SELF-INTERSECTIONS OF CLOSED GEODESICS ON A NEGATIVELY CURVED SURFACE: STATISTICAL REGULARITIES

by

Steven P. Lalley
Department of Statistics
Purdue University

ABSTRACT

For compact surfaces of negative curvature it is shown that most closed geodesics of length $\approx \ell$ have about $C\ell^2$ self-intersections, for some constant $C > 0$, and these self-intersections are approximately equidistributed on the surface. For surfaces of constant negative curvative $-\kappa$ the value of the constant C is $\kappa/2\pi^2(g-1)$, where g is the genus.

1. Introduction and Statement of Main Results

The geodesic flow on a compact, negatively curved surface is perhaps the simplest example of a smooth flow for which typical orbits exhibit "random" behavior. Periodic orbits (i.e., closed geodesics) are *not* typical – there are only countably many, and they account for only a set of measure zero in the phase space. Nevertheless, the overall randomness of the flow is reflected in the sequence of periodic orbits (see, e.g., [L$_1$], [L$_2$]) in certain ways. The purpose of this note is to record some statistical regularities in the self-intersections of closed geodesics.

Let S be a compact, C^∞ Riemannian manifold of dimension 2 with strictly negative curvature at every point. There are countably many closed geodesics $\gamma_1, \gamma_2, \dots$ (one in each free homotopy class) with lengths $\ell_1 \le \ell_2 \le \cdots$. A celebrated result of Margulis [M] states that if $\pi(t) = \max\{n\colon \ell_n \le t\}$ then as $t \to \infty$

$$\pi(t) \sim \frac{e^{ht}}{ht}$$

where $h > 0$ is the topological entropy of the geodesic flow (for surfaces with constant curvature $-1, h = 1$).

Define s_n to be the number of (transversal) self-intersections of the closed geodesic γ_n.

Theorem 1: *For each compact, negatively curved surface S there is a constant $C = C_S > 0$ such that for every $\varepsilon > 0$,*

$$\lim_{t\to\infty} \pi(t)^{-1} |\{n \le \pi(t)\colon |s_n - C\ell_n^2| < \varepsilon \ell_n^2\}| = 1 \tag{1.1}$$

If S has constant curvature – κ and genus $g \ge 2$ then

$$C_S = \frac{\kappa}{2\pi^2(g-1)}. \tag{1.2}$$

Thus, "most" closed geodesics of length $\approx \ell$ have "about" $C_S\ell^2$ self-intersections. This result is consistent with the statement that "typical" closed geodesics are similar statistically to "generic" geodesics. Observe that if one randomly threw down ℓ geodesic segments of length 1 on S then the number of intersections would be about $C\ell^2$.

Theorem 1 seems to contradict Theorem 5 of [P], which states that most closed geodesics of length $\approx \ell$ have about $\tilde{C}\ell$ self-intersections. Apparently, the formula in remark (iii), p. 212 of [P] is incorrect: the singularity in the zeta function is not a simple pole, as stated, but rather a logarithmic singularity, so the Ikehara Tauberian theorem does *not* apply.

If $s_n > 0$ define α_n to be the probability distribution on S that assigns mass $\frac{1}{s_n}$ to each point of self-intersection of γ_n; if $s_n = 0$ define α_n to be the zero measure.

Theorem 2: *For each compact, negatively curved surface S there is a Borel probability measure α on S such that for each continuous $f\colon S \to R$ and $\varepsilon > 0$*

$$(1.3) \qquad \lim_{t\to\infty} \pi(t)^{-1} \sum_{n=1}^{\pi(t)} 1\{|\int f d\alpha_n - \int f d\alpha| < \varepsilon\} = 1.$$

If S has constant curvature then $\alpha = \nu$ = normalized area measure on S.

Thus most closed geodesics on S of length about ℓ experience about $C_S\ell^2$ self-intersections *and* these self-intersections are approximately distributed according to α. In fact the proof will show that α is the projection to S of the maximum entropy invariant probability measure for the geodesic flow. Theorem 1 is proved in sec. 3, Th. 2 in sec. 4. Both theorems are in fact corollaries of the strong equidistribution result, Theorem 7, of [L_1]. This result is explained in sec. 2.

Simple closed geodesics (closed geodesics with no self-intersections) are of interest for both topological and number-theoretic reasons (see [S]). Theorem 2 shows that simple closed geodesics are atypical of closed geodesics. Results of Birman and Series [BS] suggest that at least for a noncompact surface with finite area and free fundamental group, the number of simple closed geodesics with period $\leq t$ grows no faster than polynomially in t; thus simple closed geodesics are *very* atypical. It would be interesting to have an asymptotic formula for simple closed geodesics analogous to Margulis' formula.

2. Equidistribution of Closed Geodesics

We begin by discussing Theorem 7 of [L_1], from which Theorems 1-2 will be deduced.

Let $A = A(i,j)$ be an irreducible, aperiodic $\ell \times \ell$ matrix of zeros and ones, where $\ell \geq 2$, and define

$$\Sigma_A = \{x \in \mathop{\times}_{n=-\infty}^{\infty} \{1,2,\ldots,\ell\}\colon A(x_n, x_{n+1}) = 1 \;\; \forall\, n\}.$$

The shift $\sigma\colon \Sigma_A \to \Sigma_A$ is defined by $(\sigma x)_n = x_{n+1}$. Let $r\colon \Sigma_A \to R$ be a strictly positive function which is Lipschitz relative to the metric d_ρ on Σ_A defined by

$$d_\rho(x,y) = \sum_{n=-\infty}^{\infty} 1\{x_n \neq y_n\}\rho^{|n|}$$

for some $\rho \in (0,1)$. Define the suspension space

$$\Sigma_A^r = \{(x,s)\colon x \in \Sigma_A \text{ and } 0 \leq s \leq r(x)\}$$

with the points $(x, r(x))$ and $(\sigma x, 0)$ identified. The *suspension flow* (Σ_A^r, σ_t^r), $-\infty < t < \infty$, is defined as follows. Starting at any $(x,s) \in \Sigma_A^r$, move at unit speed up the fiber

$(x, s'), s \le s' \le r(x)$, until reaching $(x, r(x))$, then jump instantaneously to $(\sigma x, 0)$ and proceed up the vertical fiber $(\sigma x, s')$, etc. Equivalently,

$$\sigma_t^r(x,s) = (x, s+t) \quad \forall\, 0 \le s \le s+t \le r(x),$$
$$\sigma_{t+t'}^r = \sigma_t^r \circ \sigma_{t'}^r.$$

Observe that the periodic orbits of the suspension flow (Σ_A^r, σ^r) are precisely those orbits which pass through some $(x, 0)$ with x a periodic sequence. Since there are countably many periodic sequences, there are countably many periodic orbits of (Σ_A^r, σ^r). These may be labelled $\gamma_1^*, \gamma_2^*, \ldots$ and there periods $\ell_1^* \le \ell_2^* \le \ldots$. The distribution ν_n^* of γ_n^* may be defined by

$$\nu_n^*(B) = \int_0^{\ell_n} 1_B(\gamma_n^*(t))dt$$

where $B \subset \Sigma_A^r$ is a Borel set. Set $\pi^*(t) = \max\{n \colon \ell_n \le t\}$.

Theorem 7 *([L$_1$]): As $t \to \infty$,*

$$\pi^*(t) \sim e^{ht}/ht$$

where h is the topological entropy of (Σ_A^r, σ^r). For each continuous $f \colon \Sigma_A^r \to R$ and each $\varepsilon > 0$, as $t \to \infty$

$$\pi^*(t)^{-1} |\{n \le \pi^*(t) \colon |\int f d\nu_n^* - \int f d\nu^*| < \varepsilon\}| \longrightarrow 1,$$

where ν^* is the maximum entropy invariant measure for (Σ_A^r, σ^r).

Now consider the geodesic flow $\Phi = \Phi_t$ on the unit tangent bundle T_1S of a compact, negatively curved surface. This flow is of Anosov type [A], hence also Axiom A, so the results of [B$_2$] (also [R]) apply. Thus, there exists a suspension flow (Σ_A^r, σ^r) and a continuous map $\pi \colon \Sigma_A^r \to T_1S$ such that

(a) *π is surjective;*

(b) *π is at most N to 1 for some $N < \infty$;*

(c) *$\pi \circ \sigma_t^r = \Phi_t \circ \pi$ for all t;*

(d) *all but finitely many of the periodic orbits $\{\overline{\gamma}_n\}$ of Φ have the property that $\pi^{-1}(\overline{\gamma}_n)$ consists of a single periodic orbit of σ^r with the same least period.*

See [L$_3$], sec. 1 for an explanation of (d). Note that the periodic orbits $\overline{\gamma}_n$ of Φ are just the lifts to T_1S of the closed geodesics γ_n. For each $\overline{\gamma}_n$ the probability distribution $\overline{\nu}_n$ may be defined by

$$\int_{T_1S} f d\overline{\nu}_n = \frac{1}{\ell_n} \int_0^{\ell_n} f(\overline{\gamma}_n(t))dt$$

for $f \colon T_1S \to R$ continuous. By (a)–(d) above, $\overline{\nu}_n$ pulls back (via π^{-1}) to $\nu_{m_n}^*$, with only finitely many exceptions. Furthermore, the maximum entropy invariant probability measure $\overline{\nu}$ for the geodesic flow Φ pulls back to the maximum entropy measure ν^* for the suspension flow σ^r. Therefore, Theorem 1 implies

Corollary 1: *For each continuous $f: T_1S \to R$ and each $\varepsilon > 0$,*

$$\lim_{t\to\infty} \frac{1}{\pi(t)} |\{n \leq \pi(t): | \int f d\overline{\nu}_n - \int f d\overline{\nu}| < \varepsilon\} = 1.$$

3. The Number of Self-Intersections

Recall that $\overline{\nu}$ is the maximum entropy invariant probability measure for the geodesic flow on T_1S, and that ν is the induced measure on S.

Lemma 1: *For any nonempty, open set $U \subset T_1S, \overline{\nu}(U) > 0$. For any orbit $\overline{\gamma}(t)$ of the flow, if $\Gamma = \{\overline{\gamma}(t): -\infty < t < \infty\}$ then $\overline{\nu}(\Gamma) = 0$.*

Note: This is part of the folklore, but the proof is not easy to find.

Proof: It suffices to prove corresponding statements for an arbitrary suspension flow, by (a)–(c) of sec. 2. Let ν^* be the maximum entropy measure for a suspension flow (Σ_A^r, σ^r); define a probability measure μ on Σ_A by

$$\mu(F) = \nu^*\{(x, s): x \in F \text{ and } 0 \leq s \leq r(x)\}.$$

Then μ is the equilibrium state (sometimes called Gibbs state) for the function $-hr(x)$, by [BR], Prop. 3 (see [B_3], Ch. 1 for the definition).

We will prove that any Gibbs state μ is nonatomic. ¿From this it follows immediately that ν^* assigns measure zero to each individual orbit of the flow. Suppose μ has an atom x, i.e., $\mu(\{x\}) = \rho > 0$. Since μ is σ-invariant ([B_1], Th. (1.2)), $\mu(\{\sigma^n x\}) = \rho \ \ \forall\ n$; as μ is a finite measure, it must be that $x = \sigma^n x$ for some $n \geq 1$. But μ is mixing ([B_1], Prop. (1.14)) so it must be that $\mu(\{x\}) = 1$ and $x = \sigma x$. This is impossible, however, because by [B_1] Th. (1.2) a Gibbs state μ gives positive mass to every cylinder set.

It remains to show that $\nu^*(W) > 0$ for every open $W \subset \Sigma_A^r$. If W is open then it contains a rectangle $R = F \times [k/m, (k+1)/m)$, where k, m are positive integers and F is a cylinder set $F = \{y \in \Sigma_A: y_n = x_n \ \ \forall\ |n| \leq n_*\}$. By [$B_1$] Th. (1.2), $\mu(F) > 0$. Since ν^* is σ_t^r-invariant,

$$\nu^*(F \times [0/m, 1/m)) = \nu^*(F \times [1/m, 2/m)) = \ldots$$

(with the obvious convention about what happens when you get to the "ceiling" $\{(x, r(x))\}$: see Fig. 1). Let k_* be the smallest integer larger than $\min_{x\in F} r(x)$; then

$$\mu(F) \leq \sum_{i=0}^{k_*-1} \nu^*(F \times [i/m, (i+1)/m)) - k_*\nu_*(R)$$

(see Fig. 1 again) so $\nu^*(R) > 0$.

Figure 1

Lemma 2: *For any geodesic $\gamma(t)$, if $G = \{\gamma(t) \colon -\infty < t < \infty\}$ then $\nu(G) = 0$.*

Proof: The maximum entropy measure $\overline{\nu}$ is ergodic for the geodesic flow on $T_1 S$ (this follows from (a)–(c) of sec. 2, because the maximum entropy measure for a suspension flow is unique and consequently ergodic). Hence, if $\nu(G) > 0$ then for $\overline{\nu}$–a.e. orbit $\overline{\varphi}(t) = (\varphi(t), \varphi'(t))$ of the geodesic flow

$$\lim_{T \to \infty} T^{-1} \int_0^T 1_G(\varphi(t)) dt = \nu(G) > 0.$$

Now $\varphi(t)$ has at most countably many transversal intersections with G, and each transversal intersection contributes zero to the integral $\int_0^T 1_G(\varphi(t)) dt$. Therefore $\varphi(t)$ must intersect G tangentially. But $\varphi(t)$ and $\gamma(t)$ are both geodesics, so it follows that for some $s \in R, \varphi(t) = \gamma(t+s) \;\; \forall\; t$. This is impossible, because by Lemma 1 $\overline{\nu}$ assigns zero mass to the orbit $\{\overline{\gamma}(t) = (\gamma(t), \gamma'(t))\}$.

Each $(x, v) \in T_1 S$ determines a geodesic $\gamma(t)$ in S emanating from $\gamma(0) = x$ and with $\gamma'(0) = \nu$. For any $\delta > 0$ let $G_\delta(x, v)$ be the segment $\{\gamma(t) \colon 0 \leq t < \delta\}$ of this geodesic, considered as a subset of S. Let $\overline{\nu} \times \overline{\nu}$ denote the product measure on $T_1 S \times T_1 S$ determined by $\overline{\nu} \times \overline{\nu}(B_1 \times B_2) = \overline{\nu}(B_1)\overline{\nu}(B_2)$. Define

$$U_\delta = \{((x, v), (x', v')) \colon G_\delta(x, v) \cap G_\delta(x', v') \neq \phi\};$$

then $\overline{\nu} \times \overline{\nu}(U_\delta)$ is the probability that two independent, randomly chosen geodesic segments of length δ cross. (Note that the $\overline{\nu} \times \overline{\nu}$-probability of a nontransversal intersection is zero, by Lemma 1.)

Lemma 3: *For each* $\delta > 0, \overline{\nu} \times \overline{\nu}(U_\delta) > 0$.

Proof: Each $U_\delta, \delta > 0$, contains a nonempty open subset of $T_1S \times T_1S$, consequently also a rectangle $A \times B$ where A, B are nonempty open subsets of T_1S. Therefore, by Lemma 1, U_δ has positive $\overline{\nu} \times \overline{\nu}$-measure.

When $\overline{\mu}$ is close to $\overline{\nu}$ in the weak-$*$ topology then $\overline{\mu} \times \overline{\mu}$ is close to $\overline{\nu} \times \overline{\nu}$ in the weak-$*$ topology, and consequently $|\overline{\mu} \times \overline{\mu}(U_\delta) - \overline{\nu} \times \overline{\nu}(U_\delta)|$ should be small. The following lemma justifies this assertion (cf. [Bi], Th. 2.1, statement (v)).

Lemma 4: *Let* ∂U_δ *denote the (topological) boundary of* U_δ *in* $T_1S \times T_1S$. *Then*

$$\overline{\nu} \times \overline{\nu}(\partial U_\delta) = 0.$$

Proof: If $((x,v),(x',v')) \in \partial U_\delta$ then an endpoint of $G_\delta(x,v)$ lies on $G_\delta(x',v')$, or an endpoint of $G_\delta(x',v')$ lies on $G_\delta(x,v)$. Consequently, to prove the lemma it suffices to show that for each $(x',v') \in T_1S$,

$$\overline{\nu}\{(x,v) \colon G_\delta(x,v) \text{ has an endpoint on } G_\delta(x',v')\} = 0.$$

For $(x,v) \in T_1S$, one endpoint of $G_\delta(x,v)$ is x; call the other x_*. By Lemma 2, $\overline{\nu}\{(x,v) \colon x \in G_\delta(x',v')\} = 0$. But $\overline{\nu}$ is invariant under time reversal, and the segment $G_\delta(x,v)$ reversed has initial endpoint x_*, so $\overline{\nu}\{(x,v) \colon x_* \in G_\delta(x,v)\} = 0$.

Geodesics, being smooth curves, look like straight lines in the small. Therefore, if $\delta > 0$ is sufficiently small then two geodesic segments $G_\delta(x,v)$ and $G_\delta(x',v')$ of length δ will have at most one transverse intersection (this also uses the compactness of T_1S). Choose δ_0 sufficiently small that this is true for all $0 < \delta < \delta_0$. For the remainder of this section, fix δ with $0 < \delta < \delta_0$.

Let γ_n be a closed geodesic with length ℓ_n. Recall that s_n is the number of (transversal) self-intersections of γ_n, and that $\overline{\nu}_n$ is the distribution of $\overline{\gamma_n} = (\gamma_n, \gamma_n')$ in T_1S. For $k = 1, 2, \ldots$ let $m_k (= m_k^{(n)})$ be the least integer $\geq k\ell_n/\delta$. Define

$$\begin{aligned} x_{j,k} &= x_{j,k}^{(n)} = \gamma_n(j\delta/k), j = 0, 1, 2, \ldots, m_k - 1; \\ v_{j,k} &= v_{j,k}^{(n)} = \gamma_n'(j\delta/k), j = 0, 1, 2, \ldots, m_k - 1; \\ \overline{\nu}_{n,k} &= \text{ uniform probability distribution on} \{(x_{j,k}, v_{j,k}) \colon j = 0, 1, \ldots, m_k - 1\} \end{aligned}$$

(i.e., $\overline{\nu}_{n,k}$ is the probability measure that puts mass $1/m_k$ on each $x_{j,k}, v_{j,k}, j = 0, 1, \ldots,$ $m_k - 1$).

Lemma 5: weak-$*$ $\lim_{k \to \infty} \overline{\nu}_{n,k} = \overline{\nu}_n$.

This is immediate from the definition of the weak-$*$ topology and elementary properties of the Riemann integral.

Lemma 6: *For each* $k = 1, 2, \ldots,$

$$(m_k^2/2k^2)(\overline{\nu}_{n,k} \times \overline{\nu}_{n,k})(U_\delta) - 2m_k/k \leq$$
$$s_n \leq (m_k^2/2k^2)(\overline{\nu}_{n,k} \times \overline{\nu}_{n,k})(U_\delta).$$

Proof: Define G_j to be the segment of γ_n from $x_{j,k}$ to $x_{j+1,k}$ for $j = 0, 1, \ldots, m_k - 2$, and G_{m_k-1} to be the segment of γ_n from $x_{m_k-1,k}$ to $x_{0,k}$. Note that if $0 \leq j \leq m_k - 2$ then $G_j = G_{\delta/k}(x_{j,k}, v_{j,k})$, whereas $G_{m_k-1} = G_r(x_{m_k-1,k}, v_{m_k-1,k})$ for some $0 < r \leq \delta/k$. Since $\delta < \delta_0$, each intersection $G_i \cap G_j, i \neq j$, consists of at most a single point. Consequently,

$$s_n = \frac{1}{2} \sum_{(i,j):i \neq j} \sum 1\{G_i \cap G_j \neq \phi\}.$$

Next, define $G_i^* = G_\delta(x_{i,k}, v_{i,k})$. If $0 \leq i < m_k - k$ then $G_i^* = G_i \cup G_{i+1} \cup \ldots \cup G_{i+k-1}$, whereas if $m_k - k \leq i \leq m_k - 1$ then $G_i^* \supset G_i \cup G_{i+1} \cup \ldots \cup G_{i+k-1}$ but $G_i^* \subset G_i \cup G_{i+1} \cup \ldots \cup G_{i+k}$ (with the convention that if $i + j \geq m_k$ then $i + j$ should be reduced mod m_k). As before, each intersection $G_i^* \cap G_j^*, |i - j| \geq k$, consists of at most one point. Each G_i is contained in k distinct G_j^*, so by the result of the previous paragraph,

$$\begin{aligned} s_n &\leq \frac{1}{2k^2} \sum_{(i,j):|i-j| \geq k} \sum 1\{G_i^* \cap G_j^* \neq \phi\} \\ &\leq \frac{m_k^2}{2k^2}(\overline{\nu}_{n,k} \times \overline{\nu}_{n,k})(U_\delta). \end{aligned}$$

Moreover, each G_i intersects only k distinct G_j^* unless $0 \leq i < k$, in which case G_i may intersect $k + 1$ distinct G_j^*. Consequently,

$$\left(\frac{1}{2k^2} \sum_{(i,j):|i-j| \geq k} \sum 1\{G_i^* \cap G_j^* \neq \phi\} \right) - s_n \leq \frac{2km_k}{2k^2} = \frac{m_k}{k}.$$

Finally,

$$\begin{aligned} &\frac{m_k^2}{2k^2}(\overline{\nu}_{n,k} \times \overline{\nu}_{n,k})(U_\delta) - \frac{1}{2k^2} \sum_{(i,j):|i-j| \geq k} \sum 1\{G_i^* \cap G_j^* \neq \phi\} \\ &= \frac{1}{2k^2} \sum_{i=0}^{m_k-1} (2k - 1) = \frac{(2k-1)m_k}{2k^2} \leq \frac{m_k}{k}. \end{aligned}$$

Proof *of* (1.1): Corollary 1 implies that for most closed geodesics γ_n with length $\ell_n \leq t$ the distribution $\overline{\nu}_n$ is close to $\overline{\nu}$ in the weak-$*$ topology. Consequently, by Lemma 5, for k sufficiently large $\overline{\nu}_{n,k}$ is close to $\overline{\nu}$ and therefore $\overline{\nu}_{n,k} \times \overline{\nu}_{n,k}$ is close to $\overline{\nu} \times \overline{\nu}$ in the weak-$*$

topology. It follows, by Lemma 4 and [Bi], Th. 2.1, that $|\overline{\nu}_{n,k} \times \overline{\nu}_{n,k}(U_\delta) - \overline{\nu} \times \overline{\nu}(U_\delta)|$ is small. Thus, by Lemma 6, s_n/ℓ_n^2 is close to

$$C = \overline{\nu} \times \overline{\nu}(U_\delta)/\delta^2 > 0.$$

Since $\delta < \delta_0$ was arbitrary in this argument, the quantity $\overline{\nu} \times \overline{\nu}(U_\delta)/\delta^2$ is independent of δ, and so C can, in principle, be evaluated by letting $\delta \to 0$.

Lemma 7: *If S has constant negative curvature then*

$$\lim_{\delta \to 0} \overline{\nu} \times \overline{\nu}(U_\delta)/\delta^2 = \frac{4}{2\pi \text{ area } (S)} \tag{3.1}$$

Observe that area $(S) = 4\pi(g-1)/\kappa$ where $g =$ genus (S) and $-\kappa =$ curvature, by the Gauss-Bonnet theorem, so Lemma 7 implies (1.2).

Proof *of Lemma 7*: The measure $\overline{\nu}$ has a simple structure for a surface S of constant curvature. Elements of $T_1 S$ may be represented as (x, θ), where $x \in S$ and $-\pi \leq \theta < \pi$; in these coordinates

$$d\overline{\nu}(x, \theta) = d\nu(x) d\theta/2\pi$$

where $d\nu(x)$ is the normalized surface area measure on S.

Recall that $\overline{\nu} \times \overline{\nu}(U_\delta)$ is the probability that two independent, randomly chosen geodesic segments of length δ will intersect. To prove (3.1) it suffices to show that for any fixed geodesic segment G_δ of length δ

$$\overline{\nu}\{(x,\theta)\colon G_\delta(x,\theta) \cap G_\delta \neq \phi\} \sim 4\delta^2/2\pi \text{ area } (S)$$

as $\delta \to 0$ (uniformly for all choices of G_δ). To compute this probability, condition on the value of θ (the angle with G_δ); then the set of x for which $G_\delta(x,\theta) \cap G_\delta \neq \phi$ is approximately a rhombus of side δ and interior angle θ (at least when δ is small). Consequently,

$$\begin{aligned} &\overline{\nu}\{(x,\theta)\colon G_\delta(x,\theta) \cap G_\delta \neq \phi\} \\ &\qquad \sim \int_{-\pi}^{\pi} \delta^2 |\sin\theta| d\theta/2\pi \text{ area } (S). \end{aligned}$$

4. Distribution of Self-Intersection Points

In sec. 3 we showed that $\overline{\nu} \times \overline{\nu}(U_\delta) = C\delta^2$ for all sufficiently small $\delta > 0$. Thus we may define Borel probability measures μ_δ on $T_1 S \times T_1 S$ by

$$\mu_\delta(B) = \overline{\nu} \times \overline{\nu}(U_\delta \cap B)/C\delta^2$$

for Borel sets $B \subset T_1S$. Since $T_1S \times T_1S$ is compact, the Helly selection theorem (cf. [RS], Th. IV.21, which they call the "Banach-Alaoglu" theorem) implies that there is a weak-$*$ convergent subsequence $\mu_{\delta_n}, \delta_n \to 0$. Define

$$\overline{\alpha} = \text{ weak-}* \lim_{n\to\infty} \mu_{\delta_n}$$

(it will not matter which subsequence δ_n is used).

The measure $\overline{\alpha}$ induces on $S \times S$ a measure α via the natural projection $p \times p\colon T_1S \to S \times S$, in particular,

$$\alpha(B) = \overline{\alpha}((p \times p)^{-1}(B)), B \subset S \text{ Borel}.$$

The measure α is supported by the "diagonal" $D = \{(x,x)\colon x \in S\}$, because μ_δ is supported by the set U_δ, and if $((x,v),(x',v')) \in U_\delta$ then distance $(x,x') \leq 2\delta$. Thus α may be regarded as a probability measure on S. We will not bother to distinguish between S and D in the subsequent arguments.

Let V be an open subset of $S \times S$; define

$$\begin{aligned} V_\delta^+ &= \{(x_1,x_2) \in S \times S\colon \text{ distance } (x_i, V) < \delta\}, \\ V_\delta^- &= \{(x_1,x_2) \in V\colon \text{ distance } (x_i, V^c) > \delta\}. \end{aligned}$$

Lemma 8: *If* $\alpha(\partial V) = 0$ *then*

$$\begin{aligned} &\lim_{n\to\infty} \mu_{\delta_n}((p \times p)^{-1}(V_{2\delta_n}^+)) = \alpha(V) \text{ and} \\ &\lim_{n\to\infty} \mu_{\delta_n}((p \times p)^{-1}(V_{\delta_n}^-)) = \alpha(V). \end{aligned}$$

bf Proof: This is a standard argument. For any $\varepsilon > 0$ there exist open $V_1 \subset V \subset V_2$ such that $\alpha(V_2 \backslash V_1) < \varepsilon$ and $V_1 \subset V_\delta^-, V_2 \supset V_\delta^+$ for some $\delta > 0$ (this follows from the dominated convergence theorem), and also such that $\alpha(\partial V_2) = \alpha(\partial V_1) = 0$. Since $\mu_{\delta_n} \to \overline{\alpha}$ it then follows ([Bi], Th. 2.1 (v)) that

$$\begin{aligned} \mu_{\delta_n}((p \times p)^{-1}(V_1)) &\longrightarrow \alpha(V_1), \\ \mu_{\delta_n}((p \times p)^{-1}(V_2)) &\longrightarrow \alpha(V_2). \end{aligned}$$

The result then follows by monotonicity, because for large $n, V_1 \subset V_{\delta_n}^- \subset V \subset V_{2\delta_n}^+ \subset V_2$.

Now let γ_n be a closed geodesic with length ℓ_n, and let $\overline{\nu}_{n,k}, m_k$, etc. be as in Lemma 6. If V is any closed subset of $S \times S$ let $s_n(V)$ be the number of self-intersection points x of γ_n such that $(x,x) \in V$.

Lemma 9: *For each* $k = 1, 2, \ldots$,

$$\begin{aligned} &(m_k^2/2k^2)(\overline{\nu}_{n,k} \times \overline{\nu}_{n,k})(U_\delta \cap (p \times p)^{-1}(V_\delta^-)) - 2m_k/k \\ &\quad \leq s_n(V) \leq (m_k^2/2k^2)(\overline{\nu}_{n,k} \times \overline{\nu}_{n,k})(U_\delta \cap (p \times p)^{-1}(V_\delta^+)). \end{aligned}$$

Proof: This is virtually the same as the proof of Lemma 6. The only novelty is in keeping track of where the geodesic segments cross. Observe that if geodesic segments G_1, G_2 of length δ intersect transversally at x such that $(x, x) \in V$ then the (initial) endpoints x_1, x_2 of G_1, G_2 are such that $(x_1, x_2) \in V_\delta^+$. Similarly, if G_1, G_2 intersection at x and the initial endpoints x_1, x_2 are such that $(x_1, x_2) \in V_\delta^-$, then $(x, x) \in V$.

Proof *of Th. 2*: Let V be any open subset of $S \times S$ such that $\alpha(\partial V) = 0$. By Th. 1, Lemma 5, and Lemma 9, for any $\varepsilon > 0$ there is a $t_\varepsilon < \infty$ large enough that for all $t \geq t_\varepsilon$ and all but at most $\varepsilon\pi(t)$ closed geodesics γ_n with length $\leq t$,

$$C\mu_\delta((p \times p)^{-1}(V_\delta^-)) - \varepsilon \leq s_n(V)/\ell_n^2 \leq C\mu_\delta((p \times p)^{-1}(V_{2\delta}^+)) + \varepsilon.$$

The quantities ε and δ are both arbitrary, but affect how large t must be so that the preceding statement is true. By letting $\delta \to 0$ through the subsequence δ_n and appealing to Lemma 8, we see that for any $\varepsilon > 0$ there exists $t_\varepsilon < \infty$ such that for all $t \geq t_\varepsilon$ and all but at most $\varepsilon\pi(t)$ closed geodesics γ_n of length $\leq t$,

$$C\alpha(V) - 2\varepsilon \leq s_n(V)/\ell_n^2 \leq C\alpha(V) + 2\varepsilon.$$

Together with Th. 1, this implies that for most closed geodesics γ_n with $\ell_n \leq t$ the distribution α_n of self-intersection points is close to α in the weak-$*$ topology.

REFERENCES

[A] Anosov, D.V. (1967). Geodesic flows on closed Riemannian manifolds of negative curvature. *Proc. Steklov Inst.* **90**.

[Bi] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, N.Y.

[BS] Birman, J. and Series, C. (1984). An algorithm for simple curves on surfaces. *J. London Math. Soc.* (2) **29**, 331–342.

[B_1] Bowen, R. (1975). *Equilibrium States and the ErgodicTheorem of Anosov Diffeomorphisms.* Springer Lecture Notes in Math. **470**

[B_2] Bowen, R. (1973). Symbolic dynamics for hyperbolic flows. *Amer. J. Math.* **95**, 429–450.

[BR] Bowen, R. and Ruelle, D. (1975). Ergodic theory of Axiom A flows. *Inv. Math.* **29**, 181–202.

[L_1] Lalley, S.P. (1989). Renewal theorems in symbolic dynamics, with applications to geodesic flows, noneuclidean tessellations and their fractal limits. *Acta Mathematica* **163**, 1–55.

[L_2] Lalley, S.P. (1987). Distribution of periodic orbits of symbolic and Axiom A flows. *Adv. Appl. Math.* **8**, 154–193.

[L_3] Lalley, S.P. (1989). Closed geodesics in homology classes on surfaces of variable negative curvature. *Duke Math. J.* **58**, 795–821.

[M] Margulis, G. (1969). Applications of ergodic theory to the investigation of manifolds of negative curvature. *Funct. Analysis Appl.* **3**, 335–336.

[P] Pollicott, M. (1985). Asymptotic distribution of closed geodesics. *Israel J. Math.* **52**, 209–224.

[R] Ratner, M. (1973). Markov partitions for Anosov flows on n-dimensional manifolds. *Israel J. Math.* **15**, 92–114.

[RS] Reed, M., and Simon, B. (1980). *Methods of Mathematical Physics I. Functional Analysis.* Academic Press, N.Y.

[S] Series, C. (1985). The geometry of Markov numbers. *Math. Intell.* **7**, 20–29.