Counting Closed Geodesics on Rank One Manifolds

Roland Gunesch
COUNTING CLOSED GEODESICS ON RANK ONE MANIFOLDS

ROLAND GUNESCH

Abstract. We establish a precise asymptotic formula for the number of homotopy classes of periodic orbits for the geodesic flow on rank one manifolds of nonpositive curvature. This extends a celebrated result of G. A. Margulis to the nonuniformly hyperbolic case and strengthens previous results by G. Knieper.

We also establish some useful properties of the measure of maximal entropy.

1. Introduction

1.1. Manifolds of rank one. Let $M$ be a compact Riemannian manifold with all sectional curvatures nonpositive. For a vector $v \in TM$, the rank of $v$ is the dimension of the vector space of parallel Jacobi fields along the geodesic tangent to $v$. The rank of $M$ is the minimal rank of all tangent vectors. Obvious consequences of this definition are that

$$1 \leq \text{rank}(M) \leq \dim(M),$$

that the rank of $\mathbb{R}^k$ with the flat metric is $k$ and that

$$\text{rank}(M \times N) = \text{rank}(M) + \text{rank}(N).$$

Every manifold whose sectional curvature is never zero is automatically of rank one. Products with Euclidean $n$-space clearly have rank at least $n + 1$. However, it is possible for a manifold to be everywhere locally

2000 Mathematics Subject Classification. 53D25 (Geodesic flows), 32Q05 (Negative curvature manifolds), 37D40 (Dynamical systems of geometric origin and hyperbolicity (geodesic and horocycle flows, etc.)), 37C27 (Periodic orbits of vector fields and flows), 37C35 (Orbit growth).


The author thanks Anatole B. Katok for suggesting this problem and for valuable comments, helpful discussions and precious support. Gerhard Knieper provided helpful discussions during the creation of this work. Parts of this work were completed while the author was at the Isaac Newton Institute for the Mathematical Sciences, Cambridge, and at the ETH Zürich.
a product with Euclidean space and still have rank one. It turns out that the rank of a manifold of nonpositive curvature is the algebraic rank of its fundamental group ([BaEb]).

Apart from manifolds of negative curvature, examples are nonpositively curved surfaces containing flat cylinders or an infinitesimal analogue of a flat cylinder, as illustrated in the following diagram.

Figure 1.1. A surface of rank one with a flat strip and a parallel family of geodesics.

In higher dimensions, examples include M. Gromov's (3-dimensional) graph manifolds ([Gro]). There is an interesting rigidity phenomenon: Every compact 3-manifold of nonpositive curvature whose fundamental group is isomorphic to that of a graph manifold is actually diffeomorphic to that graph manifold ([Sch]).

We will study properties of manifolds of rank one in this article.

1.2. Reasons to study these spaces.

1.2.1. Rank rigidity. W. Ballmann ([Bal1]) and independently K. Burns and R. Spatzier ([BuSp]) showed that the universal cover of a nonpositively curved manifold can be written uniquely as a product of Euclidean, symmetric and rank one spaces. The first two types are understood, due to P. Eberlein and others. (A general introduction to higher rank symmetric spaces is e.g. [Ebe5]; see also [BGS]. For a complete treatment of rank rigidity, see [Bal2].)

Thus, in order to understand nonpositively curved manifolds, the most relevant objects to examine are manifolds of dimension at least two with rank one. This becomes even more obvious if one considers the fact that rank one is generic in nonpositive curvature ([BBE]). Thus, in a certain sense, “almost all” nonpositively curved manifolds have rank one.

1.2.2. Limits of hyperbolic systems. Another reason to study nonpositively curved manifolds is the following. On one hand, strongly hyperbolic systems, particularly geodesic flows on compact manifolds of negative curvature, are well understood since D. V. Anosov ([Ano], [Mar2], [Mar3]). Later, P. Eberlein established a condition weaker than
negative curvature which still ensures the Anosov property of the geodesic flow ([Ebe3], [Ebe4]). Also, Ya. Pesin and M. Brin extended the notion of hyperbolicity to that of partial hyperbolicity ([Pes], [BrPe1], [BrPe2]). On the other hand, much less is known about the dynamics of systems lacking strong hyperbolicity. The open set of geodesic flows on manifolds with negative curvature is “essentially” understood (hyperbolicity is an open property), and hence the edge of our knowledge about such flows is mainly marked by the boundary of this set, which is a set of geodesic flows on manifolds of nonpositive curvature. Therefore it is important to study the dynamics of these.

However, the set of nonpositively curved manifolds is larger than just the closure of the set of negatively curved manifolds. This can be seen e.g., as follows: Some nonpositively curved manifolds, such as Gromov’s graph manifolds, contain an embedded 2-torus. Thus their fundamental group contains a copy of \( \mathbb{Z}^2 \). Hence, by Preissmann’s theorem, they do not admit any metric of negative curvature. Therefore, the investigation in this article actually deals with even more than the limits of our current knowledge of strongly hyperbolic systems.

1.3. Statement of the result. We count homotopy classes of closed geodesics ordered by length in the following sense: The number \( P_t \) of homotopy classes of periodic orbits of length at most \( t \) is finite for all \( t \). (For a periodic geodesic there may be uncountably many periodic geodesics homotopic to it, but in nonpositive curvature they all have the same length.)

Trying to find a concrete and explicit formula for \( P_t \) which is accurate for all values of \( t \) is completely hopeless, even on very simple manifolds. Nonetheless, in this article we manage to derive an asymptotic formula for \( P_t \), i.e., a formula which tells us the behavior of \( P_t \) when \( t \) is large. We will show (Theorem 5.36):

\[
P_t \sim \frac{1}{ht} e^{ht}
\]

where the notation \( f(t) \sim g(t) \) means \( \frac{f(t)}{g(t)} \to 1 \) as \( t \to \infty \). This extends a celebrated result of G. A. Margulis to the case of nonpositive curvature. It also strengthens results by G. Knieper, which were the sharpest estimates known to this date in the setup of nonpositive curvature. This is explained in more detail in the following section.

2. History

2.1. Margulis’ asymptotics. The study of the functions \( P \) and \( b_t \), where \( b_t(x) \) is the volume of the geodesic ball of radius \( t \) centered at
x, was originated by G. A. Margulis in his dissertation [Mar1]. He covers the case where the curvature is strictly negative. His influential results were published in [Mar2] and the proofs were published eventually in [Mar4]. He established that, on a compact manifold of negative curvature,

\[(2.1) \quad b_t(x) \sim c(x)e^{ht}\]

for some continuous function \(c\) on \(M\). He also showed that

\[(2.2) \quad P_t \sim c'e^{ht},\]

for some constant \(c'\). In modern notation, the exponent \(h\) is the topological entropy of the geodesic flow. See [KaHa] for a modern reference on the topic of entropy.

Margulis pointed out that if the curvature is constant with value \(K\) then the exponential growth rate equals \((n - 1)\sqrt{-K}\) and that in this case the function \(c\) is constant. In fact, \(c \equiv 1/h\). Moreover, \(c' = 1/h\) for variable negative curvature.

2.2. Beyond negative curvature; Katok’s entropy conjecture.
The vast majority of the studies that have since been done are restricted to negative curvature; see e.g. [PaPo], [BaLe], [PoSh1]. The reason is that in that case techniques from uniformly hyperbolic dynamics can be applied. From the point of view of analysis, this case is much easier to treat. However, from a geometrical viewpoint, manifolds of nonpositive curvature are a natural object to study. Already in the seventies the investigation of manifolds of nonpositive curvature became the focus of interest of geometers. (Also more general classes have been studied since, such as manifolds without focal points, i.e. where every parallel Jacobi field with one zero has the property that its length increases monotonically when going away from the zero, or manifolds without conjugate points, i.e. such that any Jacobi field with two zeroes is trivial.) In 1984 at a MSRI problem session a major list of problems which were open at the time was compiled ([BuKa]), including A. Katok’s entropy conjecture: The measure of maximal entropy is unique.

One of the first result in the direction of asymptotics of closed geodesics in nonpositive curvature is the fact that the growth rate of closed geodesics equals the topological entropy \(h\), even if the curvature is just nonpositive (instead of strictly negative). G. Knieper calculated the growth rate of closed geodesics in [Kni3]. This result can also be deduced from A. Manning’s result [Man] that the growth rate of volume equals \(h\) in nonpositive curvature.
This shows in particular that the exponent in Margulis’ asymptotics must equal $h$ (we have already written Margulis’ equations that way). A method for showing that in the case of strictly negative curvature the constant $c'$ in equation (2.2) equals $1/h$ is outlined in C. Toll’s dissertation [Tol] and published in [KaHa]. This method was developed by Margulis in his thesis [Mar1] and published in [Mar4]. The behavior of the function $c$ in the asymptotic formula (2.1) was investigated by C. Yue in [Yue1] and [Yue2]. For recent developments concerning the asymptotics of the number of homology classes see e.g. the works of N. Anantharaman ([Ana1], [Ana2]), M. Babillot and F. Ledrappier ([BaLe]), M. Pollicott and R. Sharp ([PoSh2]), and S. P. Lalley ([Lal]).

It took almost two decades after Knieper’s and Manning’s results, which in turn were published about one decade after Margulis’ results, until the next step in the analysis of asymptotics of periodic orbits on manifolds of nonpositive curvature was completed, again by Knieper.

2.3. Knieper’s multiplicative bounds. In 1996 G. Knieper proved asymptotic multiplicative bounds for volume and periodic orbits ([Kni2]) which, in the case of nonpositive curvature and rank one, were the sharpest results known until now: There exists a constant $C$ such that for sufficiently large $t$,

\[ \frac{1}{C} \leq \frac{s_t(x)}{e^{ht}} \leq C, \]

where $s_t(x)$ is the volume of the sphere of radius $t$ centered at $x$, and

\[ \frac{1}{Ct} \leq \frac{P_t}{e^{ht}} \leq C. \]

The main step in the proof of these asymptotics is the proof of Katok’s entropy conjecture. Knieper also demonstrated in [Kni1] that the measure of maximal entropy can be obtained via the Patterson-Sullivan construction ([Pat], [Sul]; see also [Kai1], [Kai2]). Moreover, for the case of higher rank Knieper obtained asymptotic information using rigidity. Namely,

\[ \frac{1}{C} \leq \frac{s_t}{t^{(\text{rank}(M)-1)/2}e^{ht}} \leq C. \]

He also estimates the number of closed geodesics in higher rank.

Knieper subsequently sharpened his results. With the same method he is able to prove that in the rank one case actually

\[ \frac{1}{C} \leq \frac{P_t}{e^{ht/t}} \leq C \]

holds (see [Kni4]). Still, the quotient of the upper and lower bounds is a constant which cannot be made close to 1.
The question whether in this setup of nonpositive curvature and rank one one can prove more precise multiplicative asymptotics—namely without such multiplicative constants—has remained open so far. In this article we establish this result.

Remark 2.1. For non-geodesic dynamical systems no statements providing asymptotics similar the ones mentioned here are known. One of the best known results is that for some prevalent set of diffeomorphisms the number of periodic orbits of period \( n \) is bounded by \( \exp(C \cdot n^{1+\delta}) \) for some \( \delta > 0 \) ([KaHu]).

But even for geodesic flows in the absence of nonpositive curvature it is difficult to count—or even find—closed geodesics. The fact that every compact manifold has even one closed geodesic was established only in 1951 by Lyusternik and Fet ([LuFe]). In the setup of positively curved manifolds and their kin, one of the strongest known results is H.-B. Rademacher’s Theorem from 1990 ([Rad1], [Rad2]) stating that every connected and simply connected compact manifold has infinitely many (geometrically distinct) closed geodesics for a \( C^r \)-generic metric for all \( r \in [2, \infty] \). See also [Rad3] for this.

For Riemannian metrics on the 2-sphere, existence of many closed geodesics took considerable effort to prove. The famous Lyusternik-Shnirelman Theorem asserts the existence of three (geometrically distinct) closed geodesics. The original proof in [LuSch] is considered to have gaps. Complete proofs were given by W. Ballmann ([Bal3]), W. Klingenberg (with W. Ballmann’s help) ([Kli]) and also J. Jost ([Jos1], [Jos2]). See also [BTZ1], [BTZ2].

J. Franks ([Fra]) established that every metric of positive curvature on \( S^2 \) has infinitely many (geometrically distinct) geodesics. This is a consequence of his results about area-preserving annulus homeomorphisms. V. Bangert managed to show existence of infinitely many (geometrically distinct) geodesics on \( S^2 \) without requiring positive curvature by means of variational methods ([Ban]).

For the case of Finsler manifolds, there actually exist examples of simply connected manifolds that possess only finitely many geometrically distinct closed geodesics. On \( S^2 \) such examples were constructed by A. B. Katok in [Kat1] as a by-product of a more general construction. Explaining this particular aspect of Katok’s construction is also the topic of [Mat]. [Zil] also studies the Katok examples.

In this article we derive asymptotics like the ones Margulis obtained. We prove them for nonpositive curvature and rank one using non-uniform hyperbolicity. Hence the same strong statement is true in considerably greater generality.
3. Geometry and dynamics in nonpositive curvature

Let $M$ be a compact rank one Riemannian manifold of nonpositive curvature. As is usual, we assume it to be connected and geodesically complete. Let $\tilde{M}$ be the unit sphere bundle of the universal covering of $M$. For $v \in \tilde{M}$ let $c_v$ be the geodesic satisfying $c'(0) = v$ (which is hence automatically parameterized by arclength). Here $c'$ of course denotes the covariant derivative of $c$. Let $g = (g^t)_{t \in \mathbb{R}}$ be the geodesic flow on $\tilde{M}$, which is defined by $g^t(v) := c'_v(t) =: v_t$.

3.1. Review of asymptotic geometry.

**Definition 3.1.** Let $\pi : TM \to M$ be the canonical projection. We say that $v, w \in \tilde{M}$ are **positively asymptotic** (written $v \sim w$) if there exists a constant $C$ such that $d(\pi g^t v, \pi g^t w) < C$ for all $t > 0$. This is evidently an equivalence relation. Similarly, $v, w \in \tilde{M}$ are **negatively asymptotic** if $-v \sim -w$.

Recall that $\text{rank}(v) := \dim \{\text{parallel Jacobi fields along } c_v\}$. Clearly the rank is constant along geodesics, i.e. $\text{rank}(c'(t)) = \text{rank}(c'(0))$ for all $t \in \mathbb{R}$.

**Definition 3.2.** We call a vector $v \in \tilde{M}$, as well as the geodesic $c_v$, **regular** if $\text{rank}(v) = 1$ and **singular** if $\text{rank}(v) > 1$. Let $\text{Reg}$ and $\text{Sing}$ be the sets of regular and singular vectors, respectively.

**Remark 3.3.** The set $\text{Reg}$ is open since rank is semicontinuous in the sense that $\text{rank}(\lim_n v_n) \geq \lim_n \text{rank}(v_n)$.

**Remark 3.4.** For every $v \in \tilde{M}$ and every $p \in \tilde{M}$ there exists some $w_+ \in S_p \tilde{M}$ which is positively asymptotic to $v$ and some $w_- \in S_p \tilde{M}$ which is negatively asymptotic to $v$. In contrast, the existence of $w_{+/-} \in T_p \tilde{M}$ which is simultaneously positively and negatively asymptotic to $v$ is rare. Moreover, if $v \sim w$ and $-v \sim -w$ then $v, w$ bound a flat strip, i.e. a totally geodesic embedded copy of $[-a, a] \times \mathbb{R}$ with Euclidean metric. Here the number $a$ is positive if $v, w$ do not lie on the same geodesic trajectory. In particular, if $\text{rank}(v) = 1$ (hence $c_v$ is a regular geodesic) then there does not exist such $w$ with $w \sim v$ and $-w \sim -v$ through any base point in the manifold outside $c_v$. In other words, if $w \sim v$ and $-w \sim -v$ on a rank 1 manifold then $w = g^t v$ for some $t$. On the other hand, if $\text{rank}(v) > 1$ (and thus $c_v$ is a singular geodesic) then $v$ and hence $c_v$ may lie in a flat strip of positive width, and in that case there are vectors $w$ with $w \sim v$ and $-w \sim -v$ at base points outside $c_v$, namely at all base points in that flat strip.
Since $\tilde{M}$ is of nonpositive curvature, it is diffeomorphic to $\mathbb{R}^n$ by the Hadamard-Cartan theorem, hence to an open Euclidean $n$-ball. It admits the compactification $\overline{M} = \tilde{M} \cup M(\infty)$ where $M(\infty)$, the boundary at infinity of $\tilde{M}$, is the set of equivalence classes of positively asymptotic vectors, i.e., $M(\infty) = SM/\sim$.

A detailed description of spaces of nonpositive curvature, even without a manifold structure, can be found in [Bal2].

3.2. Stable and unstable spaces.

Definition 3.5. Let $\mathcal{K} : T\tilde{M} \to S\tilde{M}$ be the connection map, i.e.

\[ \mathcal{K}\xi := \nabla_{d\pi\xi}Z \]

where $\nabla$ is the Riemannian connection and $Z(0) = d\pi\xi$, $\frac{d}{dt} Z(t)|_{t=0} = \xi$. We obtain a Riemannian metric on $SM$, the Sasaki metric, by setting $\langle \xi, \eta \rangle := \langle d\pi\xi, d\pi\eta \rangle + \langle \mathcal{K}\xi, \mathcal{K}\eta \rangle$ for $\xi, \eta \in T_vSM$ where $v \in SM$. Hence we can talk about length of vectors in $T\tilde{M}$.

There is a canonical isomorphism $(d\pi, \mathcal{K})$ between $T_vSM$ and the set of Jacobi fields along $c_v$. It is given by $\xi \mapsto J_\xi$ with $J_\xi(0) = d\pi \cdot \xi$, $J'_\xi(0) = \mathcal{K}\xi$. This uses the well-known fact that a Jacobi field is determined by its value and derivative at one point.

The space $T\tilde{M}$, i.e. the tangent bundle of the unit sphere bundle, admits a natural splitting

\[ T\tilde{M} = E^s \oplus E^u \oplus E^0, \]

i.e. $T_v\tilde{M} = E^s_v \oplus E^u_v \oplus E^0_v$ for all $v \in S\tilde{M}$, where

\[ E^0_v := \mathbb{R} \cdot \frac{d}{dt} g^i v \bigg|_{t=0}, \]

\[ E^s_v := \{ \xi \in T_vS\tilde{M} : \xi \perp E^0, J_\xi \text{ is the stable Jacobi field along } d\pi\xi \}, \]

\[ E^u_v := \{ \xi \in T_vS\tilde{M} : \xi \perp E^0, J_\xi \text{ is the unstable Jac. field along } d\pi\xi \}. \]

Definition 3.6. For $v \in S\tilde{M}$, define $W^s(v)$, the stable horosphere based at $v$, to be the integral manifold of the distribution $E^s$ passing through $v$. Similarly, define $W^u(v)$, the unstable horosphere based at $v$, via integrating $E^u$. The projection of $W^s$ (resp. $W^u$) to $\tilde{M}$ is again called the stable horosphere (resp. the unstable horosphere). The flow direction of course integrates to a geodesic trajectory, which one might call $W^0(v)$. The 0- and $u$-directions are jointly integrable, giving rise to an integral manifold $W^{0u}$, and similarly the 0- and $s$-directions give rise to an integral manifold $W^{0s}$. We write $B^i_\delta$ (resp. $\overline{B}^i_\delta$) for the open (resp. closed) $\delta$-neighborhood in $W^i$ ($i = u, s, 0u, 0s, 0$).
On the other hand, the \( u \)- and \( s \)-directions are usually not jointly integrable. Continuity of these foliations has been proven in this form by P. Eberlein ([Ebe2]) and J.-H. Eschenburg ([Esch]):

**Theorem 3.7.** Let \( M \) be a compact manifold of nonpositive curvature. Then the foliation \( \{W^s(v) : v \in SM\} \) of \( SM \) by stable horospheres is continuous. The same holds for the foliation \( \{W^u(v) : v \in SM\} \) of \( SM \) by unstable horospheres.

Note that due to compactness of \( M \) (hence of \( SM \)), the continuity is automatically uniform.

During the same years, Eberlein considered similar questions on Visibility manifolds ([Ebe2]). The continuity result was improved by M. Brin ([BaPe, Appendix A]) to Hölderness on the Pesin sets; see [BaPe] for the definition of these sets. For our discussion, uniform continuity is sufficient.

The following result is easier to show in the hyperbolic case (i.e. strictly negative curvature) than for nonpositive curvature, where it is a major theorem, proven by Eberlein ([Ebe1]):

**Theorem 3.8.** Let \( M \) be a compact rank one manifold of nonpositive curvature. Then stable manifolds are dense. Similarly, unstable manifolds are dense.

### 3.3. Important measures.

The Riemannian structure gives rise to a natural measure \( \lambda \) on \( SM \), called the Liouville measure. It is finite since \( M \) is compact. It is the prototypical smooth measure, i.e., for any smooth chart \( \varphi : U \to \mathbb{R}^{2n-1}, U \subset SM \) open, the measure \( \varphi_* \lambda \) on a subset of \( \mathbb{R}^{2n-1} \) is smoothly equivalent to Lebesgue measure.

The well-known variational principle (see e.g. [KaHa]) asserts that the supremum of the entropies of invariant probability measures on \( SM \) is the topological entropy \( h \). The variational principle by itself of course guarantees neither existence nor uniqueness of a **measure of maximal entropy**, i.e. one whose entropy actually equals \( h \). These two facts were established in the setup of nonpositive curvature by Knieper ([Kn1]):

**Theorem 3.9.** There is a measure of maximal entropy for the geodesic flow on a compact rank one nonpositively curved manifold. Moreover, it is unique.

The proof uses the Patterson-Sullivan construction ([Pat], [Sul]; see also [Kai1], [Kai2]). Knieper’s construction builds the measure as limit of measures supported on periodic orbits.
For the case of strictly negative curvature, the measure of maximal entropy was previously constructed (in a different way) by Margulis ([Mar3]). He used it to obtain his asymptotic results. His construction builds the measure as the product of limits of measures supported on pieces of stable and unstable leaves. The measure thus obtained is hence called the **Margulis measure**. It agrees with the **Bowen measure** which is obtained as limit of measures concentrated on periodic orbits. U. Hamenstädt ([Ham]) gave a geometric description of the Margulis measure by projecting distances on horospheres to the boundary at infinity, and this description was immediately generalized to Anosov flows by B. Hasselblatt ([Has]).

The measure of maximal entropy is adapted to the dynamical properties of the flow. In particular, we will see that the conditionals of this measure show uniform expansion/contraction with time. In negative curvature, this can be seen by considering the Margulis measure, where this property is a natural by-product of the construction. In nonpositive curvature, however, this property is not immediate. We show it in Theorem 4.6.

The measure of maximal entropy is sometimes simply called **maximal measure**. In the setup of nonpositive curvature, the name **Knieper measure** could be appropriate.

**Remark 3.10.** It is part of Katok’s entropy conjecture and shown in [Kni1] that \( m(\text{Sing}) = 0 \) (and in fact even that \( h(g|_{\text{Sing}}) < h(g) \)). In contrast, whether \( \lambda(\text{Sing}) = 0 \) or not is a major open question; it is equivalent to the famous problem of ergodicity of the geodesic flow in nonpositive curvature with respect to the Liouville measure \( \lambda \). On the other hand, ergodicity of the geodesic flow in nonpositive curvature with respect to \( m \) has been proven by Knieper.

A very useful dynamical property is mixing, which implies ergodicity. For nonpositive curvature mixing has been proven by M. Babillot ([Bab]):

**Theorem 3.11.** The measure of maximal entropy for the geodesic flow on a compact rank one nonpositively curved manifold is mixing.

We use this property in our proof of the asymptotic formula.

### 3.4. Parallel Jacobi fields.

**Lemma 3.12.** The vector \( v \in SM \) is regular if and only if \( W^u(v) \), \( W^s(v) \) and \( W^0(v) \) intersect transversally at \( v \).

Here transversality of the three manifolds means that

\[
T_vSM = T_vW^u \oplus T_vW^0 \oplus T_vW^s.
\]
Proof. $W^u(v)$ and $W^s(v)$ intersect with zero angle at $v$ if and only if there exist

$$\xi \in TW^u(v) \cap TW^s(v) \subset T_vSM.$$  

But $\xi \in TW^s(v)$ is true if and only if $J_\xi$ is the stable Jacobi field along $c_v$, and $\xi \in TW^u(v)$ is true if and only if $J_\xi$ is the unstable Jacobi field along $c_v$. A Jacobi field $J$ is both the stable and the unstable Jacobi field along $c_v$ if and only if $J$ is parallel. The nonexistence of such $J$ perpendicular to $c_v$ is just the definition of rank one. \qed

3.5. Coordinate boxes.

**Definition 3.13.** We call an open set $U \subset SM$ of diameter at most $\delta$ **regularly coordinated** if for all $v, w \in U$ there are unique $x, y \in U$ such that

$$x \in B^u_\delta(v), \ y \in B^0_\delta(x), \ w \in B^s_\delta(y).$$

In other words, $v$ can be joined to $w$ by means of a unique short three-segment path whose first segment is contained in $W^u(v)$, whose second segment is a piece of a flow line and whose third segment is contained in $W^s(w)$.

**Proposition 3.14.** If $v$ is regular then it has a regularly coordinated neighborhood.

**Proof.** Some $4\delta$-neighborhood $V$ of $v$ is of rank one. Let

$$U = B^s_\delta(g^{(-\delta,\delta)}B^u_\delta(v)).$$

This is contained in $V$ and hence of rank one. It is open since $W^0$, $W^u$ and $W^s$ are transversal by Lemma 3.12.

By construction, for any $w \in V$, there exists a pair $(x, y)$ such that

$$B^u_\delta(v) \ni x \in B^0_\delta(y), \ y \in B^s_\delta(w).$$

Assume there is another pair $(x', y')$ with this property. From

$$B^u_\delta(x) \ni v \in B^u_\delta(x')$$

we deduce $x \in B^u_{2\delta}(x')$, and from

$$B^0_\delta(x) \ni y \in B^s_\delta(w), \ w \in B^s_\delta(y'), \ y' \in B^0_\delta(x')$$

we deduce $x \in B^0_{4\delta}(x')$. Hence $x$ and $x'$ are simultaneously positively and negatively asymptotic; therefore, they bound a flat strip. Since $V$ is of rank one, there is no such strip of nonzero width in $U \subset V$. Hence $x$ and $x'$ lie on the same geodesic. Since $x \in W^u(x')$, these two points are identical.

The same argument with $u$ and $s$ exchanged shows that $y = y'$. Hence the pair $(x, y)$ is unique. \qed
3.6. The Busemann function and conformal densities.

**Definition 3.15.** Let \( b(., q, \xi) \) be the Busemann function centered at \( \xi \in \tilde{M}(\infty) \) and based at \( q \in \tilde{M} \). It is given by

\[
b(p, q, \xi) := \lim_{p_n \to \xi} (d(q, p_n) - d(p, p_n))
\]

for \( p, q \in \tilde{M} \) and is independent of the sequence \( p_n \to \xi \).

**Remark 3.16.** The function \( b \) satisfies \( b(p, q, \xi) = -b(q, p, \xi) \). Moreover,

\[
b(p, q, \xi) = \lim_{t \to \infty} (d(c_{p, \xi}(t), q) - t)
\]

where \( c_{p, \xi} \) is the geodesic parameterized by arclength with \( c_{p, \xi}(0) = p \) and \( c_{p, \xi}(t) \to \xi \) as \( t \to \infty \).

For \( \xi \) and \( p \) fixed, we have

\[
b(p, p_n, \xi) \to -\infty \quad \text{for} \quad p_n \to \xi \\
b(p, p_n, \xi) \to +\infty \quad \text{for} \quad \lim_{n} p_n \in \tilde{M}(\infty) \setminus \{\xi\}.
\]

We use the sign convention where \( b(p, q, \xi) \) is negative whenever \( p, q, \xi \) lie on a geodesic in this particular order.

**Definition 3.17.** \((\mu_p)_{p \in \tilde{M}}\) is a \( h \)-dimensional Busemann density (also called conformal density) if the following are true:

- For all \( p \in \tilde{M} \), \( \mu_p \) is a finite nonzero Borel measure on \( \tilde{M}(\infty) \).
- \( \mu_p \) is equivariant under deck transformations, i.e., for all \( \gamma \in \pi_1(M) \) and all measurable \( S \subset \tilde{M}(\infty) \) we have
  \[
  \mu_{\gamma p}(\gamma S) = \mu_p(S).
  \]
- When changing the base point of \( \mu_p \), the density transforms as follows:
  \[
  \frac{d\mu_p}{d\mu_q}(\xi) = e^{-hb(q,p,\xi)}.
  \]

In the case of nonpositive curvature, Knieper has shown in [Kni1] that \( \mu_p \) is unique up to a multiplicative factor and that it can be obtained by the Patterson-Sullivan construction.

4. The measure of maximal entropy

In section 5 we will use the fact that if \( m \) is the measure of maximal entropy then it gives rise to conditional measures \( m^u, m^{0u}, m^s \) and \( m^{0s} \) on unstable, weakly unstable, stable and weakly stable leaves which have the property that the measures \( m^{0u} \) and \( m^u \) expand uniformly with \( t \) and that \( m^s \) and \( m^{0s} \) contract uniformly with \( t \).
Remark 4.1. In [Gun] we present an alternative and more general construction of the measure of maximal entropy in nonpositive curvature and rank one which follows the principle of Margulis’ construction. Using that construction, the uniform expansion/contraction properties shown here are already a straightforward consequence of the construction. Also, that construction works for non-geodesic flows satisfying suitable cone conditions (see [Kat2] for these). On the other hand, Knieper’s approach, which substantially requires properties of rank one nonpositively curved manifolds, is shorter and therefore is the one we use in this article.

First we give Knieper’s definition of the measure of maximal entropy ([Kn1]):

Definition 4.2. Let \((\mu_p)_{p \in \tilde{M}}\) be a Busemann density. Let

\[
\Pi : S\tilde{M} \to \tilde{M}(\infty) \times \tilde{M}(\infty), \quad \Pi(v) := (v_\infty, v_-\infty)
\]

be the projection of a vector to both endpoints \(v_{\pm\infty} = \lim_{t \to \pm\infty} \pi g^tv\) of the corresponding geodesic. Then the measure of maximal entropy of a set \(A \subset S\tilde{M}\) (we can without loss of generality assume \(A\) to be regular) is given by

\[
m(A) := \int_{\xi,\eta \in \tilde{M}(\infty), \xi \neq \eta} \text{len}(A \cap \Pi^{-1}(\xi, \eta)) e^{-h(b(p,q,\xi)+b(p,q,\eta))}d\mu_p(\xi)d\mu_p(\eta),
\]

where \(q \in \pi \Pi^{-1}(\xi, \eta)\) and \(p \in \tilde{M}\) is arbitrary.

Here \text{len} is the length of the geodesic segment. Saying that \(\Pi^{-1}(\xi, \eta)\) is a geodesic already is a slight simplification, but a fully justified one since we need to deal only with the regular set.

4.1. Discussion of the conditionals. Given a vector \(v \in S\tilde{M}\) with base point \(p\), we want to put a conditional measure \(m^n\) on the stable horosphere \(b(p, .., \xi)^{-1}(0)\) given by \(v\) and centered at \(\xi := v_\infty\) (or on \(W^s(v)\), which is the unit normal bundle of \(b(p, .., \xi)^{-1}(0)\)). This conditional is determined by a multiplier with respect to some given measure on this horosphere. Note that the set of points \(q\) on the horosphere is parameterized by the set \(\tilde{M}(\infty) \setminus \{\xi\}\) via projection from \(\xi\) into the boundary at infinity, hence the multiplier depends on \(\eta := \text{proj}_\xi(q)\), i.e. is proportional to \(d\mu_x(\eta)\) for some \(x\). The canonical choice for \(x\) is \(p\). Clearly the whole horosphere has infinite \(m^n\)-measure, but \(\mu_x\) is finite for any \(x\). Thus the multiplier of \(d\mu_p\) has to have a singularity, and this has to happen at \(\eta = \xi\) since any neighborhood of \(\xi\)
is the projection of the complement of a compact piece of the horosphere. The term $e^{-hb(p,q,\eta)}$ has the right singularity (note that $\eta \to \xi$ means $q \to \xi$), and by the basic properties of the Busemann function the term $e^{-hb(p,q,\eta)}$ then converges to infinity. Therefore we investigate $m_p(q) := e^{-hb(p,q,\eta)}d\mu_p(\eta)$. First we prove that this is indeed the stable conditional measure for $dm^s$. We will parameterize $dm$ by vectors instead of their base points.

**Definition 4.3.** For $v, w \in \tilde{M}$, let

$$dm^u_v(w) := e^{-hb(\pi v, \pi w, w_\infty)} \cdot d\mu_{\pi v}(w_\infty),$$

$$dm^s_v(w) := e^{-hb(\pi v, \pi w, w_-)} \cdot d\mu_{\pi v}(w_-).$$

**Proposition 4.4.** $dm^s_v, dm^u_v$ and $dt$ are the stable, unstable and center conditionals of the measure of maximal entropy.

**Proof.** Observe that

$$dt dm^u_v(w) dm^s_v(w) = dt e^{-h(b(\pi v, \pi w, w_\infty) + b(\pi v, \pi w, w_-))} 
\cdot d\mu_{\pi v}(w_\infty) d\mu_{\pi v}(w_-)$$

$$= dt e^{-h(b(p,q,\xi) + b(p,q,\eta))} d\mu_p(\xi) d\mu_p(\eta) =: E$$

with $p := \pi v$, $q := \pi w$, $\xi := w_\infty$, $\eta := w_-$. This formula already agrees with the formula in Definition 4.2, although the meaning of the parameters does not necessarily do so: In Definition 4.2, $p$ and to some extend $q$ are arbitrary in $\tilde{M}$, while in the formula for $\tilde{E}$ they are fixed. Thus we need to show that if we change them within the range allowed in Definition 4.2, the value of $\tilde{E}$ does not change.

**Lemma 4.5.** The term $\tilde{E}$ does not change if $q$ is replaced by any point in $\tilde{M}$ on the geodesic $c_{q\xi}$ from $\eta$ to $\xi$ and $p$ by an arbitrary point in $\tilde{M}$.

**Proof.** First we show that $q$ can be allowed to be anywhere on $c_{q\xi}$: Parameterize $c_{q\xi}$ by arclength with arbitrary parameter shift in the direction from $\eta$ to $\xi$. Replacing $q = c_{q\xi}(s)$ by $q' = c_{q\xi}(s')$ changes $b(p,q,\xi)$ to $b(p,q',\xi) = b(p,q,\xi) - (s' - s)$ since we move the distance $s' - s$ closer to $\xi$. It also changes $b(p,q,\eta)$ to $b(p,q',\eta) = b(p,q,\eta) + (s' - s)$ since we move the distance $s' - s$ away from to $\eta$. Thus $\tilde{E}$ does not change under such a translation of $q$.

Now fix $q$ anywhere on $c_{q\xi}$ and replace $p$ by some arbitrary $p' \in \tilde{M}$. Note that

$$d\mu_{p'}(\xi) = e^{hb(p',p,\xi)} d\mu_p(\xi),$$

$$b(p', q, \xi) = b(p, q, \xi) + b(p', p, \xi),$$
which of course also holds with $\xi$ replaced by $\eta$. Thus

$$e^{-h(b(p',q,\xi)+b(p',q,\eta))}d\mu_{p'}(\xi)d\mu_{p'}(\eta) = e^{-h(b(p,q,\xi)+b(p,q,\eta))}d\mu_{p}(\xi)d\mu_{p}(\eta).$$

Hence $E$ also does not change if $p$ is changed to any arbitrary point. □

This also concludes the proof of Proposition 4.4. □

4.2. **Proof of uniform expansion/contraction of the conditionals.** Let $w_t$ denote $g^tw$.

**Theorem 4.6** (Uniform expansion/contraction of the conditionals). For all $t \in \mathbb{R}$ and all $v, w \in SM$ we have

$$dm^u_v(w_t) = e^{ht} \cdot dm^u_v(w),$$
$$dm^s_v(w_t) = e^{-ht} \cdot dm^s_v(w).$$

The same uniform expansion holds with $dm^u$ replaced by $dm^{0u} = dm^u dt$ and the same uniform contraction with $dm^s$ replaced by $dm^{0s} = dm^s dt$.

**Proof.**

$$dm^s_v(w_t) = e^{-hb(\pi_v,\pi w_t, w_{-\infty})}d\mu_{\pi_v}(w_{-\infty})$$
$$= e^{-h(b(\pi_v,\pi w, w_{-\infty})+b(\pi w, \pi w_t, w_{-\infty}))}d\mu_{\pi_v}(w_{-\infty})$$
$$= e^{-hb(\pi_v,\pi w, w_{-\infty})-ht}d\mu_{\pi_v}(w_{-\infty})$$
$$= e^{-ht} \cdot e^{-hb(\pi_v,\pi w, w_{-\infty})}d\mu_{\pi_v}(w_{-\infty})$$
$$= e^{-ht} \cdot dm^s_v(w).$$

Similarly, the equality $b(\pi_v, \pi w_t, w_{+\infty}) = b(\pi v, \pi w, w_{+\infty}) + t$ yields

$$dm^u_v(w_t) = e^{ht} \cdot dm^u_v(w).$$

This shows the desired uniform expansion of $m^u$ and the uniform contraction of $m^s$. From this we also immediately see the uniform expansion of $m^{0u}$ and the uniform contraction of $m^{0s}$ since $dt$ is evidently invariant under $g^t$. □

4.3. **Proof of holonomy invariance of the conditionals.** Another important property of the conditional measures on the leaves is holonomy invariance. We formulate holonomy invariance on infinitesimal unstable pieces here, but of course this is equivalent to holonomy invariance that deals with pieces of leaves of (small) positive size.

We consider positively asymptotic vectors $w, w'$ and calculate the infinitesimal $0u$-measure on corresponding leaves. We let $v, v'$ be some (arbitrary) base points used as parameters for the pieces of leaves, so that $w$ lies in the same $0u$-leaf of $v$ and similarly $w'$ in that of $v'$. The factor $dt$ is evidently invariant, so we do not have to mention it any further.
\textbf{Theorem 4.7} (Holonomy invariance of the conditionals of the measure of maximal entropy).

\[ dm_u^v(w) = dm_u^{v'}(w') \]

whenever \( v' \in W^s(v), \ w' \in W^u(w), \ w \in W^{0u}(v) \) and \( w' \in W^{0u}(v') \). In that case also \( dm^{0u}_0(w) = dm^{0u}_0(w') \) holds.

Similarly,

\[ dm^s_v(w) = dm^s_{v'}(w') \]

whenever \( v' \in W^u(v), \ w' \in W^u(w), \ w \in W^{0s}(v) \) and \( w' \in W^{0s}(v') \), and in that case also \( dm^{0s}_0(w) = dm^{0s}_0(w') \) holds.

\textit{Proof.} Note that the equation \( w' \in W^s(w) \) is equivalent to the two equations

\[ w'_\infty = w_\infty, \quad b(\pi w, \pi w', w_\infty) = 0. \]

The latter equation is equivalent to \( b(p, \pi w, w_\infty) = b(p, \pi w', w_\infty) \) for all \( p \in \tilde{M} \). Thus clearly

\[ dm_u^v(w') = e^{-h(b(\pi v', \pi w', w_\infty)} d\mu_{\pi v'}(w'_\infty) \]

\[ = e^{-h(b(\pi v', \pi w', w_\infty))} d\mu_{\pi v'}(w_\infty). \]

Now

\[ b(\pi v', \pi w', w_\infty) = b(\pi v', \pi v, w_\infty) + b(\pi v, \pi w', w_\infty) \]

\[ = b(\pi v', \pi v, w_\infty) + b(\pi v, \pi w, w_\infty) \]

and \( d\mu_{\pi v'}(w_\infty) = e^{-h(b(\pi v, \pi v', w_\infty))} d\mu_{\pi v}(w_\infty). \) Thus

\[ dm_u^v(w') = e^{-h(b(\pi v', \pi w', w_\infty) + b(\pi v, \pi v', w_\infty))} d\mu_{\pi v}(w_\infty) \]

\[ = e^{-h(b(\pi v', \pi w, w_\infty) + b(\pi v, \pi v', w_\infty))} d\mu_{\pi v}(w_\infty) \]

\[ = e^{-h(b(\pi v, \pi w, w_\infty)} d\mu_{\pi v}(w_\infty) \]

\[ = dm_u^v(w). \]

The proof for \( dm^s \) is analogous. \( \square \)

Note that \( m^{0u} \) is invariant under holonomy along \( s \)-fibers and \( m^{0s} \) under holonomy along \( u \)-fibers, but \( m^u \) is not invariant under holonomy along \( 0s \)-fibers and \( m^s \) not invariant under holonomy along \( 0u \)-fibers due to expansion (resp. contraction) in the flow direction.
5. Counting closed geodesics

In this final section we count the periodic geodesics on $M$. The method used here is a generalization of the method which, for the special case of negative curvature, was outlined in [Tol] and provided with more detail in [KaHa]. Margulis ([Mar1], [Mar4]) is the originator of that method, although the presentation in this article looks quite different.

**Definition 5.1.** Let $f = f(t, \varepsilon), \ g = g(t, \varepsilon) : [0, \infty) \times (0, 1) \rightarrow (0, \infty)$ be expressions depending on $t$ and $\varepsilon$. We are interested in the behavior for $t$ large and $\varepsilon > 0$ small.

Write

\[ f \sim g \]

if for all $\alpha > 0$ there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ there exists $t_0 \in (0, \infty)$ such that for all $t > t_0$ we have

\[ \left| \ln \frac{f(t, \varepsilon)}{g(t, \varepsilon)} \right| < \alpha. \]

Write

\[ f \asymp g \]

if there exists $K \in \mathbb{R}, \ \varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ there exists $t_0 \in (0, \infty)$ such that for all $t > t_0$ we have

\[ \left| \ln \frac{f(t, \varepsilon)}{g(t, \varepsilon)} \right| < K \varepsilon. \]

We write

\[ f \cong g \]

if there exists $f'$ with $f \sim f' \asymp g$, i.e. if there exists $K \in \mathbb{R}$ so that for all $\alpha > 0$ there exists $\varepsilon_0 > 0$ such that for all $\varepsilon \in (0, \varepsilon_0)$ there exists $t_0 > 0$ such that for all $t > t_0$ we have

\[ \left| \ln \frac{f(t, \varepsilon)}{g(t, \varepsilon)} \right| < K \varepsilon + \alpha. \]

Thus the relation "$\cong$" is implied by both "$\sim$" and "$\asymp$", which are the special cases $K = 0$ and $\alpha = 0$, respectively; but the relations "$\sim$" and "$\asymp$" are independent.

**Remark 5.2.** Similarly, in the definition of "$\sim$" and "$\cong$", the variable $t_0$ may depend on $\varepsilon$, i.e. the convergence in $t$ does not have to be uniform with respect to $\varepsilon$. All arguments in the rest of this article work without requiring this uniformity.
Remark 5.3. Obviously these relations are also well-defined if the domain of the functions \( f, g \) is \([T_0, \infty) \times (0, \gamma)\) for some \( T_0, \gamma > 0 \).

Lemma 5.4. The relations \( \sim \), \( \bowtie \) and \( \cong \) are equivalence relations.

Proof. It suffices to consider \( \cong \) since the others are special cases of it. Reflexivity and symmetry are trivial. If the functions \( f_1, f_2, f_3 : [0, \infty) \times (0, \infty) \to (0, \infty) \) satisfy \( f_1 \cong f_2 \cong f_3 \), i.e., for \( i = 1, 2 \) we have \( \forall K_i \forall \alpha > 0 \exists \varepsilon > 0 \forall \varepsilon \in (0, \varepsilon_0, i) \exists t_{0, i} > 0 \forall t > t_{0, i} : |\ln(f_i(t, \varepsilon)/f_{i+1}(t, \varepsilon))| < K_i\varepsilon + \alpha/2 \), then clearly \( \exists K_3 \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0, 3) \exists t_{0, 3} > 0 \forall t > t_{0, 3} : 
\end{equation}

\[
|\ln(f_1(t, \varepsilon)/f_3(t, \varepsilon))| < K_3\varepsilon + \alpha,
\]
which is the case we need.

Remark 5.5. In Definition 5.1, we could have written \( |\ln(f(t, \varepsilon)/g(t, \varepsilon))| \) instead of \( |\ln(f(t, \varepsilon)/g(t, \varepsilon)) - 1| \). This would be equivalent to our definition since the terms \( \ln x \) and \( |x - 1| \) differ by at most a factor 2 (indeed any \( a > 1 \)) for all \( x \) close enough to 1. The advantage of our notation is that multiple estimates can easily be transitively combined, as seen in the proof of Lemma 5.4. Also, our notation is symmetric in \( f, g \).

5.1. The flow cube. Fix any \( v_0 \in \text{Reg} \). Choose sufficiently small \( \varepsilon > 0 \) and \( \delta > 0 \) such that \( 4\delta < 2\varepsilon < \text{inj}(M) \) (the injectivity radius of \( M \)), such that \( B_{4\varepsilon}(v_0) \subset \text{Reg} \), and such that further requirements on the smallness of these which we will mention later are satisfied.

Definition 5.6. Let the flow cube be \( A := \overline{B^s(g^{[0,\varepsilon]}(B_\delta^u(v_0)))} \subset \text{Reg} \). Here \( B_\delta^u(v_0) \) is the closed unstable ball of radius \( \delta \) around \( v_0 \). We choose \( B^s = \overline{B^s(v)} \) as the closure of an open set contained in the closed stable ball of radius \( \delta \) around \( v \in g^{[0,\varepsilon]}(B_\delta^u(v_0)) \); this set, which depends on \( v \), can be chosen in such a way that it contains \( v \) and that \( A \) has the following local product structure: For all \( w, w' \in A \) there exists a unique \( \beta \in [-\varepsilon, \varepsilon] \) such that
\[
\overline{B^s(w)} \cap \overline{B_\delta^u(g^\beta w')}
\]
is nonempty, and in that case it is exactly one point. This is the local product structure in \( \text{Reg} \) described in Proposition 3.14. We call \( B^s(v) \) the stable fiber (or stable ball) in \( A \) containing \( v \).

In the following arguments, the cube \( A \) will first be fixed. In particular, \( \varepsilon \) and \( \delta \) are considered fixed (although subject to restrictions on their size). Then we make asymptotics certain numbers depending
on $t$ and $A$ as $t \to \infty$ (while $A$, hence $\varepsilon$, is fixed). Afterwards we will consider what happens to those asymptotics when $\varepsilon \to 0$.

**Definition 5.7.** Let the **depth** $\tau : A \to [0, \varepsilon]$ be defined by

$$v \in B^s(g^{\tau(v)}B^s(v_0)).$$

**Lemma 5.8.** For all $v \in A$, $w \in B^u_{28}(v) \cap A$ it is true that

$$|\tau(w) - \tau(v)| < \varepsilon^2/2.$$

**Proof.** The foliation $W^u$ is uniformly continuous by Theorem 3.7 and compactness of $SM$, and without loss of generality $\delta$ was chosen small enough. \hfill $\square$

**Lemma 5.9** (Stable fiber contraction). There is a function $\sigma = \sigma(t)$ such that

$$m^s(g^tB^s(v)) \gg \sigma(t)$$

for all $v \in A$. In particular, for all $v, w$ in $A$ we have

$$m^s(g^tB^s(v)) \gg m^s(g^tB^s(w)).$$

Moreover, the constants in the relation $\gg$ can be chosen independent of $v, w$, i.e., there exists $K > 0$, $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$. 

---

**Figure 5.1.** The flow cube $A$: an unstable neighborhood of $v_0$ (top) is iterated (center) and a stable neighborhood of that is formed (bottom).
and all flow cubes \( A = A(\varepsilon) \) there exists \( t_0 \) such that for all \( t > t_0 \) and all \( v, w \in A \) we have

\[
\left| \ln \frac{m^s(g^tB^s(v))}{\sigma(t)} \right| < K\varepsilon \quad \text{and} \quad \left| \ln \frac{m^s(g^tB^s(v))}{m^s(g^tB^s(w))} \right| < K\varepsilon.
\]

**Proof.** First we show the second claim. Observe that for any \( a \in (0, \inj(M)/2) \) (with \( a \) independent of \( \varepsilon \)), the set \( g^{[0,a]}B^s(v) \) is \( u \)-holonomic to a subset \( S \) of \( g^{[-2\varepsilon,a+2\varepsilon]}B^s(w) \). Thus

\[
m^s(B^s(v)) = m^s(B^s(w)) = \frac{m^0(S)}{m^0(g^{[0,a]}B^s(w))} \leq \frac{m^0(g^{[-2\varepsilon,a+2\varepsilon]}B^s(w))}{m^0(g^{[0,a]}B^s(w))} \leq \frac{\int_{-2\varepsilon}^{a+2\varepsilon} e^{-ht} dt}{\int_0^a e^{-ht} dt} \approx 1,
\]

since the quotient of the integrals can be bounded by \( 1 + K\varepsilon \). The inequality is symmetric in \( v \) and \( w \), proving equality. Hence \( m^s(B^s(v)) \gg m^s(B^s(w)) \), i.e. \( \exists K > 0 \), \( \varepsilon_0 > 0 \) \( \forall \varepsilon \in (0, \varepsilon_0) \) \( \exists t_0 \) \( \forall t > t_0 \) \( \forall v, w \in A \):

\[
\left| \ln \frac{m^s(g^tB^s(v))}{m^s(B^s(v))} \right| < K\varepsilon.
\]

Using uniform contraction on \( s \)-fibers (Theorem 4.6) gives

\[
\left| \ln \frac{m^s(g^tB^s(v))}{m^s(g^tB^s(w))} \right| < K\varepsilon,
\]

i.e., \( m^s(g^tB^s(v)) = m^s(g^tB^s(w)) \), showing the second claim. It immediately follows that we can define

\[
\sigma(t) := m^s(g^tB^s(v))
\]

for some arbitrary \( v \in A \), and this definition does not depend on \( v \) (up to \( \asymp \)-equivalence). The constant \( K \) in \( \gg \) is independent of \( v \). This also shows the first claim. \( \square \)

**Remark** 5.10. Uniform contraction (Theorem 4.6) then shows that \( \sigma(t) = \text{const} \cdot e^{-ht} \).

### 5.2. Expansion at the boundary.

**Definition 5.11.** For the cube \( A \) as above, we call

\[
\begin{align*}
\partial^u A := \overline{B^s}(g^{[0,\varepsilon]}(\partial B^s_\partial(v_0))) & \quad \text{the unstable end of the cube,} \\
\partial^s A := (\partial B^s)(g^{[0,\varepsilon]}(B^s_\partial(v_0))) & \quad \text{the stable end,} \\
\partial_0 A := \overline{B^s}(B^s_\partial(v_0)) & \quad \text{the back end and} \\
\partial_\varepsilon A := \overline{B^s}(g^\varepsilon(B^s_\partial(v_0))) & \quad \text{the front end of the cube.}
\end{align*}
\]
For $v \in A$ define
\[ s(v) := \sup \{ r : B^u_r(v) \subset A \} \]
to be the **distance to the unstable end** of the flow cube.

The stable and the unstable end are topologically the product of an interval, a $k$-ball and a $(k-1)$-sphere, where $k = \dim W^u(v) = \dim W^s(v) = (\dim SM - 1)/2$; hence they are connected iff $k \neq 1$, i.e. iff $M$ is not a surface.

**Lemma 5.12** (Expansion of the distance to the unstable end). There exists a monotone positive function $S : \mathbb{R} \to \mathbb{R}$ satisfying $S(t) \to 0$ as $t \to \infty$ and such that if $s(v) > S(t)$ for an element $v \in A$ which satisfies $g^t v \in A$ then
\[ B^u_\delta (g^t v) \cap A \subset g^t B^u_{S(t)}(v). \]

That means that if a point $v$ is more than $S(t)$ away from the unstable end of the cube then the the image under $g^t$ of a small $u$-disc (of size $> S(t)$) around $v$ has the property that its unstable end is completely outside $A$.

**Proof.** By nonpositivity of the curvature, $B^u_\delta$ noncontracts, i.e., for all $p, q \in B^u_\delta$ the function $t \mapsto d(g^t p, g^t q)$ is nondecreasing. This is true even infinitesimally, i.e. for unstable Jacobi fields. By convexity of Jacobi fields and rank 1, such distances also cannot stay bounded. Hence the radius of the largest $u$-ball contained in $g^t B^u_\delta$ becomes unbounded for $t \to \infty$.

Hence for all $\gamma > 0$ we can find $T_\gamma < \infty$ such that
\[
(5.1) \quad g^{T_\gamma} B^u_\delta (v) \supset B^u_\gamma (g^{T_\gamma} (v)).
\]

By compactness of $A$, this choice of $T_\gamma$ can be made independently of $v \in A$. Without loss of generality $T_\gamma$ is a strictly decreasing function of $\gamma$. Choose a function $S : [0, \infty) \to (0, \infty)$ so that $S(t) \leq \gamma$ for $t > T_\gamma$.

E.g., choose $S(\cdot) = T^{-1}$, i.e. $T_{S(t)} = t$ for $t \geq 0$. $S$ can be chosen decreasing since $T_\gamma$ can be. Therefore, given $v \in A$, if $t > T_{S(v)}$ then $s(v) > S(t)$, and thus equation (5.1) shows the claim. \qed

**Remark 5.13.** The convergence of $S$ to zero in the previous Lemma is not necessarily exponential, as opposed to the case where the curvature of $M$ is negative (i.e. the uniformly hyperbolic case). However, we do not need this property of exponential convergence.

If the smallest such $S$ would not converge to zero, it would require the existence of a flat strip of width $\lim \inf_{t \to \infty} S(t) = \lim_{t \to \infty} S(t)$, which would intersect $A$. Since a neighborhood of $A$ is regular, this cannot happen.
5.3. Intersection components and orbit segments.

**Definition 5.14.** Let \( A'_t \) be the set of \( v \in A \) with \( s(v) \geq S(t) \) and \( \tau(v) \in [\varepsilon, \varepsilon - 2\varepsilon^2] \). Thus \( A'_t \) is the set \( A \) with a small neighborhood of the unstable end and of the front end and back end removed.

**Definition 5.15.** Let \( \Phi_t \) be the set of all full components of intersection at time \( t \): If \( I \) is a connected component of \( A'_t \cap g^t(A'_t) \) then define
\[
\phi^t_I := g^{[-\varepsilon,\varepsilon]}(I) \cap A \cap g^t(A),
\]
\[
\Phi_t := \{ \phi^t_I : I \text{ is a connected component of } A'_t \cap g^t(A'_t) \}.
\]
Let \( N(A,t) := \#\Phi_t \) be the number of elements of \( \Phi_t \).

We call the set \( g^{[-\varepsilon,\varepsilon]}v \cap A \) the geometric orbit segment of length \( \varepsilon \) in \( A \) through \( v \). Similarly we speak about the orbit segment of length \( \varepsilon - 2\varepsilon^2 \) in \( A'_t \).

Let \( \Phi^s_t := \{ \phi^t_I \in \Phi_t : \phi^t_I \text{ intersects } \partial^s A'_t \} \).

**Lemma 5.16.** For every geometric orbit segment of length \( \varepsilon - 2\varepsilon^2 \) in \( A'_t \) that belongs to a periodic orbit of period in \([t - \varepsilon + 2\varepsilon^2, t + \varepsilon - 2\varepsilon^2]\) there exists a unique \( \phi^t_I \in \Phi_t \) through which the geometric orbit segment passes.

**Proof.** Existence: If \( g^t o = o \) for an orbit segment \( o \) of length \( \varepsilon - 2\varepsilon^2 \) of \( A'_t \) that belongs to a periodic orbit of period \( L \in [t - \varepsilon + 2\varepsilon^2, t + \varepsilon - 2\varepsilon^2] \) then \( o \) also intersects \( g^t A'_t \), hence some component of \( A'_t \cap g^t A'_t \).

Uniqueness: Assume that \( o \) passes through \( \phi^t_I, \phi^t_J \in \Phi_t \), i.e. \( p = o(a) \in \phi^t_I, q = o(b) \in \phi^t_J \) for \( |b-a| \leq \varepsilon \). Then \( o \) passes through \( I, J \) (the connected components corresponding to \( \phi^t_I, \phi^t_J \)) respectively. Without loss of generality, \( p, q \in A'_t \). Since \( g^t A'_t \) is pathwise connected, there is a path \( c \) in \( g^t A'_t \) from \( p \) to \( q \). Using the local product structure, we can assume that \( c \) consists of a segment in \( W^u \), followed by a segment in \( W^s \), followed by a segment in \( W^s \). By applying \( g^{-t} \), we get a path \( g^{-t} o c \) in \( A'_t \) from \( o(a-t) \in A'_t \) to \( o(b-t) \in A'_t \). The local product structure in \( A'_t \) and the fact that distances along unstable fibers become \( > 2\varepsilon \) for \( t \to \infty \) whereas distances along stable fibers become \( > 2\varepsilon \) for \( t \to -\infty \) (see Lemma 5.12) show that the \( u \)-segment and the \( s \)-segment of \( g^{-t} o c \) have length \( 0 \). Therefore \( g^{-t} o c \) and hence \( c \) is an orbit segment. This means that \( c \) lies in \( A'_t \) and in \( g^t A'_t \). Hence \( p \) and \( q \) lie in the same component, i.e. \( \phi^t_I = \phi^t_J \).

In the other direction, we have the following Lemma:

**Lemma 5.17.** For every \( \phi^t_I \in \Phi_t \setminus \Phi^s_t \) there exists a unique periodic orbit with period in \([t - \varepsilon, t + \varepsilon]\) and a unique geometric orbit segment on that orbit passing through \( \phi^t_I \).
In other words, up to a small error, intersection components correspond to periodic orbits, and of all orbit segments that belong to such a periodic orbit, just one orbit segment goes through any particular full component of intersection.

Proof. Choose $\phi^t_I$. It suffices to consider the case $t \geq 0$. Since $A'_t \subset A$ has rank one, it follows that for every $v \in A'_t$ any nonzero stable Jacobi field along $c_v$ is strictly decreasing in length, and any nonzero unstable Jacobi field is strictly increasing in length. Since the set of stable (resp. unstable) Jacobi fields is linearly isomorphic to $E^s$ (resp. $E^u$) via $(d\pi, K)^{-1}$, it follows that for all $v \in A'_t \cap g^t A'_t$:

\[ |dg^t \xi| < |\xi| \quad \forall \xi \in E^s(v) \backslash \{0\}, \]

\[ |dg^{-t} \xi| < |\xi| \quad \forall \xi \in E^u(v) \backslash \{0\}. \]

By compactness of $A'_t$ and hence of $\phi^t_I$ there exists $c < 1$ such that for all $v \in A'_t \cap g^t A'_t$:

\[ |dg^t \xi| < c|\xi| \quad \forall \xi \in E^s(v) \backslash \{0\}, \]

\[ |dg^{-t} \xi| < c|\xi| \quad \forall \xi \in E^u(v) \backslash \{0\}. \]

Hence $g^t$ restricted to $\phi^t_I$ is (apart from the flow direction) hyperbolic. By the assumption that $\phi^t_I \not\in \Phi^*_t$, stable fibers are mapped to stable fibers that do not intersect the stable end of the flow cube. Thus the first return map on a transversal to the flow is hyperbolic. Hence it has a unique fixed point.

Therefore there exists a unique periodic orbit through $\phi^t_I$. Two geometrically different (hence disjoint) orbit segments would give rise to two different fixed points. Hence the geometric segment on the periodic orbit is also unique. \hfill \Box

5.4. Intersection thickness.

Definition 5.18. Define the intersection thickness (or intersection length) $\theta : \Phi_t \to [0, \varepsilon]$ by

\[ \theta(\phi^t_I) := \varepsilon - \sup \left\{ \tau(v) : v \in g^t \left( \bigcup_{w \in A, g^t w \in I} g^{[-\varepsilon, \varepsilon]} w \cap \partial_0 A \right) \right\} \]

for such $\phi^t_I$ which intersect $\partial_+ A$ (the front end of $A$) and

\[ \theta(\phi^t_I) := \inf \left\{ \tau(v) : v \in g^t \left( \bigcup_{w \in A, g^t w \in I} g^{[-\varepsilon, \varepsilon]} w \cap \partial_0 A \right) \right\} \]

for such $\phi^t_I$ which intersect $\partial_0 A$ (the back end of $A$).
Lemma 5.19 (The average thickness is asymptotically half that of the flow box).

\[
\frac{1}{N(A,t)} \sum_{\phi^t_l \in \Phi_t} \theta(\phi^t_l) \cong \frac{\varepsilon}{2}.
\]

In other words, \( \exists K < \infty \ \forall \alpha > 0 \ \exists \varepsilon_0 > 0 \ \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \ \exists t_0 > 0 \ \forall t > t_0 : \frac{1}{N(A,t)} \left| \ln(2 \sum_{\phi^t_l \in \Phi_t} \theta(\phi^t_l)/\varepsilon) \right| < K\varepsilon + \alpha. \)

Proof. Take any full component of intersection \( \phi^t_l \in \Phi_t \). Assume that it intersects the front end of \( A \). We cut \( A \) along \( n := \lfloor 1/\varepsilon \rfloor \) pieces

\[
A_i := \left\{ v \in A : \tau(v) \in \left[ \frac{i\varepsilon}{n}, \frac{(i+1)\varepsilon}{n} \right) \right\}
\]

of equal measure \( (i = 0, \ldots, n-1) \). By the mixing property, \( m(A_i \cap g^t A_0) \) is asymptotically independent of \( i \) as \( t \to \infty \). Hence the number of full components of intersection of \( A_i \) with \( g^t A_0 \) is asymptotically independent of \( i \). Since any intersection component of \( A_i \cap g^t A_0 \) has depth \( \tau \) with \( |\tau - i\varepsilon/n| < \varepsilon/n \), we see that the average of \( \theta \) is \( \varepsilon/2 \) up to an error of order \( \varepsilon^2 \).

The same reasoning applies if \( A_0 \) is changed to \( A_{n-1} \), hence for \( \phi^t_l \) intersecting the back end of \( A \) instead of the front end. \( \square \)

Note that if we compute the measure of an intersection \( A_0 \cap g^t A_{n-1} \) for \( t \) large, the terms which are not in full components of intersection contribute only a fraction which by mixing is asymptotically zero because \( m(A^t_l) \cong m(A) \), i.e., \( \exists K < \infty \ \forall \alpha > 0 \ \exists \varepsilon_0 > 0 \ \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \ \exists t_0 > 0 \ \forall t > t_0 : |\ln(m(A^t_l)/m(A))| < K\varepsilon + \alpha. \) This follows from

\[
m(\{v \in A : s(v) < S(t)\}) \to 0 \text{ as } t \to \infty
\]

and

\[
m(\{v \in A : \tau(v) \in [0, \varepsilon^2] \cup [\varepsilon - \varepsilon^2, \varepsilon]\}) = 2\varepsilon m(A).
\]

5.5. Counting intersections.

Theorem 5.20 (Few intersection components through the stable end).

The number \( \#\Phi^*_t \) of intersection components that touch the stable end \( \partial^s A \) is asymptotically a zero proportion of the number of all boundary components:

\[
\frac{\#\Phi^*_t}{N(A,t)} \cong 0.
\]

In other words, \( \exists K < \infty \ \forall \alpha > 0 \ \exists \varepsilon_0 > 0 \ \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \ \exists t_0 > 0 \ \forall t > t_0 : \frac{\#\Phi^*_t}{N(A,t)} < K\varepsilon + \alpha. \)
Proof. Let $F := g^{[0,\varepsilon]}\overline{B_0^u}(v_0)$. First note that
$$m(\phi^t_\varepsilon) \asymp \frac{\theta(\phi^t_\varepsilon)}{\varepsilon} m^u(F) \sigma(t)$$
for $\phi^t_\varepsilon \in \Phi_t \setminus \Phi^*_t$, i.e., $\exists K, \varepsilon > 0 \forall \varepsilon \in (0, \varepsilon_0)$, $A = A(\varepsilon) \exists t_0 \forall t > t_0$:
$$\left| \ln \frac{\varepsilon m(\phi^t_\varepsilon)}{\theta(\phi^t_\varepsilon)m^u(F) \sigma(t)} \right| < K\varepsilon.$$ 

This is so since by Lemma 5.9 the stable measure of the pieces of stable fibers in $\phi^t_\varepsilon$ is equal to $\sigma(t)$ up to an error term that converges to 0 as $\varepsilon \to 0$ and since by holonomy invariance (Theorem 4.7) and by Lemma 5.12 the $m^u$-measure of $u$-leaves of $\phi^t_\varepsilon$ is the same as that of $F$, except that the thickness of the intersection is not $\varepsilon$ but $\theta(\phi^t_\varepsilon)$.

For $s$-holonomic $p, q$, i.e. $p \in W^s(q)$, the bounded subsets $B^0_r(p)$ and $B^0_r(q)$ get arbitrarily close under the flow “pointwise except at the boundary” in the following sense: there exists $R_1 := d^s(p, q)$ with $d(g^t p, B^0_R(g^t q)) \to 0$ as $t \to \infty$. Moreover, if we write $H$ for the holonomy map from $B^0_r(p)$ to $B^0_r(q)$ along stable fibers, then for $R_2 > R_1$ the convergence $d(g^t p', B^0_R(g^t H(p'))) \to 0$ as $t \to \infty$ is uniform in $p'$ for all $p' \in B^0_{R_2}(p)$. See [Gun] for a proof of these claims. Thus there exists $D^s = D^s(t) := [0, \infty) \to (0, \infty)$ with $D^s(t) \to 0$ as $t \to \infty$ such that $\phi \in \Phi^*_t$ implies $\phi \subset B^s_{D^s(t)}(\partial^s A)$.

Existence of a decomposition of $m$ into conditionals (Proposition 4.4) and their holonomy invariance (Theorem 4.7) imply that $m(B^s_D(\partial^s A)) \to 0$ as $D \to 0$.

For $\phi^t_\varepsilon \in \Phi^*_t$ define
$$\hat{\phi}^t_\varepsilon := g^{[-\varepsilon, \varepsilon]}(I) \cap B^s_{D^s(t)}(A) \cap g^t A.$$ 

This differs from $\phi^t_\varepsilon$ by extending it in the stable direction beyond the stable boundary of $A$. We could also have written $\hat{\phi}^t_\varepsilon = g^{[-\varepsilon, \varepsilon]}(I) \cap B^s_{D^s(t)}(\partial^s A) \cap g^t A$. The set $\hat{\phi}^t_\varepsilon$ is the intersection of $g^t A$ not only with $A$ itself, but with a stable neighborhood of $A$; this allows us to treat $\hat{\phi}^t_\varepsilon \in \Phi^*_t$ like the elements $\phi^t_\varepsilon \in \Phi_t$. Namely, for such $\phi^t_\varepsilon \in \Phi^*_t$, the formula $m(\hat{\phi}^t_\varepsilon) \asymp \theta(\hat{\phi}^t_\varepsilon)m^u(F) \sigma(t)/\varepsilon$ still holds, by the same argument as in the case of $\phi^t_\varepsilon \in \Phi_t$. Since $\theta(\hat{\phi}^t_\varepsilon) \leq \theta(\phi^t_\varepsilon) + \varepsilon^2$ and $\theta(\phi^t_\varepsilon)/\varepsilon \leq 1$, this shows that $m(\hat{\phi}^t_\varepsilon) \leq \text{const} \cdot e^{-ht}$. Therefore
$$\# \Phi^*_t / \# \Phi_t \leq \text{const} \cdot m(B^s_{D^s(t)}(\partial^s A)) \to 0 \text{ as } t \to \infty,$$
proving the claim. \qed
Remark 5.21. The proof of Theorem 5.20 would be much shorter and very easy if distances in the stable direction would contract uniformly, as they do for uniformly hyperbolic systems. But in our case they do not. In fact, they do not necessarily even converge to zero.

Proposition 5.22. The number \( N(A, t) \) satisfies

\[
N(A, t) \approx 2e^{ht}m(A).
\]

In other words, \( \exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \ \exists t_0 > 0 \ \forall t > t_0 : |\ln(N(A, t)/2e^{ht}m(A))| < K \varepsilon + \alpha. \)

Proof. As in the proof of Theorem 5.20, writing \( F := g^{[0, \varepsilon]}D_\delta'(v_0) \) gives the estimate

\[
m(\phi'_i) \approx \frac{\theta(\phi'_i)}{\varepsilon}m^{0u}(F)\sigma(t).
\]

Since by Lemma 5.19 the average of the \( \theta(\phi'_i) \) is asymptotically \( \varepsilon/2 \), we get

\[
\frac{1}{N(A, t)} \sum_{\phi'_i \in \Phi_t} m(\phi'_i) \approx \frac{1}{2} \sigma(t)m^{0u}(F),
\]

i.e., \( \exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \ \exists t_0 \ \forall t > t_0 : \left| \ln\frac{2\sum_{\phi'_i \in \Phi_t} m(\phi'_i)}{N(A, t)\sigma(t)m^{0u}(F)} \right| < K \varepsilon + \alpha. \)

Since the measure of \( A \cap g^tA \) is asymptotically the sum of the measures of the full components of intersection (there are \( N(A, t) \) of those), we obtain

\[
m(A \cap g^tA) \approx \frac{1}{2}N(A, t)\sigma(t)m^{0u}(F),
\]

or, in more detail, \( \exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \ \exists t_0 \ \forall t > t_0 : \left| \ln(2m(A \cap g^tA)/N(A, t)\sigma(t)m^{0u}(F)) \right| < K \varepsilon + \alpha. \)

Moreover note that by the mixing property of \( g \),

\[
m(A \cap g^tA) \approx e^{ht}m(A)\sigma(t)m^{0u}(F),
\]

i.e., \( \exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \ \exists t_0 \ \forall t > t_0 : |\ln(m(A \cap g^tA)/e^{ht}\sigma(t)m(A)m^{0u}(F))| < K \varepsilon + \alpha. \)

The claim follows from combining these two estimates. \( \Box \)
5.6. Bounds on error terms in intersection counting.

**Definition 5.23.** Define $G(t, \varepsilon) := \{\text{all geometric orbit segments in } A \text{ of periodic orbits with period in } [t - \varepsilon, t + \varepsilon]\}$. Let $G(t, \varepsilon) := \# G(t, \varepsilon)$. For better comparison we write $N(t, \varepsilon) := N(A, t) = \# \Phi_t(\varepsilon)$ where $\Phi_t = \Phi_t(\varepsilon)$ is as before the set of all full intersection components for given $t, \varepsilon$.

**Proposition 5.24.** The number of orbit segments passing through $A$ that belong to periodic orbits with period in $[t - \varepsilon, t + \varepsilon]$ is $\cong N(A, t)$. I.e.,

$$N(t, \varepsilon) \approx G(t, \varepsilon),$$

or more formally: $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), A = A(\varepsilon) \exists t_0 \forall t > t_0:

$$\left| \ln \frac{G(t, \varepsilon)}{N(t, \varepsilon)} \right| < K\varepsilon + \alpha.$$

**Proof.** By Lemma 5.16 we have a map $G(t, \varepsilon - 2\varepsilon^2) \to \Phi_t(\varepsilon)$ and by Lemma 5.17 a map $\Phi_t(\varepsilon) \setminus \Phi^*_t(\varepsilon) \to G(t, \varepsilon)$. These maps are invertible between their domains and images; hence they are injective. Thus we have

$$G(t, \varepsilon - 2\varepsilon^2) \leq N(t, \varepsilon) \leq G(t, \varepsilon).$$

Since $N(t, \varepsilon - 2\varepsilon^2) \leq G(t, \varepsilon - 2\varepsilon^2)$, it suffices to show

$$N(t, \varepsilon) \cong N(t, \varepsilon + \varepsilon^2),$$

i.e., $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), A = A(\varepsilon) \exists t_0 \forall t > t_0:

$$\left| \ln \frac{N(t, \varepsilon)}{N(t, \varepsilon + \varepsilon^2)} \right| < K\varepsilon + \alpha.$$

Partition $A$ again into $n := \lfloor 1/\varepsilon \rfloor$ pieces

$$A_i := \{v \in A : \tau(v) \in [i\varepsilon/n, (i+1)\varepsilon/n]\}$$

of equal measure ($i = 0, \ldots, n - 1$). Mixing implies that

$$m(A_0 \cap g\varepsilon^i A_{n-1}) \cong \varepsilon^2 m(A)^2,$$

i.e., $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), A = A(\varepsilon) \exists t_0 \forall t > t_0:

$$|\ln(m(A_0 \cap g\varepsilon^i A_{n-1})/\varepsilon^2 m(A)^2)| < K\varepsilon + \alpha.$$

Observe that in analogy to equation (5.2) we have

$$\varepsilon^2 m(A)^2 \cong m(A_0 \cap g^{t+\varepsilon^2} A_{n-1}) \cong \frac{1}{2} \varepsilon^2 N(t, \varepsilon)m^{0u}(F)\sigma(t),$$

i.e., $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), A = A(\varepsilon) \exists t_0 \forall t > t_0:

$$|\ln(2m(A)^2/N(t, \varepsilon)m^{0u}(F)\sigma(t))| < K\varepsilon + \alpha.$$
The full components of $A_0 \cap g^{t+\varepsilon^2} A_{n-1}$ which are newly created at the back end of $A$ by increasing $t$ to $t + \varepsilon^2$ have average thickness $\varepsilon^2 / 2$ and hence average measure $\frac{1}{2} \varepsilon m_{\text{bu}}(F) \sigma(t)$. Hence this increase of $t$ to $t + \varepsilon^2$ can produce at most $\approx \varepsilon N(t, \varepsilon)$ such full components. Thus

$$N(t + \varepsilon^2, \varepsilon) \lesssim N(t, \varepsilon) + \varepsilon N(t, \varepsilon),$$

where the notation $f_1(t, \varepsilon) \lesssim f_2(t, \varepsilon)$ means $f_1(t, \varepsilon) \leq f_2(t, \varepsilon) \approx f_3(t, \varepsilon)$ for some $f_3$, i.e., $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \exists t_0 \forall t > t_0$:

$$\ln \frac{f_1(t, \varepsilon)}{f_2(t, \varepsilon)} < K \varepsilon + \alpha.$$  

In other words, we have shown $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \exists t_0 \forall t > t_0$:

$$\ln \frac{N(t + \varepsilon^2, \varepsilon)}{N(t, \varepsilon) + \varepsilon N(t, \varepsilon)} < K \varepsilon + \alpha.$$  

It follows that

$$N(t + \varepsilon^2, \varepsilon) \approx N(t, \varepsilon).$$

Since increasing $\varepsilon$ by $\varepsilon^2$ leads to a gain in the number of full components by making more of them enter the back end of the flow cube exactly like increasing $t$ by $\varepsilon^2$ does, plus a similar increase in number by making some of them delay their departure through the front end of the flow cube, we get

$$N(t, \varepsilon + \varepsilon^2) \lesssim N(t, \varepsilon) + 2\varepsilon N(t, \varepsilon),$$

i.e., we have shown $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \exists t_0 \forall t > t_0$:

$$\ln \frac{N(t, \varepsilon + \varepsilon^2)}{N(t, \varepsilon) + 2\varepsilon N(t, \varepsilon)} < K \varepsilon + \alpha.$$  

This shows that $N(t, \varepsilon + \varepsilon^2) \approx N(t, \varepsilon)$. Hence $N(t, \varepsilon) \approx G(t, \varepsilon)$ as claimed. \hfill \Box

5.7. A Bowen-type property of the measure of maximal entropy.

**Definition 5.25.** Let $P_t$ be the number of homotopy classes of closed geodesics of length at most $t$. Let $P_t(A)$ be the number of closed geodesics of length at most $t$ that intersect $A$. Let $P_t'$ be the number of regular closed geodesics of length at most $t$.
Remark 5.26 (Terminology). When we say “closed geodesic”, we mean “periodic orbit for the geodesic flow”, i.e. with parameterization (although always by arclength and modulo adding a constant to the parameter). Thus a locally shortest curve is counted as two geodesics (i.e. periodic orbits for the geodesic flow), namely one for each direction.

$P_t', P_t(A)$ are finite because there is only one regular geodesic in each homotopy class. Clearly

$$P_t(A) \leq P_t' \leq P_t$$

for any $t$. We will show that these are in fact asymptotically equal.

Lemma 5.27.

$$P_t' \sim P_t.$$  

Proof. Singular geodesics have a smaller exponential growth rate than regular ones because the entropy of the singular set is smaller than the topological entropy ([Kn2]) whereas the entropy of the regular set equals the topological entropy. \hfill \Box

In the case that $M$ is a surface, the growth rate of $\text{Sing}$ is in fact zero, since the existence of a parallel perpendicular Jacobi field implies that the largest Liapunov exponent is zero.

Definition 5.28. Let $\mu_t$ be the arclength measure on all regular periodic orbits of length at most $t$, normalized to 1:

$$P_t := \{\text{regular closed geodesics of length } \leq t\},$$

$$P_t(A) := \{\text{geodesics in } P_t \text{ which pass through } A\},$$

$$\mu_t := \frac{1}{\#P_t} \sum_{c \in P_t} \frac{1}{\text{len}(c)} \delta_c,$$

$$\mu_t^A := \frac{1}{\#P_t(A)} \sum_{c \in P_t(A)} \frac{1}{\text{len}(c)} \delta_c.$$  

Here $\delta_c$ is the length measure on $\dot{c}$.

Theorem 5.29. Any weak limit $\mu$ of $(\mu_t)_{t>0}$ is the measure of maximal entropy. Moreover, any weak limit $\mu^A$ of $(\mu_t^A)_{t>0}$ is the measure of maximal entropy.

In other words, for any $t_k \to \infty$ such that $(\mu_{t_k}^A)_{t_k \in \mathbb{R}}$ converges weakly and for any measurable $U$ the following holds:

$$\lim_{k \to \infty} \mu_{t_k}^A(U) = m(U).$$

Similarly with $\mu^A$ replaced by $\mu$. 
Proof. Knieper showed in [Kni1] that \( m \) can be obtained as a weak limit of the measures \( \mu_t \), which are Borel probability measures supported on \( P_{t_k} \); see also [Pol]. The singular closed geodesics can be neglected because the singular set has entropy smaller than \( h \). Hence any weak limit of \( \mu_t \) equals \( m \).

Since

\[
P_t(A) \geq C e^{ht/t}
\]

([Kni1, Remark after Theorem 5.8]), any weak limit of the measures \( \mu_{t_k}^A \) concentrated on \( P_{t_k}(A) \) has entropy \( h \). Since the measure of maximal entropy is unique, any such weak limit equals \( m \).

\[\square\]

Remark 5.30. This means that we can approximate the measure of maximal entropy \( m \) of a measurable set by its \( \mu_{t_k}^A \)-measure for \( k \) sufficiently large. Moreover, when counting orbits, an arbitrarily small regular local product cube \( A \) will suffice to count periodic orbits in such a way that the fraction of those not counted will converge to zero as the period of these orbits becomes large. We use this fact in the proof of Theorem 5.33.

Corollary 5.31. \( P_t(A) \sim P_t \).

Proof. By theorem 5.29 the measure on the geodesics in \( P_t \setminus P_t(A) \) (which assigns zero measure to \( A \)) would otherwise also converge weakly to the measure of maximal entropy. \( \square \)

Definition 5.32. Let \( P_{t,\varepsilon} \) be the number of regular geodesics with length in \((t - \varepsilon, t + \varepsilon]\).

Again, \( P_{t,\varepsilon} \) is finite because there is only one regular geodesic in each homotopy class.

Theorem 5.33. The number \( P_{t,\varepsilon} \) of regular closed geodesics with prescribed length is given by the asymptotic formula

\[
P_{t,\varepsilon} \simeq \frac{\varepsilon N(A,t)}{t \cdot m(A)}.
\]

I.e. \( \exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \ A = A(\varepsilon) \exists t_0 \forall t > t_0 : \)

\[
\left| \ln \frac{P_{t,\varepsilon} \cdot t \cdot m(A)}{\varepsilon N(A,t)} \right| < K \varepsilon + \alpha.
\]

Proof. By Theorem 5.29, for a typical closed geodesic \( c \) with sufficiently large length,

\[
\frac{1}{\text{len}(c)} \delta_c(A) = \frac{1}{\text{len}(c)} \int_{c \cap A} d\text{len} \cong m(A).
\]
Here “typical” means that the number of closed geodesics of length at most $t$ that have this property is asymptotically the same as the number of all closed geodesics of length at most $t$; in other words, the ratio tends to $1$. That means: $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0)$, $A = A(\varepsilon)$ $\exists t_0 \forall t > t_0$:

$$\left| \ln \left( \frac{1}{m(A) \cdot \# P_t} \sum_{c \in P_t} \frac{1}{\text{len}(c)} \delta_c \right) \right| < K\varepsilon + \alpha.$$ 

Hence such a geodesic of length $t$ (which consists of $t/\varepsilon$ segments of length $\varepsilon$) will have asymptotically $m(A)t/\varepsilon$ segments of length $\varepsilon$ intersecting $A$. Thus

$$P_{t,\varepsilon} \equiv \frac{\varepsilon G(t, \varepsilon)}{tm(A)},$$

i.e., $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0)$, $A = A(\varepsilon)$ $\exists t_0 \forall t > t_0$:

$$\left| \ln \frac{tm(A)P_{t,\varepsilon}}{\varepsilon G(t, \varepsilon)} \right| < K\varepsilon + \alpha,$$

where $G$ is as in Definition 5.23. The statement of Proposition 5.24 then shows the claim. \hfill \Box

Remark 5.34. It suffices to consider closed orbits which are not multiple iterates of some other closed orbit for the following reason: If $H(t, k)$ is the number of periodic orbits passing through $A$ of length at most $t$ which are $k$-fold iterates, then $H(t, k) = 0$ for $k > t/\text{inj}(M)$. Thus the number of segments of $A$ which are transversed by all multiple iterates is at most $\sum_{k=2}^{\lceil t/\text{inj}(M) \rceil} kH(t, k)$. By Knieper’s multiplicative estimate (see Section 2.3), this number is at most const $\cdot \sum_{k=2}^{\lceil t/\text{inj}(M) \rceil} ke^{ht/k}$, thus at most const $\cdot t^2 e^{ht/2}$. This contributes only a zero asymptotic fraction of the segments and can thus be ignored.

Proposition 5.22 and Theorem 5.33 combined yield:

Corollary 5.35.

$$P_{t,\varepsilon}(A) \approx \frac{2\varepsilon e^{ht}}{t}.$$ 

I.e., $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0)$, $A = A(\varepsilon)$ $\exists t_0 \forall t > t_0$:

$$\left| \ln (tP_{t,\varepsilon}(A)/2\varepsilon e^{ht}) \right| < K\varepsilon + \alpha.$$ 

\hfill \Box

5.8. **Proof of the main result.** The desired asymptotic formula is now derived:

Theorem 5.36 (Precise asymptotics for periodic orbits). Let $M$ be a compact Riemannian manifold of nonpositive curvature whose rank is one. Then the number $P_t$ of homotopy classes of periodic orbits of
length at most $t$ for the geodesic flow is asymptotically given by the formula

$$P_t \sim \frac{e^{ht}}{ht}$$

where $\sim$ means that the quotient converges to 1 as $t \to \infty$.

Proof. We use the standard limiting process

$$\int_a^b f(x)dx \bowtie \sum_{i=\lfloor a/2\varepsilon \rfloor}^{\lfloor b/2\varepsilon \rfloor} 2\varepsilon f((2i+1)\varepsilon)$$

for suitable functions $f$ (in particular, if $f$ is continuous and piecewise monotone, as is the case for $f(x) = e^{hx}/x$); since this is elementary, “$\bowtie$” requires no explanation in long form here. Choose some fixed sufficiently large number $t_0 > 0$. Since we can ignore all closed geodesics of length at most $t_0$ for the asymptotics, we see that for $t > t_0$ by Corollary 5.35 we get

$$P'_t \approx P_t(A) \approx \sum_{i=\lfloor t_0/2\varepsilon \rfloor}^{\lfloor t/2\varepsilon \rfloor} P_{2i+1}\varepsilon, x \approx \sum_{i=\lfloor t_0/2\varepsilon \rfloor}^{\lfloor t/2\varepsilon \rfloor} 2\varepsilon e^{h(2i+1)\varepsilon}$$

$$\approx \int_{t_0}^{t} e^{hx}/x \, dx = \left. \frac{e^{hx}}{hx} \right|_{t_0}^{t} + \int_{t_0}^{t} \frac{e^{hx}}{hx^2} \, dx \approx \frac{e^{ht}}{ht} - \frac{e^{ht_0}}{ht_0}$$

In more detail, $\exists K < \infty \forall \alpha > 0 \exists \varepsilon_0 > 0 \forall \varepsilon \in (0, \varepsilon_0), \exists t_0 \forall t > t_0$:

$$\left| \ln \frac{htP_t'}{e^{ht}} \right| < K\varepsilon + \alpha.$$  

Note that there is no longer any dependence on $\varepsilon$ and that $P'_t \sim P_t$. Hence

$$P_t \sim \frac{e^{ht}}{ht}.$$ 

This concludes the proof. \qed

References


COUNTING CLOSED GEODESICS


Universität Hamburg, Department Mathematik, SP DGL. & Dynamische Systeme, Bundesstr. 55, 20146 Hamburg, Germany

E-mail address: Roland.Gunesch@math.uni-hamburg.de