J. DIFFERENTIAL GEOMETRY
25 (1987) 249–273

ON SURFACES WITH NO CONJUGATE POINTS

W. BALLMANN, M. BRIN & K. BURNS

A complete Riemannian manifold M has no conjugate points if any two points in its universal cover are joined by a unique geodesic. If the sectional curvature of M is nonpositive, then M has no conjugate points; the converse is not true even for compact surfaces. A natural question is to what extent properties of manifolds of nonpositive sectional curvature are valid for manifolds with no conjugate points. For example, by the Gauss-Bonnet theorem, any metric of nonpositive curvature on the torus T^2 is flat. In 1943, E. Hopf [4] proved

Theorem. *Any metric on T^2 with no conjugate points is flat.*

The best way to explain the purpose of our paper and to introduce the necessary notations is to give an outline of Hopf's argument. He considers the Riccati equation

$$u' + u^2 + K(\gamma_v(t)) = 0, \tag{0.1}$$

where γ_v is the geodesic with initial velocity v and K is the Gaussian curvature. Denote by $u_R^+(v,\cdot)$ (resp. $u_R^-(v,\cdot)$) the solution of (0.1) which satisfies $u_R^+(v,-R) = +\infty$ (resp. $u_R^-(v,R) = -\infty$). Note that $u_R^+(v,t)$ is the geodesic curvature at $\gamma_v(t)$ of the geodesic circle of radius $t+R$ centered at $\gamma_v(-R)$. As we explain in the next section (see Proposition 1.3), if a surface S has no conjugate points, then $u_R^+(v,t)$ (resp. $u_R^-(v,t)$) is defined for $t > -R$ (resp. $t < +R$) and there are well-defined limit solutions

$$u^+(v,\cdot) = \lim_{R\to\infty} u_R^+(v,\cdot), \qquad u^-(v,\cdot) = \lim_{R\to\infty} u_R^-(v,\cdot).$$

Recall that the geodesic flow g^t acts on the unit tangent bundle T_1S by $g^tv = \dot\gamma_v(t)$ and preserves the Liouville measure $d\mu = dA \times d\lambda$, where A is area in S and λ is Lebesgue measure on the unit circle. The solutions u^+ and u^- are invariant under g^t in the following sense:

$$u^\pm(g^tv,s) = u^\pm(v,s+t). \tag{0.2}$$

Received April 21, 1986. The first author was partially supported by National Science Foundation Grant #DMS-85-03742; the second and third authors were partially supported by National Science Foundation Grant #DMS-82-04024.

Since they can be constructed by a limit procedure, u^+ and u^- are measurable functions of (v, t). We integrate (0.1) to obtain

$$\int_{T_1S}\int_0^1 \frac{d}{dt}u^+(v,t)\,dt\,d\mu + \int_{T_1S}\int_0^1 (u^+(v,t))^2\,dt\,d\mu + \int_{T_1S}\int_0^1 K(\gamma_v(t))\,dt\,d\mu = 0. \tag{0.3}$$

Since μ is g^t-invariant, (0.2) implies that

$$\int_{T_1S} u^+(v,1)\,d\mu = \int_{T_1S} u^+(v,0)\,d\mu,$$

and hence the first integral in (0.3) vanishes. Also for every t

$$\int_{T_1S} K(\gamma_v(t))\,d\mu = 2\pi\int_S K(p)\,dA = 4\pi^2\chi(S),$$

where $\chi(S)$ is the Euler characteristic of S. Therefore, if S is T^2, then the second integral in (0.3) is equal to 0 and hence u^+ vanishes almost everywhere. We see from (0.1) that K vanishes, which finishes the proof.

In [4], besides proving the measurability of u^+ and u^-, Hopf remarks that they are continuous in v. A counterexample is constructed in the present paper. This contrasts with the situation in nonpositive curvature where u^+ and u^- are easily shown to be continuous.

In the case of a surface S with nonpositive curvature, the limit solutions u^+ and u^- are the geodesic curvatures of the limit circles (horocycles), whose construction we now describe. Denote by $\tilde{S}$ the universal cover of S. For $v \in T_1\tilde{S}$ let $C^+(v, R)$ be the geodesic circle with center $\gamma_v(-R)$ and radius R. As $R \to \infty$, these curves converge uniformly in v in the C^2-topology to a curve $C^+(v)$, called the unstable horocycle of v [2]. The stable horocycle is constructed in a similar way. The limit solutions u^+ and u^- and the horocycles have been extensively used in the study of dynamical properties of the geodesic flow, in particular in proving that the geodesic flow is ergodic. In 1939, Hopf [3] showed that the geodesic flow of a compact surface with (variable) negative curvature is ergodic. For a compact surface of nonpositive curvature, Pesin [6] in 1977 proved that either the geodesic flow is ergodic or the surface is flat.

In the case of a surface with no conjugate points, the circles $C^+(v, R)$ converge uniformly in v in the C^1-topology to the horocycles $C^+(v)$. We do not know whether $C^+(v, R)$ always converges to $C^+(v)$ in the C^2-topology for each v. But in the counterexample constructed in this paper, u^+ and u^- are not continuous in v, and hence $C^+(v, R)$ cannot converge uniformly in the C^2-topology. Our interest in the continuity properties of u^+ and u^- stems from a recent theorem of Knieper [5]. Using a result of Pesin [6], he proved the ergodicity of the geodesic flow for a compact nonflat surface with no conjugate points under the assumption that u^+ and u^- are continuous in v.

In this paper we construct a compact surface with no conjugate points (see Theorem 4.2) for which the limit solutions u^+ and u^- are not continuous in v (see Theorem 4.6). Nevertheless, the geodesic flow is ergodic (see Remark 4.7). We believe that the geodesic flow is ergodic for any compact nonflat surface with no conjugate points.

We now give a brief description of our example. Consider the hyperbolic plane. Remove two geodesic sectors and identify their boundary rays as indicated in Figure 1 to obtain a noncompact hyperbolic surface with two conic singularities. Replace circular neighborhoods of the singularities by almost spherical caps to get a smooth noncompact surface. We compactify a large geodesic polygon containing the caps in such a way that the resulting surface $S_{d,R}$ has curvature -1 except in the caps. Moreover, we arrange that the geodesic γ_0 passing through the centers of the caps is positively and negatively asymptotic to closed geodesics σ_+ and σ_- which do not meet the caps (see Figure 2). The caps and the distance between them are chosen so that the limit solutions u^+ and u^- coincide along γ_0. Since σ_+ and σ_- lie in the region of curvature -1, $u^+ \equiv 1$ and $u^- \equiv -1$ along these geodesics. This leads to the discontinuity of u^+ and u^-.

The difficult part of the construction is to ensure that the compactified surface has no conjugate points. This is done by controlling solutions of the Riccati equation (0.1) along every geodesic in the surface.

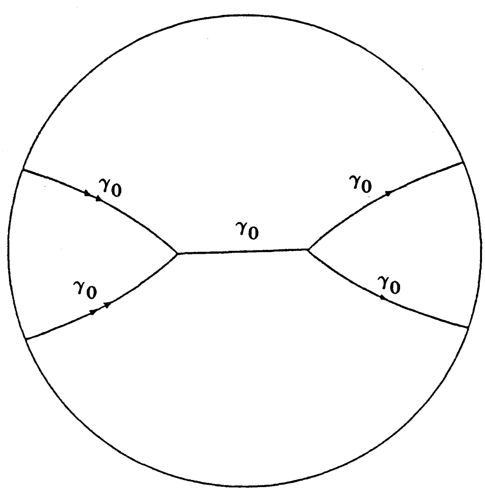

FIGURE 1

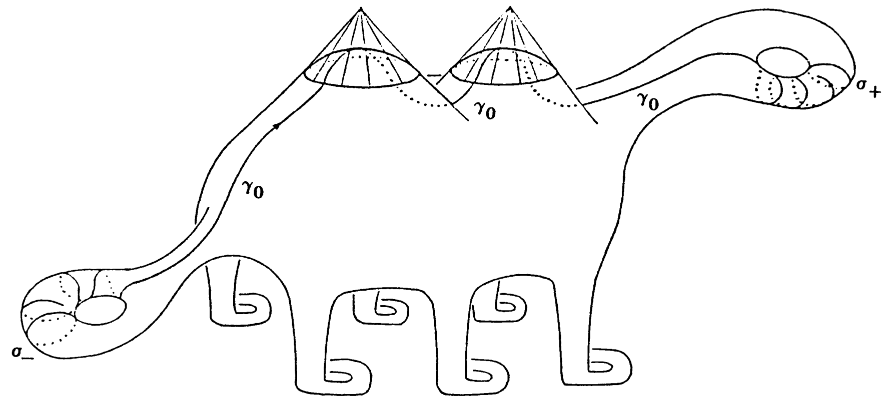

FIGURE 2

It is remarkable that even though $u^+ \equiv u^-$ along γ_0, the average curvature

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} K(\gamma_0(t))\,dt$$

is -1. Using a single almost spherical cap one can easily construct a compact surface with no conjugate points containing a closed geodesic along which $u^+ \equiv u^-$ and the average curvature is negative.

In §1 we discuss the Riccati equation and its relation with conjugate points. The construction of our example and the verification of its properties is carried out in §§2, 3, and 4. In §5 we obtain an estimate on the differential of the geodesic flow and discuss the geometry of the unit tangent bundle of a Riemannian manifold. In particular, we derive the geodesic equation and calculate certain distances in the unit tangent bundle.

1. Conjugate points and the Riccati equation

Let v be a unit vector tangent to a complete smooth surface S. We consider the Jacobi equation along the geodesic γ_v with initial velocity v:

$$y''(t) + K(\gamma_v(t))y(t) = 0. \tag{1.1}$$

If $y(t)$ is a solution of (1.1) and $N(t)$ is a continuous field of unit vectors normal to γ_v, then $y(t)N(t)$ is a Jacobi field along γ_v, i.e., a vector field obtained from a variation of γ_v through geodesics.

1.1. Definition. Two points $\gamma(t_1)$ and $\gamma(t_2)$ are conjugate along γ if there is a solution y of (1.1) with $y(t_1) = 0 = y(t_2)$. We say that S has no conjugate points if no two points are conjugate along any geodesic.

This is equivalent to the characterization of surfaces with no conjugate points given at the beginning of the introduction. In particular, any geodesic in a simply connected surface with no conjugate points is minimizing.

The Riccati equation (0.1) is obtained from the Jacobi equation (1.1) by the change of variable $u = y'/y$. The times $t_1 < t_2$ are adjacent zeros of a solution y of (1.1) if and only if the corresponding solution u of (0.1) is defined on (t_1, t_2) and $u(t) \to +\infty$ as $t \downarrow t_1$ and $u(t) \to -\infty$ as $t \uparrow t_2$. This gives us the following test for the absence of conjugate points, which we apply to the surface $S_{d,R}$ in §4.

1.2. Proposition. *If there is a solution of the Riccati equation* (0.1) *defined for all* $t \geqslant 0$, *then there are no conjugate points along the geodesic ray* $\gamma_v|[0, \infty)$. *If for every* $v \in T_1S$ *the Riccati equation* (0.1) *has this property, then* S *has no conjugate points.*

1.3. Proposition. *Let* S *be a surface with no conjugate points. Then for any* $v \in T_1S$

(1) *the solutions* $u_R^+(v, t)$ *and* $u_R^-(v, t)$ *of* (0.1) *with* $u_R^+(v, -R) = +\infty$ *and* $u_R^-(v, R) = -\infty$ *are defined for all* $t > -R$ *and all* $t < R$ *respectively*;

(2) *the limit solutions* $u^{\pm}(v, t) = \lim_{R \to \infty} u_R^{\pm}(v, t)$ *are defined for all* t;

(3) $u^+(v, t)$ (*resp.* $u^-(v, t)$) *is upper* (*resp. lower*) *semicontinuous in* v;

(4) $u^{\pm}(g^t v, s) = u^{\pm}(v, s + t)$ *for any* s *and* t (g^t *is the geodesic flow*).

Proof. (1) The solution $u_R^+(v, t)$ can fail to be defined for all $t > -R$ only if it becomes $-\infty$ after a finite time. As we saw above, this would imply the existence of conjugate points.

(2) Fix t and consider $R > |t|$. Then $u_R^+(v, t)$ and $u_R^-(v, t)$ are both defined and are decreasing and increasing functions of R respectively. By (1) we have $u_R^+(v, t) > u_R^-(v, t)$ for all R.

(3) Suppose $v_n \to v$. For each fixed t we have

$$\limsup_{n \to \infty} u^+(v_n, t) \leqslant \lim_{n \to \infty} u_R^+(v_n, t) = u_R^+(v, t).$$

But $u_R^+(v, t) \to u^+(v, t)$ as $R \to \infty$. A similar argument works for u^-.

(4) Note that $u_R^{\pm}(v, t + s) = u_{R \pm t}^{\pm}(g^t v, s)$. q.e.d.

We will often need the solutions of the Riccati equation (0.1) in the cases of constant curvature -1 and κ^2.

1.4. Propositon. (1) *If* $K(\gamma_v(t)) \equiv -1$, *then the general solution of* (0.1) *is*

$$u(t) = \frac{u(0) \cosh t + \sinh t}{u(0) \sinh t + \cosh t}.$$

If $|u(0)| < 1$, *then* u *is increasing and* $u(\operatorname{arctanh}(-u(0))) = 0$.

(2) *If* $K(v(t)) \equiv \kappa^2$, *then the general solution of* (0.1) *is*

$$u(t) = \kappa \frac{u(0) \cos(\kappa t) - \kappa \cdot \sin(\kappa t)}{u(0) \sin(\kappa t) + \kappa \cdot \cos(\kappa t)}.$$

Our estimates on the solutions of the Riccati equation will use the following lemma.

1.5. Comparison Lemma. *Let $u_i(t)$, $i = 0, 1$, be the solutions of the initial value problems*

$$u_i' + u_i^2 + K_i(t) = 0, \quad u_i(0) = w_i, \qquad i = 0, 1.$$

Suppose $w_1 \geqslant w_0$, $K_1(t) \leqslant K_0(t)$ for $t \in [0, t_0]$, and $u_0(t_0)$ is defined. Then

$$u_1(t) \geqslant u_0(t) \quad \text{for } t \in [0, t_0].$$

Proof. The difference $\Delta u(t) = u_1(t) - u_0(t)$ satisfies the linear equation

$$\Delta u' = -(u_0 + u_1)\Delta u + K_0(t) - K_1(t).$$

2. Construction of the noncompact surface

In this section we construct the smooth noncompact surface with two almost spherical caps described in the introduction. To do this we first construct a radially symmetric surface S which contains one such cap and has curvature -1 elsewhere. Then we cut S along a geodesic which does not meet the cap. Gluing together two copies of the part containing the cap gives the desired surface. To construct S, let

$$\kappa = 1000, \quad \beta = .9, \quad r = \frac{1}{\kappa} \arctan \frac{\beta}{\kappa}, \quad \delta = 10^{-20}. \tag{2.1}$$

The values of the constants are not important as long as κ is big enough and δ is small enough. Choose a monotone C^∞-function $g(s)$ such that

$$g(s) = \begin{cases} \kappa^2 & \text{for } 0 \leqslant s \leqslant r - \delta, \\ -1 & \text{for } r \leqslant s. \end{cases} \tag{2.2}$$

Let S be the radially symmetric surface whose Gaussian curvature at distance s from the center P is $g(s)$. The surface S is well defined because the solution $y(s)$ of the Jacobi equation

$$y''(s) + g(s)y(s) = 0, \quad y(0) = 0, \quad y'(0) = 1,$$

is positive for all $s > 0$. In fact,

$$y'(s) > 0 \quad \text{for } s \geqslant 0. \tag{2.3}$$

The latter follows from the fact that $r < \pi/2\kappa$. In what follows we will refer to the r-ball centered at P as the cap $\mathscr{C}$.

2.1. Proposition. *The surface S has the following properties*:

(1) *For any geodesic σ, the distance $d(\sigma(\cdot), P)$ is a convex function and is strictly convex unless σ passes through P.*

(2) *Any geodesic σ visits the cap $\mathscr{C}$ at most once and then goes away to infinity in both directions, i.e., $d(\sigma(t), P) \to \infty$ as $t \to \pm\infty$. The time that σ spends in $\mathscr{C}$ is at most $2r$, with equality if and only if σ passes through P.*

(3) *S has no conjugate points. In particular, any two points in S are connected by a unique geodesic.*

Proof. To prove (1) denote by V the unit vector field that points radially outwards from P. Then

$$d''(\sigma(t), P) = \frac{d}{dt}\langle \dot{\sigma}(t), V(\sigma(t))\rangle = \langle \dot{\sigma}(t), \nabla_{\dot{\sigma}(t)} V\rangle.$$

Let $\dot{\sigma}(t) = v + w$, where v is parallel to V and w is tangent to the circle c centered at P and passing through $\sigma(t)$. Note that $\nabla_v V = 0$ and $\langle v, \nabla_w V\rangle = 0$. Hence

$$d''(\sigma(t), P) = \langle w, \nabla_w V\rangle,$$

which is the second fundamental form of c evaluated at w. By (2.3) the second fundamental form of any circle centered at P is positive definite. Hence $d''(\sigma(t), P) \geqslant 0$ with equality if and only if $w = 0$, in which case σ passes through P.

The first assertion of (2) follows from (1). To prove that σ spends time at most $2r$ in $\mathscr{C}$, we first note that by (1) any geodesic segment joining two points of $\mathscr{C}$ lies in $\mathscr{C}$. Also if $q_1, q_2 \in \mathscr{C}$, then $\operatorname{dist}(q_1, q_2) \leqslant 2r$ with equality if and only if q_1 and q_2 are diametrically opposite. Thus (2) will follow if we show that any two points of $\mathscr{C}$ are joined by a unique geodesic in S. Suppose the contrary. Then there is a pair of points in $\mathscr{C}$ which are cut points of one another. Since the set of cut points is closed, we can find such a pair $q_1, q_2 \in \mathscr{C}$ with minimal distance. As in the proofs of Lemma 5.6 and Corollary 5.7 in [1], we see that either q_1 and q_2 lie on a closed geodesic or they are conjugate along a minimal geodesic joining them. The former contradicts (1). The latter is impossible since the curvature in S is bounded above by κ^2 and $\operatorname{dist}(q_1, q_2) \leqslant 2r < \pi/\kappa$.

To prove (3), consider a geodesic σ in S. If σ does not meet $\mathscr{C}$, the curvature along σ is -1, and hence there are no conjugate points on σ. Suppose now that σ enters $\mathscr{C}$ at $t = 0$ and exists at $t = t_e$. We will show that the solution $u(t)$ of

$$u' + u^2 + K(\sigma(t)) = 0, \qquad u(0) = 1,$$

is defined for all real t. Note that $K(\sigma(t)) \leqslant \kappa^2$ for all t. Hence by Proposition 1.4(2), Comparison Lemma 1.5, and (2.1),

$$\begin{aligned} u(2r) &\geqslant \kappa\frac{\cos(2r\kappa) - \kappa\sin(2r\kappa)}{\sin(2r\kappa) + \kappa\cos(2r\kappa)} \\ &= \frac{\kappa^2 - 2\beta\kappa^2 - \beta^2}{\kappa^2 + 2\beta - \beta^2} \geqslant -.8 - 10^{-6} > -1. \end{aligned} \tag{2.4}$$

By (2), $t_e \leqslant 2r$, and so $K(\sigma(t)) \equiv -1$ for $t \geqslant 2r$. Since $u(2r) > -1$, we see that $u(t)$ is defined for all t and in fact tends to 1 as $t \to \infty$. It follows from Proposition 1.2 that σ does not have conjugate points. q.e.d.

We now compare the effect of the cap on solutions of the Riccati equation along a diameter γ of $\mathscr{C}$ and along an arbitrary geodesic σ through $\mathscr{C}$. Parameterize γ and σ so that $\gamma(0)$ and $\sigma(0)$ are the points where the geodesics enter the cap $\mathscr{C}$. The geodesic γ leaves the cap at $t = 2r$; let t_e be the time when σ leaves $\mathscr{C}$. Denote by $u_\gamma(\lambda, t)$ and $u_\sigma(\lambda, t)$ the solutions of the Riccati equation (0.1) with initial value λ at $t = 0$ along γ and σ respectively.

2.2. Lemma. *Suppose $u_\gamma(\lambda, 2r)$ is defined. Then $u_\sigma(\lambda, t_e) \geqslant u_\gamma(\lambda, 2r)$, with equality if and only if σ is a diameter.*

Proof. Recall that the cap $\mathscr{C}$ is radially symmetric. Since the cap is convex, cf. Proposition 2.1, $t_e \leqslant 2r$. If $t_e < 2\delta$, then σ spends time less than 2δ in $\mathscr{C}$, and the lemma follows easily. Therefore we may assume that $t_e \geqslant 2\delta$. Fix any t between 0 and $t_e/2$. By the triangle inequality

$$\operatorname{dist}(\sigma(t), P) \geqslant r - t = \operatorname{dist}(\gamma(t), P),$$

where P is the center of the cap. The curvature K being a decreasing function of the distance to P, we obtain

$$K(\sigma(t)) \leqslant K(\gamma(t)), \qquad 0 \leqslant t \leqslant t_e/2. \tag{2.5}$$

A similar argument yields

$$K(\sigma(t_e - t)) \leqslant K(\gamma(2r - t)), \qquad 0 \leqslant t \leqslant t_e/2. \tag{2.6}$$

By the Comparison Lemma 1.5 and (2.5)

$$u_\sigma(\lambda, t_e/2) \geqslant u_\gamma(\lambda, t_e/2).$$

Since the curvature is positive inside the δ-interior of $\mathscr{C}$, any solution of the Riccati equation decreases there, and hence

$$u_\gamma(\lambda, t_e/2) \geqslant u_\gamma(\lambda, 2r - t_e/2), \tag{2.7}$$

with equality only if $t_e = 2r$. Finally, by the Comparison Lemma 2.5, (2.5), (2.6), and (2.7),

$$u_\sigma(\lambda, t_e) \geqslant u_\gamma(\lambda, 2r). \qquad \text{q.e.d.}$$

Later we will need the following two estimates.

2.3. Lemma. *For any diameter γ as above*

$$-.8 - 10^{-6} \leqslant u_\gamma(1, 2r) \leqslant -.8.$$

Proof. The lower estimate was already obtained in (2.4). To get the upper estimate, we set $u(\cdot) = u_\gamma(1, \cdot)$ and denote by v the solution of

$$v' + v^2 + \kappa^2 = 0, \qquad v(0) = 1.$$

Note that $K(\gamma(t)) \equiv \kappa^2$ for $\delta \leqslant t \leqslant 2r - \delta$ and $u(\delta) < 1 = v(0)$. Hence $u(2r - \delta) \leqslant v(2r - 2\delta)$. Since v is monotonically decreasing and $v(2r) > -1$,

$$|v'(t)| \leqslant |v^2(t)| + \kappa^2 \leqslant 1 + \kappa^2.$$

Therefore

$$v(2r - 2\delta) \leqslant v(2r) + (1 + \kappa^2)2\delta.$$

For $2r - \delta \leqslant t \leqslant 2r$ we have $K(\gamma(t)) \geqslant -1$. Thus

$$u' \leqslant 1 - u^2 \leqslant 1$$

and hence

$$u(2r) \leqslant v(2r) + (1 + \kappa^2)2\delta + \delta \leqslant v(2r) + 10^{-10}.$$

As in (2.4) we obtain

$$v(2r) = \frac{\kappa^2 - 2\beta\kappa^2 - \beta^2}{\kappa^2 + 2\beta - \beta^2} = -\frac{.8 \cdot 10^6 + .81}{10^6 + .99},$$

and the lemma follows from the previous inequality.

2.4. Lemma. *The geodesic curvature of the boundary circle $\partial\mathscr{C}$ is at least $\kappa^2 = 10^6$.*

Proof. Let u be the solution of the Riccati equation (0.1) along a radius of $\mathscr{C}$ with initial value $+\infty$ at P. The geodesic curvature of $\partial\mathscr{C}$ is $u(r)$. By the Comparison Lemma 1.5 and Proposition 1.4(2)

$$u(r) \geqslant \frac{\kappa \cos \kappa r}{\sin \kappa r} = \frac{\kappa^2}{\beta} = \frac{10^6}{.9}. \qquad \text{q.e.d.}$$

We now construct the simply connected surface with two caps. Let O be a point which lies at distance $r + d$ from P, where $d > 0$. Denote by γ_0 the geodesic passing through O and P and by ν the geodesic through O perpendicular to γ_0. Let S' be the part of S which is bounded by ν and contains $\mathscr{C}$. Take two copies S_1 and S_2 of S' and glue them together along their boundaries so that the points corresponding to O are identified (see Figure 3). Denote the resulting surface by S_d.

2.5. Proposition. *Any geodesic σ of S_d enters each of the caps $\mathscr{C}_1$ and $\mathscr{C}_2$ at most once, spends time at most $2r$ in each cap, and goes away to infinity in both directions, i.e. $d(\sigma(t), 0) \to \infty$ as $t \to \pm\infty$.*

Proof. By Proposition 2.1(3), σ can cross ν at most once. The proposition now follows easily from (1) and (2) of Proposition 2.1.

2.6. Remark. Since $K \geqslant -1$ in $\mathscr{C}$, the boundary circle of $\mathscr{C}$ is shorter than a geodesic circle in the hyperbolic plane of the same geodesic curvature. Hence $S \setminus \mathscr{C}$ can be obtained by identifying the boundary rays of a geodesic sector in the hyperbolic plane and replacing a circular neighborhood of the vertex by an almost spherical cap. Therefore S_d can be constructed as indicated in the introduction (see Figure 1).

We now choose d so that the limit solutions u^+ and u^- of the Riccati equation (0.1) coincide along the central geodesic γ_0. Parameterize γ_0 by arclength so that $\gamma_0(0) = 0$, the point halfway between the caps. Let u be the solution of the Riccati equation

$$u' + u^2 + K(\gamma_0(t)) = 0$$

with initial value $u(-2r - d) = 1$. We need to choose d so that

$$u(2r + d) = -1. \tag{2.8}$$

By symmetry, this is equivalent to $u(0) = 0$. It follows from Proposition 1.4 that

$$d = \operatorname{arctanh}(-u(-d)).$$

Note that for any smaller value of d there would be conjugate points along γ_0. By Lemma 2.3,

$$-.8 - 10^{-6} \leqslant u(-d) \leqslant -.8.$$

Hence

$$1 \leqslant d \leqslant 1.1. \tag{2.9}$$

3. Effect of the caps on the Riccati equation

The purpose of this section is to prove the Main Lemma 3.5 which gives lower estimates on the solutions of the Riccati equation (0.1) along geodesics as they cross the caps. As explained at the end of §1, this will be used to show that the compactification $S_{d,R}$ of S_d constructed in the next section has no conjugate points.

As before denote by γ_0 the central geodesic. Set $v_0 = \dot\gamma_0(0) \in T_0S_d$. For any v let $u_v(\lambda, t_0, \cdot)$ be the solution of (0.1) with value λ at $t = t_0$. If γ_v meets at least one cap, let $\alpha = \alpha(v)$ be the time when γ_v first enters a cap and $\omega = \omega(v)$ the time when γ_v last exits a cap. If γ_v crosses both caps, let $t_1 = t_1(v)$ be the time when γ_v exits the first cap it enters and $t_2 = t_2(v)$ the time γ_v enters the other cap. These notations are illustrated in Figure 3.

Observe that

$$\alpha(v_0) = -2r - d, \quad t_1(v_0) = -d, \quad t_2(v_0) = d, \quad \omega(v_0) = 2r + d,$$

and, by our choice of d, see (2.8),

$$u_{v_0}(1, -2r - d, 2r + d) = -1.$$

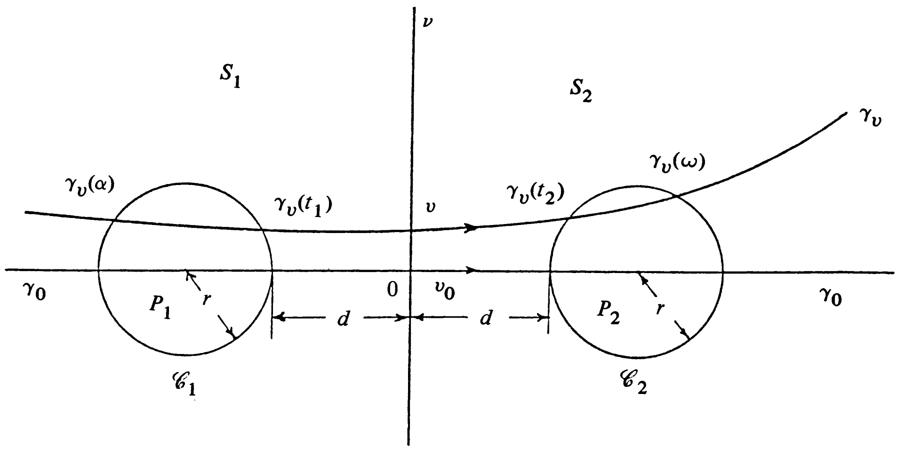

FIGURE 3

The following propositions show that $\gamma_0 = \gamma_{v_0}$ and its reparameterizations are affected by the caps more than any other geodesic.

3.1. Proposition. *Suppose* $\lambda \leqslant 1$, $u_{v_0}(\lambda, -2r-d, -d) > -1$, *and* $u_{v_0}(\lambda, -2r-d, 2r+d)$ *is defined. Then*

$$u_v(\lambda, \alpha(v), \omega(v)) \geqslant u_{v_0}(\lambda, -2r-d, 2r+d),$$

with equality if and only if γ_v *and* γ_0 *are geometrically the same geodesic.*

Proof. If γ_v crosses only one cap, Lemma 2.2 shows that

$$u_v(\lambda, \alpha(v), \omega(v)) \geqslant u_{v_0}(\lambda, -2r-d, -d) > -1.$$

On the other hand

$$u_{v_0}(\lambda, -2r-d, 2r+d) \leqslant u_{v_0}(1, -2r-d, 2r+d) = -1,$$

by the choice of d. Therefore

$$u_v(\lambda, \alpha(v), \omega(v)) > u_{v_0}(\lambda, -2r-d, 2r+d).$$

Suppose now that γ_v crosses both caps. By Lemma 2.2,

$$u_v(\lambda, \alpha(v), t_1(v)) \geqslant u_{v_0}(\lambda, -2r-d, -d) > -1.$$

Since the caps are convex, $t_2(v) - t_1(v) \geqslant 2d$ with equality if and only if γ_v and γ_0 are geometrically the same geodesic. Therefore $\gamma_v(t)$ is in the area of constant curvature -1 when $t_1(v) \leqslant t \leqslant t_1(v) + 2d$. Hence

$$u_v(\lambda, \alpha(v), t_1(v) + 2d) \geqslant u_{v_0}(\lambda, -2r-d, d) > -1.$$

Since $\lambda \leqslant 1$ and $K \geqslant -1$, $u_v(\lambda, \alpha(v), t) \leqslant 1$ for all $t \geqslant \alpha(v)$. Hence $-1 < u_v(\lambda, \alpha(v), t_1(v) + 2d) \leqslant 1$ and thus $u_v(\lambda, \alpha(v), \cdot)$ is nondecreasing on the interval $[t_1(v) + 2d, t_2(v)]$. Therefore

$$u_v(\lambda, \alpha(v), t_2(v)) \geqslant u_{v_0}(\lambda, -2r-d, d),$$

with equality if and only if γ_v and γ_0 are geometrically the same geodesic. The statement of the proposition now follows from Lemma 2.2 applied to the second cap. q.e.d.

We will now give a quantitative comparison of the solutions of (0.1) along γ_0 and along geodesics close to γ_0. Any geodesic sufficiently close to γ_0 must cross both caps and the geodesic ν perpendicular to γ_0 at $0 = \gamma_0(0)$. By symmetry, it suffices to consider only geodesics which cross ν in the same direction as γ_0.

3.2. Proposition. *There are constants* $\varepsilon_0 > 0$ *and* λ_0, $0 < \lambda_0 < 1$, *such that, for any unit tangent vector* v *with footpoint on* ν *and* $\operatorname{dist}(v, v_0) \leqslant \varepsilon_0$,

$$u_v(\lambda, \alpha(v), \omega(v)) \geqslant -1 + 99(\lambda - 1) + 5 \cdot 10^5 \operatorname{dist}^2(v, v_0)$$

if $\lambda_0 \leqslant \lambda \leqslant 1$.

Proof. We will show that

$$u_v(1, \alpha(v), \omega(v)) \geqslant -1 + 5 \cdot 10^5 \cdot \operatorname{dist}^2(v, v_0) \tag{3.1}$$

and

$$u_v(\lambda, \alpha(v), \omega(v)) \geqslant u_v(1, \alpha(v), \omega(v)) + 99(\lambda - 1) \tag{3.2}$$

for any v close enough to v_0 and λ close enough to 1. The proposition follows immediately from these claims.

To prove (3.1) we keep track of the solution $u_v(1, \alpha(v), \cdot)$ of the Riccati equation along γ_v as γ_v crosses the first cap, then the region between the caps and finally the second cap. By Lemma 2.2,

$$u_v(1, \alpha(v), t_1(v)) \geqslant u_{v_0}(1, -2r - d, -d). \tag{3.3}$$

Lemma 3.3 below implies that

$$t_2(v) - t_1(v) \geqslant 2d + 1.8 \cdot 10^6 \operatorname{dist}^2(v, v_0) \tag{3.4}$$

if v is close enough to v_0 and the footpoint of v lies on ν. By Lemma 2.3 and symmetry,

$$.8 < u_{v_0}(1, -2r - d, d) < .81.$$

Hence

$$u'_{v_0}(1, -2r - d, d) = -K - u^2 > 1 - (.81)^2 > .3.$$

By continuity, if v is close enough to v_0, then $u'_v(1, \alpha(v), t) > .3$ for all $t \in [t_1(v) + 2d, t_2(v)]$. Therefore, by (3.3) and (3.4)

$$u_v(1, \alpha(v), t_2(v)) \geqslant u_{v_0}(1, -2r - d, d) + .54 \cdot 10^6 \cdot \operatorname{dist}^2(v, v_0).$$

We have

$$\begin{aligned} u_v(1,\alpha(v),\omega(v)) &\geqslant u_{v_0}\big(u_{v_0}(1,-2r-d,d) \\ &\qquad + .54\cdot 10^6\cdot \operatorname{dist}^2(v,v_0),d,2r+d\big) \\ &\geqslant u_{v_0}(1,-2r-d,2r+d)+.54\cdot e^{-4r}\cdot 10^6\cdot \operatorname{dist}^2(v,v_0) \\ &= -1+5\cdot 10^5\cdot \operatorname{dist}^2(v,v_0), \end{aligned}$$

where the first inequality follows from Lemma 2.2. The second inequality follows from Lemma 3.4 below, since $u_{v_0}(1,-2r-d,t)\leqslant 1$ for all t. The third inequality follows from (2.8) and the fact that $r<10^{-6}$ by (2.1). This proves claim (3.1).

Since $u_{v_0}(1,-2r-d,t)\geqslant -1$ for all t, Lemma 3.4 below implies that

$$\begin{aligned} \frac{\partial u_{v_0}}{\partial\lambda}(1,\alpha(v),\omega(v)) &\leqslant \exp(-2(4r+2d)\cdot(-1)) \\ &\leqslant \exp\big(2(4\cdot 10^{-6}+2\cdot 1.1)\big)<99. \end{aligned}$$

Claim (3.2) follows by continuity.

3.3. Lemma. *Let $v(s)$ be a curve of unit tangent vectors with footpoints on ν and such that $v(0)=v_0$ and $\|(d/ds)v(s)|_{s=0}\|=1$. Then*

$$\frac{d}{ds}(t_2(v(s))-t_1(v(s)))=0,$$

$$\frac{d^2}{ds^2}(t_2(v(s))-t_1(v(s)))>2\cdot 10^6.$$

Proof. The first assertion is true because γ_0 is the shortest connection between the caps. Also

$$\frac{d}{ds}t_1(v(s))\bigg|_{s=0}=\frac{d}{ds}t_2(v(s))\bigg|_{s=0}=0 \tag{3.5}$$

since γ_0 is the shortest connection between ν and each of the caps. Next we compute $(d^2/ds^2)t_2(v(s))|_{s=0}$.

Let

$$\tilde v(s)=\frac{t_2(v(s))}{d}v(s).$$

Then, by (3.5),

$$\frac{d}{ds}\tilde v(s)\bigg|_{s=0}=\frac{d}{ds}v(s)\bigg|_{s=0}\overset{\text{def}}{=}X. \tag{3.6}$$

Let σ_s be the geodesic with initial velocity $\tilde v(s)$. Then the point $\sigma_s(d)$ lies on the boundary of the second cap. Set $J_s(t)=(\partial/\partial s)\sigma_s(t)$. By (3.6), since $v(s)$ is a curve of unit vectors with footpoints on ν,

$$J_0(t)=(a\cosh t+b\sinh t)n(t),\qquad 0\leqslant t\leqslant d,$$

where $n(t)$ is the normal field to $\sigma_0 = \gamma_0$ and a and b are the horizontal and vertical components of X respectively. Then, by assumption, $a^2 + b^2 = \|X\|^2 = 1$. Since $t_2(v(s))$ is the length of $\sigma_s|[0, d]$, the formula for the second variation of arclength [1] gives

$$\begin{aligned}\frac{d^2}{ds^2} t_2(v(s))\bigg|_{s=0} &= \langle \nabla_{J_0} J_s, \dot\sigma_0\rangle\Big|_0^d + \int_0^d \langle J_0', J_0'\rangle\, dt - \int_0^d \langle R(J_0, \dot\sigma_0)\dot\sigma_0, J_0\rangle\, dt \\ &= \langle \nabla_{J_0} J_s, \dot\sigma_0\rangle\Big|_0^d + \langle J_0, J_0'\rangle\Big|_0^d \\ &= k(a\cosh d + b\sinh d)^2 + (a^2 + b^2)\sinh d\cosh d \\ &\quad + ab(\sinh^2 d + \cosh^2 d - 1),\end{aligned}$$

where k is the geodesic curvature of the boundary circle $\partial\mathscr{C}$. Similarly,

$$\begin{aligned}\frac{d^2}{ds^2}(-t_1(v(s)))\bigg|_{s=0} &= k(a\cosh d - b\cosh d)^2 + (a^2 + b^2)\sinh d\cosh d \\ &\quad + ab(1 - \sinh^2 d - \cosh^2 d).\end{aligned}$$

By Lemma 2.4 and (2.9), $k \geqslant 10^6$, $d \geqslant 1$.

3.4. Lemma. *Let h: $\mathbf{R} \to \mathbf{R}$ be a continuous function, and let $u(\lambda, t_0, \cdot)$ be the solution of*

$$u' + u^2 + h(t) = 0, \qquad u(\lambda, t_0, t_0) = \lambda.$$

Suppose $t_1 > t_0$ and $u(\lambda, t_0, t_1)$ is defined. Denote by M and m the maximum and minimum respectively of $u(\lambda, t_0, t)$ on $[t_0, t_1]$. Then

$$e^{-2(t_1 - t_0)M} \leqslant \frac{\partial u}{\partial\lambda}(\lambda, t_0, t_1) \leqslant e^{-2(t_1 - t_0)m}.$$

Proof. Let $w(t) = (\partial u/\partial\lambda)(\lambda, t_0, t)$. Then $w(t_0) = 1$ and

$$w'(t) + 2w(t)\cdot u(\lambda, t_0, t) = 0.$$

Hence

$$\frac{w(t_1)}{w(t_0)} = \exp\left\{-2\int_{t_0}^{t_1} u(\lambda, t_0, t)\, dt\right\}. \qquad \text{q.e.d.}$$

The following lemma gives the final estimates on the solutions of the Riccati equation along geodesics which cross at least one of the caps.

3.5. Main Lemma. *There exists ε_1, $0 < \varepsilon_1 < 10^{-4}$, such that*

(1) *if v is a unit tangent vector with footpoint on ν,* $\operatorname{dist}(v, \pm v_0) < \varepsilon_1$, *and*

$$1 - 5\cdot 10^3 \operatorname{dist}^2(v, \pm v_0) \leqslant \lambda \leqslant 1,$$

then

$$u_v(\lambda, \alpha(v), \omega(v)) \geqslant -1 + 5\cdot 10^3 \operatorname{dist}^2(v, \pm v_0);$$

(2) *if* v *is a unit tangent vector with footpoint on* ν, $\operatorname{dist}(v, \pm v_0) \geqslant \varepsilon_1$, γ_v *intersects both caps, and*

$$\lambda_1 \stackrel{\text{def}}{=} 1 - 5 \cdot 10^3 \varepsilon_1^2 \leqslant \lambda \leqslant 1,$$

then $u_v(\lambda, \alpha(v), \omega(v)) \geqslant -\lambda_1$;

(3) *if* γ_v *intersects only one cap and* $\lambda_1 \leqslant \lambda \leqslant 1$, *then* $u_v(\lambda, \alpha(v), \omega(v)) \geqslant -\lambda_1$.

Proof. Proposition 3.2 implies that statement (1) is true for any $\varepsilon_1 \leqslant \varepsilon_0$.

We will now show that there is λ_2, $0 < \lambda_2 < 1$, such that if the footpoint of v lies on ν, $\operatorname{dist}(v, \pm v_0) \geqslant \varepsilon_0$, and γ_v intersects both caps, then $u_v(\lambda, \alpha(v), \omega(v)) \geqslant -\lambda_2$. If such a λ_2 does not exist, then there is a sequence $\{v_n\}$ of such vectors with

$$u_{v_n}\left(1 - \frac{1}{n}, \alpha(v_n), \omega(v_n)\right) \leqslant -1 + \frac{1}{n}.$$

Since the caps are compact, $\{v_n\}$ has a limit vector v for which $u_v(1, \alpha(v), \omega(v)) \leqslant -1$. By Proposition 3.1, $v = \pm v_0$, which is a contradiction.

By Lemmas 2.2 and 2.3, there exists λ_3, $0 < \lambda_3 < 1$, such that $u_v(\lambda_3, \alpha(v), \omega(v)) \geqslant -\lambda_3$ for every vector v as in (3). Now choose ε_1 so that

$$0 < \varepsilon_1 < \min\left(\varepsilon_0, 10^{-4}\right) \quad \text{and} \quad \lambda_1 = 1 - 5 \cdot 10^3 \varepsilon_1^2 \geqslant \max(\lambda_2, \lambda_3). \quad \text{q.e.d.}$$

Although we will not use it, we point out an immediate consequence of the Main Lemma 3.5, Proposition 2.5, and Proposition 1.2.

3.6. Corollary. *The noncompact surface* S_d *has no conjugate points.*

4. Construction of the compact surface $S_{d,R}$

Set $R = \operatorname{arctanh}(\lambda_1) + 5$, where λ_1 is defined in Main Lemma 3.5. Recall that, by Proposition 1.4(1), the solution u of the equation $u' + u^2 - 1 = 0$ with initial value $-\lambda_1$ reaches 0 at $t = \operatorname{arctanh}(\lambda_1)$. Therefore

$$u(t) > 0 \quad \text{for } t > R - 5. \tag{4.1}$$

Let B_R be the ball in S_d of radius R centered at $\gamma_0(0) = 0 = \gamma_0 \cap \nu$ (see Figure 4).

Figure 4 is in the hyperbolic plane. The surface S_d with the caps removed is obtained from the unshaded area by identifying the indicated pairs of geodesic rays. Easy geometric arguments in the hyperbolic plane show that there exist arcs of hyperbolic geodesics

$$\nu'_+, c'_1, d'_1, c'_2, \cdots, d'_{n-1}, c'_n, \nu'_-,$$

$$\nu''_+, c''_1, d''_1, c''_2, \cdots, d''_{n-1}, c''_n, \nu''_-$$

with the following properties indicated in Figure 4:

(1) none of the arcs intersects B_R;

(2) adjacent geodesics are orthogonal;

(3) ν'_+, ν''_+, ν'_-, and ν''_- are orthogonal to γ_0 and each has length arccosh 2;

(4) length(c'_i) = length(c''_i).

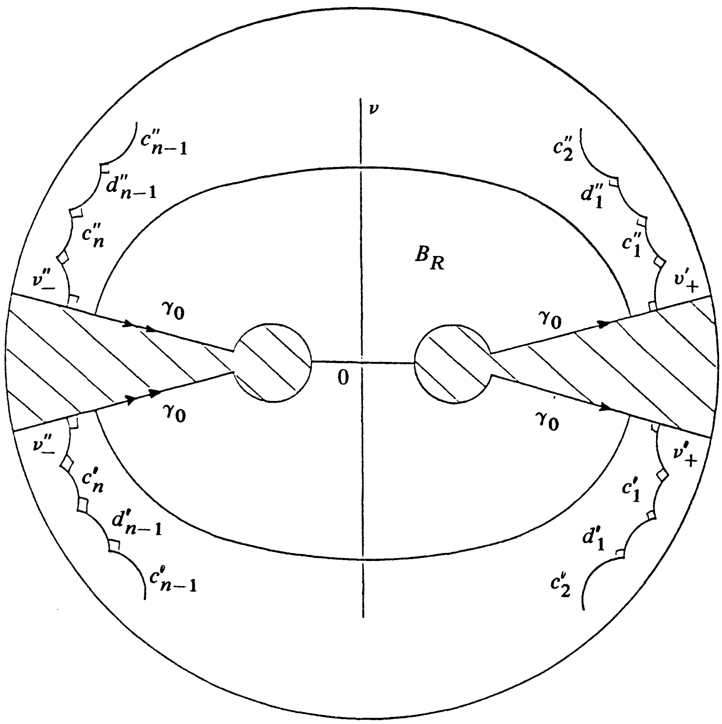

FIGURE 4

Consider the region in S_d bounded by these geodesic arcs. By identifying c'_i with c''_i we obtain the surface bounded by the closed geodesics $\nu_+, d_1, \cdots, d_{n-1}, \nu_-$ shown in Figure 5. Along each of the boundary geodesics we attach a handle with curvature -1. The handle attached to ν_+ is constructed in the following way. Take two regular hexagons in the hyperbolic plane with all interior angles $\pi/2$. Note that the side of such a hexagon has length arccosh 2 $= \frac{1}{2}$ length(ν_+). Identify the corresponding sides of the hexagons to get a pair of pants as shown in Figure 6. Now identify the closed geodesics marked σ_+ to

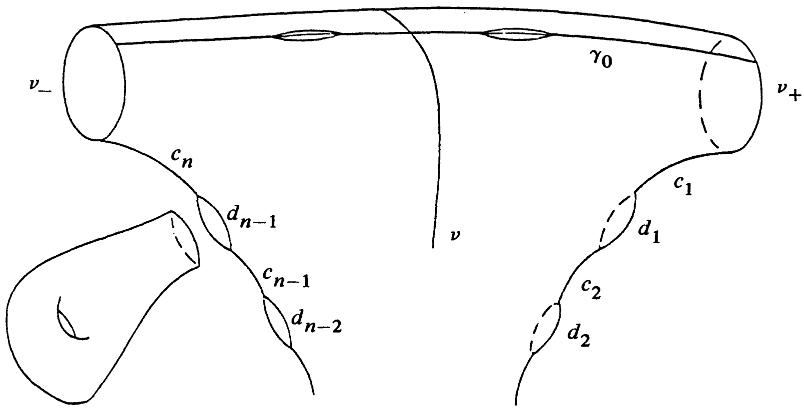

FIGURE 5

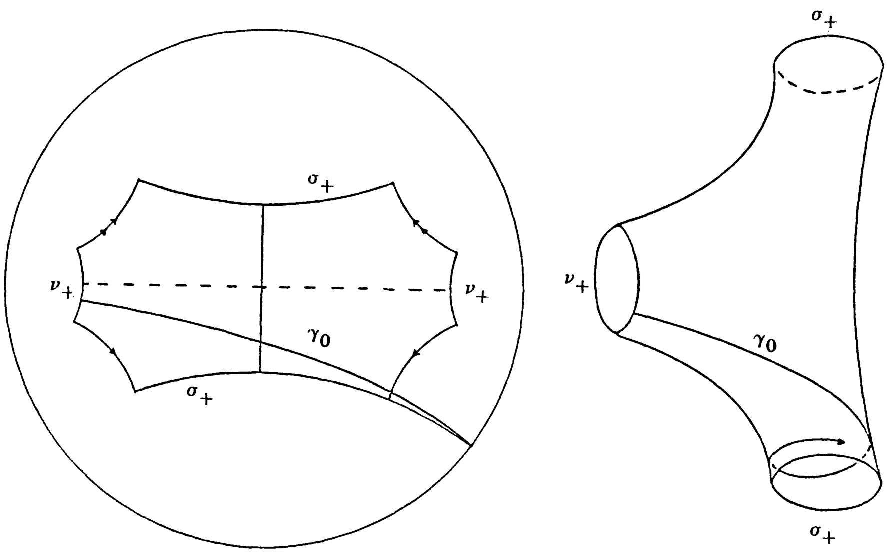

FIGURE 6

obtain a handle. We attach this handle to our surface along ν_+ in such a way that γ_0 is asymptotic to σ_+ as shown in Figure 6.

It is clear from our construction that

(1) γ_0 enters the handle and never leaves it;

(2) the injectivity radius at every point in the handle, in particular at every point on γ_0, is greater than $1/2$;

(3) for any ball B of radius $1/2$ in the handle there is an isometry ϕ of B into the hyperbolic plane, and $d\phi$ maps vectors asymptotic to γ with footpoints in B to asymptotic vectors.

In the same way we attach a handle along ν_- so that γ_0 is negatively asymptotic to the corresponding closed geodesic σ_- and the analogs of (1), (2), and (3) hold. We call the resulting surface $S_{d,R}$.

4.1. Lemma. *For any* $s, t > 3$ *or* $s, t < -3$, $\operatorname{dist}(\dot\gamma_0(s), -\dot\gamma_0(t)) > 1/2$.

Proof. Let $s, t > 3$ and let B be the closed ball in $S_{d,R}$ of radius $1/2$ about $\gamma_0(s)$. It is clearly enough to consider the case when the shortest curve connecting $\dot\gamma_0(s)$ and $-\dot\gamma_0(t)$ has footpoints in B. Since the injectivity radius at $\gamma_0(s)$ is greater than $1/2$, there is an isometric embedding φ of B into the hyperbolic plane. By property (3) above, the vectors $d\varphi\dot\gamma_0(s)$ and $d\varphi\dot\gamma_0(t)$ are asymptotic. The lemma now follows from Lemma 5.5.

4.2. Theorem. *The surface* $S_{d,R}$ *has no conjugate points.*

Proof. For a unit vector v denote by γ_v the geodesic determined by $\dot\gamma_v(0) = v$ and by u_v the solution of the initial value problem

$$u_v' + u_v^2 + K(\gamma_v(t)) = 0, \qquad u_v(0) = 1.$$

If $\gamma_v(0)$ is not in the caps, we will show that $u_v(t)$ is defined and $|u_v(t)| \leqslant 1$ for all positive t. By Proposition 1.2, this implies that there are no conjugate points along $\gamma_v|[0, \infty)$. The theorem now follows easily.

Since $K \geqslant -1$ and $u_v(0) = 1$, we see that $u_v(t) \leqslant 1$ for all $t \geqslant 0$ for which $u_v(t)$ is defined.

We follow γ_v as it moves along $S_{d,R}$. If $\gamma_v|[0, \infty)$ never crosses the caps, then $u_v(t) \equiv 1$ for $t \geqslant 0$, and the claim is clear. Otherwise there is a first time $\alpha_1 > 0$ when γ_v enters a cap. By Proposition 2.5, there is a first time $z_1 > \alpha_1$ when γ_v leaves the ball B_R (see Figure 4). Denote by $\omega_1 \in (\alpha_1, z_1)$ the last time γ_v is in a cap before leaving B_R. If γ_v crosses the caps again, denote by $\alpha_2 > z_1$ the first time it enters a cap after z_1 and by $a_2 \in (z_1, \alpha_2)$ the last time it reenters B_{R-1} before α_2. Let z_2 be the first time γ_v leaves B_R after α_2 and let ω_2 be the last time before z_2 at which γ_v is in a cap. By proceeding inductively we obtain a (possibly finite) sequence of times

$$\alpha_1 < \omega_1 < z_1 < a_2 < \alpha_2 < \cdots < a_n < \alpha_n < \omega_n < z_n < \cdots.$$

Thus γ_v crosses the caps for the nth time during the interval $[\alpha_n, \omega_n]$. Denote by ν_R the diameter of B_R orthogonal to γ_0. If γ_v crosses ν_R at a time $\tau_n \in (\alpha_n, \omega_n)$, then set

$$d_n = \operatorname{dist}(\dot\gamma_v(\tau_n), \pm v_0),$$

and let $d_n = \infty$ if $\gamma_v([\alpha_n, \omega_n]) \cap \nu_R = \varnothing$.

We prove by induction in n that u_v is defined on $[0, \alpha_n]$ and satisfies either

$$(4.2) \qquad u_v(\alpha_n) \geqslant 1 - 5 \cdot 10^3 d_n^2 \quad \text{if } d_n < \varepsilon_1$$

or

$$(4.3) \qquad u_v(\alpha_n) \geqslant 1 - 5 \cdot 10^3 \cdot \varepsilon_1^2 = \lambda_1 \quad \text{if } d_n \geqslant \varepsilon_1.$$

This implies that $u_v(t)$ is defined for all positive t even if the sequence $\{\alpha_n\}$ terminates. In the latter case, let α_k be the last term of the sequence. Then $K(\gamma_v(t)) \equiv -1$ for $t \geqslant \omega_k$. By Main Lemma 3.5 and the inductive hypothesis, $u_v(\omega_k) \geqslant -1$, and so $u_v(t)$ is defined for all $t \geqslant 0$.

Since the curvature is identically equal to -1 on $\gamma_v([0, \alpha_1])$, we see that $u_v(t) \equiv 1$ for $t \in [0, \alpha_1]$, and the inductive hypothesis holds for $n = 1$. We now prove that the inductive hypothesis holds for $n+1$ if it holds for n, provided α_{n+1} is defined. By Main Lemma 3.5, $u_v(\omega_n) \geqslant -1$. Since the curvature is identically equal to -1 on $\gamma_v([\omega_n, \alpha_{n+1}])$, we see that u_v is defined on $[0, \alpha_{n+1}]$. Consider the following four cases.

Case 1: $d_n \geqslant \varepsilon_1$, $d_{n+1} \geqslant \varepsilon_1$. Since $\gamma_v(\omega_n)$ lies within distance 2 from the center 0 of B_R, we have $z_n - \omega_n > R - 2$. By the inductive hypothesis and Main Lemma 3.5, $u_v(\omega_n) \geqslant -\lambda_1$ and therefore $u_v(z_n) > 0$ by (4.1). Hence $u(a_{n+1}) > 0$. Similarly $\alpha_{n+1} - a_{n+1} > R - 3$ and thus $u_v(\alpha_{n+1}) \geqslant \lambda_1$. This proves the inductive hypothesis in Case 1.

We will need the following lemmas to treat the remaining cases. If $d_k < \varepsilon_1$, set

(4.4) $$T_k = -\ln 4 - 1 - \ln d_k.$$

Note that $T_k > 6$ since $\varepsilon_1 < 10^{-4}$ (see Main Lemma 3.5).

4.3. Lemma. *If $d_k < \varepsilon_1$ and $\tau_k - T_k \leqslant t \leqslant \tau_k + T_k$, then*

(1) $\operatorname{dist}(\dot\gamma_v(t), \dot\gamma_{v_0}(t - \tau_k)) < 1/4$ *if* $\dot\gamma_v(\tau_k)$ *is close to* $+v_0$, *i.e.*, $d_k = \operatorname{dist}(\dot\gamma_v(\tau_k), +v_0)$;

(2) $\operatorname{dist}(\dot\gamma_v(t), \dot\gamma_{-v_0}(t - \tau_k)) < 1/4$ *if* $\dot\gamma_v(\tau_k)$ *is close to* $-v_0$, *i.e.*, $d_k = \operatorname{dist}(\dot\gamma_v(\tau_k), -v_0)$.

Proof. By symmetry it suffices to prove (1) for $t \geqslant \tau_k$. Let $v(\theta)$, $0 \leqslant \theta \leqslant d_k$, be the shortest connection between $v_0 = v(0)$ and $\dot\gamma_v(\tau_k)$ in the unit tangent bundle. We claim that for any s, $0 \leqslant s \leqslant T_k$, and all θ

$$\operatorname{dist}\big(\dot\gamma_{v(\theta)}(s), \dot\gamma_{v_0}(s)\big) < 1/4.$$

If not, there is a smallest $s_0 > 0$ where this fails for some θ_0. Since $\operatorname{dist}(\dot\gamma_{v(\theta)}(s), \dot\gamma_{v_0}(s)) \leqslant 1/4$ for $0 \leqslant s \leqslant s_0$ and any θ, the geodesics $\gamma_{v(\theta)}|[0, s_0]$ cross at most one cap. Hence by (2.1),

$$\int_0^{s_0} \left|1 - K\big(\gamma_{v(\theta)}(s)\big)\right| ds \leqslant 2s_0 + 2r \cdot \kappa^2 \leqslant 2s_0 + 1.8.$$

Hence, by Lemma 5.1 and (4.4),

$$\frac{1}{4} = \operatorname{dist}\big(\dot\gamma_{v(\theta_0)}(s_0), \dot\gamma_{v_0}(s_0)\big) \leqslant \int_0^{\theta_0} \left\| dg^t(v(\theta)) \cdot \frac{dv(\theta)}{d\theta} \right\| d\theta$$
$$\leqslant d_k\left(\frac{1}{2}(2s_0 + 1.8)\right) < \frac{1}{4}.$$

This contradiction proves our claim and the lemma follows.

4.4. Lemma. *If $d_k < \varepsilon_1$ and a_{k+1} is defined, then $\tau_k + T_k \leqslant a_{k+1}$.*

Proof. If $T_k \leqslant R - 1/4$, then by the previous lemma and the triangle inequality, $\gamma_v(t) \in B_R$ for $t \in [\tau_k, \tau_k + T_k]$, and so $\tau_k + T_k \leqslant z_k < a_{k+1}$. If $T_k > R - 1/4$, then $\operatorname{dist}(\gamma_0(t), 0) > R - 1/4$ for $R - 1/4 < t \leqslant T_k$ since γ_0 leaves B_R and never returns. Therefore, by the previous lemma and the triangle inequality, $\operatorname{dist}(\gamma_v(\tau_k + t), 0) > R - 1/2$ for $R - 1/4 < t \leqslant T_k$. The statement of the lemma follows since a_{k+1} is the time when γ_v reenters B_{R-1}.

4.5. Lemma. *Let $u(-\lambda, \cdot)$ be the solution of $u' + u^2 - 1 = 0$, $u(-\lambda, 0) = -\lambda$. Then*

$$u(-1 + 5 \cdot 10^3 \varepsilon^2, T) > 0$$

if $T > -\ln \varepsilon - \ln 50$.

Proof. By Proposition 1.4(1), $u(-\lambda, t) > 0$ if $t > \operatorname{arctanh}(\lambda)$.

Case 2: $d_n < \varepsilon_1$, $d_{n+1} < \varepsilon_1$. We will show now that the numbers T_n, T_{n+1} defined in (4.4) satisfy

$$\tau_n + T_n < \tau_{n+1} - T_{n+1}. \tag{4.5}$$

If not, then, by Lemma 4.3, $t = \tau_n + T_n$ satisfies

$$\operatorname{dist}\big(\dot\gamma_v(t), \dot\gamma_{+v_0}(t - \tau_n)\big) < 1/4, \tag{4.6}$$

and either

$$\operatorname{dist}\big(\dot\gamma_v(t), \dot\gamma_{+v_0}(t - \tau_{n+1})\big) < 1/4 \tag{4.7}$$

or

$$\operatorname{dist}\big(\dot\gamma_v(t), \dot\gamma_{-v_0}(t - \tau_{n+1})\big) < 1/4, \tag{4.8}$$

where (4.7) corresponds to the case $d_{n+1} = \operatorname{dist}(\dot\gamma_v(\tau_{n+1}), -v_0)$ and (4.8) corresponds to the case $d_{n+1} = \operatorname{dist}(\dot\gamma_v(\tau_{n+1}), +v_0)$. We now show that (4.6) and (4.7) are contradictory. Note that together with the triangle inequality they imply

$$\operatorname{dist}\big(\gamma_{v_0}(t - \tau_n), \gamma_{v_0}(t - \tau_{n+1})\big) < 1/2.$$

But $t - \tau_n > 6$ and $t - \tau_{n+1} < 0$ by (4.4) and Lemma 4.4. Hence the above inequality cannot hold since $\gamma_{v_0} \| [-R, R]$ minimizes distances and the closed geodesics σ_+ and σ_- to which γ_{v_0} is asymptotic lie far apart from each other. On the other hand, (4.6), (4.8) are also contradictory. For together with the triangle inequality they imply that the distance between $\dot\gamma_{v_0}(t - \tau_n)$ and $\dot\gamma_{-v_0}(t - \tau_{n+1}) = -\dot\gamma_{v_0}(\tau_{n+1} - t)$ is less than $1/2$. This contradicts Lemma 4.1 since $t - \tau_n > 6$ and, by Lemma 4.4, $\tau_{n+1} - t \geqslant \tau_{n+1} - a_{n+1} > R - 2 > 3$.

By Main Lemma 3.5 and the inductive hypothesis (4.2),

$$u_v(\omega_n) \geqslant -1 + 5 \cdot 10^3 d_n^2. \tag{4.9}$$

By Lemma 4.4, $K(\gamma_v(t)) \equiv -1$ for $t \in [\omega_n, \tau_n + T_n]$, and hence $u_v(t)$ is increasing on this interval because $|u_v(\omega_n)| \leqslant 1$. Since the caps are very small (cf. (2.1)), (2.9) implies that $\omega_n < \tau_n + 1.2$. Therefore

$$\begin{aligned} u_v(\tau_n + T_n) &\geqslant u_v(\omega_n + T_n - 1.2) \\ &= \frac{u_v(\omega_n)\cosh(T_n - 1.2) + \sinh(T_n - 1.2)}{u_v(\omega_n)\sinh(T_n - 1.2) + \cosh(T_n - 1.2)} \end{aligned} \tag{4.10}$$

by Proposition 1.4(1). It follows easily from (4.9) and (4.4) that $\tanh(T_n - 1.2) > -u_v(\omega_n)$. Hence the numerator in (4.10) is positive. The denominator is also positive since $u_v(\omega_n) \geqslant -1$. Hence

$$u_v(\tau_n + T_n) > 0.$$

Recall (4.5) that $\tau_{n+1} - T_{n+1} > \tau_n + T_n$ and note that $K(\gamma_v(t)) \equiv -1$ for $\tau_n + T_n \leqslant t \leqslant \tau_{n+1} - T_{n+1}$. Hence $u_v(\tau_{n+1} - T_{n+1}) > 0$. A similar calculation now shows that

$$u_v(\alpha_{n+1}) \geqslant 1 - 5 \cdot 10^3 \cdot d_{n+1}^2.$$

This completes the inductive step in Case 2.

Case 3: $d_n < \varepsilon_1$, $d_{n+1} \geqslant \varepsilon_1$. Inequality (4.10) shows that $u_v(\tau_n + T_n) > 0$. On the other hand, $\tau_n + T_n < a_{n+1}$, by Lemma 4.4, and therefore

$$u_v(a_{n+1}) > u_v(\tau_n + T_n) > 0.$$

Now an argument similar to Case 1 shows that $u_v(\alpha_{n+1}) \geqslant \lambda_1$.

Case 4: $d_n \geqslant \varepsilon_1$, $d_{n+1} < \varepsilon_1$. This case follows from Case 3 by reversing time. This finishes the proof of Theorem 4.2.

Orient σ_+ and σ_- so that γ_0 is positively asymptotic to σ_+ and negatively asymptotic to σ_-.

4.6. Theorem. *The function $v \to u_v^+(0)$ is discontinuous at each tangent vector $\dot\sigma_+(s)$ of σ_+. The function $v \to u_v^-(0)$ is discontinuous at each tangent vector of σ_-.*

Proof. We prove only the first statement. Since γ_0 is asymptotic to σ_+, there exists a sequence $\{w_n\}$ of vectors tangent to γ_0 converging to $w = \dot\sigma_+(s)$. By construction, $u_{w_n}^+(0) = -1$ for all large enough n. On the other hand, $u_w^+(0) = 1$.

4.7. Remark. The geodesic flow of $S_{d,R}$ is ergodic.

To see this, note that $u^-(v) = -u^+(-v)$ and that the Liouville measure μ is invariant under the involution $v \to -v$. Hence

$$\int_{T_1 S_{d,R}} u^+(v)\, d\mu(v) = \frac{1}{2}\int_{T_1 S_{d,R}} (u^+(v) - u^-(v))\, d\mu(v) > 0$$

since it can be seen from the proof of Theorem 4.2 that $u^+(v) - u^-(v) > 0$ for all $v \in T_1 S_{d,R}$ not tangent to γ_0. Hence the geodesic flow has positive entropy (see Theorem 7.8 in [7]) and is therefore ergodic by Theorem 9.4 in [6].

5. Appendix

First we give an estimate on the differential of the geodesic flow g^t.

5.1. Lemma. *Let M be a complete Riemannian manifold. For a unit tangent vector u set $R_u w = R(w, u)u$. Then for each unit vector v the differential at v of the geodesic flow g^t satisfies*

$$\|dg^t(v)\| \leqslant \exp\left\{\frac{1}{2}\int_0^t \|\mathrm{Id} - R_{g^s v}\|\, ds\right\}.$$

Proof. Let Y be a Jacobi field along γ_v. Then

$$\begin{aligned}\left(\langle Y, Y\rangle + \langle Y', Y'\rangle\right)' &= 2\langle Y', Y + Y''\rangle \\ &= 2\left\langle Y', (\mathrm{Id} - R_{g^s v})Y\right\rangle \leqslant 2\|Y\| \cdot \|Y'\| \cdot \|\mathrm{Id} - R_{g^s v}\| \\ &\leqslant \left(\langle Y, Y\rangle + \langle Y', Y'\rangle\right)\|\mathrm{Id} - R_{g^s v}\|. \qquad \text{q.e.d.}\end{aligned}$$

We now study the geometry of the unit tangent bundle $T_1 M$ of a complete Riemannian manifold M with the metric

$$\langle X, Y\rangle = \langle X_H, Y_H\rangle + \langle X_V, Y_V\rangle,$$

where H and V denote the horizontal and vertical components, respectively.

5.2. Proposition (*Sasaki* [8]). *Let $v(\cdot)$ be a curve of vectors in $T_1 M$. Denote by $c(t)$ the footpoint of $v(t)$. Then $v(\cdot)$ is a geodesic in $T_1 M$ if and only if*

$$v'' = -\langle v', v'\rangle v, \tag{5.1}$$

$$\dot{c}' = R(v', v)\dot{c}, \tag{5.2}$$

where prime denotes covariant differentiation along c, $\dot{c}(t)$ is the tangent vector of $c(\cdot)$ at t, and R is the Riemann curvature tensor on M.

Proof. The energy of the curve $v\colon [a, b] \to T_1 M$ is given by

$$E(v) = \frac{1}{2}\int_a^b \dot{c}^2\, dt + \frac{1}{2}\int_a^b (v')^2\, dt.$$

The second integral is the energy of the curve $\tilde{v}$ in the unit sphere at $c(0)$ obtained by parallel translating the vectors $v(t)$ along c to $v(0)$.

Assume v is a critical point of E. By considering variations of v which leave fixed the curve c of footpoints and $v(a)$ and $v(b)$, we see that $\tilde{v}$ is a geodesic segment in the unit sphere. Hence v satisfies (5.1).

Now consider an arbitrary variation $v(\varepsilon, \cdot)$ of $v(\cdot)$ with footpoint $c(\varepsilon, \cdot)$ which keeps $v(\cdot, a)$ and $v(\cdot, b)$ fixed. Then

$$0 = \frac{dE}{d\varepsilon} = \int_a^b \left[\left\langle \frac{D}{d\varepsilon}\dot{c}, \dot{c} \right\rangle + \left\langle \frac{D}{d\varepsilon}v', v' \right\rangle \right] dt$$

$$= \int_a^b \left[\left\langle \frac{D}{dt}\frac{\partial c}{\partial \varepsilon}, \dot{c} \right\rangle + \left\langle \frac{D}{dt}\frac{D}{d\varepsilon}v, v' \right\rangle + \left\langle R\left(\frac{\partial c}{\partial \varepsilon}, \frac{\partial c}{\partial t} \right) v, v' \right\rangle \right] dt$$

$$= \int_a^b \left[\frac{d}{dt}\left(\left\langle \frac{\partial c}{\partial \varepsilon}, \dot{c} \right\rangle \right) - \left\langle \frac{\partial c}{\partial \varepsilon}, \dot{c}' \right\rangle + \frac{d}{dt}\left(\left\langle \frac{D}{d\varepsilon}v, v' \right\rangle \right) \right.$$

$$\left. - \left\langle \frac{D}{d\varepsilon}v, v'' \right\rangle + \left\langle R(v', v)\dot{c}, \frac{\partial c}{\partial \varepsilon} \right\rangle \right] dt.$$

Note that

$$\int_a^b \frac{d}{dt}\left\langle \frac{\partial c}{\partial \varepsilon}, \dot{c} \right\rangle dt = 0 = \int_a^b \frac{d}{dt}\left\langle \frac{D}{d\varepsilon}v, v' \right\rangle dt$$

since $c(\varepsilon, a)$, $c(\varepsilon, b)$, $v(\varepsilon, a)$, and $v(\varepsilon, b)$ are kept fixed. Also $\langle (D/d\varepsilon)v, v'' \rangle = 0$ since v'' and v are collinear by (5.1) and $\langle (D/d\varepsilon)v, v \rangle = 0$ because v is a curve of unit vectors. Hence (5.2) is satisfied.

Conversely, our computations show that v is a critical point of E if it satisfies (5.1) and (5.2). q.e.d.

Note that (5.1) and (5.2) imply $\langle v', v' \rangle' = 0 = \langle \dot{c}, \dot{c} \rangle'$ and therefore

$$(5.3) \qquad \|v'\| = \text{const}, \qquad \|\dot{c}\| = \text{const}.$$

We now describe geodesics in the unit tangent bundle T_1S of a surface S. We use the notation of Proposition 5.2.

5.3. Corollary. *Let $v(\cdot)$ be a geodesic in T_1S. Then exactly one of the following three possibilities holds*:

(1) *The geodesic v is vertical, i.e., $\dot{c}(t) \equiv 0$, and v is an arc of the unit circle at $c(0)$.*

(2) *The geodesic v is horizontal, i.e., $v'(t) \equiv 0$, and c is a geodesic in S.*

(3) *If $\|v'\| \neq 0$ and $\|\dot{c}\| \neq 0$, then*

$$(5.4) \qquad v''(t) = -\|v'\|^2 \cdot v(t),$$

$$(5.5) \qquad k(t) = \frac{\|v'\|}{\|\dot{c}\|} K(c(t)),$$

where $k(t)$ is the geodesic curvature of c with respect to the normal $N(t)$ such that $(\dot{c}(t), N(t))$ and $(v(t), v'(t))$ have the same orientation.

Proof. Assertions (1), (2), and (5.4) follow directly from Proposition 5.2. To prove (5.5) note that by (5.2)

$$k = \frac{\langle \dot{c}', N\rangle}{\|\dot{c}\|^2} = \frac{K(c(t))}{\|\dot{c}\|^2}\big(\langle v, \dot{c}\rangle\langle v', N\rangle - \langle v', \dot{c}\rangle\langle v, N\rangle\big)$$

$$= \frac{\|v'\|}{\|\dot{c}\|}K(c(t)). \qquad \text{q.e.d.}$$

We use the previous corollary to calculate the distance between v and $-v$, where v is a unit tangent vector of the hyperbolic plane H.

5.4. Lemma. *Let* $v \in T_1H$. *Then* $\operatorname{dist}(v, -v) = \pi$.

Proof. Clearly $\operatorname{dist}(v, -v) \leqslant \pi$ since the semicircles from v to $-v$ in the unit circle at the footpoint of v have length π. Now let $w(t)$, $0 \leqslant t \leqslant 1$, be a nonvertical geodesic in T_1H joining v to $-v$. We will show now that the length of w is at least π. We can assume that the curve c of the footpoints of w has length at most π. By (5.5), c is a geodesic circle of finite radius r traversed n times, $n = 1, 2, \cdots$, and

$$\operatorname{length}(c) = 2\pi n \sinh r \leqslant \pi.$$

Denote by P_c the parallel translation along c. Then clearly

$$\sphericalangle(P_c v, v) = \int_0^1 k(t) \cdot \|\dot{c}(t)\|\, dt \mod 2\pi,$$

where $k(t) \equiv \coth r$ is the geodesic curvature of c. Hence

$$\sphericalangle(P_c v, v) = 2\pi n \cosh r - 2\pi j \in [0, 2\pi),$$

where

$$0 \leqslant j = [n \cosh r] \leqslant [n(\sinh r + 1)] \leqslant [1/2 + n] = n.$$

Therefore

$$\sphericalangle(P_c v, -v) = 2\pi n \cosh r - \pi(2j+1) \in [-\pi, \pi).$$

By the Pythagoras theorem,

$$(\operatorname{length}(v))^2 \geqslant (2\pi n \sinh r)^2 + (2\pi n \cosh r - \pi(2j+1))^2 \overset{\text{def}}{=} Q(\cosh r).$$

The above estimate on j implies that the quadratic function $Q(x)$ is greater than π^2 for $x = \cosh r > 1$.

5.5. Lemma. *Let v and w be asymptotic unit vectors in the hyperbolic plane. Then*

(1) $\operatorname{dist}(v, w) \leqslant \sqrt{2}\operatorname{dist}(\pi v, \pi w)$,

(2) $\operatorname{dist}(v, -w) \geqslant \pi/(1 + \sqrt{2})$.

Proof. Let $\operatorname{dist}(\pi v, \pi w) = l$ and let c be the geodesic with $c(0) = \pi v$, $c(l) = \pi w$. Denote by $v(s)$ the unit vector at $c(s)$ asymptotic to v and by $\sigma(s, \cdot)$ the geodesic with initial tangent vector $v(s)$. Then $\sigma(s, 0) = c(s)$,

$v(s) = (\partial\sigma/\partial t)(s,0)$, and $Y_s(\cdot) = (\partial\sigma/\partial s)(s,\cdot)$ is a Jacobi field along $\sigma(s,\cdot)$. The horizontal component of $(dv/ds)(s)$ is $(dc/ds)(s) = Y_s(0)$ and the vertical component is

$$\nabla_{dc/ds} v(s) = \nabla_{\partial\sigma/\partial s}\frac{\partial\sigma}{\partial t}(s,0) = \nabla_{\partial\sigma/\partial t}\frac{\partial\sigma}{\partial s}(s,0) = \frac{D}{dt}Y_s(0) = Y_s'(0).$$

Since $Y_s(\cdot)$ is a stable Jacobi field and the curvature is -1,

$$\left\|\frac{dv}{ds}(s)\right\|^2 = \|Y_s(0)\|^2 + \|Y_s'(0)\|^2 \leqslant 2\|Y_s(0)\|^2 = 2\left\|\frac{dc}{ds}(s)\right\|^2,$$

which proves (1).

To prove (2) we may assume that $\operatorname{dist}(\pi v, \pi w) \leqslant \pi/(1+\sqrt{2})$. By Lemma 5.4, $\operatorname{dist}(w,-w) = \pi$. Therefore, by the triangle inequality and (1),

$$\operatorname{dist}(v,-w) \geqslant \operatorname{dist}(w,-w) - \operatorname{dist}(v,w) \geqslant \pi - \sqrt{2}\,\frac{\pi}{1+\sqrt{2}} = \frac{\pi}{1+\sqrt{2}}.$$

References

[1] J. Cheeger & D. Ebin, *Comparison theorems in Riemannian geometry*, North-Holland, Amsterdam, 1975.

[2] E. Heintze & H. C. Im Hof, *Geometry of horospheres*, J. Differential Geometry **12** (1977) 481–491.

[3] E. Hopf, *Statistik der geodätischen Linien in Mannigfaltigkeiten negativer Krümmung*, Leipziger Berichte **91** (1939) 261–304.

[4] ______, *Closed surfaces without conjugate points*, Proc. Nat. Acad. Sci. U.S.A. **34** (1948) 47–51.

[5] G. Knieper, *Mannigfaltigkeiten ohne konjugierte Punkte*, Inaugural dissertation, Bonn, 1985.

[6] Ya. B. Pesin, *Geodesic flows on closed Riemannian manifolds without focal points*, Math. USSR Izv. **11** (1977) 1195–1228.

[7] ______, *Geodesic flows with hyperbolic behaviour of the trajectories and objects connected with them*, Russian Math. Surveys **36** (1981) 1–59.

[8] S. Sasaki, *On the differential geometry of tangent bundles of Riemannian manifolds*, Tôhoku J. Math. **10** (1958) 338–354.

UNIVERSITY OF MARYLAND