Nova Série
BOLETIM
DA SOCIEDADE BRASILEIRA DE MATEMÁTICA

Bol. Soc. Bras. Mat., Vol. 21, No. 2, 205-229

An analogue of Mertens' theorem for closed orbits of Axiom A flows

Richard Sharp(1)

Abstract. For an Axiom A flow restricted to a basic set we prove an analogue of Mertens' theorem of prime number theory. The result is also established for the geodesic flow on a non-compact, finite area surface of constant negative curvature. Applying this to the modular surface yields some asymptotic formulae concerning quadratic forms.

0. Introduction

In recent years a number of papers have pointed to similarities between the distributional properties of prime numbers and those of hyperbolic dynamical systems. In particular Parry and Pollicott [12] have proved an analogue of the prime number theorem for Axiom A flows. Precisely, they proved that for a (topologically) weak-mixing Axiom A flow, with closed orbits τ of least period $\lambda(\tau)$ and entropy h,

$$\mathrm{card}\{\tau : N(\tau) \leq x\} \sim \frac{x}{\log x},$$

where $N(\tau) = e^{h\lambda(\tau)}$ with a modified asymptotic formula for flows which are not weak-mixing.

This paper is motivated by Mertens' theorem of prime number theory ([**8**], pp. 349-353), which states that

$$\prod_{p \leq x} (1 - \frac{1}{p}) \sim \frac{e^{-\gamma}}{\log x}$$

where the product is taken over all primes $p \leq x$ and γ is Euler's constant. We

---

(1) Supported by S.E.R.C. grant no. 88001623.
Received 20 November 1990.

prove that for an Axiom A flow φ

$$\prod_{N(\tau)\le x}\left(1-\frac{1}{N(\tau)}\right)\sim\frac{e^{-\gamma}}{\mathrm{Res}(\zeta_\varphi,1)\log x}$$

where ζ_φ is the Ruelle zeta function for φ. The proof, as one would expect, relies heavily on the symbolic dynamics of Bowen [5] and the thermodynamic formalism of Ruelle [15]. One should also note that the proof is elementary (just as the proof of Mertens' theorem is elementary) making no use of the deeper results about the analytic properties of ζ_φ obtained in [12] nor the associated Tauberian theorems.

We also establish a similar result in the case of the geodesic flow on a non-compact, finite area surface of constant negative curvature. To do this we make use of the prime geodesic theorem of Sarnak and Woo (cf. [16]). A classical example of such a surface is the modular surface and applying our theorem in this case, in conjunction with the relationship between closed geodesics on the modular surface and equivalence classes of quadratic forms, leads to some asymptotic formulae of a number theoretic character.

The author would like to thank William Parry for originally suggesting the problem and for his advice and encouragement during the course of this work and also Mark Pollicott for some helpful comments.

1. Shifts of finite type and their suspensions

Let A be an aperiodic $k\times k$ zero-one matrix (i.e. for some $n, A^n(i,j)>0$ for all $1\le i,j\le k$) and define

$$\Sigma_A=\{x\in\{1,\ldots,k\}^{\mathbb{Z}}: A(x_n,x_{n+1})=1,\quad \text{for all}\quad n\in\mathbb{Z}\}.$$

Give $\{1,\ldots,k\}$ the discrete topology and Σ_A the product topology. With respect to this topology Σ_A is compact and zero-dimensional, with a basis for the topology being given by finite unions of closed-open cylinders

$$[x_0,\ldots,x_{n-1}]^m=\{y: y_{i+m}=x_i,\quad 0\le i\le n-1\}.$$

(We write $[x_0,\ldots,x_{n-1}]$ for $[x_0,\ldots,x_{n-1}]^0$.) The shift of finite type $\sigma:\Sigma_A\to\Sigma_A$ is defined by $(\sigma x)_i=x_{i+1}$ and is a homeomorphism with respect to the given topology.

For $f\in C(\Sigma_A)$ and $0<\theta<1$ we define

$$\mathrm{var}_n f=\sup\{|f(x)-f(y)|: x,y\in\Sigma_A, x_i=y_i\quad \text{for}\quad |i|\le n\}$$

and $|f|_\theta=\sup(\mathrm{var}_n f/\theta^n)$.

The space $F_\theta=\{f\in C(\Sigma_A): |f|_\theta<\infty\}$ is a Banach space with respect to the norm

$$\|f\|_\theta=|f|_\infty+|f|_\theta$$

where $|\,.\,|_\infty$ is the uniform norm.

Define the pressure $P: C(\Sigma_A)\to\mathbb{R}$ by

$$P(f)=\sup\{h_m(\sigma)+\int f\,dm:\ m\ \text{a}\ \sigma-\text{invariant probability measure}\}$$

where $h_m(\sigma)$ denotes measure theoretic entropy. If $f\in F_\theta$, this supremum is attained for a unique, ergodic probability measure μ called the *equilibrium state* of f and μ has the following Gibbsian property. There exist positive constants C_1,C_2 such that for any $n\in\mathbb{N}$ and $x\in\mathrm{Fix}_n:=\{x:\sigma^n x=x\}$,

$$C_1\mu([x_0,\ldots,x_{n-1}])\le\exp\left(f^n(x)-nP(f)\right)\le C_2\mu([x_0,\ldots,x_{n-1}])\quad(1.1)$$

where

$$f^n(x)=f(x)+f(\sigma x)+\ldots+f(\sigma^{n-1}x).$$

(Bowen [6].)

If $f\in F_\theta$ is real and strictly positive, define the f suspension space $\Sigma_{A,f}$ to be

$$\{(x,t): x\in\Sigma_A\quad\text{and}\quad 0\le t\le f(x)\}$$

with $(x,f(x))$ and $(\sigma x,0)$ identified. Define a flow, σ^f, on this space by $\sigma_t^f(x,s)=(x,s+t)$, remembering identifications, i.e., a vertical flow under the graph of f. The entropy of σ^f is the entropy of σ_1^f. Using a result of Abramov [1], it is possible to show that this is the unique $h=h(\sigma^f)$ such that $P(-hf)=0$ and that if μ is the unique equilibrium state of $-hf$ then the Lebesgue extension of μ is the unique measure of maximal entropy for σ^f. The flow σ^f is said to be topologically weak-mixing if there does not exist a non-trivial solution to $F\sigma_t^f=e^{iat}F$ with $a>0$ and $F\in C(\Sigma_{A,f})$. If σ^f is not weak-mixing such an a is called an eigenfrequency for σ^f.

Let τ denote a generic closed σ^f orbit and let $\lambda(\tau)$ be its least period. Define its norm $N(\tau)$ to be $e^{h\lambda(\tau)}$. We define the zeta function for σ^f by

$$\zeta_{\sigma f}(s) = \prod_{\tau}(1 - N(\tau)^{-s})^{-1} = \exp \sum_{\tau} \sum_{k=1}^{\infty} \frac{1}{k} N(\tau)^{-ks}.$$

Each such τ corresponds to n distinct elements of Fix_n (for some $n \in \mathbb{N}$), $\{\xi, \sigma\xi, \ldots, \sigma^{n-1}\xi\}$ say, and $\lambda(\tau) = f^n(\xi)$. Also, if $n|m$ ($m = ni$ say), these n elements are also in Fix_m and $f^m(\xi) = if^n(\xi)$. Hence

$$\zeta_{\sigma f}(s) = \exp \sum_{n=1}^{\infty} \frac{1}{n} \sum_{\xi \in \mathrm{Fix}n} \exp - \mathrm{shf}^n(\xi).$$

$\zeta_{\sigma f}(s)$ is analytic and non-zero for $\mathrm{Re}(s) > 1$ and has a simple pole at $s = 1$ (Ruelle [15]). Furthermore, we have the following proposition.

Proposition 1. (**Ruelle [15], Parry [11].**)

$$Z(s) = \exp \sum_{n=1}^{\infty} \frac{1}{n} \left(\sum_{\xi \in \mathrm{Fix}n} \exp - \mathrm{shf}^n(\xi) - e^{nP(-\,\mathrm{shf})} \right)$$

converges uniformly to a non-zero analytic function in a neighbourhood U of $s = 1$ and $\zeta_{\sigma f}(s)$ can be analytically extended to $U - \{1\}$ by defining

$$\zeta_{\sigma f}(s) = Z(s)/(1 - e^{P(-\,\mathrm{shf})}).$$

We note that

$$\begin{aligned} \mathrm{Res}(\zeta_{\sigma f}, 1) &= \lim_{s \to 1} \frac{s-1}{1 - e^{P(-\,\mathrm{shf})}} Z(1) \\ &= - \left[\frac{de^{P(-\,\mathrm{shf})}}{ds} \right]_{s=1}^{-1} Z(1) \qquad (1.2) \\ &= \frac{1}{h \int f d\mu} Z(1) \end{aligned}$$

where μ is the equilibrium state of $-hf$. This uses the fact that

$$\left[\frac{dP(-\,\mathrm{shf})}{ds} \right]_{s=1} = -h \int f d\mu$$

(Ruelle [15]).

2. Axiom A flows

Let M be a compact Riemannian manifold and let φ be a C^1-flow on M. A compact φ-invariant set Λ containing no fixed points is said to be *hyperbolic* if the tangent bundle restricted to Λ can be written as the Whitney sum of three $D\varphi$-invariant continuous sub-bundles

$$T_\Lambda M = E + E^s + E^u$$

where E is the one-dimensional bundle tangent to the flow and there exist constants $C, \lambda > 0$ such that

(a) $\left\| D_{\varphi_t}(v) \right\| \le Ce^{-\lambda t} \|v\|$ for $v \in E^s, t \ge 0$

(b) $\left\| D_{\varphi_{-t}}(v) \right\| \le Ce^{-\lambda t} \|v\|$ for $v \in E^u, t \ge 0.$

A hyperbolic set Λ is said to be *basic* if

(i) the periodic orbits of φ restricted to Λ are dense in Λ

(ii) φ restricted to Λ is topologically transitive (i.e., Λ contains a dense orbit)

(iii) there exists an open set $U \supset \Lambda$ such that

$$\Lambda = \bigcap_{t \in \mathbb{R}} \varphi_t(U).$$

The non-wandering set Ω is defined by

$$\Omega = \{x \in M \colon \forall \text{ open } V \ni x \ \exists \text{ a sequence } t_i \uparrow \infty \text{ with } \varphi_{t_i}(V) \cap V \neq \varnothing\}.$$

The flows satisfies *Axiom A* if Ω is a disjoint union of a finite number of basic sets and hyperbolic fixed points. In what follows we will consider φ restricted to a basic set which is non-trivial (i.e., consists of more than one closed orbit).

Topological weak-mixing for φ is defined in the same way as for σ^f in the previous section. Bowen [3] has shown that either φ is not weak-mixing or φ is mixing with respect to the measure of maximal entropy.

As in the suspended flow case we define the zeta function for φ by

$$\prod_{\tau} (1 - N(\tau)^{-s})^{-1}$$

where the product is taken over all closed φ-orbits τ of least period $\lambda(t)$ and

$$N(\tau) = e^{h(\varphi)\lambda(\tau)}.$$

We can relate Axiom A flows to suspended flows by means of the following result due to Bowen [5].

Proposition 2. *If φ is an Axiom A flow restricted to a (non-trivial) basic set V then there exists a suspension of a shift of finite type $\Sigma_{A,f}$ (where $f \in F_\theta$ for some $0 < \theta < 1$), and a Hölder continuous map $\pi: \Sigma_{A,f} \to \Lambda, \pi\sigma_t^f = \varphi_t\pi$, where π is surjective, finite-one, measure-preserving with respect to the measures of maximal entropy and one-one a.e. with respect to the measure of maximal entropy for σ^f.*

We call σ^f the principal suspension. As a consequence of the above, φ is topologically weak-mixing if and only if σ^f is topologically weak-mixing. Clearly if σ^f is topologically weak-mixing then φ is topologically weak-mixing. On the other hand, if φ is topologically mixing then, by the comment above, it is mixing with respect to its measure of maximal entropy, hence σ^f is mixing with respect to its measure of maximal entropy, so it must be topologically weak-mixing. It also follows that $h(\varphi) = h(\sigma^f)$, since π is finite-one.

We wish to count the number of closed φ-orbits. We shall do this by means of the following proposition of Bowen [5] which is a refinement of the work of Manning for the diffeomorphism case [10].

Proposition 3. **(Bowen-Manning.)** In addition to the principal suspension, there exist suspensions of shifts of finite type $\Sigma_{A_i,f_i}, f_i \in F_\theta,\ i = 1, \ldots, p, \ldots, q$, Hölder continuous maps $\pi_i: \Sigma_{A_i,f_i} \to \Lambda,\ \pi_i\sigma_t^{f_i} = \varphi_t\pi_i$, where

(i) π_i is finite-one

(ii) π_i is not surjective

(iii) if $\nu(.,x)$ denotes the number of closed orbits of least period x then

$$\nu(\varphi, x) = \nu(\sigma^f, x) + \sum_{i=1}^{p} \nu(\sigma^{f_i}, x) - \sum_{i=p+1}^{q} \nu(\sigma^{f_i}, x).$$

By (i) and (ii) we have that $h(\sigma^{f_i}) < h(\varphi), i = 1, \ldots, p, \ldots, q$ and from

(iii) we obtain

$$\zeta_\varphi(s) = \zeta_{\sigma^f}(s)\, \frac{\prod_{i=1}^{p} \zeta_{\sigma^{f_i}}(hs/h(\sigma^{f_i}))}{\prod_{i=p+1}^{q} \zeta_{\sigma^{f_i}}(hs/h(\sigma^{f_i}))} \tag{2.1}$$

and

$$\mathrm{Res}(\zeta_\varphi, 1) = \mathrm{Res}(\zeta_{\sigma^f}, 1)\, \frac{\prod_{i=1}^{p} \zeta_{\sigma^{f_i}}(h/h(\sigma^{f_i}))}{\prod_{i=p+1}^{q} \zeta_{\sigma^{f_i}}(h/h(\sigma^{f_i}))} \tag{2.2}$$

Finally we remark that Parry and Pollicott [12] have shown that for a weak-mixing Axiom A flow

$$\pi(x) = \mathrm{card}\{\tau: N(\tau) \le x\} \sim \frac{x}{\log x}$$

and for an Axiom A flow that is not weak-mixing with least positive eigenfrequency a

$$\pi(x) \sim \frac{2\pi h/a}{\log x} \sum_{e^{2\pi nh/a} \le x} e^{2\pi nh/a}.$$

In either case there exist positive constants A, B such that

$$A\frac{x}{\log x} \le \pi(x) \le B\frac{x}{\log x} \quad \text{for all large } x \tag{2.3}$$

(Bowen [4]).

3. The main theorem

We will prove our result first for suspended flows and then use the results first for suspended flows and then use the results of the previous section to carry it over to Axiom A flows.

Let $\Sigma_{A,f}, \sigma^f$ be as in section 1 and set

$$a = \inf\{f(\xi): \xi \in \Sigma_A\},$$
$$b = \sup\{f(\xi): \xi \in \Sigma_A\}$$

and

$$y = \frac{\log x}{h}.$$

We begin by considering

$$K(x) = \sum_{N(\tau)\le x} \sum_{k=1}^{[y/\lambda(\tau)]} \frac{1}{k} N(\tau)^{-k} = \sum_{n=1}^{\infty} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}'_n} \exp -hf^n(\xi).$$

where Fix'_n denotes the set of $\xi \in \mathrm{Fix}_n$ with $f^n(\xi) \le y$.

First note that if $f^n(\xi) \le y$ for some $\xi \in \mathrm{Fix}_n$ then $na \le y$, so $n \le [y/a]+1$ (here [ ] denotes integral part). Thus we have in fact

$$K(x) = \sum_{n=1}^{[y/a]+1} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}'_n} \exp -hf^n(\xi).$$

We split the range of summation into

$$1 \le n \le [y/\int f\,d\mu] \quad \text{and} \quad [y/\int f\,d\mu]+1 \le n \le [y/a]+1$$

and note that if $f^n(\xi) > y$ for some $\xi \in \mathrm{Fix}_n$ then $nb > y$, so $n \ge [y/b]+1$. Thus

$$\begin{aligned} K(x) = &\sum_{n=1}^{[y/\int f d\mu]} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}_n} \exp -hf^n(\xi) \\ &+ \sum_{n=[y/\int f d\mu]+1}^{[y/a]+1} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}'_n} \exp -hf^n(\xi) \\ &- \sum_{n=[y/b]+1}^{[y/\int f d\mu]} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}''_n} \exp -hf^n(\xi) \end{aligned}$$

(where Fix''_n denotes the set of $\xi \in \mathrm{Fix}_n$ with $f^n(\xi) > y$)

$$= \sum_{n=1}^{[y/\int f d\mu]} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}_n} exp - hf^n(\xi) + A(x) - B(x),$$

say.

Now

$$\begin{aligned} \sum_{n=1}^{[y/\int f d\mu]} \frac{1}{n} &= \log[y/\int f\,d\mu] + \gamma + o(1) \\ &= \log\log x + \log \frac{1}{h\int f\,d\mu} + \gamma + o(1) \end{aligned}$$

where γ is the Euler's constant and Proposition 1 gives

$$\sum_{n=1}^{[y/\int f d\mu]} \frac{1}{n} \left\{ \sum_{\xi\in \mathrm{Fix}_n} exp - hf^n(\xi) - 1 \right\} = \log Z(1) + o(1).$$

Combining this and (1.2) yields

$$K(x) = \log\log x + \gamma + \log \mathrm{Res}(\zeta_{\sigma f}, 1) + A(x) - B(x).$$

Our aim is now to show that $A(x) = o(1), B(x) = o(1)$. We do this by means of the next two lemmas. Choose $0 < \varepsilon < min(\int f\,d\mu - a, b - \int f\,d\mu)$, and write

$$A(x) = A_1(x) + A_2(x),$$
$$B(x) = B_1(x) + B_2(x),$$

where

$$A_1(x) = \sum_{n=[y/(\int f d\mu-\varepsilon)]+1}^{[y/a]+1} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}'_n} \exp -hf^n(\xi),$$

$$A_2(x) = \sum_{n=[y/(\int f d\mu)]+1}^{[y/(\int f d\mu-\varepsilon)]} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}'_n} \exp -hf^n(\xi),$$

$$B_1(x) = \sum_{n=[y/b]+1}^{[y/(\int f d\mu+\varepsilon)]} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}''_n} \exp -hf^n(\xi),$$

and

$$B_2(x) = \sum_{n=[y/(\int f d\mu+\varepsilon)]+1}^{[y/\int f d\mu]} \frac{1}{n} \sum_{\xi\in \mathrm{Fix}''_n} \exp -hf^n(\xi).$$

By the ergodic theorem $f^n(\eta)/n \to \int f\,d\mu$ as $n \to \infty$ for $\mu - a.e.\ \eta$, so we can choose N so large that for every $n \ge N$

$$\mu\left(\left\{\eta \in \Sigma_A : \left|\frac{f^n(\eta)}{n} - \int f\,d\mu\right| > \frac{1}{2}\varepsilon\right\}\right) < \varepsilon \tag{3.1}$$

and put

$$N' = \max\left(N, \left[\frac{2|f|_\theta}{\varepsilon(1-\theta)}\right] + 1\right).$$

Lemma 1. *For $x \geq e^{N'hb}$, i.e. for $y \geq N'b$,*

$$A_1(x) < \left(\frac{\int f d\mu}{a} - 1 + \frac{2\int f d\mu}{N'b}\right) C_2\ \varepsilon,$$
$$B_1(x) < \left(\frac{b}{\int f d\mu} - 1 + \frac{2}{N'}\right) C_2\ \varepsilon$$

where C_2 is defined by (1.1).

Proof. Suppose $[y/(\int f d\mu - \varepsilon)] + 1 \leq n \leq [y/a] + 1$ and $f^n(\xi) \leq y$, for some $\xi \in \mathrm{Fix}_n$. Then

$$\frac{f^n(\xi)}{n} < \int f d\mu - \varepsilon.$$

For n in this range we have

$$n \geq \frac{N'b}{\int f d\mu - \varepsilon} \geq N' > \frac{2|f|_\theta}{\varepsilon(1-\theta)},$$

so for every $\eta \in [\xi_0, \xi_1, \dots, \xi_{n-1}]$,

$$\left|\frac{f^n(\eta)}{n} - \frac{f^n(\xi)}{n}\right| \leq \frac{|f|_\theta}{n}(1 + \theta + \cdots + \theta^{n-1}) \leq \frac{|f|_\theta}{n(1-\theta)} < \frac{1}{2}\ \varepsilon$$

and so

$$\frac{f^n(\eta)}{n} < \frac{f^n(\xi)}{n} + \frac{1}{2}\ \varepsilon < \int f d\mu - \frac{1}{2}\ \varepsilon.$$

Thus $\xi \in \mathrm{Fix}_{n,\varepsilon}$ where $\mathrm{Fix}_{n,\varepsilon}$ denotes the set of those $\xi \in \mathrm{Fix}_n$ such that for every $\eta \in [\xi_0, \xi_1, \dots, \xi_{n-1}]$,

$$\left|\frac{f^n(\eta)}{n} - \int f d\mu\right| > \frac{1}{2}\ \varepsilon.$$

Hence for

$$[y/(\int f d\mu - \varepsilon)] + 1 \leq n \leq [y/a] + 1,$$

$$\sum_{\xi \in \mathrm{Fix}'_n} \exp -hf^n(\xi) \leq \sum_{\xi \in \mathrm{Fix}_{n,\varepsilon}} \exp -hf^n(\xi)$$
$$\leq C_2 \mu\left(\bigcup_{\xi \in \mathrm{Fix}_{n,\varepsilon}} [\xi_0, \xi_1, \dots, \xi_{n-1}]\right) \quad \text{by (1.1)}$$
$$< C_2 \varepsilon \quad \text{by (3.1).}$$

A similar argument gives that for $[y/b] + 1 \leq n \leq [y/(\int f d\mu + \varepsilon)]$,

$$\sum_{\xi \in \mathrm{Fix}''_n} \exp -hf^n(\xi) < C_2\ \varepsilon.$$

Thus

$$A_1(x) \leq \left\{\sum_{n=[y/(\int f d\mu - \varepsilon)]+1}^{[y/a]+1} \frac{1}{n}\right\} C_2\ \varepsilon$$
$$\leq \frac{[y/a] - [y/(\int f d\mu - \varepsilon)] + 1}{[y/(\int f d\mu - \varepsilon)] + 1} C_2\ \varepsilon$$
$$\leq \left\{\frac{\int f d\mu - \varepsilon}{a} - 1 + \frac{2(\int f d\mu - \varepsilon)}{y}\right\} C_2\ \varepsilon$$
$$\leq \left\{\frac{\int f d\mu}{a} - 1 + \frac{2\int f d\mu}{N'b}\right\} C_2\ \varepsilon$$

and

$$B_1(x) \leq \left\{\sum_{n=[y/b)]+1}^{[y/(\int f d\mu + \varepsilon)]} \frac{1}{n}\right\} C_2\ \varepsilon$$
$$\leq \frac{[y/(\int f d\mu + \varepsilon)] - [y/b]}{[y/b)] + 1} C_2\ \varepsilon$$
$$\leq \left\{\frac{b}{\int f d\mu + \varepsilon} - 1 + \frac{2b}{y}\right\}$$
$$\leq \left\{\frac{b}{\int f d\mu} - 1 + \frac{2}{N'}\right\} C_2\ \varepsilon.$$

Lemma 2. *For all $x > 1$*

$$A_2(x) \leq \left(\frac{\varepsilon}{\int f d\mu - \varepsilon} + \frac{2h\int f d\mu}{\log x}\right) C_2,$$
$$B_2(x) \leq \left(\frac{\varepsilon}{\int f d\mu} + \frac{2h\int f d\mu}{\log x}\right) C_2.$$

Proof. By (1.1), for every $n \geq 1$

$$\sum_{\xi \in \mathrm{Fix}_n} \exp -hf^n(\xi) \leq C_2.$$

Hence

$$A_2(x) \le \left\{ \sum_{n=[y/\int f d\mu]+1}^{[y/(\int f d\mu - \varepsilon)]} \frac{1}{n} \right\} C_2$$
$$\le \frac{[y/(\int f d\mu - \varepsilon)] - [y/\int f d\mu]}{[y/\int f d\mu]+1} C_2$$
$$\le \left\{ \frac{\int f d\mu}{\int f d\mu - \varepsilon} - 1 + \frac{2\int f d\mu}{y} \right\} C_2$$
$$= \left\{ \frac{\varepsilon}{\int f d\mu - \varepsilon} + \frac{2h \int f d\mu}{\log x} \right\} C_2$$

and

$$B_2(x) \le \left\{ \sum_{n=[y/(\int f d\mu + \varepsilon)]+1}^{[y/\int f d\mu]} \frac{1}{n} \right\} C_2$$
$$\le \frac{[y/\int f d\mu] - [y/(\int f d\mu + \varepsilon)]}{[y/(\int f d\mu + \varepsilon)]+1} C_2$$
$$\le \left\{ \frac{\int f d\mu + \varepsilon}{\int f d\mu} - 1 + \frac{2\int f d\mu}{y} \right\} C_2$$
$$\le \left\{ \frac{\varepsilon}{\int f d\mu} + \frac{2h \int f d\mu}{\log x} \right\} C_2.$$

Since we may choose $\varepsilon > 0$ as small as we please, the two lemmas combine to give $A(x) = o(1), B(x) = o(1)$, and so

$$K(x) = \log\log x + \gamma + \log \operatorname{Res}(\zeta_{\sigma f}, 1) + o(1).$$

Now

$$\sum_{N(\tau)\le x} \log\left(\frac{1}{1 - N(\tau)^{-1}}\right) = \sum_{N(\tau)\le x} \sum_{k=1}^{\infty} \frac{1}{k} N(\tau)^{-k}$$

so

$$0 < \sum_{N(\tau)\le x} \log\left(\frac{1}{1 - N(\tau)^{-1}}\right) - K(x)$$
$$= \sum_{N(\tau)\le x} \sum_{k=[y/\lambda(\tau)]+1}^{\infty} \frac{1}{k} N(\tau)^{-k}$$
$$\le \sum_{N(\tau)\le x} \frac{\log N(\tau)}{\log x} \sum_{k=2}^{\infty} \frac{1}{k} N(\tau)^{-k}$$
$$= \frac{1}{\log x} \sum_{N(\tau)\le x} \frac{\log N(\tau)}{N(\tau)(N(\tau)-1)}$$

and this last term tends to 0 as $x \to \infty$, since

$$\sum_{\tau} \frac{\log N(\tau)}{N(\tau)(N(\tau)-1)}$$

converges. (To see this, write the sum as a Stieltjes integral with respect to $\pi(x)$, partially integrate and apply (2.3).) Hence

$$\sum_{N(\tau)\le x} \log\left(\frac{1}{1 - N(\tau)^{-1}}\right) = \log\log x + \gamma + \log \operatorname{Res}(\zeta_{\sigma f}, 1) + o(1).$$

Now, let φ be an Axiom A flow and let $f, f_1, \ldots, f_p, \ldots, f_q$ be as in Propositions 2 and 3. By (2.1) and (2.2)

$$\sum_{\substack{N(\tau)\le x \\ \tau \text{ a } \varphi\text{-orbit}}} \log\left(\frac{1}{1 - N(\tau)^{-1}}\right) - \sum_{\substack{N(\tau)\le x \\ \tau \text{ a } \sigma^f\text{-orbit}}} \log\left(\frac{1}{1 - N(\tau)^{-1}}\right) \to \log \frac{\operatorname{Res}(\zeta_\varphi, 1)}{\operatorname{Res}(\zeta_{\sigma f}, 1)}$$

as $x \to \infty$ and so for closed φ-orbits we also have

$$\sum_{N(\tau)\le x} \log\left(\frac{1}{1 - N(\tau)^{-1}}\right) = \log\log x + \gamma + \log \operatorname{Res}(\zeta_\varphi, 1) + o(1). \quad (3.2)$$

Now note

$$0 < \sum_{N(\tau)\le x} \log\left(\frac{1}{1 - N(\tau)^{-1}}\right) - \sum_{N(\tau)\le x} \frac{1}{N(\tau)}$$
$$= \sum_{N(\tau)\le x} \sum_{k=2}^{\infty} \frac{1}{k} N(\tau)^{-k}$$
$$\le \sum_{N(\tau)\le x} \frac{1}{2N(\tau)(N(\tau)-1)}.$$

It is easy to see, by the same argument as above, that this last sum converges as $x \to \infty$ and, since

$$\sum_{N(\tau)\le x} \left\{ \log\left(\frac{1}{1-N(\tau)^{-1}}\right) - \frac{1}{N(\tau)} \right\}$$

is increasing, this ensures the convergence of

$$\sum_{\tau} \left\{ \log\left(\frac{1}{1-N(\tau)^{-1}}\right) - \frac{1}{N(\tau)} \right\} \tag{3.3}$$

As an immediate consequence of (3.2) and (3.3) we have the following theorem.

Theorem 1. *For an Axiom A flow φ (restricted to a non-trivial basic set)*

$$\prod_{N(\tau)\le x} \left(1 - \frac{1}{N(\tau)}\right) \sim \frac{e^{-\gamma}}{\operatorname{Res}(\zeta_\varphi, 1)\log x}$$

and

$$\sum_{N(\tau)\le x} \frac{1}{N(\tau)} = \log\log x + B + o(1)$$

where the constant B is given by

$$B = \gamma + \log \operatorname{Res}(\zeta_\varphi, 1) - \sum_{\tau} \left\{ \log\left(\frac{1}{1-N(\tau)^{-1}}\right) - \frac{1}{N(\tau)} \right\}$$

Remarks. (i) Since this paper was first written, Mark Pollicott has pointed out to the author that, in the weak-mixing case, a considerably shorter but non-elementary proof of Theorem 1 is possible using a complex Tauberian theorem due to Agmon [2]. To apply this result one needs the additional information that if φ is weak-mixing then, apart from the simple pole at $s = 1$, $\zeta_\varphi(s)$ is analytic and non-zero in a neighbourhood of $\operatorname{Re}(s) \ge 1$ [12]. An advantage of this approach is that it sharpens the error term in the expression for $\Sigma_{N(\tau)\le x} 1/N(\tau)$ from $o(1)$ to $o(1/\log x)$. This comment also applies to the results in the next two sections.

(ii) As we noted in section 2, for a weak-mixing Axiom A flow,

$$\pi(x) \sim x/\log x \sim \operatorname{li} x$$

where

$$\operatorname{li} x = \prod_{2}^{\infty} \frac{1}{\log t} dt.$$

It is interesting to have information about the error term in this theorem, i.e. $\Psi(x) = \pi(x) - \operatorname{li} x$. It is known that for the geodesic flow on a manifold of constant negative curvature, $\Psi(x) = O(x^\alpha)$ for some $0 < \alpha < 1$ [9]. On the other hand, for a suspended flow with the suspending function depending on only finitely many coordinates this can never be the case [14]. Our result has some bearing on the average behaviour of $\Psi(x)$. We have, using the fact that $\operatorname{li} x = x/\log x + O(x/(\log x)^2)$,

$$\begin{aligned}\sum_{N(\tau)\le x} \frac{1}{N(\tau)} &= \int_2^x \frac{1}{t} d\pi(t) = \frac{\pi(x)}{x} + \int_2^x \frac{\pi(t)}{t^2} dt \\ &= \int_2^x \frac{1}{t\log t} dt + \int_2^x \frac{\Psi(t)}{t^2} dt + O\left(\frac{1}{\log x}\right) \\ &= \log\log x + \int_2^x \frac{\Psi(t)}{t^2} + O\left(\frac{1}{\log x}\right)\end{aligned}$$

(modulo a constant) and comparing this with our result reveals that

$$\int_2^\infty \frac{\Psi(t)}{t^2} dt$$

converges.

(iii) The asymptotic formulae of Theorem 1 also hold for an Axiom A diffeomorphism φ. To see this consider the time-1 suspension flow. Then $N(\tau)$ has the same value whether τ is regarded as a closed orbit of the diffeomorphism or of the flow. The function $\zeta_\varphi(s)$ is defined by $\zeta_\varphi(s) = \zeta(e^{-hs})$ where $\zeta(z)$ is the usual Artin-Mazur zeta function for φ:

$$\zeta(z) = \exp \sum_{n=1}^{\infty} \frac{z^n}{n} \operatorname{card}\{x\colon \varphi^n x = x\}.$$

We now turn our attention to the situation considered by Parry and Pollicott in [13]. Let $\tilde{\varphi}\colon \widetilde{M} \to \widetilde{M}$ be an Axiom A flow and let G be a finite group of diffeomorphisms which acts freely on $\widetilde{M}$ and commutes with $\tilde{\varphi}$. This gives rise to a flow φ on the quotient manifold $M = \widetilde{M}/G$, defined by $\varphi_t(Gx) = G(\tilde{\varphi}_t x)$. It can be shown that φ is also an Axiom A flow. We shall suppose that $\tilde{\Lambda}$ is a (non-trivial) basic set for $\tilde{\varphi}$, then $\Lambda = \tilde{\Lambda}/G$ is a basic set for φ. As usual, we shall consider $\tilde{\varphi}, \varphi$ restricted to $\tilde{\Lambda}$, Λ respectively.

For any closed φ-orbit, τ, let $\tilde{\tau}_1, \ldots, \tilde{\tau}_n$ be the closed $\tilde{\varphi}$-orbits which lie above τ. Then $n||G|$ and

$$\lambda(\tilde{\tau}_i) = \frac{|G|}{n}\lambda(\tau), \qquad i = 1, \dots, n.$$

For each $\tilde{\tau}_i$ there exists a unique *Frobenius element* $[\tilde{\tau}_i] \in G$ such that $[\tilde{\tau}_i]x = \tilde{\varphi}_{\lambda(\tau)}x$ for every $x \in \tilde{\tau}_i$. If $g\tilde{\tau}_i = \tilde{\tau}_j, g \in G$ then $[\tilde{\tau}_j] = g[\tilde{\tau}_i]g^{-1}$ so that the *Frobenius class*, the conjugacy class of $[\tilde{\tau}_i]$ depends only on τ.

Choose $g \in G$ and let $C = C(g)$ be its conjugacy class. Let R_χ be an irreducible representation of G with irreducible character χ, and define

$$L(s,\chi) = \prod_\tau \det\left(I - \frac{R_\chi(\tau)}{N(\tau)^s}\right)^{-1}$$

where $R_\chi(\tau) = R_\chi([\tilde{\tau}])$ for any $\tilde{\tau}$ lying above τ. This product converges for $Re(s) > 1$. Clearly, if χ_0 denotes the trivial character, $L(s,\chi_0) = \zeta_\varphi(s)$. On the other hand, if $\chi \neq \chi_0$, then $L(s,\chi)$ is analytic in a neighbourhood of $Re(s) = 1$ [13].

By the orthogonality relation for characters and (3.2)

$$\begin{aligned}
\sum_{\substack{N(\tau)\le x\\ [\tilde{\tau}]\in C}} \log\left(1 - \frac{1}{N(\tau)}\right) &= \sum_{\substack{N(\tau)\le x\\ [\tilde{\tau}]\in C}} \sum_{k=1}^{\infty} \frac{1}{k}N(\tau)^{-k} \\
&= \frac{|C|}{|G|} \sum_{N(\tau)\le x} \sum_{k=1}^{\infty} \frac{1}{k}N(\tau)^{-k} \\
&\quad + \frac{|C|}{|G|} \sum_{\chi\neq\chi_0} \chi(g^{-1}) \sum_{N(\tau)\le x} \sum_{k=1}^{\infty} \frac{\chi([\tilde{\tau}])}{k} N(\tau)^{-k} \\
&= \frac{|C|}{|G|}\log\log x + \frac{|C|}{|G|}\gamma + \frac{|C|}{|G|}\log \mathrm{Res}(\zeta_\varphi, 1) \\
&\quad + \frac{|C|}{|G|} \sum_{\chi\neq\chi_0} \chi(g^{-1}) \log L(1,\chi) + o(1).
\end{aligned}$$

Hence we have:

Corollary 1.

$$\prod_{\substack{N(\tau)\le x\\ [\tilde{\tau}]\in C}} \left(1 - \frac{1}{N(\tau)}\right) \sim \frac{e^{-(|C|/|G|)\gamma}}{A(\log x)^{|C|/|G|}}$$

where

$$A = \mathrm{Res}(\zeta_\varphi, 1)^{|C|/|G|} \left\{ \prod_{\chi\neq\chi_0} e^{-\chi(g^{-1})} L(1,\chi) \right\}^{|C|/|G|}$$

and

$$\sum_{\substack{N(\tau)\le x\\ [\tilde{\tau}]\in C}} \frac{1}{N(\tau)} = \frac{|C|}{|G|}\log\log x + \text{ constant} + o(1)$$

(This last formula is deduced in the same way as in Theorem 1.)

4. Non-compact quotients

Let $\mathbb{H}^+$ denote the upper half plane $\{x + iy \in \mathbb{C} : y > 0\}$ equipped with the Poincaré metric $ds^2 = (dx^2 + dy^2)/y^2$. Let Γ be a discrete subgroup of $\mathrm{PSL}(2,\mathbb{R})$, then Γ acts on $\mathbb{H}^+$ as linear fractional transformations, $z \to (az+b)/(cz+d)$, and these transformations are isometries of $\mathbb{H}^+$ with respect to the Poincaré metric. The surface S is the quotient space $\mathbb{H}^+/\Gamma$. With respect to the induced metric, this surface has curvature -1.

Let T_1S be the unit tangent bundle of S and let φ be the geodesic flow on T_1S, i.e. for $(x,v) \in T_1S, \varphi_t(x,v)$ is the point reached by starting at (x,v) and flowing for time t along the unique unit-speed geodesic through x in the direction v. Such flows have $h(\varphi) = 1$. There is an exact correspondence between closed geodesics on S and closed φ-orbits. We define the norm of a closed geodesic τ of length $\lambda(\tau)$ to be $N(\tau) = e^{\lambda(\tau)}$ and define

$$\zeta_\Gamma(s) = \prod_\tau (1 - N(\tau)^{-s})^{-1}.$$

If S is compact then φ satisfies Axiom A and the analysis of the previous section applies (with $\zeta_\Gamma(s) = \zeta_\varphi(s)$). We now consider the case where S is not compact but has finite area (with respect to the Riemann measure). It remains true in this situation that $\zeta_\Gamma(s)$ is analytic and non-zero for $\mathrm{Re}(s) > 1$ with a simple pole at $s = 1$.

Let Δ be the Laplace-Beltrami operator on S, and let

$$0 < \lambda_1 \le \lambda_2 \le \dots \le \lambda_k \le \frac{3}{16}$$

be the discrete eigenvalues of $-\Delta$ in $[0,3/16]$. Write

$$\alpha_j = \frac{1}{2} + \sqrt{(\frac{1}{4} - \lambda_j)}, \quad j = 1, \dots, k,$$

so that $\frac{1}{2} \le \alpha_j < 1$. We have the following result.

Proposition 4. **(Sarnak and Woo, cf. [16].)** Let $\pi(x)$ denote the number of closed geodesics τ on S with $N(\tau) \le x$, then

$$\pi(x) = \operatorname{li} x + \operatorname{li} x^{\alpha_1} + \dots + \operatorname{li} x^{\alpha_k} + O(x^{3/4}(\log x)^2).$$

Since $\operatorname{li} x = x/\log x + O(x/(\log x)^2)$, we have

$$\pi(x) = x/\log x + O(x/(\log x)^2). \tag{4.1}$$

Theorem 2. *For a non-compact, finite area surface of constant negative curvature $S = \mathbb{H}^+/\Gamma$, we have*

$$\sum_{N(\tau)\le x} \frac{1}{N(\tau)} = \log\log x + B + O\left(\frac{1}{\log x}\right)$$

B constant, and

$$\prod_{N(\tau)\le x} \left(1 - \frac{1}{N(\tau)}\right) \sim \frac{e^{-\gamma}}{\operatorname{Res}(\zeta_\Gamma, 1)\log x}.$$

Proof. Let N_0 be the norm of the shortest geodesic on M. Then

$$\begin{aligned}\sum_{N(\tau)\le x} \frac{1}{N(\tau)} &= \int_{N_0}^{x} \frac{1}{t} d\pi(t) \\ &= \int_{N_0}^{x} \frac{\pi(t)}{t^2} dt + \frac{\pi(x)}{x} \\ &= \log\log x - \log\log N_0 + \int_{N_0}^{x} \frac{\pi(t) - t/\log t}{t^2} dt + \frac{\pi(x)}{x}\end{aligned}$$

By (4.1),

$$\int_{N_0}^{x} \frac{\pi(t) - t/\log t}{t^2} dt = \int_{N_0}^{\infty} \frac{\pi(t) - t/\log t}{t^2} dt + O\left(\frac{1}{\log x}\right)$$

and $\pi(x)/x = O(1/\log x)$, so we have

$$\sum_{N(\tau)\le x} \frac{1}{N(\tau)} = \log\log x + B + O\left(\frac{1}{\log x}\right)$$

where

$$B = \int_{N_0}^{\infty} \frac{\pi(t) - t/\log t}{t^2} dt - \log\log N_0.$$

By the same argument as in the previous section, the series

$$\sum_\tau \left\{\log\left(1 - \frac{1}{N(\tau)}\right) + \frac{1}{N(\tau)}\right\}$$

is convergent, with sum F say. Hence

$$\log \prod_{N(\tau)\le x} \left(1 - \frac{1}{N(\tau)}\right) = -\log\log x - B + F + o(1)$$

and so to deduce the result we only need to show that

$$F - B = -\gamma - \log \operatorname{Res}(\zeta_\Gamma, 1).$$

We shall use the fact that

$$\gamma = -\int_0^\infty e^{-u} \log u du.$$

If $\delta \ge 0$ it is easy to see that

$$0 < -\log(1 - N(\tau)^{-1-\delta}) - N(\tau)^{-1-\delta} \le \frac{1}{2} N(\tau)^{-1}(N(\tau) - 1)^{-1}.$$

Hence the series

$$F(\delta) = \sum_\tau \left\{\log\left(1 - N(\tau)^{-1-\delta}\right) + N(\tau)^{-1-\delta}\right\}$$

converges uniformly for all $\delta \ge 0$ and so $F(\delta) \to F(0) = F$ as $\delta \to 0$.

Now suppose $\delta > 0$. Then

$$F(\delta) = G(\delta) - \log \zeta_\Gamma(1 + \delta)$$

where

$$G(\delta) = \sum_\tau N(\tau)^{-1-\delta}.$$

Write

$$L(x) = \sum_{N(\tau)\le x} \frac{1}{N(\tau)} = \log\log x + B + E(x), \qquad E(x) = O\left(\frac{1}{\log x}\right).$$

Then

$$\sum_{N(\tau)\le x} N(\tau)^{-1-\delta} = x^{-\delta} L(x) + \delta \int_{N_0}^{x} t^{-1-\delta} L(t) dt$$

If we let $x \to \infty, x^{-\delta}L(x) \to 0$, so

$$G(\delta) = \delta \int_{N_0}^{\infty} t^{-1-\delta} L(t)dt$$
$$= \delta \int_{N_0}^{\infty} t^{-1-\delta}(\log\log t + B)dt + \delta \int_{N_0}^{\infty} t^{-1-\delta}E(t)dt.$$

Put $t = e^{u/\delta}$. Then

$$\delta \int_{1}^{\infty} t^{-1-\delta}\log\log t\,dt = \int_0^{\infty} e^{-1}\log(u/\delta)du = -\gamma - \log\delta$$

and

$$\delta \int_1^{\infty} t^{-1-\delta}dt = 1.$$

Hence

$$G(\delta) + \log\delta - B + \gamma = \delta \int_{N_0}^{\infty} t^{-1-\delta}E(t)dt - \delta \int_1^{N_0} t^{-1-\delta}(\log\log t + b)dt.$$

Now if $T = \exp(1/\sqrt{\delta})$,

$$\left|\delta \int_{N_0}^{\infty} t^{-1-\delta}E(t)dt\right| \leq \text{const. } \delta \int_{N_0}^{T} t^{-1}dt + \frac{\text{const. } \delta}{\log T}\int_T^{\infty} t^{-1-\delta}dt$$
$$\leq \text{ const. } \delta\log T + \text{ const. } \frac{T^{-\delta}}{\log T}$$
$$\leq \text{ const.} \sqrt{\delta} + \text{const. } e^{-\sqrt{\delta}}\sqrt{\delta} \to 0 \quad \text{as} \quad \delta \to 0.$$

Also,

$$\left|\delta \int_1^{N_0} t^{-1-\delta}(\log\log t + B)dt\right| <$$
$$< \delta \int_1^{N_0} t^{-1}(|\log\log t| + |B|)dt \to 0 \quad \text{as} \quad \delta \to 0.$$

Hence

$$G(\delta) + \log\delta \to B - \gamma \quad \text{as} \quad \delta \to 0,$$

but

$$\log\zeta_\Gamma(1+\delta) + \log\delta \to \log\operatorname{Res}(\zeta_\Gamma, 1) \quad \text{as} \quad \delta \to 0,$$

so

$$F(\delta) \to B - \gamma - \log\operatorname{Res}(\zeta_\Gamma, 1), \quad \text{i.e.} \quad F = B - \gamma - \log\operatorname{Res}(\zeta_\Gamma, 1)$$

and the proof is complete.

Remark. The above proof was inspired by [**8**], pp. 349-353.

Now suppose that $W = \mathbb{H}^+/\Gamma(W)$ and $S = \mathbb{H}^+/\Gamma(S)$, where $\Gamma(W), T(S)$ are discrete co-finite area subgroups of $PSL(2,\mathbb{R})$, with $\Gamma(W)$ a normal subgroup of $\Gamma(S)$ such that $G = \Gamma(S)/\Gamma(W)$ is finite. Then S is the quotient of W by the group G and a result analogous to Corollary 1 holds. As in the Axiom A case, lying above each closed geodesic τ in S there are a finite number of closed geodesics $\tilde{\tau}_1, \dots, \tilde{\tau}_n$ in W. As before, each $\tilde{\tau}_i$ gives rise to a unique Frobenius element $[\tilde{\tau}_i] \in G$, the conjugacy class of which depends only on τ. If $n = |G|$ then we say that τ splits completely. This is the case if and only if $[\tilde{\tau}_i]$ is the identity, $1 \leq i \leq n$.

Once again let R_χ be an irreducible representation of G with irreducible character χ and define

$$L(s,\chi) = \prod_\tau \det\left(I - \frac{R_\chi([\tilde{\tau}])}{N(\tau)^s}\right)^{-1}$$

where $\tilde{\tau}$ lies over τ. If χ_0 denotes the trivial character, then $L(s,\chi) = \zeta_{\Gamma(S)}(s)$ and for $\chi \neq \chi_0$, $L(s,\chi)$ is analytic in a neighbourhood of $s = 1$. Applying the analysis of Section 3, we obtain the following.

Corollary 2. *Let $g \in G$ and $C = C(g)$ be its conjugacy class, then*

$$\prod_{\substack{N(\tau)\leq x \\ [\tilde{\tau}]\in C}} \left(1 - \frac{1}{N(\tau)}\right) \sim \frac{e^{-(|C|/|G|)\gamma}}{A(\log x)^{|C|/|G|}}$$

where

$$A = \operatorname{Res}(\zeta_{\Gamma(S)}, 1)^{|C|/|G|}\left\{\prod_{\chi\neq\chi_0} e^{-\chi(g^{-1})}L(1,\chi)\right\}^{|C|/|G|}$$

and

$$\sum_{\substack{N(\tau)\leq x \\ [\tilde{\tau}]\in C}} \frac{1}{N(\tau)} = \frac{|C|}{|G|}\log\log x + \text{ constant } + o(1).$$

5. The modular surface and quadratic forms

A classical example of a non-compact, finite area surface of constant negative curvature is the modular surface, $\mathbb{H}^+/\Gamma$ where $\Gamma = PSL(2,\mathbb{Z})$. (In fact its area is $2\pi^2/3$.) We shall now describe the elegant relationship between geodesics on the modular surface and quadratic forms (cf. Sarnak [16]).

We consider quadratic forms that are primitive and indefinite, for example

$$Q(x, y) = ax^2 + bxy + cy^2$$

where $(a, b, c) = 1$ and the discriminant $d = b^2 - 4ac$ satisfies $d > 0, d \equiv 0, 1 \pmod 4$ and is not a perfect square. Denote the set of such d by D. Two such forms Q, Q' are called equivalent if we can transform one to the other by a substitution

$$x' = \alpha x + \beta y$$
$$y' = \gamma x + \delta y$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$ and $\alpha\delta - \beta\gamma = 1$. This relation partitions forms into classes and it is clear that two forms from the same class have the same discriminant. Gauss showed that the number of classes with a given discriminant $d > 0$ is finite, this number is denoted by $h(d)$ [7].

The substitutions which preserve Q are called the automorphs of Q. All the automorphs of Q may be written in terms of solutions of Pell's equation

$$t^2 - du^2 = 4, \qquad t, u \in \mathbb{Z}, \qquad (t, u) \neq (0, 0)$$

by choosing

$$\alpha = \frac{1}{2}(t - bu) \qquad \beta = -cu$$
$$\gamma = au \qquad \delta = \frac{1}{2}(t + bu).$$

Let (t_0, u_0) be the solution for which $\varepsilon_d = \frac{1}{2}(t_0 + u_0\sqrt{d})$ is least, then all solutions (t, u) are generated by

$$\frac{1}{2}(t + u\sqrt{d}) = \varepsilon_d^n, \qquad n \in \mathbb{Z}.$$

In [7] Gauss noticed that

$$\sum_{\substack{d \in D \\ d \leq x}} h(d) \log \varepsilon_d = \frac{\pi^2 x^{3/2}}{18\varsigma(3)} + O(x \log x)$$

(here $\varsigma(s)$ is the Riemann zeta function) and this was later proved by Siegel [18]. One would like to be able to separate the quantities $h(d)$ and $\log \varepsilon_d$ to obtain an asymptotic formula for $\Sigma_{d \in D, d \leq x} h(d)$, but this appears to be difficult and remains an unsolved problem. However, using the next proposition, Sarnak obtained an

asymptotic expression for $h(d)$ summed over the sets $D_x = \{d \in D : \varepsilon_d \leq x\}$ [16].

Proposition 5. *There is a bijection between closed geodesics on the modular surface and equivalence classes of quadratic forms. Furthermore a closed geodesic corresponding to an equivalent class with discriminant d has length* $2 \log \varepsilon_d$.

Applying this correspondence to Proposition 4 and using the fact that the Laplace-Beltrami operator on the modular surface has no eigenvalues in [0, 3/16], Sarnak obtained

$$\sum_{d \in D_x} h(d) = \mathrm{li}(x^2) + O\left(x^{3/2}(\log x)^2\right),$$

Remark. The function $d \to \varepsilon_d$ seems fairly irregular and nothing much better than $\sqrt{d} \leq \varepsilon_d \leq e^d$ is known.

We now combine Proposition 5 with the results described in Theorem 2 to give two new asymptotic formulae involving $h(d)$ and ε_d.

Proposition 6.

$$\prod_{d \in D_x} \left(1 - \varepsilon_d^{-2}\right)^{h(d)} \sim \frac{e^{-\gamma}}{2\,\mathrm{Res}(\varsigma_\Gamma, 1) \log x}$$

and

$$\sum_{d \in D_x} h(d)\varepsilon_d^{-2} = \log\log x + \text{ constant } + O\left(\frac{1}{\log x}\right).$$

Let $p \geq 3$ be a prime and let $\Gamma(p)$ be the principal congruence subgroup of Γ of level p, i.e.

$$\Gamma(p) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod p \right\}.$$

The surface $\mathbb{H}^+/\Gamma(p)$ is a finite regular covering of $\mathbb{H}^+/\Gamma$ and the covering group $G = \Gamma/\Gamma(p) \simeq \mathrm{PSL}(2, \mathbb{Z}/p\mathbb{Z})$. In this situation a geodesic on $\mathbb{H}^+/\Gamma$ splits completely if and only if $d \in D_p = \{d \in D : p \mid u_0\}$ (where $\varepsilon_d = \frac{1}{2}(t_0, u_0\sqrt{d})$). Applying Corollary 2 and noting that

$$|G| = \frac{(p^2 - 1)p}{2},$$

we have:

Corollary 3.

$$\prod_{d\in D_{p,x}} \left(1-\varepsilon_d^{-2}\right)^{h(d)} \sim \frac{e^{-2\gamma/p(p^2-1)}}{A4^{1/p(p^2-1)}(\log x)^{2/p(p^2-1)}}$$

where $D_{p,x}=\{d\in D_p : \varepsilon_d\le x\}$,

$$A=\mathrm{Res}(\zeta_\Gamma,1)^{2/p(p^2-1)}\left\{\prod_{\chi\ne\chi_0} e^{-\dim\chi}L(1,\chi)\right\}^{2/p(p^2-1)}$$

(the product being taken over all non-trivial irreducible characters of $\mathrm{PSL}(2,\mathbb{Z}/p\mathbb{Z})$ *and*

$$\sum_{d\in D_{p,x}} h(d)\varepsilon_d^{-2}=\frac{2}{p(p^2-2)}\log\log x+\text{ constant } o(1).$$

Remark. Similar results may be obtained for quadratic forms over the ring of integers in $\mathbb{Q}(\sqrt{-D})$ where $D\in\mathbb{Z}$ is positive and square free by means of the correspondence between equivalence classes of such forms and geodesic on certain arithmetic three manifolds (Sarnak [17]).

References

1. L. M. Abramov, *On the entropy of a flow*, Transl. AMS. **49** (1966), 167-170.
2. S. Agmon, *Complex variable Tauberians*, Trans. Amer. Math. Soc. **74** (1953), 444-481.
3. R. Bowen, *Periodic orbits for hyperbolic flows*, Amer. J. Math. **94** (1972), 1-30.
4. ________, *The equidistribution of closed geodesics*, Amer. J. Math. **94** (1972), 413-423.
5. ________, *Symbolic dynamics for hyperbolic flows*, Amer. J. Math. **95** (1973), 429-459.
6. ________, "Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms," Lecture Notes in Mathematics 470. Springer-Verlag, Berlin, 1975.
7. C. F. Gauss, "Disquistiones Arithmeticae," Yale University Press,. New Haven and London, 1966
8. G. H. Hardy and E. M. Wright, "An Introduction to the Theory of Numbers," Oxford University Press. Oxford, 1938
9. D. A. Hejhal, *The Selberg trace formula and the Riemann zeta function*, Duke Math. J. **43** (1976), 441-482.
10. A. Manning, *Axiom A diffeomorphisms have rational zeta functions*, Bull. London Math. Soc. **3** (1971), 215-220.
11. W. Parry, *Bowen's equidistribution theory and the Dirichlet density theorem*, Ergodic Theory Dyn. Syst. **4** (1984), 117-134.

12. W. Parry and M. Pollicott, *An analogue of the prime number theorem for closed orbits of Axiom A flows*, Ann. of Math. **118** (1983), 573-591.
13. ________, *The Chebotarov theorem for Galois coverings of Axiom A flows*, Ergodic Theory Dyn. Syst. **6** (1986), 133-148.
14. M. Pollicott, *Meromorphic extensions of generalised zeta functions*, Invent. math. **85** (1986), 147-164.
15. D. Ruelle, "Thermodynamic Formalism," Addison-Wesley, Reading, Mass.. 1978
16. P. Sarnack, *Class numbers of indefinite binary quadratic forms*, J. Number Theory **15** (1982), 229-247.
17. P. Sarnack, *The arithmetic and geometry of some hyperbolic three manifolds*, Acta math. **151** (1983), 253-295.
18. C. L. Siegel, *The average measure of quadratic forms with given determinant and signature*, Ann. of Math. **45** (1944), 667-685.

Richard Sharp
Mathematics Institute
University of Warwick
Coventry, CV47AL
England