J. DIFFERENTIAL GEOMETRY
11 (1976) 321-333

RIEMANNIAN MANIFOLDS WITHOUT FOCAL POINTS

JOHN J. O'SULLIVAN

Introduction

One of several equivalent conditions which can be used to characterize riemannian manifolds without focal points is the condition that the length of all nontrivial, initially vanishing Jacobi fields be strictly increasing for $t > 0$. Consequently any manifold of nonpositive sectional curvature has no focal points. However, the existence of some positive sectional curvatures does not necessarily destroy the nonfocality hypothesis. For examples, both compact and noncompact, of manifolds without focal points whose sectional curvatures change sign, see Gulliver [6].

In this paper we consider only those manifolds without focal points with the property that $\|Y(t)\| \to \infty$ as $t \to \infty$ for all nontrivial Jacobi fields Y satisfying $Y(0) = 0$. It is known that the existence of a lower bound for the sectional curvatures of the manifold implies this property. See [4, Theorem 2.1] and [2, Proposition 2.9]. In particular, therefore, any manifold which is a riemannian covering of a compact manifold without focal points has this property.

In § 1 we establish some properties of the geodesics on a simply connected manifold without focal points. Given a geodesic ray k and a point p, we show (Proposition 3) that there exists a unique geodesic ray which begins at p and is asymptotic to k. In Propositions 2 and 4 we study the functions $d(h(t), k)$ and $d(h(t), k(t))$ for geodesic rays h, k which either intersect or are asymptotic. These two propositions can be obtained for nonpositively curved manifolds by using convexity properties of the distance function. On the other hand, for geodesics which neither intersect nor are asymptotic, we show that analogies with the case of nonpositive curvature break down. In particular, $d(h(t), k)$ may have a critical point which is a local maximum.

§ 2 is devoted to proving the so-called "flat strip theorem" for simply connected manifolds without focal points whose dimension is ≥ 2. This theorem was proved in the two-dimensional case by Green [4].

From recent work of Gromoll & Wolf [5] and Yau [11], it follows that a solvable subgroup of the fundamental group of a compact manifold M of nonpositive sectional curvature is a Bieberbach group, and that such a group has a strong influence on the geometry of the manifold, i.e., it gives rise to an

Received September 11, 1974. Supported in part by National Science Foundation grant GP36418X.

isometric and totally geodesic immersion into M of a compact euclidean space form. In § 3 we extend these results to compact manifolds without focal points. See Theorem 2 and Corollaries 1, 2. We also determine (Proposion 5) the structure of the set of critical points of the square displacement function of an isometry which maps each point into the complement of its cut locus.

We are indebted to P. Eberlein for the proof of Lemma 2 and also for simplifying the proof of Proposition 2. We would also like to thank J. Milnor and B. Smyth for several discussions and helpful suggestions.

0. Preliminaries

Throughout this paper all riemannian manifolds are assumed to be connected, complete, C^∞, and of dimension ≥ 2. A riemannian manifold M is said to have *no focal points* if for any imbedded open geodesic segment $c: (-a, a) \to M$, the exponential map $\exp: c^\perp \to M$ is everywhere nonsingular, where $0 < a \leq \infty$, and $c^\perp$ is the normal bundle of c. By [8, Proposition 4] this condition is equivalent to the condition that $\|Y(t)\|$ be strictly increasing for $t > 0$ where Y is any nontrivial, initially vanishing Jacobi field. In particular, no nontrivial Jacobi field vanishes more than once, and so a manifold without focal points is also without conjugate points. Consequently, if M is simply connected, all maximal geodesics $c: (-\infty, \infty) \to M$ are imbedded, and it follows from [7, Theorem A] that for such c, $\exp: c^\perp \to M$ is a covering map and therefore a diffeomorphism. It follows that for any point p not on c there is a unique geodesic which passes through p and intersects c at right algles.

It is not known whether there are manifolds without focal pints which admit a nontrivial Jacobi field Y with $Y(0) = 0$ and $\|Y(t)\|$ bounded for $t > 0$. However such Jacobi fields cannot exist if the sectional curvatures of the manifold are bounded from below by $-\kappa^2$ for some $\kappa > 0$. See [4, Theorem 2.1] and [2, Proposition 2.9]. From now on we restrict our attention to those manifolds without focal points which satisfy the condition that $\|Y(t)\| \to \infty$ as $t \to \infty$ for all nontrivial, initially vanishing Jacobi fields Y. Note that any manifold which is a riemannian covering of a compact manifold without focal points satisfies this condition.

The condition that the length function of any nontrivial, initially vanishing Jacobi field be strictly increasing for $t > 0$ and unbounded as $t \to \infty$ can be used to establish a uniform growth result for such Jacobi fields. We now state this result without proof. A proof is contained in the proof of [3, Lemma 4].

Proposition 1. *Let p be any point of a manifold M without focal points, and R any positive number. Then $\exists\, T > 0$, $T = T(p, R)$, such that $\|Y(t)\| > R\,\|Y'(0)\|$ for $t > T$, where Y is any nontrivial Jacobi field along any unit speed geodesic ray going out from p and $Y(0) = 0$.*

1. Pairs of geodesics

In this section we will examine the behavior of the functions $d(h(t), k(t))$ and $d(h(t), k)$ where h and k are geodesics on a simply connected manifold M without focal points. Note that all geodesics of M are length minimizing since M is simply connected and has no conjugate points. We begin with the following lemma.

Lemma 1. *Let h and k be two unit speed geodesic rays going out from a point p of a simply connected manifold M without focal points. Suppose that the angle at p between h and k is not equal to 0 or π. Let $c\colon [0,1] \to M$ be the constant speed geodesic segment joining $h(a)$ to $k(a)$ for any fixed $a > 0$, and let $m = \min_{s\in[0,1]} d(p, c(s))$. Then $d(h(t), k(t)) < d(h(a), k(a))$ for $a - m < t < a$.*

Proof. For each $s \in [0, 1]$, there is a unique vector $Z(s) \in T_p(M)$ such that $c(s) = \exp_p Z(s)$. Set $X(s) = Z(s)/\|Z(s)\|$, and consider the variation $r\colon [0, \infty) \times [0, 1] \to M$ defined by $r(u, v) = \exp uX(v)$. The vector fields $r_u = r_*\dfrac{\partial}{\partial u}$ and $r_v = r_*\dfrac{\partial}{\partial v}$ are orthogonal because for each fixed v, r_u is the unit tangent vector to a geodesic ray going out from p. Further for each v, $\|r_v\|^2$ is a strictly increasing function of u because r_v is a nontrivial Jacobi field vanishing at $u = 0$. Now let $\varphi(s) = d(p, c(s))$. Then $c(s) = r(\varphi(s), s)$, and $\dot c(s) = \varphi'(s) r_u(\varphi(s), s) + r_v(\varphi(s), s)$. For $0 < l < m$, $g(s) = r(\varphi(s) - l, s)$ is a curve joining $h(a - l)$ to $k(a - l)$, and $\dot g(s) = \varphi'(s) r_u(\varphi(s) - l, s) + r_v(\varphi(s) - l, s)$. Therefore

$$
\begin{aligned}
L(g) &= \int_0^1 [(\varphi'(s))^2 + \|r_v(\varphi(s) - l, s)\|^2]^{1/2} ds \\
&< \int_0^1 [(\varphi'(s))^2 + \|r_v(\varphi(s), s)\|^2]^{1/2} ds = d(h(a), k(a)) \,.
\end{aligned}
$$

Consequently $d(h(a - l), k(a - l)) < d(h(a), k(a))$, and the lemma follows.

Proposition 2. *Let M be simply connected, and let h, k be two distinct unit speed geodesic rays going out from a point p of M. Then for $t > 0$, both $d(h(t), k)$ and $d(h(t), k(t))$ are strictly increasing and tend to infinity as $t \to \infty$.*

Proof. The case where $\dot h(0) = -\dot k(0)$ is trivial, so we assume that the angle at p between h and k is not π. Now $d(h(t), k) = L(g_t)$ where g_t is the unique geodesic segment which joins $h(t)$ to k and is perpendicular to k. An easy first variation argument gives $L'(g_t) = 0$ if and only if g_t is also perpendicular to h. Since $\exp\colon g_t^\perp \to M$ is a diffeomorphism, this cannot happen. Hence L' never vanishes and, since L' cannot be negative for small values of t, it follows that $L'(g_t) > 0$ for all $t > 0$. Hence $d(h(t), k)$ is strictly increasing. Now let $0 < t_1 < t_2$. By Lemma 1, there exists a nontrivial subinterval $[a, t_2)$ of $[t_1, t_2)$ such that $d(h(u), k(u)) < d(h(t_2), k(t_2))$ for $a < u < t_2$. Let $[b, t_2)$ be the largest subinterval of $[t_1, t_2)$ with the property. To show that $d(h(t_1), k(t_1)) < d(h(t_2), k(t_2))$ it clearly suffices to prove that $d(h(b), k(b)) < d(h(t_2), k(t_2))$. If $g_t\colon [0, 1] \to M$ is the constant speed geodesic segment joining $h(t)$ to $k(t)$, then $d(p, g_t(s))$ depends continuously on both t and s and vanishes only when $t = 0$. Hence we can find an $\varepsilon > 0$ such that $\min_{s \in [0,1]} d(p, g_{b+\varepsilon}(s)) > \varepsilon$. But then it follows from Lemma 1 that $L(g_b) < L(g_{b+\varepsilon})$, and since $L(g_{b+\varepsilon}) < d(h(t_2), k(t_2))$ by choice of ε, we must have $d(h(b), k(b)) < d(h(t_2), k(t_2))$. Therefore $d(h(t_1), k(t_1)) < d(h(t_2), k(t_2))$, and so $d(h(t), k(t))$ is strictly increasing. It remains to show that $d(h(t), k)$ and $d(h(t), k(t))$ both tend to infinity as $t \to \infty$. It is sufficient to show that $d(h(t), k(t)) \to \infty$ since it is then immediate, via the triangle inequality, that $d(h(t), k) \to \infty$. We now suppose that $\exists\, \beta$ such that $d(h(t), k(t)) < \beta$, $\forall\, t > 0$. For any real number R, it follows from Proposition 1 that $\exists\, T > 0$ such that $\|Y(t)\| > R\,\|Y'(0)\|$ for $t > T$ where Y is any nontrivial Jacobi field along any geodesic ray going out from p and $Y(0) = 0$. Choose $a > T + \beta$, let $g_a\colon [0, 1] \to M$ be the constant speed geodesic segment which joins $h(a)$ to $k(a)$, and consider the variation r defined during the proof of Lemma 1. Then for each $s \in [0, 1]$, $\varphi(s) = d(p, g_a(s)) > T$. Moreover, it follows from the definition of r that $Y_s(u) = r_v(u, s)$ is a Jacobi field satisfying $Y_s(0) = 0$, $\|Y_s'(0)\| = \|\dot{X}(s)\|$ where $\dot{X}(s) = \dfrac{dX}{ds}$. Therefore

$$L(g_a) \geq \int_0^1 \|r_v(\varphi(s), s)\|\, ds > R \int_0^1 \|\dot{X}(s)\|\, ds \geq R\theta\,,$$

where $0 < \theta < \pi$, and $\cos\theta = \langle \dot{h}(0), \dot{k}(0)\rangle$. The last inequality holds, since $X(s)$ is a curve in the unit sphere of $T_p(M)$ which joins $\dot{h}(0)$ to $\dot{k}(0)$. Consequently, if we choose $R > \beta/\theta$, we get $L(g_a) > \beta$. This contradiction shows that $\lim_{t\to\infty} d(h(t), k(t)) = \infty$. The proof of the proposition is complete.

Definition. Let M be simply connected and without focal points. Two unit speed geodesic rays $h, k\colon [0, \infty) \to M$ are said to be *asymptotic* if $\exists\, \alpha$ such that $d(h(t), k(t)) \leq \alpha$ for all t.

Proposition 3. *Let $p \in M$ and let k be any unit speed geodesic ray. Then there is a unique geodesic ray going out from p, which is asymptotic to k.*

Proof. (i) *Existence.* For $p \in k$, the subray of k which begins at p is clearly asymptotic to k. For $p \notin k$, let $\{p_n\}$ be a sequence of points whose limit is p, and let $\{t_n\}$ be a positive sequence which tends to infinity. Since M is simply connected, there is a unique unit speed geodesic ray $h_n\colon [0, \infty) \to M$ beginning at p_n and passing through $k(t_n)$. For each n, $\dot{h}_n(0)$ is a unit tangent vector at p_n so, by passing to a subsequence if necessary, we may assume that $\dot{h}_n(0)$ converges to a unit vector Z at p. Let h be the geodesic ray defined by $h(t) = \exp_p tZ$ for $t \geq 0$. We will show that h is asymptotic to k. Let c_n (resp. c) be the geodesic segment joining p_n (resp. p) to $k(0)$. For each n we can define a geodesic variation r_n by considering geodesic rays which go out from

$k(t_n)$ and intersect c_n. Fix $t > 0$. Then it follows from Lemma 1 that for all sufficiently large n we can fined a curve joining $h_n(t)$ to $k(t)$, whose length is less than the length of c_n. Therefore for such n, $d(h_n(t), k(t)) < d(p_n, k(0))$, and taking limits as $n \to \infty$ we obtain $d(h(t), k(t)) \leq d(h(0), k(0))$. Hence h is asymptotic to k.

(ii) *Uniqueness*. The uniqueness of h is immediate from Proposition 2.

Proposition 4. *Let h_0, k_0 be two asymtotic geodesic rays. Extend h_0, k_0 to maximal geodesics $h, k\colon (-\infty, \infty) \to M$ respectively. Then both $d(h(t), k(t))$ and $d(h(t), k)$ are nonincreasing functions of t.*

Proof. For $a \in \boldsymbol{R}$ set $h_a(t) = h(a + t)$, $k_a(t) = k(a + t)$. Then the geodesic rays h_a and k_a are asymptotic, and so by Proposition 3, $d(h_a(t), k_a(t)) \leq d(h_a(0), k_a(0))$ for $t \geq 0$. Setting $t = b - a$ where $b \geq a$, we obtain $d(h(b), k(b)) \leq d(h(a), k(a))$, and therefore $d(h(t), k(t))$ is a nonincreasing function. Now if $a \in \boldsymbol{R}$, we can certainly reparametrize k so that the perpendicular geodesic from $h(a)$ to k intersects k at the point $k(a)$. Hence, if $b \geq a$, we have $d(h(b), k) \leq d(h(b), k(b)) \leq d(h(a), k(a)) = d(h(a), k)$, and so $d(h(t), k)$ is nonincreasing. This completes the proof of the proposition.

If h and k are two distinct maximal geodesics on a simply connected manifold without focal points, then either

1. h and k intersect, or
2. $\exists$ a subray of h which is asymptotic to a subray of k (if this is the case, we can parametrize h and k by arc length so that $d(h(t), k(t))$ is a nonincreasing function of t, and also use the term asymptotic to describe two maximal geodesics which can be so parametrized), or
3. h and k neither intersect nor are asymptotic.

For any pair h, k of distinct maximal geodesics on a simply connected manifold of nonpositive sectional curvature, it follows from the second variation formula that each of the functions $d^2(h(t), k)$ and $d^2(h(t), k(t))$ is convex. If the geodesics neither intersect nor are asymptotic, it follows that each of these functions assumes its absolute minimum at $t = t_0$ say, is nondecreasing for $t > t_0$, nonincreasing for $t < t_0$ and tends to infinity as $t \to \pm\infty$. However, there is no analogue to Propositions 2 and 4 for geodesics of this type on an arbitrary simply connected manifold without focal points. In fact, $d(h(t), k)$ may have a local maximum. To see this let M be a simply connected surface without focal points whose Gaussian curvature K is positive at some point p. Let U be a geodesically convex neighborhood of p where the curvature is everywhere positive. Let $c\colon [0, 1] \to M$ be a constant speed geodesic segment which lies entirely in U. Let h, k be maximal unit speed geodesics such that $h(0) = c(0)$, $k(0) = c(1)$ and such that h and k are both perpendicular to c. Let $c_t\colon [0, 1] \to M$ be the constant speed geodesic segment which joins $h(t)$ to k and is perpendicular to k. A first variation argument shows that $\dfrac{d}{dt}(L(c_t))$ is equal to $\cos(\pi - \psi(t))$ for $t < 0$ and equal to $\cos\psi(t)$ for $t > 0$ where $\psi(t)$ is the interior angle at $h(t)$ of the geodesic quadrilateral bounded by c_t, c, h and k. For t in some interval $(-\varepsilon, \varepsilon)$ this quadrilateral lies completely in U, and it follows from the Gauss-Bonnet formula that for $0 < |t| < \varepsilon$, $\psi(t)$ is an obtuse angle and consequently $d(h(t), k)$ has a local maximum at $t = 0$. A similar argument shows that $d(h(t), k(t))$ can also have a local maximum. It is however true that both $d(h(t), k)$ and $d(h(t), k(t))$ tend to infinity as $t \to \pm\infty$. To see this, suppose there is a sequence $t_n \to +\infty$ (or $-\infty$) such that $d(h(t_n), k(t_n))$ remains bounded. Then using the variations r_n obtained by considering the geodesic rays going out from $h(0)$, which intersect the geodesic segment joining $h(t_n)$ to $k(t_n)$, we could construct, along some geodesic ray going out from $h(0)$, a nontrivial bounded Jacobi field which vanishes initially, and this would contradict Proposition 1. Therefore $d(h(t), k(t)) \to \infty$ as $t \to \pm\infty$, and we can now use the triangle inequality to obtain the same result for $d(h(t), k)$.

2. The flat strip theorem

Again let M be simply connected and without focal points. Let h, k be two distinct maximal unit speed geodesics of M such that $d(h(t), k(t))$ is bounded for all $t \in \boldsymbol{R}$. By Proposition 4 we can parametrize h and k by arc length so that both of the functions $d(h(t), k)$ and $d(h(t), k(t))$ are constant, and if we choose our parametrization in such a manner that the geodesic γ, which joins $h(0)$ to k and is perpendicular to k, intersects k at $k(0)$, then both of these functions are equal to the same constant, b say. It thus follows easily from the first variation formula that γ_t, the geodesic which joins $h(t)$ to k and is perpendicular to k, is also perpendicular to h. Also γ_t must intersect k at $k(t)$ for otherwise, since $d(h(t), k) = d(h(t), k(t)) = b$, there would be two perpendiculars from $h(t)$ to k in contradiction of the fact that M has no focal points. Now let $X(t)$ be the unique unit vector in the normal bundle $h^{\perp}$ of h, such that $k(t) = \exp bX(t)$. Define $r: \boldsymbol{R} \times [0, b] \to M$ by $r(t, u) = \exp uX(t)$. Since $\exp: h^{\perp} \to M$ is a diffeomorphism, r defines an imbedding of the rectangular strip $\boldsymbol{R} \times [0, b] \subset \boldsymbol{R}^2$ into M. Let Q denote the resulting strip of surface in M. The purpose of this section is to prove the following theorem.

Theorem 1. *Let h and k be two maximal unit speed geodesics of a simply connected manifold M without focal points. Suppose that $d(h(t), k(t))$ is bounded $\forall\, t \in \boldsymbol{R}$. Then, with respect to the metric induced from M, the strip of surface Q defined above is flat and totally geodesic.*

Before proceeding with the proof of the theorem we first prove the following lemma.

Lemma 2. *Let N be any simply connected manifold without conjugate points. Let p and q be two points of N with $d(p, q) = \alpha > 0$. Let $c_n: [0, \alpha] \to N$ be a sequence of smooth curves satisfying*

(i) *$c_n(0) = p_n$ and $c_n(\alpha) = q_n$ where $p_n \to p$ and $q_n \to q$ as $n \to \infty$,*

(ii) *$\|\dot{c}_n(t)\| \leq 1$ for all n and for all t,*

(iii) $\{c_n(t)\}$ *converges for each* t.
Then $c_n(t) \to \gamma(t)$ *for every* $t \in [0, \alpha]$ *where* γ *is the unique unit speed geodesic segment such that* $\gamma(0) = p$, $\gamma(\alpha) = q$.

Proof. Let $q(t) = \lim_{t\to\infty} c_n(t)$. Then

$$d(p, q(t)) = \lim_{t\to\infty} d(p_n, c_n(t)) \leq \limsup_{n\to\infty} L(c_n)|_{[0,t]} = \limsup_{n\to\infty} \int_0^t \| \dot{c}_n(s) \| \, ds \leq t$$

by Assumption (ii). Similarly $d(q(t), q) = \lim_{n\to\infty} d(c_n(t), q_n) \leq \alpha - t$. But now we have $\alpha = d(p, q) \leq d(p, q(t)) + d(q(t), q) \leq \alpha$. It therefore follows that $q(t)$ lies on the geodesic segment joining p to q, and $d(p, q(t)) = t$. Hence $q(t) = \gamma(t)$, and the lemma is proved.

Proof of Theorem 1. All notation is as defined at the beginning of this section. The curves $u = 0$ and $u = b$ in Q, i.e., h and k are geodesics of M. We now show that all of the curves $u =$ constant are geodesics of M. Let t_n be a positive sequence which tends to infinity. We can define a variation r_n with $k(t_n)$ as vertex as follows: For each $s \in [0, b]$ there is a unique unit vector $Z(s)$ tangent to M at $k(t_n)$ such that the geodesic ray of M determined by $Z(s)$ passes through $\gamma(s)$ where γ is the unit speed geodesic segment starting at $h(0)$ and ending at $k(0)$. Let $r_n\colon [0, \infty) \times [0, b] \to M$ be given by $r_n(t, s) = \exp_{k(t_n)} tZ(s)$. Let $\varphi_n(s) = d(k(t_n), \gamma(s))$, and let $a > 0$. Then for all sufficiently large n the curve $c_n\colon [0, b] \to M$, where $c_n(s) = r_n(\varphi_n(s) - a, s)$ is well defined and smooth. Further,

(i) $c_n(0) = p_n$, where $\lim_{n\to\infty} p_n = h(a)$ and $c_n(b) = k(a)$,

(ii) by an argument used in the proof of Lemma 1, $\| \dot{c}_n(s) \| \leq \| \dot{\gamma}(s) \| = 1$,

(iii) for each fixed $s \in [0, b]$, $r_n(t, s)$, $0 \leq t \leq \varphi_n(s)$, is the unit speed geodesic segment joining $k(t_n)$ to $\gamma(s)$. Therefore $\lim_{n\to\infty} c_n(s) = g_s(a)$ where g_s is the unit speed geodesic ray such that $g_s(0) = \gamma(s)$ and g_s is asymptotic to $k|_{[0,\infty)}$. Thus the curves c_n satisfy the hypotheses of Lemma 2, and it therefore follows that $\lim_{n\to\infty} c_n(s) = \gamma_a(s)$ for all $s \in [0, b]$ where γ_a is the unit speed geodesic segment starting at $h(a)$ and ending at $k(a)$. Combining this with (iii) above, we see that for $t \geq 0$ each curve $u = u_0$ in Q is a geodesic ray of M, namely, the unit speed geodesic ray which starts at $r(0, u_0)$ and is asymptotic to $k|_{[0,\infty)}$. A similar argument gives the same result for $t \leq 0$, and so it follows that for each u_0 the curve $u = u_0$ in Q is a maximal unit speed geodesic of M. Now the curves $t =$ const. and $u =$ const. in Q are two mutually orthogonal systems. Also, they are geodesics of M and therefore of Q, since Q bears the metric induced from M. Hence it follows from the Gauss-Bonnet formula that Q is flat. Again since the curves $u =$ const. are geodesics of M, the unit vector field $Y(t) = r_* \frac{\partial}{\partial u}(t, u)$ is a Jacobi field of M for each u. But Eberlein has shown (see [2, Corollary 3.3]) that a bounded perpendicular Jacobi field on M must be parallel. Hence $\nabla_{r_*\partial/\partial t} Y = 0$ where ∇ is the riemannian connection on M. Therefore $\nabla_{r_*\partial/\partial t} r_* \dfrac{\partial}{\partial u} = 0$ and this combined with the fact that the curves $t =$ const. and $u =$ const. are geodesics of M implies that Q is totally geodesic. This completes the proof of the theorem.

Remark. This result was proved for surfaces without focal points by Green. See [4, Theorem 4.1].

3. Isometries and the fundamental group

Let M be any riemannian manifold and let f be an isometry of M. We say that f translates the maximal geodesic h if $f(h(t)) = h(t + \alpha)$ for some $\alpha \neq 0$ and for all $t \in \boldsymbol{R}$. If h is a length minimizing geodesic, we will call h an *axis* of f. The isometry f is said to be of *small displacement* if it carries each point into the complement of its cut locus. Then for all $p \in M$, the geodesic segment c joining p to $f(p)$ such that $L(c) = d(p, f(p))$ is unique. We can lift f to an isometry $\tilde{f}$ of the universal covering space $\tilde{M}$ as follows. Let π be the covering projection, let $x \in \pi^{-1}(p)$ and let $\tilde{c}$ be the lift of c whose initial point is x. Then $\tilde{f}(x)$ is defined to be the final point of $\tilde{c}$. It is easily checked that $\tilde{f}$ is an isometry of small displacement, and that $d(x, \tilde{f}(x)) = d(\pi x, f(\pi x))$ for all $x \in \tilde{M}$. For any isometry f of small displacement, the function $\delta_f \colon M \to \boldsymbol{R}$, where $\delta_f(p) = d(p, f(p))$, is smooth except at fixed points of f, and δ_f^2 is smooth everywhere. Further, p is a critical point of δ_f^2 if and only if p is a fixed point of f, or f translates the unique maximal geodesic, which passes through p and $f(p)$, and is length minimizing between these two points. See [9, Chapter 1].

We will say that a closed subset F of a manifold M is *geodesically convex* if for each $p, q \in F$, every geodesic segment joining p to q lies completely in F. It is known that such subsets are totally geodesic submanifolds (perhaps with boundary) of M. See [9, Lemma 1.3.2] or [1, Theorem 1.6].

Proposition 5. *Let M be a manifold without focal points, and f an isometry of small displacement. Denote by* Crit (f) *the set of critical points of δ_f^2. Then*

(i) Crit (f) *is the set of points where δ_f assumes its absolute minimum, and is a closed connected totally geodesic submanifold of M possibly with boundary,*

(ii) Crit (f) *is isometric to $N \times \boldsymbol{R}$ if M is simply connected and f has no fixed points, where N is a closed connected totally geodesic submanifold of M possibly with boundary.*

Proof. Clearly Fix (f), the set of fixed points of f, is contained in Crit (f). Let p be a fixed point of f, and suppose that $\exists\, q \in$ Crit (f) $-$ Fix (f). Let $\tilde{f}$ be the lift of f to $\tilde{M}$, and let $x \in \pi^{-1}(p)$, $y \in \pi^{-1}(q)$. Then x is a fixed point of $\tilde{f}$, and the unique maximal geodesic k, which passes through y and $\tilde{f}(y)$, is an axis of $\tilde{f}$. Let h be the geodesic which passes through x and intersects k at right angles. Then $\tilde{f} \circ h$ also passes through x, intersects k at right angles, and $\tilde{f} \circ h \neq h$ since k is an axis of $\tilde{f}$. This is impossible since $\tilde{M}$ has no focal points, and so we can conclude that either

(a) Crit $(f) =$ Fix (f), or

(b) f has no fixed points, and Crit (f) consists of all points q such that f translates the unique maximal geodesic, which passes through q and $f(q)$ and is length-minimizing between these two points.

For Case (a) it is immediate that Crit (f) is the set of points where δ_f assumes its absolute minimum. Further, if p and q are fixed points of f, and c is any geodesic segment joining p to q, then we can lift c to a geodesic segment $\tilde{c}$ in $\tilde{M}$ such that both end points of $\tilde{c}$ are fixed by $\tilde{f}$. Since $\tilde{M}$ is simply connected without focal points, it follows that $\tilde{f}$ fixes every point of $\tilde{c}$, and so by projecting back to M we conclude that f fixes every point of c. It follows that Crit (f) is a closed connected geodesically convex subset of M, and this establishes the proposition when f has fixed points.

Now suppose that f has no fixed points and that $p \in$ Crit (f). Then there is a unique f-invariant geodesic h which passes through p and $f(p)$ and is length minimizing between these two points. Lift h to a geodesic $\tilde{h}$ of $\tilde{M}$. Then $\tilde{h}$ is an axis of $\tilde{f}$, and it follows from [8, Lemma 1] that for $x \in \tilde{M}$, $d(x, \tilde{f}x)$ assumes its absolute minimum on $\tilde{h}$. Consequently, since $d(x, \tilde{f}(x)) = d(\pi(x), f \circ \pi(x))$, δ_f must assume its absolute minimum at p. Now let $q \in$ Crit (f), $q \neq p$, and let c be any geodesic segment joining p to q. If c is a segment of the geodesic h, then it is immediate that δ_f is constant on c, and so $c \subset$ Crit (f). If c is not a segment of h, let $\tilde{c}$ be a lift of c to $\tilde{M}$, and let $\tilde{g}$ (resp. $\tilde{k}$) be the axis of $\tilde{f}$ which passes through the initial point (resp. final point) of $\tilde{c}$. Then $d(\tilde{g}(t), \tilde{k}(t))$ is bounded for all t, and so by Theorem 1, $\tilde{g}$ and $\tilde{k}$ form the boundary of a flat totally geodesic rectangular strip of surface and this strip is clearly $\tilde{f}$-invariant. Therefore $d(x, \tilde{f}(x))$ must be constant for $x \in \tilde{c}$, and projecting back to M it follows that δ_f is constant on c. This shows that when f has no fixed points, Crit (f) is also a closed connected geodesically convex subset of M, and completes the proof of Part (i) of the Proposition.

If f has no critical points, then Part (ii) is vacuously true. If Crit $(f) \neq \emptyset$, define a vector field X on Crit (f) by setting $X(p)$ equal to the unit tangent vector at p to the unique length minimizing geodesic segment joining p to $f(p)$. Since an axis of f passes through each point of Crit (f), it is an easy consequence of Theorem 1 that X is parallel. Therefore, using the de Rham decomposition theorem, we can split Crit (f) isometrically as $N \times \boldsymbol{R}$, where N is a closed connected totally geodesic submanifold of M possibly with boundary and for each $z \in N$, the leaf $\{z\} \times \boldsymbol{R}$ is an axis of f. This completes the proof of the proposition.

Lemma 3. *Let M be a simply connected manifold without focal points, F a closed connected geodesically convex subset of M, and G a finite group of isometries of F. Then there is a point $q \in F$ such that $\varphi(q) = q$ for all $\varphi \in G$.*

Proof. Let q_0 be any point of F, let $G = \{\varphi_1, \varphi_2, \cdots, \varphi_l\}$, and consider the continuous G-invariant function $f\colon F \to \boldsymbol{R}$ where $f(p) = \max\{d(\varphi_i q_0, p)\colon 1 \leq i \leq l\}$. Since M is simply connected, it follows that for each $\alpha > 0$, $\{p \in F\colon f(p) \leq \alpha\}$ lies in a compact set, and therefore f assumes its minimum value, α_0 say. Hence $f^{-1}(\alpha_0)$, the set of points where f assumes it minimum, is nonempty, and we see easily that it is equal to $\bigcap_{i=1}^{l} B(\varphi_i q_0, \alpha_0) \cap F$ where for $x \in M$, $B(x, \alpha_0)$ is the ball of radius α_0 centered at x. Now [3, Lemma 1] states that on any simply connected manifold without focal points the function $y \to d^2(x, y)$ is strictly convex for each x. Consequently the minimal set of f must be a single point. To see this, suppose that q_1 and q_2 were two distinct points in $f^{-1}(\alpha_0)$. Then there would be some $\varepsilon > 0$ such that $c|_{(0,\varepsilon)}$, where $c\colon [0, 1] \to M$ is the constant speed geodesic segment with $c(0) = q_1$ and $c(1) = q_2$, would lie in the interior of each of the balls $B(\varphi_i q_0, \alpha_0)$, and this would contradict the minimality of α_0. Since $f^{-1}(\alpha_0)$ is invariant under the action of G on F, the lemma is established.

Note. This lemma remains valid when G is a compact group of isometries of M. See [10, Theorem 3.4.1].

The fundamental group. We now study solvable subgroups of the fundamental group of a compact riemannian manifold M without focal points, and show that such groups have a strong influence on the geometry of M. Let $\tilde{M}$ be the universal covering space of M, and let $\sigma \in \pi_1(M)$ be a nontrivial deck transformation. Since Proposition 5 is valid for any isometry of $\tilde{M}$ and since σ has no fixed points, it follows that Crit (σ) consists of all points in $\tilde{M}$ through which an axis of σ passes, and this set is not empty since M is compact.

Theorem 2. *Let M be a compact riemannian manifold without focal points, $\tilde{M}$ its universal covering space, and G a nontrivial solvable subgroup of $\pi_1(M)$.*

(i) *Then $\tilde{M}$ contains a nonempty closed connected G-invariant totally geodesic subspace N which is isometric to $\boldsymbol{R}^m \times N^*$, where $\boldsymbol{R}^m$ is m-dimensional euclidean space and N^* may have boundary. Further if $\sigma \in G$, the action of σ on $N = \boldsymbol{R}^m \times N^*$ is given by $\sigma(u, v) = (\sigma u, v)$, $\forall\, u \in \boldsymbol{R}^m$, $v \in N^*$.*

(ii) *Then G is torsion free and is a finite extension of a free abelian normal subgroup of rank m.*

Corollary 1. *Let M be a compact riemannian manifold without focal points. Then there exists a solvable subgroup G of $\pi_1(M)$ if and only if there exists an isometric and totally geodesic immersion into M of a compact euclidean space form Q where $\pi_1(Q) = G$.*

Corollary 2. *Let M be a compact riemannian manifold without focal points. Then $\pi_1(M)$ is solvable if and only if M is flat.*

Proof of Theorem 2. G must be torsion free since otherwise there would exist closed geodesics on $\tilde{M}$. We first prove the theorem for the case where G is abelian. If $\sigma_1 \in G$, then Crit (σ_1) is nontrivial, σ_1-invariant and isometric to $\boldsymbol{R} \times N_1$ where the action of σ_1 on $\boldsymbol{R} \times N_1$ is given by $\sigma_1(u, v) = (u + t_1, v)$

for some constant t_1 and for all $u \in \boldsymbol{R}$, $v \in N_1$. Since G is torsion free and abelian, G is free. Hence, if G has rank 1, we can choose σ_1 to be a generator of G, and the theorem is proved. Otherwise, $\exists\, \sigma_2 \in G$ such that $\{\sigma_1, \sigma_2\}$ is a subgroup of G with rank 2. Since G is abelian, Crit (σ_1) is invariant under the action of σ_2 on $\tilde{M}$, and so σ_2 is an isometry of Crit (σ_1). Let $\alpha = \inf\{d(x, \sigma_2 x) : x \in \text{Crit}\,(\sigma_1)\}$. We claim that $\exists\, z \in \text{Crit}\,(\sigma_1)$ such that $d(z, \sigma_2 z) = \alpha$. To see this let x_n be any sequence in Crit (σ_1) such that $d(x_n, \sigma_2 x_n) \to \alpha$. Let F be a compact fundamental domain in $\tilde{M}$. Then there are sequences $y_n \in F$ and $\tau_n \in \pi_1(M)$ such that $x_n = \tau_n y_n$. Since the y_n all lie in a compact set, we may assume that $y_n \to y$. Since $d(y_n, \tau_n^{-1}\sigma_2\tau_n y_n) = d(x_n, \sigma_2 x_n) \to \alpha$, it follows easily from the triangle inequality that all of the points $\tau_n^{-1}\sigma_2\tau_n y$ lie within a ball of finite radius centered at y. Consequently, since $\pi_1(M)$ acts freely and properly discontinuously on $\tilde{M}$, there is a subsequence of the τ_n such that $\tau^{-1}\sigma_2\tau$ is independent of the choice of τ in this subsequence. By passing to the subsequence, we may assume that $\tau_n^{-1}\sigma_2\tau_n$ is independent of n, and since $y_n \to y$ we get $d(y, \tau_n^{-1}\sigma_2\tau_n y) = \alpha$ for all n. Since an axis of σ_1 passes through each point x_n, $d(x_n, \sigma_1 x_n) = \beta$ for all n, where $\beta = \inf\{d(x, \sigma_1 x) : x \in \tilde{M}\}$. Hence $d(y_n, \tau_n^{-1}\sigma_1\tau_n y_n) = \beta$, and by repeating the above argument we arrive at a subsequence of the τ_n such that $\tau^{-1}\sigma_1\tau$ is independent of the choice of τ in this subsequence. Then for any member τ of the subsequence we have $d(y, \tau^{-1}\sigma_1\tau y) = d(\tau y, \sigma_1\tau y) = \beta$, and therefore $\tau y \in \text{Crit}\,(\sigma_1)$. But $d(\tau y, \sigma_2\tau y) = d(y, \tau^{-1}\sigma_2\tau y) = \alpha$, and so if we put $z = \tau y$, our claim is established. Now Crit (σ_1) is totally geodesic and σ_2-invariant, so the geodesic segments which join z to $\sigma_2 z$ and $\sigma_2 z$ to $\sigma_2^2 z$ both lie completely in Crit (σ_1). They must be part of the same maximal geodesic since otherwise the triangle inequality would give $d(y, \sigma_2 y) < d(z, \sigma_2 z)$ for any point y in the interior of the segment which joins z to $\sigma_2 z$, and this would contradict the minimality of α. Therefore an axis of σ_2 passes through z, and we have shown that Crit $(\sigma_1) \cap$ Crit (σ_2) is not empty. Denote this set by Crit (σ_1, σ_2), and let $[\sigma_1, \sigma_2]$ be the subgroup of G generated by σ_1 and σ_2. Crit (σ_1, σ_2) is obviously invariant by $[\sigma_1, \sigma_2]$, and since the intersection of two geodesically convex sets is again geodesically convex, we can conclude that Crit (σ_1, σ_2) is a closed connected totally geodesic submanifold of $\tilde{M}$ possibly with boundary. For $i = 1, 2$ we can define a vector field V_i on Crit (σ_1, σ_2) by setting $V_i(x)$ equal to the unit tangent vector at x to the unique geodesic segment joining x to $\sigma_i x$. Let x be in the interior of Crit (σ_1, σ_2), X a unit vector perpendicular to $V_1(x)$ and tangent to Crit (σ_1, σ_2), and c a geodesic segment which lies completely in Crit (σ_1, σ_2) with $\dot{c}(0) = X$. Then by Theorem 1 the strip of surface obtained by considering all axes of σ_1 which intersect c is totally geodesic, and therefore by Synge's lemma the vector field V_1 is parallel along c. Thus V_1 (and, by a similar argument, V_2) is a parallel vector field on Crit (σ_1, σ_2). Further, $V_1(x)$ and $V_2(x)$ are linearly independent at each point x of Crit (σ_1, σ_2), since $[\sigma_1, \sigma_2]$ has rank 2. It now follows, via the de Rham decomposition theorem, that Crit (σ_1, σ_2) is isometric to $\boldsymbol{R}^2 \times N_2$ where for each $v \in N_2$ the leaf $\boldsymbol{R}^2 \times \{v\}$ is the flat totally geodesic $[\sigma_1, \sigma_2]$-invariant plane obtained by taking the union of all axes of σ_1 which intersect the axis of σ_2 passing through the point $(0, v)$. It is clear that $[\sigma_1, \sigma_2]$ acts on each leaf by euclidean translations. If G has rank 2, we obtain the theorem by choosing σ_1 and σ_2 so that $G = [\sigma_1, \sigma_2]$. Otherwise, we can continue this argument inductively, and show that if $H = \{\sigma_1, \sigma_2, \cdots, \sigma_k\}$ is a subgroup of G of rank k, then $\bigcap_{i=1}^{k} \operatorname{Crit}(\sigma_i)$ is nonempty and isometric to $\boldsymbol{R}^k \times N_k$, and further that this subspace is invariant under the action of H on $\tilde{M}$, with $\sigma \in H$ mapping $(u, v) \in \boldsymbol{R}^k \times N_k$ into $(u + u_1, v)$ where $u_1 \in \boldsymbol{R}^k$ depends only on σ. Therefore G must be finitely generated with rank $m \leq \dim M$. We now obtain the theorem for the case where G is abelian by choosing $\{\sigma_1, \sigma_2, \cdots, \sigma_m\}$ to be a generating set for G and considering $\bigcap_{i=1}^{m} \operatorname{Crit}(\sigma_i)$.

Nonabelian solvable subgroups. Once the theorem is established for abelian subgroups of the fundamental group of a compact manifold without focal points, the proof of [11, Theorem 1] can be used to show that a solvable subgroup G of $\pi_1(M)$ is a Bieberbach group of rank $m \leq \dim M$. Let A be an abelian normal subgroup of G such that G/A is finite, and let $W = \boldsymbol{R}^m \times W^*$ be the A-invariant subspace of $\tilde{M}$ constructed above. Let τ be a nontrivial element of G, σ a nontrivial element of A, and let $z \in W$. Since A is normal in G, an axis of $\tau^{-1}\sigma\tau$ passes through z, and so an axis of σ passes through τz. Therefore $\tau z \in W$, and W is G-invariant. Let V_τ be the vector field on W defined by setting $V_\tau(z)$ equal to the unit tangent vector at z to the geodesic segment joining z to τz. Now for $\sigma \in A$, $V_\sigma(z)$ is tangent to the euclidean leaf $\boldsymbol{R}^m$ which passes through z, and the set of all such $V_\sigma(z)$ generates the tangent space to this leaf. Since A is normal in G and $\tau_* V_\sigma = V_{\tau\sigma\tau^{-1}}$, the action of G on W preserves the splitting $W = \boldsymbol{R}^m \times W^*$. In other words, for $\tau \in G$, $u \in \boldsymbol{R}^m$ and $v \in W^*$, $\tau(u, v) = (\tau_1 u, \tau_2 v)$, where τ_1 is an isometry of $\boldsymbol{R}^m$ and τ_2 is an isometry of W^*. In particular, G acts by isometries on W^*, and since A acts trivially on this space, we can regard G/A as a finite group of isometries of W^*. Let $N^* = \{v \in W^* : \varphi v = v \forall\, \varphi \in G/A\}$. By Lemma 3, N^* is nonempty, and since $\tilde{M}$ is simply connected, N^* must contain the geodesic segment joining any two of its points. It is now easy to checked that the subspace $N = \boldsymbol{R}^m \times N^*$ of W is a closed connected totally geodesic G-invariant subspace of $\tilde{M}$, and that $\tau(u, v) = (\tau u, v)$ for all $\tau \in G$, $u \in \boldsymbol{R}^m$ and $v \in N^*$. This completes the proof of the theorem.

Proof of Corollary 1. Let $i: Q \to M$ be such an immersion. Since M has no conjugate points, there is only one geodesic segment joining each pair of points in its universal covering space. Therefore there is a unique (not necessarily smoothly) closed geodesic segment in each homotopy class at any point p of M. Consequently, the induced homomorphism $\hat{i}: \pi_1(Q) \to \pi_1(M)$ is injective. Conversely, let G be a solvable subgroup of $\pi_1(M)$, and let $\boldsymbol{R}^m$ be any euclidean leaf of the G-invariant subspace of $\tilde{M}$ constructed in Theorem 2. Then $\pi(R^m)$, where $\pi: \tilde{M} \to M$ is the covering map, is an isometrically and totally geodesically immersed compact euclidean space form with fundamental group G.

Proof of Corollary 2. If $\pi_1(M)$ is solvable, it is a Bieberbach group. Therefore there is a riemannian covering of M by a compact manifold whose fundamental group is abelian, and it follows from [8, Theorem A] that M is flat. The converse is immediate.

Added in proof. A recent result of M. S. Goto (*Manifolds without focal points*, to appear in J. Differential Geometry) shows that on any manifold without focal points the length of a nontrivial initially vanishing Jacobi field is unbounded. Consequently, the results of this paper, except for Theorem 2 and its corollaries, are valid for all manifolds without focal points. The results of the paper have also been obtained independently by J. H. Eschenburg (*Manifolds with bounded asymptotes*, to appear).

References

[1] J. Cheeger & D. Gromoll, *On the structure of complete manifolds of nonnegative* curvature, Ann. of Math. **96** (1972) 413–443.
[2] P. Eberlein, *When is a geodesic flow of Anosov type*? I, J. Differential Geometry **8** (1973) 437–463.
[3] ——, *When is a geodesic flow of Anosov type*? II, J. Differential Geometry **8** (1973) 565–577.
[4] L. W. Green, *Surfaces without conjugate points*, Trans. Amer. Math. Soc. **76** (1954) 529–546.
[5] D. Gromoll & J. Wolf, *Some relations between the metric structure and the algebraic structure of the fundamental group in manifolds of non-positive curvature*, Bull. Amer. Math. Soc. **77** (1971) 545–552.
[6] R. Gulliver, *On the variety of manifolds without conjugate points*, Trans. Amer. Math. Soc. **210** (1975) 185–201.
[7] R. Hermann, *Focal points of closed submanifolds of Riemannian spaces*, Nederl. Akad. Wetensch. Proc. Ser. A, **66** (1963) 613–628.
[8] J. J. O'Sullivan, *Manifolds without conjugate points*, Math. Ann. **210** (1974) 295–311.
[9] V. Ozols, *Critical points of the displacement function of an isometry*, J. Differential Geometry **3** (1969) 411–432.
[10] S. T. Yau, Thesis, Berkeley, 1970.
[11] ——, *On the fundamental group of compact manifolds of non-positive curvature*, Ann. of Math. **93** (1971) 579–585.

INSTITUTE FOR ADVANCED STUDY
PENNSYLVANIA STATE UNIVERSITY