$$(w_1, 0), (w_2, 0), (t_1 + t_3, s_1), (t_2 + t_3, s_2),$$

(recall that $s_1 = h_1 + h_3$, $s_2 = h_2 + h_3$).

Let $v = s_1(t_2 + t_3) - s_2(t_1 + t_3)$. In order to produce a lattice element with second coordinate of 1 we must have $(s_1, s_2) = 1$. Let $r = \gcd(w_1, w_2)$. Every lattice element of the form $(d, 0)$ in the subgroup generated by the third and fourth element satisfies $v|d$. Thus in order to generate $(1, 0)$ it is necessary and sufficient that $(v, r) = 1$. $\square$

Corollary 4.3. *The surface $S(p/q, \alpha)$ is a primitive cover.*

4.1 Connectedness of the space of primitive covers

Let $\mathcal{P}_d(\beta) \subset \mathcal{M}_d(\beta)$ denote the primitive torus covers in $\mathcal{M}_d(\beta)$. The main result of the section is the following:

Theorem 4.4. *$\mathcal{P}_d(1, 1)$ is connected.*

Theorem 1.1 now follows from Theorem 2.4, Corollary 4.3 and Theorem 4.4.

Proof of Theorem 4.4. In this section we will assume that the surface has area d, and the base torus has area 1. First note that primitivity is invariant under continuous deformations. The condition of primitivity of degree d is open: if $(M_n, \omega_n) \to (M, \omega)$ is a sequence of degree d covers converging to a primitive cover, then if (M_n, ω_n) factored through a cover of smaller degree, then so would (M, ω). The condition of primitivity is also closed, since the conditions of Lemma 4.2 are clearly closed conditions. We use the coordinates of Proposition 3.2. For any point S_1 in $\mathcal{P}_d(1, 1)$, We can clearly continuously deform the base lattice (v_1, v_2) until $v_1 = (1, 0)$, $v_2 = (0, 1)$, i.e. the base torus is the standard torus. We can also continuously deform (δ_1, δ_2) to some fixed point (ϵ, ϵ). Thus any surface is connected by a continuous path to a cover $\pi : S \to \mathbb{T}^2$, such that $\pi(z_1) = 0$, $\pi(z_2) = (\epsilon, \epsilon)$. Recall that we have a covering map $p : \mathcal{P}_d(1, 1) \to \mathcal{T}^2$, where $\mathcal{T}^2 \cong SL(2, \mathbb{R}) \ltimes \mathbb{R}^2 / SL(2, \mathbb{Z}) \ltimes \mathbb{Z}^2$ is the space of tori with two marked points, one of which is always at the origin. Hence any point in $\mathcal{P}_d(1, 1)$ can be connected by a continuous path to a point of the $p^{-1}(\mathbb{T}^2, (\epsilon, \epsilon))$ (i.e. the fiber of p above the point $(\mathbb{T}^2, (\epsilon, \epsilon)) \in \mathcal{T}^2$). We will now show that any two points in $p^{-1}(\mathbb{T}^2, (\epsilon, \epsilon))$ can be connected by a continuous path in $\mathcal{P}_d(1, 1)$. Note that the points of $p^{-1}(\mathbb{T}^2, (\epsilon, \epsilon))$ are parametrized by the discrete parameters $(w_1, w_2, s_1, s_2, h_3, t_1, t_2, t_3, \sigma)$ of Proposition 3.2.

We now consider the following three kinds of continuous paths connecting points in $p^{-1}(\mathbb{T}^2, (\epsilon, \epsilon))$:

The horizontal kernel foliation. Let $\gamma_h : \mathcal{T}^2 \to \mathcal{T}^2$ be the path $\gamma_h(t) = (\mathbb{T}^2, (\epsilon + t, \epsilon))$, so that $\gamma_h(0) = \gamma_h(1) = (\mathbb{T}^2, (\epsilon, \epsilon))$. Since the covering p is unbranched away from the points in $\mathcal{T}^2$ where the two marked points coincide, γ_h lifts to a path $\tilde{\gamma_h}$ on $\mathcal{P}_d(1,1)$ connecting two points of the fiber $p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon))$. Let $F_h : p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon)) \to p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon))$ denote the map taking $\tilde{\gamma_h}(0)$ to $\tilde{\gamma_h}(1)$. It is clear that F_h preserves the leaves of the horizontal foliations, hence it preserves the heights and widths of the cylinders, i.e F_h fixes the parameters $w_1, w_2, s_1, s_2, h_3, \sigma$. The twists are changed by $\pm||v_1||$ (and thus the t_i are changed by ± 1), with the sign determined by σ. More precisely, $F_h(\dots,t_1,t_2,t_3,\sigma) = (\dots,t_1+\sigma,t_2+\sigma,t_3-\sigma,\sigma)$.

The vertical kernel foliation. Let $\gamma_v : \mathcal{T}^2 \to \mathcal{T}^2$ be the path $\gamma_v(t) = (\mathbb{T}^2,(\epsilon,\epsilon+t))$, so that $\gamma_v(0) = \gamma_v(1) = (\mathbb{T}^2,(\epsilon,\epsilon))$. Let $\tilde{\gamma_v}$ denote any lift of γ_v to $\mathcal{P}_d(1,1)$, and let $F_v : p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon)) \to p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon))$ denote the map sending $\tilde{\gamma_v}(0)$ to $\tilde{\gamma_v}(1)$.

Without loss of generality we assume $w_1 < w_2$. Suppose first that $\sigma = +1$. Then as long as $h_3 > 1$, F_v sends h_3 to $h_3 - 1$, sends h_i to $h_i + 1, i = 1,2$ and fixes all the other coordinates. In other words the heights of two narrow cylinders increase by 1, and the height of the wide cylinder decreases by 1. If $h_3 < 1$ then we completely collapse the wide cylinder, and the cylinder decomposition may change. There are in fact two possibilities, depending on the values of the twist parameters. If (when the wide cylinder has height 0) the top edge of C_1 touches the bottom edge of C_1, then the new cylinders have width (w_1, w_2, w_1+w_2) as before, except that we switched σ to $-\sigma$, and h_3 becomes ϵ. Also h_1 goes from $n+\epsilon$ to $n+1-\epsilon$ for some integer n and similarly for h_2.

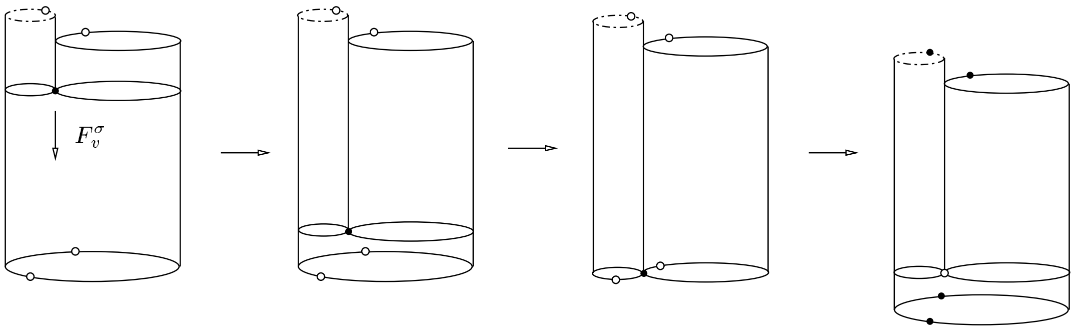

Figure 3. Applying repeatedly F_v^σ does not change the cylinder decomposition if $t_3 - \sigma\delta_1 < w_1$, but σ changes sign.

Suppose on the other hand the bottom edge of C_1 is disjoint from the top edge of C_1. this corresponds to the twist satisfying $t_3 - \sigma\delta_1 > w_1$. Then at the stage at which $h_3 = 0$, the surface decomposes into cylinders C_1 and C_2. Each boundary component of C_1 is the single saddle returning to a zero, while each boundary component of C_2

consists of one of these saddles as well as a pair of saddle connections joining the two zeroes. Then as we continue the deformation, C_2 becomes the wide cylinder and the cylinders have widths $(w_2 - w_1, w_1, w_2)$, (see Figure 4). In this case σ stays $+1$ and if original narrow cylinders had heights $h_1 = n_1 + \epsilon$ and $h_2 = n_2 + \epsilon$ and the wide cylinder had height $h_3 = 1 - \epsilon$, then after deformation, the heights are $h_1 = n_1 + 1 + \epsilon$, $h_2 = \epsilon$ and $h_3 = n_2 + 1 - \epsilon$.

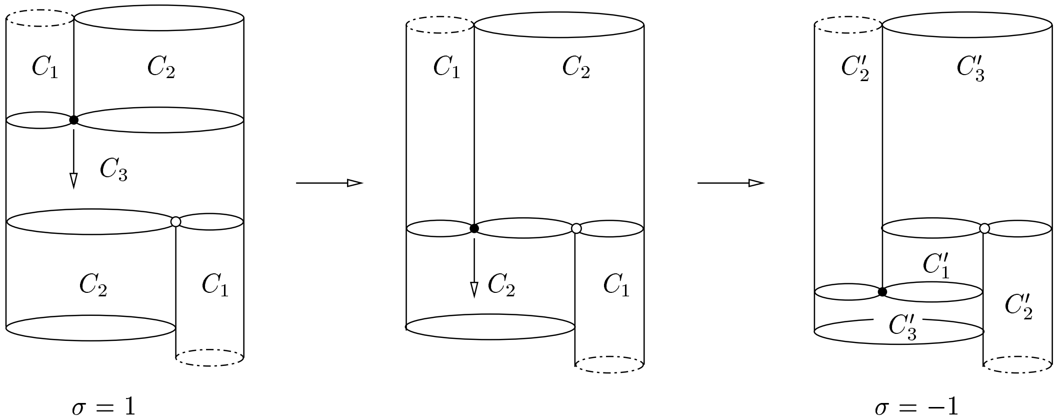

Figure 4. Applying repeatedly F_v^σ changes the cylinder decomposition and σ, if $t_3 - \sigma\delta_1 > w_1$.

In the case $\sigma = -1$, the same analysis applies with F_v replaced by F_v^{-1}.

The $SL(2,\mathbb{Z})$ action. Given an element $g \in SL(2,\mathbb{Z})$ we now define a map $F_g : p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon)) \to p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon))$. There is a slight technical complication since the action of $g \in SL(2,\mathbb{Z}) \subset SL(2,\mathbb{R})$ preserves $\mathbb{T}^2$ but not $(\epsilon,\epsilon) \in \mathbb{T}^2$. We solve this as follows: let (ϵ',ϵ') be a point in $\mathbb{R}^2$ such that $\|g(\epsilon',\epsilon')\| < 1/3$; then let $\gamma_1(t) = (\mathbb{T}^2,(1-t)(\epsilon,\epsilon)+t(\epsilon',\epsilon'))$, $\gamma_2(t) = g(t)(\mathbb{T},(\epsilon',\epsilon'))$ where $g(t) \in SL(2,\mathbb{R})$ is a path connecting the identity to $g \in SL(2,\mathbb{Z})$, and $\gamma_3(t) = (\mathbb{T}^2,(1-t)g(\epsilon',\epsilon')+t(\epsilon,\epsilon))$. Then let γ be the composition of γ_1, γ_2 and γ_3; this is a closed path on $\mathcal{T}^2$. As above, let $\tilde{\gamma}$ be any lift of γ, and let $F_g : p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon)) \to p^{-1}(\mathbb{T}^2,(\epsilon,\epsilon))$ denote the map sending $\tilde{\gamma}(0)$ to $\tilde{\gamma}(1)$. (A priori the map F_g may depend on some choices; however it will be independent of the choices in the case where we use it).

We now continue the proof of Theorem 4.4. If $w_1 < w_2$, then by applying a suitable power of F_h we may achieve the situation that if the height of the wide cylinder is collapsed to 0, the top and bottom edge of C_1 are disjoint. Specifically, we apply F_h so that $t_3 - \sigma\delta_1 > w_1$. Then we apply F_v^σ until the cylinder decomposition changes.

We obtain a surface with three narrower cylinders. We can now repeat this procedure until $w_1 = w_2$. Thus after doing the moves repeatedly one can assume that the given torus covering has two cylinders of equal width w.

Assuming this, we proceed by applying a power of F_h so that any vertical trajectory moves from the first cylinder into the second one (the cylinders are of equal width now) and the other way round after returning to the third cylinder. After developing the surface in the plane as in Figure 5, we obtain a parallelogram R of width w and area d (see the right of Figure 5) and opposite sides identified, together with two slits whose endpoints are the singular points. That is, the new surface consists of a torus $\mathbb{R}^2/L$ with 2 slits, (where $L \subset \mathbb{Z} \oplus \mathbb{Z}$ is a lattice). We can assume the base of one slit is at the origin and the base of the other at an integer point. After applying a suitable element of the vertical kernel foliation, we can arrange for both slits have length $\epsilon\sqrt{2}$.

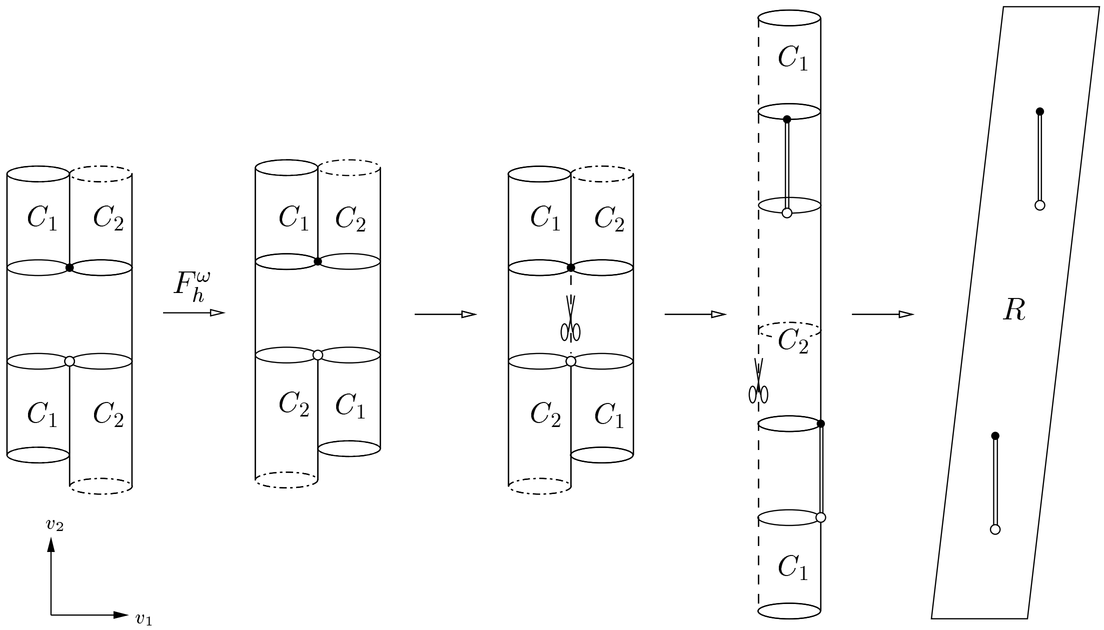

Figure 5. Deform and develop a given surface with two narrow cylinders of equal width. The twist in the torus on the right depends on the twists of the cylinders C_1 and C_2 (not shown on the left).

Lemma 4.5. *Let P_n denote the set of 2 by 2 integral matrices of determinant n. Then representatives for the double cosets $SL(2,\mathbb{Z})\backslash P_n/SL(2,\mathbb{Z})$ consist of the diagonal matrices $\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$ where $d_1 d_2 = n$ and $d_1 | d_2$.*

Proof. This is the well known elementary divisor theorem, see [12]. $\square$

Corollary 4.6. *Let L be a sublattice of $\mathbb{Z} \oplus \mathbb{Z}$ of index n. Then there exists $g \in SL(2,\mathbb{Z})$ such that $L' = gL$ is a lattice such that the fundamental domain for L' is a rectangle with the length of the short side dividing the length of the long side and the product of side lengths being n.*

Proof of corollary: This is just Lemma 4.5 in a different language. $\square$

We now apply the map F_g, where $g \in SL(2,\mathbb{Z})$ is as in Corollary 4.6. (We note that in this situation, F_g does not depend on the choice of paths, because the effect of a different path is to rotate the slit by an angle of 2π.) Let S' denote the resulting surface. Then S' is a rectangular torus $\mathbb{R}^2/L'$, with two slits, one of which is assumed to be at the origin. Hence the one slit is from $(0,0)$ to (ϵ,ϵ), and the other is from (u_1,u_2) to $(u_1+\epsilon, u_2+\epsilon)$, where $u_i \in \mathbb{Z}$. (see Figure 6).

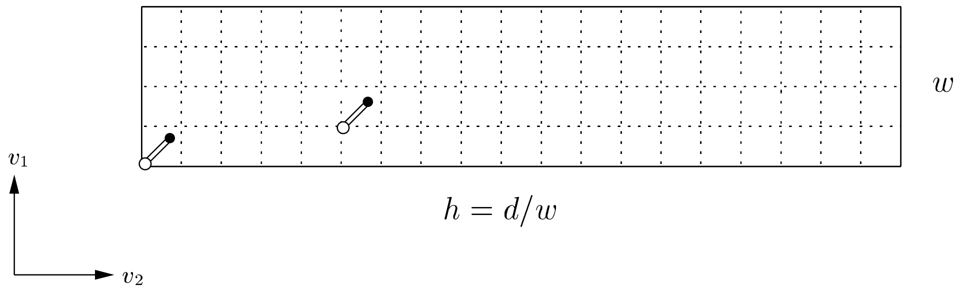

Figure 6. The torus $\mathbb{R}^2/L'$ with two slits developed in $\mathbb{R}^2$.

Since S' was obtained from the original surface S by a continuous path, and S was assumed primitive, S' is also primitive. Let $u = (u_1,u_2)$ be the vector connecting the 2 slits. Thus the absolute homology of the surface is generated by L' and u and the slit curve. The holonomy along the two sides of the slit curve add to 0. Thus, by Lemma 4.1, $(d_1,0)$, $(0,d_2)$ and u generate $\mathbb{Z} \oplus \mathbb{Z}$. But $d_1|d_2$. This is a contradiction unless $d_1 = 1$. Indeed if $d_1 > 1$, reducing modulo d_1 we get that $u \pmod{d_1}$ generates the finite group $(\mathbb{Z}/d_1\mathbb{Z}) \oplus (\mathbb{Z}/d_1\mathbb{Z})$: but that finite group is not generated by any one element.

So $d_1 = 1$, and hence the surface S' consists of a vertical strip of width 1, height d with 2 slits, one from $(0,0)$ to (ϵ, ϵ) and one from $(0, u_2)$ to $(\epsilon, u_2 + \epsilon)$. Since S' is primitive, $(1,0)$, $(0, u_2)$, and $(0, d)$ generate $\mathbb{Z} \oplus \mathbb{Z}$, hence u_2 and d are relatively prime. Hence there exists $k \in \mathbb{Z}$ such that $ku_2 = 1 \pmod{d}$. Let n such that $ku_2 - 1 = nd$. Now after we apply F_g where $g = \begin{pmatrix} u_2 & 1 \\ nd & k \end{pmatrix}$ we get a surface S_0 which is again $\mathbb{R}^2/L'$ with two slits, one from $(0,0)$ to (ϵ, ϵ) and one from $(0,1)$ to $(\epsilon, 1+\epsilon)$. Hence we have joined any primitive cover to a fixed cover S_0 by a continuous path of primitive covers. This completes the proof of Theorem 4.4. $\square$

4.2 Counting primitive covers

Let $N_d^P(1,1)$ denote the number of primitive covers of degree d of a surface of genus 2 that are branched over 2 points of the standard torus and $N_d^P(2)$, the number of primitive covers of degree d of a surface of genus 2 branched over a single point.

Let v be the quantity

$$s_2(t_1 + t_3) - (t_2 + t_3)s_1 = (t_1 + t_3, s_1) \times (t_2 + t_3, s_2).$$

Lemma 4.7. *Assume $(s_1, s_2) = 1$. For fixed w_1, w_2 the number of (t_1, t_2, t_3) satisfying $0 \le t_1 < w_1$, $0 \le t_2 < w_2$, $0 \le t_3 < w_1 + w_2$, $r|w_1, r|w_2, r|v$ is*

$$\frac{w_1 w_2 (w_1 + w_2)}{r}.$$

Proof. Define a linear map

$$L : \mathbb{Z}/r\mathbb{Z} \times \mathbb{Z}/r\mathbb{Z} \times \mathbb{Z}/r\mathbb{Z} \to \mathbb{Z}/r\mathbb{Z}$$

by

$$L(t_1, t_2, t_3) = v \ mod(r).$$

Since $(s_1, s_2) = 1$, the map L is onto. Thus $|kerL| = r^2$. Now dividing the intervals $[0, w_1]$, $[0, w_2]$, $[0, w_1 + w_2]$ into subintervals of length r, in each triple of subintervals, we have exactly r^2 solutions of $r|v$. Since there are $w_1 w_2 (w_1 + w_2)/r^3$ triples of intervals, the lemma follows. $\square$

Lemma 4.8. *We have $N_2^P(1,1) = 4$. If $d \geq 3$,*

$$N_d^P(1,1) = \sum_{r|d} \mu(r) \left(\sum_{\substack{(s_1,s_2)=1 \\ s_1u_1+s_2u_2=\frac{d}{r}}} r^2 u_1 u_2 (u_1+u_2) \min(s_1,s_2) \right), \tag{7}$$

where $\mu(\cdot)$ is the Möbius function.

Proof. We use the coordinates of Proposition 3.2, with $v_1 = (1,0)$, $v_2 = (0,1)$.

Note that in the proposition the torus was not standard, while in counting covers, we can assume that the torus is standard. Since we are counting primitive covers, we can in view of Lemma 4.2 from now on assume

$$(s_1, s_2) = 1$$

By Lemma 4.7 the number of covers that satisfy $r|w_1, r|w_2, r|v$ is

$$\sum_{\substack{s_1w_1+s_2w_2=d \\ r|w_1, r|w_2}} \frac{w_1w_2(w_1+w_2)}{r} \min(s_1,s_2) + \sum_{r|w} \sum_{2w=d} 2w^2$$

(where the last term is coming from the fixed points of τ. Indeed $(s_1, s_2) = 1$ and $s_1 = s_2$ together imply $s_1 = s_2 = 1$, and since $t_1 = t_2$, $v = 0$. This means that there is no condition on the t_i other than $0 \leq t_i < w_i$. Also $w_1 = w_2 = w$, hence $d = 2w$. Thus the number of fixed points of τ satisfying $(s_1, s_2) = 1$ and $r|w_1$, $r|w_2$, $r|v$ is $2w^2 = \frac{d^2}{2}$ if $r|\frac{d}{2}$ and 0 otherwise.)

After substituting $w_i = ru_i$ the sum becomes

$$\left(\sum_{\substack{(s_1,s_2)=1 \\ s_1u_1+s_2u_2=\frac{d}{r}}} r^2 u_1 u_2 (u_1+u_2) \min(s_1,s_2) \right) + \chi(r,d)\frac{d^2}{2} \tag{8}$$

where $\chi(r,d) = 1$ if $r|\frac{d}{2}$ and $\chi(r,d) = 0$ otherwise. Now, using the Möbius inversion formula, we get (7). (Note that the fact that the contribution of the fixed points of τ cancels can be seen directly since there are no primitive covers which are fixed points of τ unless $d = 2$.) $\square$

Lemma 4.9. *As $d \to \infty$,*

$$N_d^P(1,1) = \frac{d^4}{3} \sum_{r|d} \frac{\mu(r)}{r^2} + o(d^4).$$

Proof. See Appendix A. $\square$

Remark: In fact it can be derived from the results of [2] and [1] that

$$N_d^P(1,1) = \frac{1}{3} d^3(d-1) \sum_{r|d} \frac{\mu(r)}{r^2}.$$

Lemma 4.10. *Let $N_d^P(2)$ denote the number of primitive genus 2 covers of the standard torus branched over a single point. Then, $N_2^P(2) = 0$, $N_3^P(2) = 3$ and for $d \geq 4$,*

$$N_d^P(2) = \sum_{r|d} \mu(r) \left(\sum_{\substack{h_1u_1+h_2u_2=d/r \\ (h_1,h_2)=1 \\ u_1<u_2}} ru_1u_2 + \frac{1}{3} \sum_{u_1+u_2+u_3=d/r} d \right). \tag{9}$$

Proof. We begin with counting the number of primitive covers with 2 cylinders in $\mathcal{H}(2)$, using the coordinates of Proposition 3.6. Again we are counting covers over the standard torus. For the cover to be primitive we must have $(h_1, h_2) = 1$, which we now assume. We must also have $(w_1, w_2, (t_1, h_1) \times (t_2, h_2)) = 1$. As in the lemma, the number of twists such that $r|w_1, r|w_2, r|(t_1h_2 - t_2h_1)$ is

$$\frac{w_1w_2}{r}.$$

Thus the number of primitive covers of degree d with 2 cylinders is:

$$\sum_{r|d} \mu(r) \sum_{\substack{h_1u_1+h_2u_2=d/r \\ (h_1,h_2)=1 \\ u_1<u_2}} ru_1u_2. \tag{10}$$

Note that there are no covers in degree 2.

We do the same counting of the number of primitive covers with one cylinder, using the coordinates of Proposition 3.7. We must have $h = 1$ in order for the cover to be primitive. As before we make this assumption. We then have $d = w = l_1+l_2+l_3$, where the l_i are the lengths of the curves. For the cover to be primitive we must also have $(l_1, l_2, l_3) = 1$. Then arguing as before, using the symmetry of the l_i, the number of primitive covers is for $d \geq 4$,

$$\frac{1}{3}\sum_{r|d}\mu(r)\sum_{u_1+u_2+u_3=d/r} d. \tag{11}$$

(Note that there are no primitive covers with $l_1 = l_2 = l_3$ unless $d = 3$). Now the lemma follows immediately from (10) and (11). □

Lemma 4.11. *As $d \to \infty$,*

$$N_d^P(2) = \frac{3}{8}d^4\sum_{r|d}\frac{\mu(r)}{r^2} + o(d^3).$$

Proof. See Appendix A. □

Remark: In fact it can be shown that

$$N_d^P(2) = \frac{3}{8}d^2(d-2)\sum_{r|d}\frac{\mu(r)}{r^2}.$$

4.3 The constant $s_1(d)$.

We normalize the measure μ on the space of tori $SL(2,\mathbb{R})/SL(2,\mathbb{Z})$ to be 1. The measure $\tilde{\mu}$ of $P_d(2)$ is then given by the degree of the covering map $p : P_d(2) \to SL(2,\mathbb{R})/SL(2,\mathbb{Z})$ which is $N_d^P(2)$. Moreover for a surface of area 1 which is a d fold cover of a torus, the torus has area $\frac{1}{d}$. Thus with this normalization the measure μ of $\mathcal{T}^2$ is $1/d$. Again the measure $\tilde{\mu}$ of $P_d(1,1)$ is multiplied by the degree of the corresponding covering map. Altogether we have

$$\tilde{\mu}(P_d(1,1)) = \frac{1}{d}N_d^P(1,1) \tag{12}$$

while

$$\tilde{\mu}(P_d(2)) = N_d^P(2).$$

Given a surface in $\mathcal{H}(1,1)$ and a saddle connection γ joining the two zeroes of small length, if there is no saddle connection homologous to γ, we can deform the surface by letting the length of the saddle connection go to 0. The resulting surface lies in

$\mathcal{H}(2)$. We keep the absolute homology constant during the deformation. Conversely, given a surface in $\mathcal{H}(2)$, and a vector γ of length ϵ, we can break up the double zero into two simple zeroes together with a saddle connection joining the two zeroes such that the holonomy of the flat structure along the saddle is γ. Because the total angle at the double zero is 6π, there are in fact 3 ways of doing this (we refer the reader to [6] for details). Since this process preserves the absolute homology, the next lemma is immediate:

Lemma 4.12. *Given a surface in $P_d(1,1)$ the surface given by collapsing a single saddle connection lies in $P_d(2)$. Conversely, given a surface in $S \in P_d(2)$, and a short vector $\gamma \in \mathbb{R}^2$, there are 3 surfaces with two simple zeroes joined by a single saddle connection of holonomy γ, which have the property that after collapsing γ we get S. All three surfaces lie in $P_d(1,1)$ (because the absolute homology is preserved).*

Proposition 4.13. *The constant $s_1(d)$ in Theorem 1.1 can be expressed in terms of the number of d-fold branched covers of the torus with prescribed branching. In fact, we have,*

$$s_1(d) = 3\frac{dN_d^P(2)}{N_d^P(1,1)}$$

where $N_d^P(2)$ is the number of d-fold primitive covers of a surface of genus 2 over a torus, branched over a single point, and $N_d^P(1,1)$ is the number of d-fold primitive covers over the torus with simple order 2 branching at two distinct points.

Proof of Proposition 4.13. We let f be the characteristic function of a disc of radius ϵ in $\mathbb{R}^2$. The Siegel Veech formula and the connectedness of $P_d(1,1)$ (Theorem 4.4) says that

$$s_1(d)\pi\epsilon^2 = \frac{\int_{\mathcal{P}_d(1,1)} \hat{f}\, d\tilde{\mu}}{\tilde{\mu}(P_d(1,1))}. \tag{13}$$

Now Lemma 4.12 says that the numerator is $3\pi\epsilon^2\tilde{\mu}(P_d(2))$ and so

$$s_1(d) = 3\frac{\tilde{\mu}(P_d(2))}{\tilde{\mu}(P_d(1,1))}$$

which by (12) is

$$3\frac{dN_d^P(2)}{N_d^P(1,1)},$$

where $N_d^P(2)$ is given by (9) and $N_d^P(1,1)$ is given by (7).

Proof of Theorem 1.3 for $s_1(d)$. In view of Lemma 4.11, Lemma 4.9, and §3.3,

$$\lim_{d\to\infty} s_1(d) = \lim_{d\to\infty} 3\frac{dN_d^P(2)}{N_d^P(1,1)} = 3(\frac{3/8}{1/3}) = \frac{27}{8} = s_1(1,1).$$

□

4.4 The constant $c(d)$.

Proposition 4.14. *For $d \geq 3$, we have*

$$c(d) = \frac{d}{N_d^P(1,1)} \sum_{r|d} \mu(r) \sum_{\substack{(s_1,s_2)=1 \\ s_1u_1+s_2u_2=\frac{d}{r}}} u_1u_2(u_1+u_2) \min(s_1, s_2) \left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \frac{1}{(u_1+u_2)^2}\right), \tag{14}$$

where $N_d^P(1,1)$ is given by (7),

Proof. We let f be the characteristic function of a disc of radius ϵ in $\mathbb{R}^2$. The Siegel Veech formula says that:

$$c(d)\pi\epsilon^2 = \frac{\int_{\mathcal{P}_d(1,1)} \hat{f}\, d\tilde{\mu}}{\tilde{\mu}(P_d(1,1))} = \zeta(2)\frac{\int_{\mathcal{P}_d(1,1)} \tilde{f}\, d\tilde{\mu}}{\tilde{\mu}(P_d(1,1))}. \tag{15}$$

Now $\hat{f}$ counts the number of cylinders of (imprimitive) closed geodesics with length at most ϵ, and $\tilde{f}$ counts the number of primitive cylinders of closed geodesics of length at most ϵ. (The relation

$$\zeta(2)\int_{\mathcal{P}_d(1,1)} \tilde{f}\, d\tilde{\mu} = \int_{\mathcal{P}_d(1,1)} \hat{f}\, d\tilde{\mu}$$

was proved in a more general setting in [3], see the remark following the proof of Theorem 5.1(b)).

We wish to evaluate the numerator in the above expression. As above, we use the coordinates of Proposition 3.2. Clearly if $\tilde{f}(S) \neq 0$, S contains a closed curve of length at most ϵ. Since s_1, s_2 are positive integers and $\|v_2\| \geq \|v_1\|$, there is a uniform lower bound on the length of any curve which is is not parallel to v_1. Hence, for ϵ sufficiently small, the three candidates curves of length at most ϵ are the widths of the three cylinders. Hence, (assuming $w_1 \le w_2$) for sufficiently small ϵ,

$$\tilde{f}(S) = \begin{cases} 0, & \text{if } w_1\|v_1\| > \epsilon \\ 1, & \text{if } w_2\|v_1\| > \epsilon > w_1\|v_1\| \\ 2, & \text{if } (w_1+w_2)\|v_1\| > \epsilon > w_2\|v_1\| \\ 3, & \text{if } \epsilon > (w_1+w_2)\|v_1\| \end{cases}$$

Let $\chi : \mathbb{R}^2 \to \mathbb{R}$ be defined by:

$$\chi(v) = \begin{cases} 0, & \text{if } w_1\|v\| > \epsilon\sqrt{d} \\ 1, & \text{if } w_2\|v\| > \epsilon\sqrt{d} > w_1\|v\| \\ 2, & \text{if } (w_1+w_2)\|v\| > \epsilon\sqrt{d} > w_2\|v\| \\ 3, & \text{if } \epsilon\sqrt{d} > (w_1+w_2)\|v\| \end{cases}$$

Then

$$\int_{\mathbb{R}^2} \chi(v)\, dv = \pi\epsilon^2 d \left(\frac{1}{w_1^2} + \frac{1}{w_2^2} + \frac{1}{(w_1+w_2)^2} \right). \tag{16}$$

For a unimodular lattice $\Delta \subset \mathbb{R}^2$, let Δ' denote the primitive vectors in Δ, and let $\tilde{\chi}(\Delta) = \sum_{v\in\Delta'} \chi(v)$. Then, by the Siegel formula,

$$\int_{SL(2,\mathbb{R})/SL(2,\mathbb{Z})} \tilde{\chi}(\Delta)\, d\mu(\Delta) = \frac{1}{\zeta(2)} \int_{\mathbb{R}^2} \chi(v)\, dv \tag{17}$$

and μ is normalized Haar measure on the space of unimodular tori $SL(2,\mathbb{R})/SL(2,\mathbb{Z})$. Note that if we fix all the parameters except (v_1, v_2),

$$\tilde{f}((v_1,v_2),\dots) = \tilde{\chi}(\mathbb{Z}v_1\sqrt{d} \oplus \mathbb{Z}v_2\sqrt{d}) \tag{18}$$

and $\mathbb{Z}v_1\sqrt{d} \oplus \mathbb{Z}v_2\sqrt{d}$ is a unimodular lattice. We now compute $\int_{\mathcal{P}_d(1,1)} \tilde{f}(S)\, d\nu_1(S)$ by parametrising S as in Proposition 3.2 and Lemma 4.8, and computing the integral over (v_1, v_2) first. Then, in view of (18), (17), and (16) we obtain

$$\int_{\mathcal{P}_d(1,1)} \tilde{f}(S)\, d\nu_1(S) =$$

$$= \frac{\pi\epsilon^2 d}{d\zeta(2)} \sum_{r|d} \mu(r) \sum_{\substack{(s_1,s_2)=1 \\ s_1u_1+s_2u_2=\frac{d}{r}}} u_1u_2(u_1+u_2)\min(s_1,s_2) \left(\frac{1}{u_1^2} + \frac{1}{u_2^2} + \frac{1}{(u_1+u_2)^2} \right) \tag{19}$$

where in the sum $w_i = ru_i$, and the factor of d in the denominator comes from the integral over the δ_i in Proposition 3.2. Now the lemma follows from (19), (15) and (12). $\square$

The proof of Theorem 1.3 for $c(d)$ is postponed until Appendix A.

4.5 The constant $s_2(d)$.

Proposition 4.15. *Assume $d \geq 3$. We have*

$$s_2(d) = \frac{d}{N_d^P(1,1)} \left[\sum_{r|d} \mu(r) \left(\sum_{\substack{h_1u_1+h_2u_2=d/r \\ (h_1,h_2)=1 \\ u_1<u_2}} ru_1u_2 \right) + \sum_{\substack{w|d \\ w\neq d}} \varphi(d/w) \sum_{r|w} \frac{\mu(r)}{r} w^2 + \frac{d\varphi(d)}{2} \right] \tag{20}$$

where $N_d^P(1,1)$ is given by (7), and φ is the Euler function.

Proof. We let f be the characteristic function of a disc of radius ϵ in $\mathbb{R}^2$. The Siegel-Veech formula says that:

$$s_2(d)\pi\epsilon^2 = \frac{\int_{\mathcal{P}_d(1,1)} \hat{f}\, d\tilde{\mu}}{\tilde{\mu}(P_d(1,1))}. \tag{21}$$

Now $\hat{f}$ counts the number of saddle connections of multiplicity 2 of length at most ϵ. We wish to evaluate the asymptotics of the numerator as $\epsilon \to 0$.

We use the coordinates of Proposition 3.2, except we choose to work with surfaces of area d (which cover a torus of area 1). We first show that the contribution to the integral from the part where the base torus is degenerate has lower order $O(\epsilon^4)$. For suppose $||v_1|| \leq \epsilon$ so that the measure of the set of such v_1 is $O(\epsilon^2)$. The saddle connections cannot be horizontal for they would bound cylinders. Thus it must happen that each crosses a cylinder which means that the distance across a cylinder must be bounded by ϵ. Since this is true for two cylinders, we have another factor of $O(\epsilon^2)$, coming from the measure of the heights of the cylinders, proving the claim.

Since $s_i = h_i + h_3$ are integers, there are two ways to produce two saddle connections of length at most ϵ: either the pair of short saddle connections cross the wide cylinder C_3 ($h_3 \leq \epsilon$) or one saddle crosses C_1 ($h_1 \leq \epsilon$) and the other crosses C_2 ($h_2 \leq \epsilon$). (This is true because any curve wrapping around a cylinder would have length at least ϵ). In either case we have $\delta_2 \leq \epsilon$.

Case 1: The pair crosses C_3. Recall x is the representation of the zero on the bottom so that C_1 lies to the left, and x' is on the top so that C_1 lies to the right. Let y, y' the other representation of the zeroes on the bottom and top respectively. There are two subcases here. The first subcase is if one saddle connection joins x on the bottom to y' on the top. In order for there to be a second saddle connection joining y to x' parallel of the same length to the first, we must have $w_1 = w_2$ and $t_3 = w_1$. By primitivity $(s_1, s_2) = 1$. Also by primitivity we have $(s_2(t_1 + t_3) - s_1(t_2 + t_3), w_1) = 1$ or $(s_2t_1 - s_1t_2, w_1) = 1$. Also the projections of the two zeroes to the base torus must be at most ϵ apart. There is also a factor of $1/2$ from the action of τ. Then the contribution to the integral for this subcase is

$$\frac{\pi\epsilon^2}{2} \sum_{\substack{(s_1+s_2)w=d \\ (s_1,s_2)=1}} \sum_{\substack{(s_1t_2-s_2t_1,w)=1 \\ 0\leq t_i<w}} 1 = \frac{\pi\epsilon^2}{2} \sum_{\substack{w|d \\ w\neq d}} \varphi(d/w) \sum_{r|w} \frac{\mu(r)}{r} w^2$$

where φ is the Euler function.

In the second subcase, one saddle joins x on the bottom to x' on the top and the other saddle connection across C_3 joins y to y'. They are parallel of the same length. The saddle connections are homologous and together form a dividing curve. If we cut the surface along the dividing curve we have a pair of tori glued together. Conversely, given a pair of tori with marked point at the origin on each, and a vector γ, we can realize the surface as the tori glued along the slits. Namely, we take the vector γ based at the origin, slit each torus along γ and glue the slit tori. The tori have widths w_i, heights h_i and twists t_i. Since the area must be d we have

$$h_1w_1 + h_2w_2 = d.$$

Moreover the lattices of the two tori together must generate a unimodular lattice since the surface itself was a primitive cover. Thus we must have a similar pair of conditions we encountered in the case of two cylinders in $\mathcal{H}(2)$; namely, we must have $(h_1, h_2) = 1$ and $(w_1, w_2, t_1h_2 - t_2h_1) = 1$. On the other hand, w_1 and w_2 are arbitrary, but interchanging the two tori yields the same surface. Hence there is a term exactly as in the two cylinder case $\mathcal{H}(2)$ with $w_1 < w_2$, and also a term when $w_1 = w_2$ which is exactly the term in subcase 1. Again using Theorem 4.4 we find that the contribution to the integral from this subcase is given by

$$\pi\epsilon^2 \sum_{r|d} \mu(r) \sum_{\substack{h_1u_1+h_2u_2=d/r \\ (h_1,h_2)=1 \\ u_1<u_2}} ru_1u_2 + \frac{\pi\epsilon^2}{2} \sum_{\substack{w|d \\ w\neq d}} \varphi(d/w) \sum_{r|w} \frac{\mu(r)}{r} w^2.$$

Case 2: One saddle crosses C_1, the other crosses C_2. We must have $t_1 = t_2 = 0$ in order for the saddle to be short. Again since $s_i = h_i + h_3$, and by primitivity $(s_1, s_2) = 1$, we must have $s_1 = s_2 = 1$. Also because of primitivity, $(w_1, w_2) = 1$. The area $s_1 w_1 + s_2 w_2 = d$, hence $w_1 + w_2 = d$. Also t_3 can be arbitrary. There is a factor of $1/2$ due to the action of τ. Hence the contribution to the integral is

$$\frac{1}{2}\pi\epsilon^2 \sum_{\substack{w_1+w_2=d \\ (w_1,w_2)=1}} (w_1 + w_2) = \frac{\pi\epsilon^2 d}{2}\varphi(d).$$

The Proposition follows by adding the contributions for the three cases. $\square$

The proof of Theorem 1.3 for $s_2(d)$ is postponed until Appendix A.

A Some asymptotic formulas

Proof of Lemma 3.5. It is easy to see that the second term in (4) is $o(d^4)$. Hence,

$$N_d(1,1) = \sum_{r|d} \sum_{\substack{(s_1,s_2)=1 \\ s_1w_1+s_2w_2=d/r}} r w_1 w_2 (w_1 + w_2) \min(s_1, s_2) + o(d^4). \tag{22}$$

We compute the asymptotics as $d \to \infty$. Suppose d, r satisfy $d/r \le d^{1/5}$. Then $w_i, s_i \le d^{1/5}$ and so each term is at most $2dd^{4/5}$. There are at most $dd^{4/5}$ terms so the contribution to the sum is at most $2d^2d^{8/5}$. The contribution to the sum from just the term $r = 1$ is of the order d^4, as we shall see. Thus the asymptotics of the contribution from those r for which $r \le d^{4/5}$ is the same as the total contribution. In particular $d/r \to \infty$ as $d \to \infty$. For these terms, we will calculate the contribution, each with an error that is $o(d/r)^4)$ so that the total error is $o(d^4)$ as $d \to \infty$.

For each (s_1, s_2), choose a smallest $w_2^0 \le s_1$ such that

$$s_1 w_1 + s_2 w_2 = d/r$$

has a solution (w_1^0, w_2^0). Then for any pair (w_1, w_2) such that $w_1 s_1 + w_2 s_2 = d/r$, subtracting (w_1^0, w_2^0), we find

$$s_1(w_1 - w_1^0) + s_2(w_2 - w_2^0) = 0$$

and since $(s_1, s_2) = 1$, there must be λ so that

$$(w_1, w_2) = (w_1^0, w_2^0) + \lambda(-s_2, s_1)$$

with

$$0 \leq \lambda \leq \frac{d/r}{s_1 s_2}. \tag{23}$$

Note that $(w_1 w_2)(s_1 s_2) = (w_1 s_1)(w_2 s_2) \leq (d/r)^2$. Hence,

$$w_1 w_2 \leq \frac{(d/r)^2}{s_1 s_2}. \tag{24}$$

Also $w_1 + w_2 \leq 2\frac{(d/r)}{\min(s_1,s_2)}$. Hence, for fixed (s_1, s_2),

$$\sum_{s_1 w_1 + s_2 w_2 = d/r} w_1 w_2 (w_1 + w_2) \min(s_1, s_2) \leq 2\frac{(d/r)^4}{s_1^2 s_2^2} \tag{25}$$

(where we have used (23) to bound the number of terms in the sum). Hence, the contribution to (22) from the terms with $s_1 s_2 \geq (d/r)/\log(d/r)$ is $o(d/r)^4$. (We could use $(d/r)/f(d/r)$ for any function $f(d/r)$ which is $o(d/r)$.)

From now on we assume (s_1, s_2) satisfy

$$s_1 s_2 \leq (d/r)/\log(d/r).$$

This implies that $d/r - s_1 w_1^0 = o(d/r)$, and so the sum (22) over such pairs (s_1, s_2) becomes

$$\sum_{r|d} \sum_{\substack{(s_1,s_2)=1 \\ s_1 s_2 \leq (d/r)/\log(d/r)}} \sum_{\lambda=0}^{\frac{d/r}{s_1 s_2}} r\left(\frac{d/r}{s_1} - \lambda s_2\right)(\lambda s_1)\left(\frac{d/r}{s_1} - \lambda s_2 + \lambda s_1\right)\min(s_1, s_2) + o((d/r)^4).$$

By replacing the inner sum with an integral over λ, the sum becomes

$$\sum_{r|d} \frac{(d/r)^4}{12} \sum_{\substack{(s_1,s_2)=1 \\ s_1 s_2 \leq (d/r)/\log(d/r)}} r\left(\frac{1}{s_1^3 s_2^2} + \frac{1}{s_1^2 s_2^3}\right)\min(s_1, s_2) + o((d/r)^4). \tag{26}$$

We see immediately that for $r = 1$, the sum is of the order d^4, justifying the assumption that $r \leq d^{4/5}$.

Now we wish to further evaluate the sum (26). Recall, by definition, that the multiple ζ function $\zeta(m_1, m_2)$ is defined by

$$\zeta(m_1, m_2) = \sum_{s_1 < s_2}^{\infty} \frac{1}{s_1^{m_1} s_2^{m_2}}.$$

We will need to compute $\zeta(2,2)$ and $\zeta(1,3)$. Such formulas are well known (see e.g. [8]), but we give a short proof here for completeness. First we have

$$\zeta(2)^2 = 2\zeta(2,2) + \zeta(4). \tag{27}$$

We also have

$$\zeta(2)^2 = \sum_{x,y} \frac{(x+y)^3}{x^2y^2(x+y)^3} = \sum \frac{(x^2+y^2)(x+y) + 2x^2y + 2xy^2}{x^2y^2(x+y)^3} =$$
$$\sum \left(\frac{1}{x^2} + \frac{1}{y^2}\right) \frac{1}{(x+y)^2} + 2\sum \left(\frac{1}{y} + \frac{1}{x}\right) \frac{1}{(x+y)^3} = 2\zeta(2,2) + 4\zeta(1,3).$$

Together with (27) this gives

$$\zeta(2,2) + \zeta(1,3) = \frac{\zeta(2)^2}{2} - \frac{\zeta(4)}{4} = \zeta(4).$$

Thus

$$\sum_{s_1,s_2} \left(\frac{1}{s_1^3 s_2^2} + \frac{1}{s_1^2 s_2^3}\right) \min(s_1, s_2) = 2(\zeta(2,2) + \zeta(1,3) + \zeta(4)) = 4\zeta(4).$$

If we therefore perform the same sum conditioned over (s_1, s_2) we introduce a factor of $\frac{1}{\zeta(4)}$ and so the sum is 4. We therefore find from (26) that as $d \to \infty$,

$$N_d(1,1) = \sum_{r|d} \frac{d^4}{3r^3} + o(d^4)$$

and so

$$\nu(\mathcal{H}(1,1)) = \lim_{D\to\infty} \frac{10}{D^5} \sum_{d=1}^{D} \left(\sum_{r|d} \frac{d^4}{3r^3} + o(d^4) \right).$$

If we let $d = rq$, the double sum is evaluated by

$$\sum_{r\le D} \sum_{q \le D/r} rq^4 + o(D^5) = \sum_{r\le D} \frac{r}{5}\left(\frac{D}{r}\right)^5 + ro\left(\frac{D}{r}\right)^5 + o(D^5) = \frac{\zeta(4)}{5} D^5 + o(D^5)$$

and so

$$\nu(\mathcal{H}(1,1)) = \frac{\pi^4}{135}.$$

Proof of Lemma 3.9. The number of covers of degree d with 2 cylinders is

$$\sum_{\substack{h_1w_1+h_2w_2=d\\ w_1<w_2}} w_1w_2 = \sum_{r|d} \sum_{\substack{(s_1,s_2)=1\\ s_1w_1+s_2w_2=\frac{d}{r}}} \frac{1}{2}w_1w_2.$$

Arguing as before, we can assume $d/r \to \infty$ as $d \to \infty$, $s_1s_2 \leq \frac{d/r}{\log(d/r)}$ and find solutions of the form $w_1 = w_1^0 - \lambda s_2$, $w_2 = w_2^0 + \lambda s_1$, where $\frac{s_1w_1^0}{(d/r)} \to 1$ as $d \to \infty$ and $0 \leq \lambda \leq \frac{d}{rs_1s_2}$. Thus the above sum is

$$\frac{1}{2}\sum_{r|d}\left(\sum_{(s_1,s_2)=1}\sum_{\lambda=0}^{\frac{d}{rs_1s_2}}\left(\frac{d}{rs_1}-\lambda s_2\right)\lambda s_1 + o((d/r)^3)\right)$$

which if we replace by an integral over λ becomes

$$\frac{1}{2}\sum_{r|d}\sum_{(s_1,s_2)=1}\frac{1}{s_1^2s_2^2}\frac{d^3}{r^3} + o(d^3)$$

which as we let $d \to \infty$ becomes

$$\frac{1}{12}\sum_{r|d}\frac{d^3}{r^3}\frac{\zeta(2)^2}{\zeta(4)} + o(d^3).$$

Up to lower order terms, the number of covers of degree d with one cylinder is

$$\frac{1}{3}\sum_{h|d}\sum_{l_1+l_2+l_3=\frac{d}{h}}\frac{d}{h} = \sum_{r|d}\frac{d^3}{6r^3} + o(d^3).$$

Thus

$$\nu(\mathcal{H}(2)) = \lim_{D\to\infty}\frac{8}{D^4}\sum_{d=1}^{D}\sum_{r|d}\frac{d^3}{r^3}\left(\frac{\zeta(2)^2}{12\zeta(4)}+\frac{1}{6}\right)$$

which is equal to

$$2\zeta(4)\left(\frac{\zeta(2)^2}{12\zeta(4)}+\frac{1}{6}\right) = \frac{\zeta(2)^2}{6}+\frac{\zeta(4)}{3} = \frac{\pi^4}{120}.$$

Proof of Lemma 4.9. We compute the asymptotics as $d \to \infty$ of (8): Exactly as in Lemma 3.5 the first term is $\frac{d^4}{3r^2} + o(d)^4$. The second term is obviously $o(d^4)$.

Proof of Lemma 4.11. The sums here are exactly those found in Lemma 3.9. Thus the first in (9), corresponding to the 2-cylinder covers is asymptotic to

$$\frac{d^3}{12}\frac{\zeta(2)^2}{\zeta(4)}\sum_{r|d}\frac{\mu(r)}{r^2} + o(d^3) = \frac{5}{24}d^3\sum_{r|d}\frac{\mu(r)}{r^2} + o(d^3).$$

The last term in (9), corresponding to the 1-cylinder covers with a symmetry, is clearly $o(d^3)$. Finally, the second term in (9), corresponding to 1-cylinder covers without a symmetry is asymptotic to

$$\frac{1}{6}\sum_{r|d}\mu(r)d(d/r)^2 + o(d^3) = \frac{1}{6}d^3\sum_{r|d}\frac{\mu(r)}{r^2} + o(d^3).$$

This implies the lemma. $\square$

Proof of Theorem 1.3 for $c(d)$. By taking s_1, u_2 (or s_2, u_1) small, it is clear that the inner sum in (19), is of order $(d/r)^3$ as $d \to \infty$. However the sum over both u_1, u_2 small has strictly smaller order, so the contribution from the third term is of lower order than the contribution from the first two; that is, of order $o((d/r)^3)$. Thus by symmetry,

$$\int_{P_d(1,1)} \hat{f}\,d\nu = 2\pi\epsilon^2\sum_{r|d}\mu(r)\sum_{\substack{(s_1,s_2)=1\\ s_1u_1+s_2u_2=d/r}}\frac{u_1u_2(u_1+u_2)}{u_2^2}\min(s_1,s_2) + o((d/r)^3).$$

The inner sum is dominated by terms for which u_1, s_2 are large and s_1, u_2 are small. In particular, we can assume $s_1 = \min(s_1, s_2)$ and $u_1 + u_2 = u_1$ up to terms of lower order. We perform Möbius inversion on pairs (s_1, s_2) so that the integral becomes

$$2\pi\epsilon^2\sum_{r|d}\sum_{k|\frac{d}{r}}\mu(r)\mu(k)\sum_{s_1u_1+s_2u_2=\frac{d}{rk}}\frac{u_1^2}{u_2}ks_1 + o((d/r)^3) =$$

$$2\pi\epsilon^2\sum_{r|d}\sum_{k|\frac{d}{r}}\sum_{t|\frac{d}{rk}}\mu(r)\mu(k)\sum_{\substack{(s_1,u_2)=1\\ s_1u_1+s_2u_2=\frac{d}{rkt}}}\frac{u_1^2}{u_2}ks_1 + o((d/r)^3).$$

Replacing the sum with an integral as above, we find that

$$\int_{P_d(1,1)} \hat{f} d\nu = 2\pi\epsilon^2 \sum_{r|d} \sum_{k|\frac{d}{r}} \sum_{t|\frac{d}{rk}} \mu(r)\mu(k) \sum_{(s_1,u_2)=1} \frac{d^3}{3k^2 r^3 t^3 s_1^2 u_2^2} + o((d/r)^3).$$

Now we set $y = rt$ and find

$$\int_{P_d(1,1)} \hat{f} d\nu = \frac{2d^3\pi\epsilon^2}{3} \sum_{(s_1,u_2)=1} \sum_{r|y|\frac{d}{k}} \mu(r)\mu(k) \frac{1}{k^2 y^3 s_1^2 u_2^2} + o((d/r)^3).$$

Now $\sum_{r|y} \mu(r) = 0$ unless $y = 1$, in which case it is 1. Thus

$$\int_{P_d(1,1)} \hat{f} d\nu = \frac{2d^3\pi\epsilon^2}{3} \sum_{k|d} \sum_{(s_1,u_2)=1} \frac{\mu(k)}{k^2} \frac{1}{s_1^2 u_2^2} + o(d^3).$$

Thus by (15) and Lemma 4.9 we find that

$$\lim_{d\to\infty} c(d) = \lim_{d\to\infty} \frac{\int_{P_d(1,1)} \hat{f} d\nu}{\nu(P_d(1,1))} = 5.$$

□

Proof of Theorem 1.3 for $s_2(d)$**.** It is easy to see that only the first subcase of case 1 contributes as $d \to \infty$.

In view of Proposition 4.15, the computation of the asymptotics in Lemma 4.11, for the number of 2 cylinder covers in $\mathcal{H}(2)$ and Lemma 4.9, and §3.3,

$$\lim_{d\to\infty} s_2(d) = \frac{5/24}{1/3} = \frac{5}{8} = s_2(1,1).$$

□

References

[1] S. Bloch and A. Okounkov, *The Character of the Infinite Wedge Representation*, alg-geom/9712009.

[2] R. Dijkgraaf, *Mirror symmetry and elliptic curves*, The Moduli Space of Curves, R. Dijkgraaf, C. Faber, G. van der Geer (editors), Progress in Mathematics, **129**, Birkhäuser, 1995.

[3] A. Eskin, H. Masur. "Asymptotic formulas on flat surfaces." *Ergodic Th. Dyn Syst.* **21** (2001) 443-478

[4] A. Eskin, A. Okounkov. "Asymptotics of numbers of branched coverings of a torus and volumes of moduli spaces of holomorphic differentials" Inventiones Mathematicae, **145:1** (2001), 59–104.

[5] A. Eskin, A. Okounkov, A. Zorich, "Volumes of spaces of Abelian and quadratic differentials", to appear

[6] A. Eskin, H. Masur, A. Zorich, "Moduli Spaces of Abelian Differentials, the Principal Boundary, Counting Problems and the Siegel-Veech Constants, to appear

[7] S.G. Dani and G.A. Margulis. Limit distributions of orbits of unipotent flows and values of quadratic forms. *Advances in Soviet Math.* **16**(1993), 91–137.

[8] A. B. Goncharov. Multiple polylogarithms, cyclotomy and modular complexes. *Math. Res. Lett.* **5** (1998), no. 4, 497–516.

[9] E. Gutkin. Billiards in polygons. Physica D, **19** (1986), 311-333.

[10] E. Gutkin, C. Judge. Private communication.

[11] E. Gutkin, C. Judge. Affine mappings of translation surfaces: geometry and arithmetic. *Duke Math. J* **103** (2000), 191-213

[12] N. Jacobson, *The theory of rings.* Amer. Math. Soc. Colloq. Publ. **37**, Amer. Math. Soc., Providence, RI, 1943.

[13] A. Katok, A. Zemlyakov. Topological transitivity of billiards in polygons. Math. Notes, **18** (1975), 760-764.

[14] H. Masur, S. Tabachnikov, Flat structures and rational billiards, to appear, Handbook on Dynamical systems

[15] M. Ratner. Strict measure rigidity for unipotent subgroups of solvable groups. *Invent. Math.* **101**(1990), 449–482.

[16] M. Ratner. On measure rigidity of unipotent subgroups of semisimple groups. *Acta. Math.* **165**(1990), 229–309.

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

[18] M. Ratner. Raghunathan's topological conjecture and distributions of unipotent flows. *Duke Math. J.* **63**(1991), 235–290.

[19] N. Shah. Limit distributions of expanding translates of certain orbits on homogeneous spaces. *Proc. Indian Acad. Sci. (Math Sci)* **106**(2), (1996), pp. 105–125.

[20] W. Veech. Teichüller curves in moduli space. Eisenstein series and and an application to triangular billiards. Invent. Math., **97** (1989), 553-583.

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

[22] 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.

[23] A. Zorich *Square tiled surfaces and Teichmüller volumes of the moduli spaces of Abelian differentials*, in collection "Rigidity in Dynamics and Geometry", M. Burger, A. Iozzi (Editors), Springer Verlag, 2002, 459–471.

Alex Eskin: Department of Mathematics, University of Chicago, Chicago IL 60637, USA; E-mail: eskin@math.uchicago.edu

Howard Masur: Department of Mathematics, University of Illinois at Chicago, Chicago IL 60607, USA; E-mail: masur@math.uic.edu

Martin Schmoll: Institut für Mathematik, Technische Universität Berlin, 10623 Berlin, Germany; E-mail: schmoll@math.tu-berlin.de