J. DIFFERENTIAL GEOMETRY
9 (1974) 443-450

THE EXISTENCE OF GEODESICS JOINING TWO GIVEN POINTS

WILLIAM B. GORDON

1. Introduction

Let M be a manifold (always assumed to be C^∞, finite dimensional and without boundary), and let $g = (g_{ij})$ be a riemannian metric on M. Since we shall be varying metrics, the following terminology will be convenient.

Definition. An open domain (=connected open subset) D of M will be said to be *g-connected* if every homotopy class of paths in D joining two given points in D contains a (smooth) geodesic segment whose length is minimum *for that class* of paths lying entirely in D. (This geodesic need not furnish a minimum arc length for the corresponding homotopy class of paths in M. See § 5C for an example.)

We shall also have occasion to speak of the *g-completeness* of M, meaning that M is complete in the riemannian sense with respect to g. A standard result in riemannian geometry asserts that the g-completeness of a manifold implies its g-connectivity.

The purpose of this paper is to present some criteria for the g-connectivity of domains. The search for such criteria is largely motivated by the following problem. Let V be a smooth "potential" on M, and consider the conservative dynamical system

$$(*) \qquad \frac{D^2x}{dt^2} + \nabla V = 0 \, .$$

It is well-known that the trajectories to (*) with total energy h are re-parametrized geodesics with respect to the *Jacobi metric* $\tilde{g}_{ij} = (h - V)g_{ij}$. (See e.g. [6] for a rigorous account of this theorem.) Hence the $\tilde{g}$-connectivity of a domain D implies that every pair of points in D can be joined by a trajectory of (*) with total energy h. But the components of $V^{-1}(-\infty, h)$ are not likely to be $\tilde{g}$-complete, and therefore the standard results of riemannian geometry only guarantee the $\tilde{g}$-connectivity of "small" domains (such as normal balls). We would like to be able to construct $\tilde{g}$-connected domains which are reasonably large and physically meaningful.

Finally, we mention that since the geodesics (or trajectories) whose existence

Communicated by R. S. Palais, February 20, 1973.

is asserted in the theorems below are obtained as paths at which certain "energy" or "action" integrals are minimized, one might reasonably expect that such trajectories are machine computable by the use of direct methods in the calculus of variations. (Cf. [4], [7].)

2. Statement of results

2.1. Definitions. For ease of exposition all maps and functions will be assumed to be of class C^∞. Recall that a map f between manifolds is said to be *proper* iff $f^{-1}(K)$ is compact whenever K is compact. Hence a (real-valued) function f is proper iff $f^{-1}[a, b]$ is compact for every closed interval $[a, b]$. If D is an open domain with compact closure, then a function f on D is proper iff $|f(p)| \to \infty$ as $p \to \partial D$. For domains whose closure is not compact, this condition for propriety is necessary but not sufficient.

A function f defined on an open domain D is said to be *convex* (*resp. strictly convex*) on D iff the second covariant differential (Hessian) $\nabla^2 f(p)$ is positive semidefinite (resp. positive definite) at every point p of D. (We are of course assuming the existence of a fixed given metric g. The definition of convexity in terms of local coordinate representations is given in § 3–A.) Equivalently, f is convex (resp. strictly convex) on D iff $f''(x(t)) \geq 0$ (resp. > 0) for every geodesic arc $x = x(t)$ in D.

Note that an open domain D which supports a strictly convex function cannot contain a nonconstant periodic geodsic. Also, if N is a compact subset of such a domain, then any nonconstant geodesic which enters N must eventually leave N, (but may later return to N). The existence of a convex function on a manifold has implications for the structure of the manifold; see [2], and also [1] and [3] for further applications and examples of this and related notions of convexity.

2.2. Theorems and remarks. We are now in a position to state our theorems. Again, it will always be assumed that M is a smooth riemannian manifold endowed with a given metric g. It will *not* be assumed that M is g-complete.

Theorem 1. *An open domain D of M is g-connected if it supports a proper positive convex function.*

Theorem 2. *An open domain D of M is g-connected if it supports a positive convex function f such that*

(i) *$f(p) \to \infty$ as $p \to \partial D$,*

(ii) *for every real number c and closed bounded set B in D, $B \cap f^{-1}[0, c]$ is compact.*

Remarks. (1) None of the domains D described in the theorems need to have compact closure. Also, the domains D need not be homeomorphic to euclidean space. For example, an annular region in R^2 can support a proper positive function which is convex with respect to certain metrics. (Example:

The hyperboloid of revolution $x^2 + y^2 = 1 + z^2$, is diffeomorphic to an open annulus. If we give the hyperboloid the metric induced by its embedding in $\boldsymbol{R}^3$, it turns out that $f(p) =$ (squared distance from p to the axis of rotation) is proper, positive and convex. Cf. § 5 for other examples.)

(2) If the domain D in Theorem 2 has compact closure, then condition (i) implies that f is proper; i.e., Theorem 1 applies so that condition (ii) becomes redundent. Condition (ii) also becomes redundant if M is g-complete, for in this case closed and bounded sets are necessarily compact.

(3) It is easy to show that every point p has a neighborhood D satisfying the hypothesis of Theorem 1. (In a coordinate patch centered at p, take $f(x) = -\log(c^2 - |x|^2)$. Then for sufficiently small c, f is proper, positive and convex on the domain $|x| < c$.)

2.3. Illustration. The proof of Theorem 2 is a trivial modification of the proof of Theorem 1, and a large part of the geometrical content of the hypotheses of Theorem 1 is provided by the following illustration.

A simple example of a domain which is g-connected but not g-complete is the open unit disk in $\boldsymbol{R}^2$, where g is the standard euclidean metric. Suppose we remove a pie shaped piece from the disk, as indicated in Fig. 1, and thus destroy its g-connectivity. Let f be a proper function on this domain which assumes the value $+\infty$ at the boundary. Consider the geodesic (straight line) $x = x(t)$ running from A to E, as shown in Fig. 1.

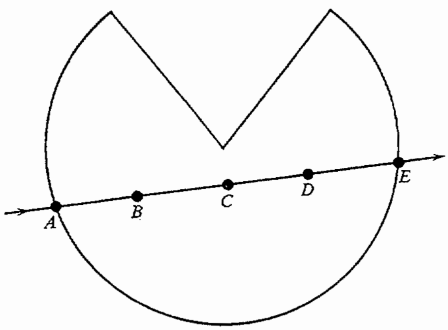

Fig. 1

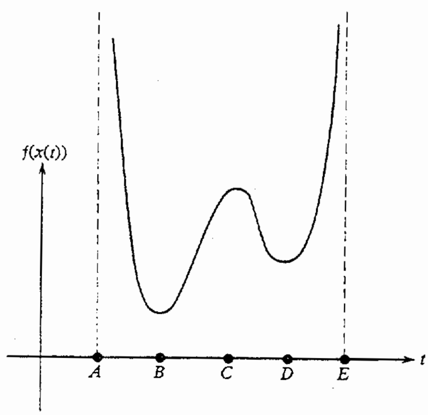

Fig. 2

The graph of $f(x(t))$ is shown in Fig. 2. We see that $f(x(t))$ assumes the value $+\infty$ at A, decreases to a local minimum at B, increases again and assumes a rather large value at a point C near the boundary, etc. Obviously, a function with such a graph cannot be convex.

3. Preliminaries to the proof

3.1. Geometric preliminaries. Our proofs will use the following construction.

Lemma 1. *Let M be a (not necessarily complete) riemannian manifold with riemann metric $g = (g_{ij})$, and f any proper function on M. Then M is necessarily complete with respect to the metric $\tilde{g} = (\tilde{g}_{ij})$ where*

$$\tilde{g}_{ij} = g_{ij} + f_i f_j \,, \qquad (f_i = \partial f / \partial x^i) \,. \tag{1}$$

A proof is given in [5], where the proposition is used to prove that a riemannian manifold is complete iff it supports a proper function whose gradient is bounded in modulus. (Note that $\tilde{g} = g + df \otimes df$ is the metric which g induces on the graph of f; i.e., the proposition states that the graph of a proper function is complete with respect to the graph metric.)

We shall employ the usual conventions of tensor calculus. In particular we use the summation convention, and the inverse matrix to (g_{ij}) will be denoted by (g^{ij}). The following identities can be verified by straightforward calculations:

$$\tilde{g}^{ij} = g^{ij} - (1 + |\nabla f|^2)^{-1} f^i f^j \,, \tag{2}$$

$$\tilde{\Gamma}^i_{jk} = \Gamma^i_{jk} + (1 + |\nabla f|^2)^{-1} f_{jk} f^i \,, \tag{3}$$

where $f^i = g^{ir} f_r$, $|\nabla f|^2 = g^{ij} f_i f_j = f^i f_i$, Γ^i_{jk} and $\tilde{\Gamma}^i_{jk}$ are the Christoffel symbols associated with g and $\tilde{g}$ respectively, and

$$f_{jk} = \frac{\partial^2 f}{\partial x^j \partial x^k} - \Gamma^i_{jk} f_i \,.$$

Note that (f_{jk}) are the coefficients of $\nabla^2 f$, so that f is convex (resp. strictly convex) iff the eigenvalues of the matrix (f_{jk}) are all nonnegative (resp. positive).

3.2. Function-analytic preliminaries. The following facts about Sobolev spaces are well-known; in particular we refer to [8], [9]. For any smooth map $x = x(t)$ from $[0, 1]$ to $\boldsymbol{R}^N$ we define the norm $\|\cdot\|_1$ by $\|x\|_1^2 = \int_0^1 \{|x(t)|^2 + |\dot{x}(t)|^2\} dt$. Let H^1 denote the hilbert space obtained by completing $C^\infty([0, 1], \boldsymbol{R}^N)$ with respect to $\|\cdot\|_1$. The weak H^1-topology is stronger than the uniform topology, i.e., weak H^1 convergence implies C^0 convergence. (In fact, H^1 consists of absolutely continuous maps with L^2 derivatives.) Let the manifold M be isometrically embedded in $\boldsymbol{R}^N$ (always possible by a theorem of Nash), and for any pair of points p, q on M let $\Omega_{p,q}$ denote the set of all paths $x = x(t)$ in H^1 which lie on M and for which $x(0) = p, x(1) = q$. Define the "energy" functional E on $\Omega_{p,q}$ by

$$E(x) = \int_0^1 |\dot{x}|^2 dt = \int_0^1 g_{ij}(x(t))\dot{x}^i(t)\dot{x}^j(t)dt \ .$$

Then using some well-known generalities about hilbert spaces one obtains the following proposition.

Lemma 2. *Let $\{x_n\}$ be a sequence of paths in $\Omega_{p,q}$ such that $E(x_n) \leq$ constant, and suppose also that all the x_n lie in some compact subset K of M. Then there exist a path $x = x(t)$ belonging to $\Omega_{p,q}$ and lying in K, and a subsequence $\{x'_n\}$ of $\{x_n\}$ such that*

(i) $x'_n \xrightarrow{weak,\,H^1} x$ *(and hence* $x'_n \xrightarrow{C^0} x$*),*

(ii) $E(x) \leq \overline{\lim}\ \{E(x_n)\}$ $(=\limsup \{E(x_n)\})$.

Remark. A sequence $\{x_n\}$ in $\Omega_{p,q}$ on which E is bounded always contains a subsequence which converges to a path in $\boldsymbol{R}^N$. The requirement that the $\{x_n\}$ lie in some compact K is necessary because of the possible lack of completeness of M; i.e., if M is not a closed submanifold of $\boldsymbol{R}^N$, then one cannot conclude that a subsequence in $\Omega_{p,q}$ which is bounded in H^1-norm contains a subsequence which converges to a path on M.

4. Proof of Theorem 1

(i) Let $\Omega_{p,q}(D)$ denote the space of all curves $x = x(t)$ which belong to $\Omega_{p,q}$ and lie in D. We shall construct a curve $x_\infty \in \Omega_{p,q}(D)$ at which $E|\Omega_{p,q}(D)$ attains a minimum value. It is well-known that such a curve is a geodesic, and that its arc length is also minimum for all curves belonging to $\Omega_{p,q}(D)$.

Let f be a proper positive convex function on D, and for each positive integer n let $g^{(n)} = (g^{(n)}_{ij})$ where

$$g^{(n)}_{ij} = g_{ij} + \frac{1}{n} f_i f_j \ .$$

Let $E^{(n)}$ be the "energy" corresponding to $g^{(n)}$; i.e.,

$$E^{(n)}(x) = \int_0^1 \left\{ g_{ij}\dot{x}^i\dot{x}^j + \frac{1}{n} f_i f_j \dot{x}^i \dot{x}^j \right\} dt = E(x) + \frac{1}{n}\int_0^1 \langle \dot{x}, \nabla f\rangle^2 dt \ .$$

Now according to Lemma 1 the domain D (considered as a manifold) is $g^{(n)}$-complete, and therefore $g^{(n)}$-connected. Therefore for every n there exists a curve $x^{(n)} = x^n(t)$ in $\Omega_{p,q}(D)$ which minimizes $E^{(n)}$. (In fact, there exists such a curve for each homotopy class of paths joining p to q in D. In the sequel we shall choose each of the $x^{(n)}$ to belong to the same fixed homotopy class.)

(ii) Without loss of generality we can assume $f(p) \leq f(q)$. Let $K = f^{-1}[0, f(q)]$. Then the propriety of f implies that K is compact.

(iii) Let $(f_{ij}^{(n)})$ represent the coefficients of the Hessian of f with respect to the metric $g^{(n)}$. Replacing f by $f/\sqrt{n}$ in formulas (1), (2) and (3), one easily obtains the relation

$$f_{ij}^{(n)} = \frac{n}{n + |\nabla f|^2} f_{ij} .$$

Hence f is convex with respect to each of the metrics $g^{(n)}$. But each of the curves $x^{(n)}$ is a geodesic with respect to $g^{(n)}$, so that the real valued function $t \to f(x^{(n)}(t))$ is convex. Therefore, for each $n, 0 \leq f(x^{(n)}(t)) \leq f(q), (0 \leq t \leq 1)$. That is, each $x^{(n)}$ lies in the compact set K.

(iv) From the definitions and the minimizing properties of the curves $x^{(n)}$ we get

$$(4) \qquad E(x) \leq E^{(n+1)}(x) \leq E^{(n)}(x) \qquad \text{for all } x \text{ in } \Omega_{p,q}(D) ,$$

$$(5) \qquad E^{(n+1)}(x^{(n+1)}) \leq E^{(n+1)}(x^{(n)}) \leq E^{(n)}(x^{(n)}) .$$

Therefore $\{E^{(n)}(x^{(n)})\}$ is a decreasing sequence, and

$$E(x^{(n)}) \leq E^{(n)}(x^{(n)}) \leq \text{constant} .$$

(v) Having established that the $x^{(n)}$'s all lie in some compact set K and that $E(x^{(n)})$ is bounded, we can now apply Lemma 2. Therefore by passing to a subsequence we can assume that the $x^{(n)}$ converge to some $x_\infty \in \Omega_{p,q}(D)$ in the weak H^1-topology. We are also given that

$$(6) \qquad E(x_\infty) \leq \overline{\lim} \{E(x^{(n)})\} .$$

We have to show that

$$(7) \qquad E(x_\infty) \leq E(x) \qquad \text{for all } x \in \Omega_{p,q}(D), (x \sim x_\infty) .$$

Remark. The fact that K is a compact subset of the open set D implies that K does not intersect ∂D. This is important since otherwise x_∞ might be a broken geodesic with corners abutting on ∂D.

(vi) To establish (7) we use (4), (5) and (6) to obtain

$$E(x_\infty) \leq \overline{\lim} \{E(x^{(n)})\} \leq \overline{\lim} \{E^{(n)}(x_n)\} \leq \overline{\lim} \{E^{(n)}(x)\} = E(x) .$$

This completes the proof of Theorem 1. The proof of Theorem 2 differs from that of Theorem 1 only in some minor details.

5. Examples

We conclude by giving a few low dimensional examples of domains D, which satisfy the conditions of Theorems 1 and 2.

1. Again we consider the hyperboloid M whose isometric embedding in $\boldsymbol{R}^3$ is given by $x^2 + y^2 = 1 + z^2$. As we have already mentioned, the function $f = x^2 + y^2$ is proper, positive, and convex; hence M provides an example of a domain homeomorphic to an open annulus which supports a function satisfying the conditions of Theorem 1.

Let $r_0 > 1$, and let D be the domain (of M) given by $z^2 < r_0^2 - 1$. Let F be a function on D given by $F(p) = b - \log(r_0^2 - f(p))$, where the constant b is chosen to make F positive. An easy calculation shows that $F_{ij} = (r_0^2 - f)^{-1}f_{ij} + (r_0^2 - f)^{-2}f_i f_j$. Hence F is proper, positive, and convex on D.

Note that f is strictly convex on the domains $z > 0$ and $z < 0$, so that neither of these domains contain periodic trajectories. On the other hand, M contains a periodic geodesic around its waist $z = 0$. More generally, if a domain which supports a convex function f contains a periodic geodesic, then this geodesic must lie on a hypersurface $f =$ *constant*. (Cf. [2], [3].)

2. Let M be the standard unit circle $\boldsymbol{S}^1, p \in \boldsymbol{S}^1$ and $D = S^1 - \{p\}$. Then it is easily shown that D supports a proper positive convex function.

If we cross this example with $\boldsymbol{R}^1$, we obtain an example of a domain D satisfying the conditions of Theorem 2; i.e., M is the cylinder $\boldsymbol{S}^1 \times \boldsymbol{R}^1$, and D is the cylinder with a generating line removed.

3. Finally, we give an example which gives content to the parenthetical remark following the definition of g-connectedness in § 1. We construct a domain D in a (compact, simply connected) manifold M with the following properties.

(i) There exist two (distinct) points p, q in D and a (unique) geodesic γ joining p to q in D, whose arc length is minimum for all paths *in D* joining p to q and homotopic to γ.

(ii) There exists a geodesic γ' joining p to q *in M*, which is homotopic *in M* to γ and whose arc length is strictly less than that of γ.

To this end, let $\boldsymbol{S}^2$ be the standard 2-sphere whose isometric embedding in $\boldsymbol{R}^3$ is given by $x^2 + y^2 + z^2 = 1$. Let ε be a small positive number, and let p, q be the two points on $\boldsymbol{S}^2$, which lie on the two planes $x = 0$ and $z = -\varepsilon$. Let D be the domain $z < -\varepsilon/2$, and γ be the short great circle arc in D, which joins p to q. One can easily construct a proper positive convex function on D (which depends on z alone), so that D is g-connected. Now let M be the topological sphere which is obtained by removing the closed domain $z \geq -\varepsilon/3$ and replacing it with a cap C which is smoothly attached to the remaining part of $\boldsymbol{S}^2$. If the cap C is sufficiently flat, then one easily obtains a minimizing geodesic γ' which joins p to q, whose arc length is less than that of γ and which is necessarily homotopic to γ (in M) since M is simply connected.

References

[1] S. Alexander & R. L. Bishop, *Convex-supporting domains on spheres,* Illinois J. Math. **18** (1974) 37–47.

[2] R. L. Bishop & B. O'Neill, *Manifolds of negative curvature,* Trans. Amer. Math. Soc. **145** (1969) 1–49.

[3] W. B. Gordon, *Convex functions and harmonic maps,* Proc. Amer. Math. Soc. **33** (1972) 433–437.

[4] ——, *Physical variational principles which satisfy the Palais-Smale condition,* Bull. Amer. Math. Soc. **78** (1972) 712–716.

[5] ——, *An analytical criterion for the completeness of Riemannian manifolds,* Proc. Amer. Math. Soc. **37** (1973) 221–225.

[6] ——, *On the equivalence of second order systems occurring in the calculus of variations,* Arch. Rational Mech. Anal. **50** (1973) 118–126.

[7] ——, *Conservative dynamical systems involving strong forces,* in preparation.

[8] R. S. Palais, *Foundations of non-linear global analysis,* Benjamin, New York, 1968.

[9] R. S. Palais et al., *Seminar on the Atiyah-Singer index theorem,* Annals of Math. Studies, No. 57, Princeton University Press, Princeton, 1965.

U. S. Naval Research Laboratory