J. DIFFERENTIAL GEOMETRY
16 (1981) 421–429

THE REGULAR FOCAL LOCUS

JAMES J. HEBDA

1. This paper examines Riemannian manifolds containing a submanifold whose first focal locus consists entirely of focal points having the same multiplicity. It is shown that a compact manifold admitting such a submanifold, provided the homomorphism on the fundamental groups induced by the inclusion map is surjective, is expressible as the topological union of two disk bundles, whose cores are the submanifold and its cut locus, which, in this case, is also a submanifold. Under more stringent conditions, simple-connectivity and the transversality condition defined below, it is proved that the ambient manifold, the submanifold, and its cut locus all have integral cohomology rings of compact rank-one symmetric spaces. In some sense, this work generalizes that of F. W. Warner in [12] and [13].

Throughout, W denotes a complete Riemannian manifold of dimension $m + r$, and $f\colon M \to W$ a smooth submanifold of W of dimension m. M may be considered to be a subset of W, and, in this case, the inclusion map f will be suppressed. Recall, f is proper if the inverse image under f of every compact set is compact. $N(M)$ will denote the total space of the normal bundle of M in W. $\exp\colon N(M) \to W$ will denote the exponential map obtained by restricting the exponential map of W to $N(M)$. A focal point is a critical point of exp, and its multiplicity is the dimension of the kernel of the differential $\exp_*$ at that point. An M-geodesic is a geodesic of W initially perpendicular to M. An M-Jacobi field is a Jacobi field along an M-geodesic which is the transverse vector field to a variation of M-geodesics. Observe that a focal point occurs where a nontrivial M-Jacobi field vanishes. For a proper submanifold, a cut point along an M-geodesic is the point after which the M-geodesic no longer minimizes the distance to the submanifold. The cut locus of M in $N(M)$ consists of those points which correspond to cut points along M-geodesics under exp. [2] is a good reference on M-Jacobi fields and cut loci (there called minimum loci) of submanifolds.

Received November 10, 1980. This paper is part of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Chicago, 1979.

2. A *ray* in the vector bundle $N(M)$ is the set of positive multiples of a nonzero vector.

Definition. A focal point Z in $N(M)$ is a *regular focal point* if there exists a neighborhood V of Z in $N(M)$ such that for every ray r meeting V, there is at most one focal point in $r \cap V$.

The following two theorems are generalizations of Theorems 3.1 and 3.2 in [12]. The proofs go through with obvious modifications to the focal point situation.

Theorem 2.1. *The set of regular focal points is an open and dense subset of the set of all focal points of M in $N(M)$. Furthermore it can be given the structure of a submanifold of $N(M)$, which has codimension* 1 *and is transverse to every ray meeting it.*

Theorem 2.2. *At regular focal points having multiplicity at least two, the kernel of* $\exp_*$ *is contained in the tangent space to the focal locus.*

3. Definition. M has a *very regular first focal locus* if the multiplicity of the first focal point along every M-geodesic is constant and, in case the multiplicity is one, $\ker(\exp_*)$ is contained in the tangent space to the focal locus at every first focal point.

Remark. By Gauss's lemma, this condition is satisfied if both the multiplicity of the first focal point and the distance to the first focal point along every M-geodesic are constant.

Theorem 3.1. *Suppose W is a connected compact Riemannian manifold, and M is a connected compact submanifold having a very regular first focal locus such that $f_*\colon \pi_1(M) \to \pi_1(W)$ is onto. If $s - 1$ is the multiplicity of the first focal points of M, then the cut locus K of M in W is a submanifold of W with codimension s, and W may be expressed as a topological union $D_M \cup_\varphi D_K$, where D_M and D_K are disk bundles over M and K respectively, and $\varphi\colon \partial D_M \to \partial D_K$ is a diffeomorphism.*

Proof. Let F denote the first focal locus of M in $N(M)$. The cut locus of M in $N(M)$ coincides with F. (The proof of this fact is analogous to Theorem 5.11 in [3] where we use the Morse theory of the space of paths with initial point in M and fixed terminal point. Observe that the surjectivity of f_* replaces the simple-connectivity condition. For the multiplicity-one case, see [5].) Hence F is compact by the compactness of M and W, and the cut locus K of M in W inherits a submanifold structure from the submersion $\exp\colon F \to K$ having codimension s in W.

Since the rays in $N(M)$ are transverse to F, and the kernel of $\exp_*$ is contained in the tangent space to F, it follows that an M-geodesic when it first meets K is not tangent to K. Thus by compactness we can find an $\varepsilon_0 > 0$ such that no M-geodesic prior to meeting K is tangent to the ε-tube about K for any

$0 < \varepsilon \leqslant \varepsilon_0$. (The ε-tube is the set of points of W a distance ε from K.) It follows that every M-geodesic meets the ε_0-tube exactly once before meeting K; otherwise it would be tangent to some ε-tube with $0 < \varepsilon < \varepsilon_0$. Set

$$D_M = \{Z \in N(M)\colon \exp(tZ) \text{ does not meet the } \varepsilon_0\text{-tube for } t \in [0, 1).\},$$

and let D_K be the set of points of W whose distance to K is at most ε_0. Then D_M and D_K are diffeomorphic to the normal disk bundle over M and K respectively. Define a diffeomorphism $\varphi\colon \partial D_M \to \partial D_K$ by letting $\varphi(Z)$ be the intersection of the ε_0-tube with the M-geodesic determined by Z before it meets K, namely $\varphi(Z) = \exp(Z)$. Clearly W is the topological union of D_M and D_K. q.e.d.

Conversely, there is the following.

Theorem 3.2. *Suppose W is a compact manifold expressible as $D_M \cup_\varphi D_K$, where D_M and D_K are disk bundles over compact submanifolds M and K respectively, and $\varphi\colon \partial D_M \to \partial D_K$ is a diffeomorphism. Assume the codimension s of K in W is at least two. Then* (1) *$f_*\colon \pi_1(M) \to \pi_1(W)$ is onto, and* (2) *there exists a Riemannian metric on W such that M has a very regular first focal locus, and K is the cut locus of M.*

Proof. Statement (1) follows from a general position argument since the codimension of K is at least two, and M is a deformation retract of $W - K$.

In [9, pp. 236–238] Omori constructs a Riemannian metric on W having the properties that K is the cut locus of M, and the distance from M to the cut locus is constant. From the first variation formula it therefore follows that every geodesic meeting M perpendicularly meets K perpendicularly and vice versa, so that every M-geodesic can be reparametrized to be a K-geodesic and vice versa. Moreover, M-Jacobi fields are precisely K-Jacobi fields. The number of linearly independent K-Jacobi fields vanishing at points of K is $s - 1$. Thus using the M-Jacobi field characterization of focal points we see that M has a regular focal locus which occurs at a constant distance. Therefore M has a very regular first focal locus. (See the remark before Theorem 3.1.)

4. For a manifold which is decomposable as the union of two disk bundles, there are two long exact cohomology sequences relating the cohomology groups of the manifold to those of the cores of the two disk bundles. With the notation of Theorem 3.1 and letting $r = \operatorname{codim}(M)$ and $s = \operatorname{codim}(K)$, observe that $H^i(D_M) \cong H^i(M)$ via a deformation retraction, and $H^i(W, D_M) \cong H^i(D_K, \partial D_K) \cong H^{i-s}(K)$ via exclusion and the Thom isomorphism. (Here we take the coefficients of the cohomology groups to be the integers modulo 2, or the integers in case D_K in an oriented bundle over K.) Starting from the long exact cohomology sequence of the pair (W, D_M) and using the above isomorphisms, one obtains the exact cohomology sequence

$$(4.1a) \qquad \leftarrow H^i(M) \xleftarrow{f^*} H^i(W) \leftarrow H^{i-s}(K) \xleftarrow{\delta} H^{i-1}(M) \leftarrow .$$

Similarly, interchanging the roles of M and K, there is the exact sequence

$$(4.1b) \qquad \leftarrow H^i(K) \xleftarrow{g^*} H^i(W) \leftarrow H^{i-r}(M) \xleftarrow{\delta} H^{i-1}(K) \leftarrow .$$

Furthermore, the homomorphisms in these sequences are homomorphisms of $H^*(W)$-modules. ($H^*(M)$ and $H^*(K)$ are $H^*(W)$-modules under the ring homomorphisms f^* and g^* induced by the inclusion maps. See [4, p. 321].)

An easy application of these sequences is the following.

Theorem 4.1. *Let W and M satisfy the hypothesis of Theorem* 3.1. (1) *If* $m < \frac{1}{2}\dim(W)$ *and* $s > \frac{1}{2}\dim(W)$, *then* $m + s = \dim(W)$. (2) *If* $m + 1 < \frac{1}{2}\dim(W)$ *and* $s > \frac{1}{2}\dim(W)$, *then M and its cut locus K have isomorphic $Z/(2Z)$ cohomology rings and W has the $Z/(2Z)$ cohomology ring of a sphere bundle over M.*

Proof. Theorem 3.1 applies. Let $k = \dim(K)$ and recall $r = \operatorname{codim}(M)$. We want to show $r = s$.

Conditions (1) are equivalent to $r > m$ and $s > k$. Suppose $r < s$. Then $m > k$, and thus $H^m(W) = 0$ by exactness of (4.1b). On the other hand, $m - s + 1 < 0$ since $m < r < s$, and thus $H^m(W) = H^m(M) \neq 0$ by exactness of (4.1a). We obtain a similar contradiction assuming $r > s$. Thus $r = s$.

If (2) holds, then $m = k$ and $r = s$ by (1). If $i \leqslant m$, then $i - s + 1 \leqslant m - r + 1 < 0$. Thus, by (4.1a), f^* is an isomorphism if $i \leqslant m$. Similarly, g^* in (4.1b) is an isomorphism for $i \leqslant k = m$. Since these isomorphisms are induced by the inclusion maps, they preserve the ring structure. Therefore M and K have isomorphic rings.

From the exactness of (4.1b), when $i > m$, it follows that $H^{i-r}(M) \to H^i(W)$ is an isomorphism. Since this isomorphism is a $H^*(W)$-module homomorphism, it follows that

$$u \cup : H^{i-r}(W) \to H^i(W)$$

is an isomorphism for $m < i \leqslant \dim(W)$ and $u \in H^r(W)$ the generator. Thus W has the $Z/(2Z)$ cohomology of an r-dimensional sphere bundle over M.

Remark. One can use integral cohomology in Theorem 4.1 if W, M, and K are orientable.

A topological union $D_M \cup_\varphi D_K$ of disk bundles is said to satisfy the *transversality condition* if the image under φ of every fiber of the sphere bundle ∂D_M over M is transverse to every fiber of the sphere bundle ∂D_K over K. (Recall two submanifolds are transverse if the span of their tangent spaces at points of intersection is the whole tangent space of the ambient manifold.)

Remark. If either M or K is a point, the transversality condition is automatically satisfied.

Example. Consider two copies of the solid torus $S^1 \times D^2$ as disk bundles over S^1 via projection on the first factor. Consider the identity map I of $S^1 \times S^1$ and the twist map $T: S^1 \times S^1 \to S^1 \times S^1$ defined by $T(x, y) = (y, x)$. Then $S^1 \times D^2 \cup_I S^1 \times D^2$ does not satisfy the transversality condition, whereas $S^1 \times D^2 \cup_T S^1 \times D^2$ does.

Theorem 4.2. *Suppose W is a simply-connected compact connected Riemannian manifold, and that M is a compact connected submanifold of W having a very regular first focal locus. Let K be the cut locus of M. If the transversality condition holds for the decomposition of Theorem* 3.1, *then one of the following holds*:

(1) *W is homeomorphic to a sphere, and M and K are either both points or both diffeomorphic to spheres.*

(2) *W, M, and K have the same homotopy type as complex projective spaces.*

(3) *W, M, and K have integral cohomology rings of quaternionic projective spaces.*

(4) *W has the integral cohomology ring of the Cayley plane, and one of M and K is a point, the other being homeomorphic to S^8.*

Proof. In the notation of Theorem 3.1, identify $\partial D_M \cong \partial D_K$ with a submanifold X of W. Let $k = \dim(K)$, and recall $s = \operatorname{codim}(K)$ and $r = \operatorname{codim}(M)$. X is a sphere bundle over M with fiber S^{r-1} and also a sphere bundle over K with fiber S^{s-1}. Denote the projection maps of these bundles by p_M and p_K respectively, and let j_M and j_K denote the inclusion maps of the respective fibers S^{r-1} and S^{s-1} into X.

Consider the map $p: X \to M \times K$ defined by the two projections p_M and p_K. The transversality condition implies that p is a submersion. Hence the image is open. It is also closed since X is compact. Thus p is a submersion of X onto $M \times K$. (The only time $M \times K$ is disconnected is when $r = 1$. However p still meets both components.) Therefore p is a fibration of X over $M \times K$ with fiber $F = j_M(S^{r-1}) \cap j_K(S^{s-1})$. Similarly the compositions $p_M \circ j_K: S^{s-1} \to M$ and $p_K \circ j_M: S^{r-1} \to K$ are fibrations with the same fiber F. We want to show F is connected.

The simple connectivity of W, a general position argument, and a deformation retraction of $W - K$ onto M shows that M is simply-connected if $s \geqslant 3$. Furthermore, if $s \geqslant 3$, the exact homotopy sequence for the fibration p_K shows that $p_{K*}: \pi_1(X) \to \pi_1(K)$ is an isomorphism. Also X is connected if $r \geqslant 2$. Therefore consideration of the exact homotopy sequence for the fibration p,

$$\pi_1(X) \xrightarrow{p_*} \pi_1(M \times K) \to \pi_0(F) \to \pi_0(X),$$

shows that F is connected if $s \geqslant 3$ and $r \geqslant 2$. Similarly, F is connected if $r \geqslant 3$ and $s \geqslant 2$.

If $r = 2$ and $s = 2$, F is either 0- or 1-dimensional. If 1-dimensional, F is the entire fiber of the fibration $p_K \circ j_M$ and hence is connected. If 0-dimensional, let n be the number of points in F. First of all, the fibrations $p_M \circ j_K$ and $p_K \circ j_M$ are n-fold covers of M and K by S^1. X is an orientable S^1-bundle over M, since M being 1-dimensional is an orientable submanifold of the simply-connected space W and hence has an oriented normal bundle. Thus $X \cong M \times S^1$ as bundles over M. (M is just S^1.) A computation using the Van Kampen's theorem applied to the decomposition of W by D_M and D_K with $X = D_M \cap D_K$ shows $\pi_1(W) \cong Z/(nZ)$. Thus, since W is simply-connected, $n = 1$ and F is connected.

If $r = 1$ and F is disconnected, the fibration $p_K \circ j_M$ is a fibration of S^0 over K. Thus K consists of one point, and F consists of two. Thus p_M is a double cover of M, and hence W is not simply-connected. Therefore if $r = 1$, F is connected.

We have also shown that the normal bundles of M and K are orientable. This is so because M and K are themselves always orientable being either simply-connected or 1-dimensional, and an orientable submanifold of an orientable manifold has an orientable normal bundle. Hence we will use integer coefficients in the exact sequences (4.1a) and (4.1b).

Now $p_M \circ j_K$ and $p_K \circ j_M$ are fibrations of spheres over compact manifolds M and K respectively with a compact connected fiber F. The only ways that a sphere can be fibered over a non-contractible space with a compact connected fiber F are if F is a point, a homotopy 1-sphere, a homotopy 3-sphere, or a homotopy 7-sphere [11]. Now K or M fail to be noncontractible, when they are points. If one of M or K is a point, we are in the situation examined by Warner in [13]. His classification applies and fits into our scheme.

Assume neither M nor K is a point. We examine the possibilities for F.

Suppose F is a point. Then p, $p_M \circ j_K$ and $p_K \circ j_M$ are diffeomorphisms. Hence we have the following pair of commutative diagrams. (Both of the top arrows are the diffeomorphism $p \circ ((p_M \circ j_K)^{-1}, (p_K \circ j_M)^{-1})$.)

$$\begin{tikzcd} X \arrow[r, "\cong"] \arrow[d, "p_M"'] & S^{s-1} \times S^{r-1} \arrow[d] \\ M & S^{s-1} \arrow[l, "\cong", "p_M \circ j_K"'] \end{tikzcd} \qquad \begin{tikzcd} X \arrow[r, "\cong"] \arrow[d, "p_K"'] & S^{s-1} \times S^{r-1} \arrow[d] \\ K & S^{r-1} \arrow[l, "\cong", "p_K \circ j_M"'] \end{tikzcd}$$

Since the fibers are preserved, we can extend these diffeomorphisms to homeomorphisms defined on the disk bundles by a mapping cylinder construction. Hence the following commutative diagrams:

$$\begin{tikzcd} D_M \arrow[r, "\cong"] \arrow[d] & S^{s-1} \times D^r \arrow[d] \\ M & S^{s-1} \arrow[l, "\cong"] \end{tikzcd} \qquad \begin{tikzcd} D_K \arrow[r, "\cong"] \arrow[d] & D^s \times S^{r-1} \arrow[d] \\ K & S^{r-1} \arrow[l, "\cong"] \end{tikzcd}$$

Notice that $\varphi: \partial D_M \to \partial D_K$ becomes the identity map of $S^{s-1} \times S^{r-1}$ under these identifications. Therefore there is induced the following homeomorphism:

$$W \cong D_M \cup_\varphi D_K \cong S^{s-1} \times D^r \cup_{id} D^s \times S^{r-1} \cong S^{r+s-1}.$$

Suppose F is a homotopy $(d-1)$-sphere where $d = 2, 4$, or 8. By a standard Gysin sequence argument [6], since M and K are the base space of a fibering of a sphere by a homotopy sphere, the integral cohomology rings of M and K are truncated polynomial rings with a generator in dimension d. Also d divides m, k, r, s, and $\dim(W)$. Let $m' = m/d$ and $k' = k/d$.

Easy computations involving the long exact sequences (4.1a) and (4.1b) with knowledge of the cohomology of M and K show that

$$H^i(W) = \begin{cases} Z & 0 \leqslant i \leqslant \dim(W) \text{ and } i \equiv 0 \pmod{d}, \\ 0 & \text{otherwise.} \end{cases}$$

In fact, $H^*(W)$ has as a free basis the set $\{1, x, \cdots, x^{m'}, u, ux, \cdots, ux^{k'}\}$, where $x \in H^d(W)$ is a generator, and $u \in H^s(W)$ is the generator which comes from restricting the Thom class in $H^s(D_K, X) \cong H^0(K)$ to W. (In proving this one needs to use the fact that the homomorphism $H^{i-s}(K) \to H^i(W)$ in (4.1a) is an $H^*(W)$-module homomorphism.)

In order to show that the integral cohomology ring of W is a truncated polynomial ring, we must show that $u = x^{m'+1}$. This requires a geometric construction.

We construct a space $\hat{M} \subset W$ by attaching an s-disk to M along the map $p_M \circ j_K$ as follows. Let D^s be a fiber of the disk bundle D_K over K. Set $S^{s-1} = D^s \cap \partial D_K$. Let $\hat{M} = D^s \cup \{\text{all rays in } D_M \text{ meeting } S^{s-1}\}$. The cohomology of $\hat{M}$ has as a free basis the set $\{1, x \mid \hat{M}, \cdots, (x \mid \hat{M})^{m'}, u \mid \hat{M}\}$. $u \mid \hat{M} \in H^s(\hat{M})$ is the generator by the definition of the Thom class. Now $\hat{M}$ is a Poincaré duality space, in fact, a topological manifold (see appendix). Hence the Poincaré duality argument on p. 207 of [13] shows that $(u \mid \hat{M}) = \pm(x \mid \hat{M})^{m'+1}$. Hence $u \mid \hat{M} = \pm x^{m'+1} \mid \hat{M}$. Therefore $u = \pm x^{m'+1}$ since $H^s(W) \to H^s(\hat{M})$ is an isomorphism.

We have proved that the integral cohomology rings of W, M and K are truncated polynomial rings with a generator in dimension d. For $d = 2$, such spaces have the homotopy type of complex projective spaces [7, p. 537]. For $d = 4$, this is the integral cohomology ring of quaternionic projective spaces. For $d = 8$, Adem [1] shows $x^3 = 0$. This contradicts our expression for $H^*(W)$ assuming neither M nor K is a point. Thus one of M or K is a point, and Warner's proof shows that the other is a homology 8-sphere. However, as we say above, it is also simply-connected. Thus it is homeomorphic to S^8.

Remark. In view of Theorem 3.2, one can construct exotic examples of spaces satisfying the hypothesis of Theorem 4.2. Witness the Kuiper-Eells quaternionic planes [8].

APPENDIX. We show that the space $\hat{M}$ constructed in the proof of Theorem 4.2 is a topological manifold and hence satisfies Poincaré duality.

$\hat{M}$ is obtained by pasting a disk D^s onto M via a fiber bundle map $\hat{p}$: $\partial D^s \to M$ with fiber the homotopy sphere Σ^{d-1}, $d = 2$, 4, or 8. In order to show that $\hat{M}$ is a topological manifold, it suffices to show that for every $x \in M \subset \hat{M}$, x has a neighborhood in $\hat{M}$ which is a manifold. (We certainly do not have to worry about the points in the interior of the disk.)

Let $x \in M$, and let U be a neighborhood of x in M so that there is a homeomorphism $\hat{p}^{-1}(U) = U \times \Sigma^{d-1}$. Consider the set $\{ty\colon t \in (0,1],\ y \in \hat{p}^{-1}(U)\} \cong U \times \Sigma^{d-1} \times (0,1]$ which is open in D^s. This projects down to a neighborhood $\hat{U}$ of x in $\hat{M}$ for which there is a homeomorphism $\hat{U} \cong U \times C\Sigma^{d-1}$ where $C\Sigma^{d-1}$ is the cone on Σ^{d-1}. If $d = 2$ or 8, Σ^{d-1} is homeomorphic to the sphere S^{d-1}. Thus $C\Sigma^{d-1}$ is homeomorphic to CS^{d-1} which is the interior of the d-disk, which is a manifold. (We take open cones $C\Sigma^{d-1} = (\Sigma^{d-1} \times (0,1])/(\Sigma^{d-1} \times \{1\})$.)

If $d = 4$, because the double suspension of a homotopy 3-sphere is homeomorphic to S^5 [10], it follows that $R \times C\Sigma^3$ is a manifold, for this set can be embedded as an open set of the double suspension of Σ^3. Since M is not a point, $\dim(U) \geqslant 1$, and so $U \times C\Sigma^3$ is a manifold.

This completes the proof that $\hat{M}$ is a topological manifold.

References

[1] J. Adem, *Relations on reduced powers*, Proc. Nat. Acad. Sci. U.S.A. **39** (1953) 636–638.

[2] R. L. Bishop & R. J. Crittenden, *Geometry of manifolds*, Academic Press, New York, 1964.

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

[4] A Dold, *Lectures on algebraic topology*, Springer, 1972.
[5] J. J. Hebda, *Conjugate and cut loci and the Cartan-Ambrose-Hicks theorem*, Indiana Math. J.
[6] S. T. Hu, *Homotopy theory*, Academic Press, New York, 1959.
[7] W. Klingenberg, *Manifolds with restricted conjugate loci*, Ann. of Math. **78** (1963) 527–547.
[8] N. Kuiper & J. Eells, *Manifolds which are like projective planes*, Inst. Hautes Études Sci. Publ. Math. No. 14 (1961).
[9] H. Omori, *A class of Riemannian metrics on a manifold*, J. Differential Geometry **2** (1968) 233–252.
[10] L. C. Siebenmann, *Are non-triangulable manifolds triangulable*?, Topology of Manifolds, Proc. University of Georgia Topology of Manifolds Institute, 1969.
[11] K. Srinivasacharyulu, *Sur les structures differentiable et les variations de structure complexes*, Sémaire Ehresmann, iv., 1962 – 1963.
[12] F. W. Warner, *The conjugate locus of a Riemannian manifold*, Amer. J. Math. **87** (1965) 575–604.
[13] ______, *Conjugate loci of constant order*, Ann. of Math. **86** (1967) 192–212.

UNIVERSITY OF CHICAGO