J. DIFFERENTIAL GEOMETRY
16 (1981) 685–697

TOTALLY FOCAL EMBEDDINGS: SPECIAL CASES

SHEILA CARTER & ALAN WEST

In this paper we take up again the study of totally focal embeddings introduced in [2]. In the previous paper we showed that a compact totally focal hypersurface must be just a round sphere. In this paper we show that if an embedding of $\mathbf{S}^2$ in $\mathbf{R}^4$ is totally focal, then again $\mathbf{S}^2$ must be embedded as a round sphere. We also show that if an embedding of $\mathbf{S}^1$ in $\mathbf{R}^n$ is totally focal, then $\mathbf{S}^1$ is embedded as a round circle. These two results do not seem very easy to prove. The first result depends on some general results which perhaps pave the way for further study. The second result seems to require *ad hoc* methods.

The general results we need are not all about totally focal embeddings. We have devoted a separate section to some general theorems about submanifolds of $\mathbf{R}^n$ which we require but do not seem to be generally known. They concern the focal set and although we do not use them in the generality presented here, we think they are of independent interest.

We know that there are plenty of totally focal embeddings which are more interesting than just round spheres. See [3] and [4] for examples. However it seems desirable to sort out what can happen in the simple cases before trying to discuss these. With this in mind it seems curious that the proofs we present here are not very straightforward, and it is irritating that, at present, we still cannot discover whether there is any totally focal embedding of $\mathbf{S}^2$ in $\mathbf{R}^5$ more interesting than just a round sphere.

0. Notation and conventions

Let $M \to \mathbf{R}^n$ be a smooth connected compact m-dimensional manifold without boundary embedded in n-dimensional Euclidean space; its normal bundle is $N(M) \subset M \times \mathbf{R}^n$ where $(p, x) \in N(M)$ means that the line in $\mathbf{R}^n$ through x and p is normal to the manifold M at p. This is a slightly different convention from that used in [2]. The projection π: $N(M) \to M$ is given by

Received November 4, 1980, and, in revised form, November 19, 1981.

$\pi(p, x) = p$, and the end-point map η: $N(M) \to \mathbf{R}^n$ is given by $\eta(p, x) = x$. We will write $N_p = \{x \in \mathbf{R}^n: (p, x) \in N(M)\}$, and observe that η: $\pi^{-1}(p) \to N_p$ is an isomorphism.

The critical points of η, called critical normals, form the set Γ. The *focal* set is $\eta(\Gamma)$. We put $F_p = \eta\{\pi^{-1}(p) \cap \Gamma\}$ and observe that F_p is the zero-set of a polynomial of degree m.

For any $x \in \mathbf{R}^n$ the distance function L_x: $M \to \mathbf{R}$ is given by $L_x(p) = \| p - x\|^2$. L_x is nondegenerate if and only if $x \notin \eta(\Gamma)$, and p is a critical point of L_x if and only if $(p, x) \in N(M)$. The index of p is also the index of $(p, x) \in N(M)$. Thus each normal in $N(M)$ has an index, and this index is a constant on each connected component of $N(M) \setminus \Gamma$. Let us denote the set of normals in $N(M) \setminus \Gamma$ which have index k by $N^k(M)$. We put $N_p^k = \{x: (p, x) \in N^k(M)\}$.

As in [2] we say that M is totally focal if $\eta^{-1} \circ \eta(\Gamma) = \Gamma$. This means that every distance function L_x, $x \in \mathbf{R}^n$, is either nondegenerate or has only degenerate critical points. This is also equivalent to saying that $N_p \cap \eta(\Gamma) = F_p$ for all $p \in M$.

1. Extension to projective space

First of all we extend our ideas to projective space $\mathbf{P}^n$. We suppose that $\mathbf{R}^n \subset \mathbf{P}^n$ is embedded in a standard way, and identify $x \in \mathbf{R}^n$ with $[x, 1] \in \mathbf{P}^n$. Thus the "hyperplane at infinity" is an embedding of $\mathbf{P}^{n-1} \subset \mathbf{P}^n$, and we write $[z]$ instead of $[z, 0]$ where $z \in \mathbf{S}^{n-1}$. If $z \in S^{n-1}$, the height function H_z is given by $H_z(p) = \langle z \cdot p\rangle$. The idea is to utilize the fact that height functions are in a certain sense the limits of distance functions L_x as the point x tends to infinity.

The normal bundle is extended in the obvious way to $\tilde{N}(M) \subset M \times \mathbf{P}^n$, and the end-point map also extends to give η: $\tilde{N}(M) \to \mathbf{P}^n$ if we put $\eta(p, [z]) = [z]$. We define $\tilde{\Gamma}$ to be the set of critical points of this extended map η. Thus we can talk about critical normals and focal points at infinity. We define $\tilde{N}_p \subset \mathbf{P}^n$ and $\tilde{F}_p = \eta(\tilde{\Gamma} \cap \pi^{-1}(p))$ by analogy with N_p and F_p. The following theorem is crucial when we apply these ideas to our problem.

Theorem 1.1. *Let $M \subset \mathbf{R}^n$ be a smooth compact manifold without boundary embedded in Euclidean n-space. Then $\eta^{-1} \circ \eta(\Gamma) = \Gamma$ if and only if $\eta^{-1} \circ \eta(\tilde{\Gamma}) = \tilde{\Gamma}$.*

To prove this it is clearly sufficient to prove the following more general result.

Theorem 1.2. *Let $M \subset \mathbf{R}^n$ be a smooth compact manifold without boundary embedded in Euclidean n-space. Then $\overline{\eta(\Gamma)} = \eta(\tilde{\Gamma})$ in $\mathbf{P}^n$.*

The proof of this theorem uses a classical result from differential geometry which we quote as a lemma.

Lemma 1.3. *Let $U \subset \mathbf{R}^n$ be a smooth hypersurface without boundary such that at each point of U one of the principal curvatures is zero and has constant multiplicity k. Then U has a k-dimensional foliation, each leaf of which is closed in U, and is an open subset of some k-plane in $\mathbf{R}^n$, and such that on each leaf the normal direction to U is constant.*

The proof of this lemma can be found, for example, in [6].

Proof of Theorem 1.2. Consider an ε-neighborhood of M. If ε is sufficiently small, the boundary of this is a manifold $M^* \subset \mathbf{R}^n$ embedded in $\mathbf{R}^n$ with codimension 1. The focal set of M^* in $\mathbf{P}^n$ will be $\eta(\tilde{\Gamma}) \cup M$. Since M is compact, it is closed and does not intersect the hyperplane at infinity. So to show that $\overline{\eta(\Gamma)} = \eta(\tilde{\Gamma})$ it is sufficient to show that $\overline{\eta(\Gamma)} \cup M = \eta(\tilde{\Gamma}) \cup M$. That is, it is sufficient to prove the result when M is replaced by M^*.

So we assume that M has codimension 1 in $\mathbf{R}^n$, and therefore there is a unique normal direction $[v_p]$ at each $p \in M$.

Now observe that since M is compact, so is $\tilde{N}(M)$ and thus as $\tilde{\Gamma}$ is closed in $\tilde{N}(M)$, $\tilde{\Gamma}$ and $\eta(\tilde{\Gamma})$ are compact. Hence $\Gamma \subset \tilde{\Gamma}$ implies that $\overline{\eta(\Gamma)} \subset \eta(\tilde{\Gamma})$. Therefore it is only necessary to show that $\eta(\tilde{\Gamma}) \subset \overline{\eta(\Gamma)}$.

Suppose we can find $[z] \in \eta(\tilde{\Gamma})$ where $[z] \notin \overline{\eta(\Gamma)}$. We will obtain a contradiction. For such a z, let $(p,[v_p]) = (p,[z]) \in \tilde{\Gamma}$. There must exist a neighborhood of this point which does not intersect Γ. We deduce that there is a neighborhood U_p of p in M such that for all $q \in U_p$, $(q,[v_p])$ is a critical normal, and all these critical normals must have the same multiplicity k, say.

Let $W = \{p \in M: [v_p] = [z]\}$. Clearly W is compact. Now put $U = \bigcup \{U_p: p \in W\}$ and observe that we can apply Lemma 1.3 to U and $W \subset U$. For any $p \in W$ let V be the leaf through p as given by the lemma. The normal direction is constant over V and therefore $V \subset W$. Since V is closed in U and W is compact, we deduce that V is compact. This is a contradiction since by the lemma V is an open subset of some k-plane. This contradiction shows that $[z]$ cannot exist and therefore $\eta(\tilde{\Gamma}) \subset \overline{\eta(\Gamma)}$.

2. General results on focal sets

In this section we collect together some results about the focal sets of embedded manifolds which we require. We do not actually require them in the generality given here, but they seem to be interesting in their own right.

In this section we are essentially only concerned with local results so we assume that M is a smooth connected submanifold of $\mathbf{R}^n$ but is not necessarily compact or closed.

Theorem 2.1. *Let $M \subset \mathbf{R}^n$ have normal bundle $N(M)$, and let $N^m(M)$ be the set of normals of index $m = \dim M$. Then for each $p \in M$, $N^m(M) \cap \pi^{-1}(p)$ is either empty or is an open convex subset of $\pi^{-1}(p)$.*

The proof depends on a technical lemma from the theory of algebraic plane curves. We will state and prove this lemma before beginning the proof of Theorem 2.1.

Lemma 2.2. *Let $\phi(x, y)$ be a polynomial of degree m on $\mathbf{R}^2$ such that $\phi(x,0)$ is not identically zero. Suppose that for all a, $\phi(a, y) = 0$ has m real roots (as a function of y). Suppose further that $y = 0$ is a root of multiplicity l of $\phi(0, y) = 0$. Then $x = 0$ is a root of multiplicity $\geqslant l$ for $\phi(x,0) = 0$.*

Proof of Lemma 2.2. Let us first assume that ϕ is an irreducible polynomial. We can find $\varepsilon > 0$ such that $x = 0$ is the only root of $\phi(x,0) = 0$ with $|x| \leqslant \varepsilon$. Also since the roots of $\phi(a, y) = 0$ must be continuous functions of a, we can choose ε so that there is a $\delta > 0$ such that if $|a| \leqslant \varepsilon$ then $\phi(a, y) = 0$ has exactly l roots with $|y| \leqslant \delta$. Further since ϕ is irreducible, we can choose ε so that if $0 < |a| \leqslant \varepsilon$ these roots are all distinct. This last remark comes from observing that the common zeros of ϕ and $\partial\phi/\partial y$ must be finite so we can choose ε, δ so that $(0,0)$ is the only common zero in the rectangle $\{(x, y): |x| \leqslant \varepsilon, |y| \leqslant \delta\}$. These facts are easily derived from standard results about algebraic curves [8].

Thus there will be l continuous functions $f_i: [-\varepsilon, \varepsilon] \to [-\delta, \delta]$, $i = 1, \cdots, l$ such that for any $a \in [-\varepsilon, \varepsilon]$, $a \neq 0$, $y = f_1(a), \cdots, f_l(a)$ are the distinct and nonzero roots of $\phi(a, y) = 0$ for which $|y| \leqslant \delta$. Out of the $2l$ values $f_i(\varepsilon)$, $f_i(-\varepsilon)$, $i = 1, \cdots, l$ we may suppose that at least l are positive (otherwise we replace y by $-y$), and if $\gamma > 0$ is chosen to be less than all these positive values, a repeated application of the mean value theorem to the intervals $[-\varepsilon, 0]$ and $[0, \varepsilon]$ shows that there are at least l distinct roots of $\phi(x, \gamma) = 0$ with $|x| < \varepsilon$. Applying Rolle's theorem to $\phi(x, \gamma)$ in suitable intervals we deduce that for any r, $0 < r < l$, there is some x, $|x| < \varepsilon$, with $(\partial^r\phi/\partial x^r)(x, \gamma) = 0$. Since ε and δ may be chosen arbitrarily small we deduce that $(\partial^r\phi/\partial x^r)(0,0) = 0$. This means that $x = 0$ is a root of $\phi(x,0) = 0$ with multiplicity at least l. Thus the result is proved if ϕ is irreducible.

Now suppose $\phi = \phi_1 \cdot \phi_2 \cdot \cdots \cdot \phi_r$ is the decomposition of ϕ into its irreducible factors (over $\mathbf{R}$). Then $y = 0$ must be a root of $\phi_i(0, y) = 0$ with multiplicity l_i, putting $l_i = 0$ if it is not a root, with $l_1 + l_2 + \cdots + l_r = l$. Applying the result to each irreducible factor we obtain the required result applied to ϕ.

Proof of Theorem 2.1. Clearly since $\eta \,|\, \pi^{-1}(p): \pi^{-1}(p) \to N_p$ is an isomorphism, we only need show that N_p^m is an open convex set of N_p. Take $x_1, x_2 \in N_p^m$. These together with p determine a plane in N_p, which we can extend to a real projective plane in $\mathbf{P}^n$. We work entirely in this plane.

The projective line through x_1 and x_2 meets the line at infinity at a point which we will call x_∞, and these points separate the line into three closed segments x_1x_2, x_1x_∞ and x_2x_∞. If $x \neq x_\infty$ is a point on this line, we will let px denote the closed segment of the line through p and x, which does not intersect the line at infinity.

Now every line through p intersects $\tilde{F}_p$ in exactly m (real) points, counting multiplicities, and $x \in N_p^m$ means $x \notin \tilde{F}_p$ and there are m points in $px \cap \tilde{F}_p$. Thus we are given that there are m points in $px_1 \cap \tilde{F}_p$ and in $px_2 \cap \tilde{F}_p$, and we wish to show that $x \notin \tilde{F}_p$ and there are m points in $px \cap \tilde{F}_p$ whenever $x \in x_1x_2$. The method depends on the fact that as x varies the number of points in $px \cap \tilde{F}_p$ can change only if x passes through a point of $\tilde{F}_p$. Thus it is sufficient to show that $x_1x_2 \cap \tilde{F}_p = \varnothing$.

To show that $x_1x_2 \cap \tilde{F}_p$ is empty we show that there must be m points in $\{x_1x_\infty \cup x_2x_\infty\} \cap \tilde{F}_p$ if we count multiplicities correctly. So let x vary in x_1x_∞, and write k_x for the number of points in which px intersects $\tilde{F}_p$, counting multiplicities. Thus $k_x = m$ when $x = x_1$ or x_2 and k_x can be reduced only when x passes through a point of $\tilde{F}_p$. However the lemma implies the stronger result that if k_x is reduced by losing l coincident points, then x must pass through at least l coincident points in the intersection of x_1x_∞ with $\tilde{F}_p$. Thus for any $x \in x_1x_\infty$ if px intersects $\tilde{F}_p$ in k_x points with x itself accounting for l_x coincident points, then xx_1 must intersect $\tilde{F}_p$ in k_x points with x itself accounting for l_x coincident points, then xx_1 must intersect $\tilde{F}_p$ in at least $m - k_x + l_x$ points including multiplicities.

We now let x tend to x_∞ along x_1x_∞, and observe that the segment px will tend to a segment, call it L_1, of the line through p and x_∞. We deduce that if this line segment L_1 intersects $\tilde{F}_p$ at k_1 points with perhaps l of them being accounted for by l coincident points at x_∞, then x_1x_∞ must intersect $\tilde{F}_p$ in at least $m - k_1 + l$ points. Similarly we can apply the same argument to x_2x_∞, letting x tend to x_∞ and the segment px tend to the segment L_2 of the line through p and x_∞. We deduce that x_2x_∞ intersects $\tilde{F}_p$ in at least $m - k_2 + l$ points. However $L_1 \cup L_2$ is a complete projective line through p, and so intersects $\tilde{F}_p$ in exactly m points. Thus $k_1 + k_2 - l = m$. Also the line through x_1 and x_2 can intersect $\tilde{F}_p$ in no more than m points, none of them at x_1 or x_2. So if k_0 is the number of points in which x_1x_2 intersects $\tilde{F}_p$, we get $(m - k_1 + l) + (m - k_2 + l) + k_0 - l \leqslant m$. This gives $m + k_0 \leqslant m$. So $k_0 = 0$, and this proves that $x \in N_p^m$ whenever $x \in x_1x_2$ with $x_1, x_2 \in N_p^m$. Hence N_p^m is convex.

Notice that this general theorem can be modified to apply to immersions. Note also that we have proved rather more than was said in the theorem. In fact any projective line which intersects N_p^m must intersect $\tilde{F}_p$ in exactly m (real) points if we take into account their multiplicities.

The following lemma is fairly obvious, but we need it later and the proof is short.

Lemma 2.3. *Let $M \subset \mathbf{R}^n$ and let $\tilde{\Lambda} \subset \mathbf{P}^n$ be a projective k-plane. Suppose that for all $p \in M$, $\tilde{\Lambda} \subset \tilde{F}_p$. Then M lies on a sphere with focal set $\tilde{\Lambda}$.*

Note. If $\tilde{\Lambda}$ does not lie entirely in the hyperplane at infinity, then $\tilde{\Lambda}$ is the closure in $\mathbf{P}^n$ of a k-plane $\Lambda \subset \mathbf{R}^n$, and a sphere with focal set $\tilde{\Lambda}$ is a $(n-k-1)$-sphere with centre in Λ lying in an $(n-k)$-plane perpendicular to Λ. If $\tilde{\Lambda} \subset \mathbf{P}^n \setminus \mathbf{R}^n$, then a sphere with focal set $\tilde{\Lambda}$ is an $(n-k-1)$-plane whose normal directions are the points of $\tilde{\Lambda}$.

Proof. We first prove the theorem when $k = 0$ and $\tilde{\Lambda}$ is just a point $c \in \mathbf{R}^n$. In this case for any $p \in M$, (p, c) is a normal in $N(M)$. Thus every point $p \in M$ is a critical point of the distance function L_c. Since M is connected, this means that L_c must be a constant. Thus M lies on a hypersphere centre c.

Now consider the case when $\tilde{\Lambda}$ is a point $[a] \in \mathbf{P}^n \setminus \mathbf{R}^n$. (We can think of $a \in \mathbf{S}^{n-1}$ as determining a direction.) Again $(p, [a])$ is a normal in $\tilde{N}(M)$, and this means that every point $p \in M$ is a critical point of the height function H_a. Thus H_a is a constant, so M lies on a hyperplane perpendicular to a.

In the general case we take a fixed point on M, say p_0, and observe that we have proved that for any $x \in \tilde{\Lambda}$, M lies on the hypersphere (or hyperplane) through p_0 and with centre x. The intersection of all these hyperspheres is a sphere with focal set $\tilde{\Lambda}$.

Theorem 2.4. *Let $M \subset \mathbf{R}^n$, $\dim M = m$, and let $\tilde{\Pi}$ be a projective hyperplane in $\mathbf{P}^n$. Suppose that for all $p \in M$, $\tilde{F}_p \subset \tilde{\Pi}$. Then M is an open subset of a round m-sphere or a flat m-plane in $\mathbf{R}^n$.*

Proof. Write $\Pi = \tilde{\Pi} \cap \mathbf{R}^n$. We note that M must intersect Π transversally, because if $p \in M \cap \Pi$ and there is a projective line through p which is normal there to both M and Π, then the focal points on this line cannot be at p and so cannot lie on $\tilde{\Pi}$. Thus $M \setminus \Pi$ is dense in M.

Now if $p \in M \setminus \Pi$, then $\tilde{N}_p \cap \tilde{\Pi}$ is a projective hyperplane of $\tilde{N}_p$ and contains $\tilde{F}_p$. Since $\tilde{F}_p$ is an algebraic hypersurface of degree m in $\tilde{N}_p$, this means that $\tilde{F}_p$ must be just this hyperplane counted m times. Thus $M \setminus \Pi$ is totally umbilic, and M is also totally umbilic since $M \setminus \Pi$ is dense in M. When $m > 1$, we can use the standard result that M must then be an open subset of a round m-sphere or a flat m-plane [5]. The case $m = 1$ needs a special proof.

In this case we can assume that M is a curve given in the conventional way by a vector function $\mathbf{r}(s)$ of the arc-length. As usual we write $\mathbf{r}' = \mathbf{t}$ and, if $\mathbf{t}' \neq 0$, $\mathbf{t}' = \kappa\mathbf{n}$ where $\kappa > 0$. We will deal with the case when κ is not identically zero since clearly if $\mathbf{t}' = 0$ everywhere, then the curve M is a straight line. Let

M_0 be the set of points in M where $\mathbf{t}' = 0$. We first consider $M \setminus M_0$ where $\kappa > 0$, and write $\rho = \kappa^{-1}$ as usual. Let $r \in M \setminus M_0$, then $\tilde{F}_r$ does not belong to the hyperplane at infinity, so it is enough to suppose $F_r \subset \Pi$, which is then equivalent to $\tilde{F}_r \subset \tilde{\Pi}$. If we define $B_r \subset \mathbf{R}^n$ by $\mathbf{b} \in B_r$ if and only if $\langle \mathbf{t} \cdot \mathbf{b} \rangle = \langle \mathbf{n} \cdot \mathbf{b} \rangle = 0$, then $F_r = \mathbf{r} + \rho\mathbf{n} + B_r$. If $n = 3$ then, of course, B_r is just the 1-dimensional subspace determined by the binormal $\mathbf{b}$. The analogue of $\mathbf{n}' = \tau\mathbf{b} - \kappa\mathbf{t}$ in the case $n \geqslant 3$ can be written $\mathbf{n}' + \kappa\mathbf{t} \in B_r$; it is derived in a similar way.

The object now is to prove that F_r is constant by showing that $\mathbf{r} + \rho\mathbf{n}$ is constant, and then that B_r is constant.

We let Π be given by $\mathbf{x} \in \Pi$ if and only if $\langle \mathbf{x} - \mathbf{c} \cdot \mathbf{a} \rangle = 0$, where $\mathbf{a}$ is a unit vector. Then we are given that for all $\mathbf{b} \in B_r$, $\langle \mathbf{r} + \rho\mathbf{n} + \mathbf{b} - \mathbf{c} \cdot \mathbf{a} \rangle = 0$. This is equivalent to $\langle \mathbf{r} + \rho\mathbf{n} - \mathbf{c} \cdot \mathbf{a} \rangle = 0$ and $\langle \mathbf{t} \cdot \mathbf{b} \rangle = \langle \mathbf{n} \cdot \mathbf{b} \rangle = 0 \Rightarrow \langle \mathbf{b} \cdot \mathbf{a} \rangle = 0$. Thus differentiating this condition we see that

$$\langle \mathbf{t} + \rho'\mathbf{n} + \rho\mathbf{n}' \cdot \mathbf{a} \rangle = 0,$$

and $\langle \mathbf{t} \cdot \mathbf{u} \rangle + \kappa\langle \mathbf{n} \cdot \mathbf{b} \rangle = \langle \mathbf{n} \cdot \mathbf{u} \rangle + \langle \mathbf{n}' \cdot \mathbf{b} \rangle = 0 \Rightarrow \langle \mathbf{u} \cdot \mathbf{a} \rangle = 0$. Using the fact that $\mathbf{b}$ and $\mathbf{n}' + \kappa\mathbf{t} \in B_r$ and so $\langle \mathbf{b} \cdot \mathbf{a} \rangle = \langle \mathbf{n}' + \kappa\mathbf{t} \cdot \mathbf{a} \rangle = 0$, this becomes

$$\langle \rho(\mathbf{n}' + \kappa\mathbf{t}) + \rho'\mathbf{n} \cdot \mathbf{a} \rangle = \rho'\langle \mathbf{n} \cdot \mathbf{a} \rangle = 0,$$

and $\langle \mathbf{t} \cdot \mathbf{u} - \langle \mathbf{n}' \cdot \mathbf{b} \rangle \mathbf{n} \rangle = \langle \mathbf{n} \cdot \mathbf{u} - \langle \mathbf{n}' \cdot \mathbf{b} \rangle \mathbf{n} \rangle = 0 \Rightarrow \langle \mathbf{u} \cdot \mathbf{a} \rangle = 0$. But this also implies $\langle \mathbf{u} - \langle \mathbf{n}' \cdot \mathbf{b} \rangle \mathbf{n} \cdot \mathbf{a} \rangle = 0$, so we conclude that

$$\mathbf{b} \in B_r \Rightarrow \langle \mathbf{n}' \cdot \mathbf{b} \rangle \langle \mathbf{n} \cdot \mathbf{a} \rangle = \langle \mathbf{n}' + \kappa\mathbf{t} \cdot \mathbf{b} \rangle \langle \mathbf{n} \cdot \mathbf{a} \rangle = 0.$$

Taking $\mathbf{b} = \mathbf{n}' + \kappa\mathbf{t}$ we deduce that

$$\rho'\langle \mathbf{n} \cdot \mathbf{a} \rangle = \|\mathbf{n}' + \kappa\mathbf{t}\|^2 \langle \mathbf{n} \cdot \mathbf{a} \rangle = 0.$$

Now if $r \in M \setminus (M_0 \cup \Pi)$, then $\langle \mathbf{r} - \mathbf{c} \cdot \mathbf{a} \rangle \neq 0$. So $\langle \mathbf{n} \cdot \mathbf{a} \rangle \neq 0$. Hence if $r \in M \setminus (M_0 \cup \Pi)$, then $\kappa' = \|\mathbf{n}' + \kappa\mathbf{t}\| = 0$. Thus $\kappa' = 0$ and $(\mathbf{n} + \kappa\mathbf{r})' = 0$ for all $r \in M \setminus (M_0 \cup \Pi)$. This implies $\kappa' = 0$ and $(\mathbf{n} + \kappa\mathbf{r})' = 0$ for all $r \in M \setminus M_0$, since $M \setminus \Pi$ is dense in M. Using the principle that any nonconstant smooth function on a connected manifold must have regular values and observing that κ^2 is a smooth function on M with $(\kappa^2)' = 0$, we deduce that κ is a nonzero constant on M. Thus M_0 is empty, and $\mathbf{n} + \kappa\mathbf{r}$ is also constant on M. Hence $\mathbf{r} + \rho\mathbf{n}$ is constant on M.

Now consider the function $\langle \mathbf{t} \cdot \mathbf{d} \rangle^2 + \langle \mathbf{n} \cdot \mathbf{d} \rangle^2$ for any (constant) $\mathbf{d} \in \mathbf{R}^n$. Using $\mathbf{n}' = -\kappa\mathbf{t}$, $\mathbf{t}' = \kappa\mathbf{n}$ we see that this has zero derivative and hence is constant. So if we fix a point $r_0 \in M$, and let $\mathbf{t}_0, \mathbf{n}_0$ be the tangent and principal normal at that point, then $\mathbf{d} \in B_{r_0}$ if and only if $\langle \mathbf{t}_0 \cdot \mathbf{d} \rangle^2 + \langle \mathbf{n}_0 \cdot \mathbf{d} \rangle^2 = 0$ which means $\langle \mathbf{t} \cdot \mathbf{d} \rangle^2 + \langle \mathbf{n} \cdot \mathbf{d} \rangle^2 = 0$ and so $\mathbf{d} \in B_r$ for all $r \in M$. Thus B_r

is a constant $(n-2)$-dimensional subspace of $\mathbf{R}^n$. So $F_r = \mathbf{r} + \rho\mathbf{n} + B_r$ is a constant $(n-2)$-plane; call it Λ. We can now use Lemma 2.3 to conclude that M lies on a round circle with focal set Λ. This concludes the proof of Theorem 2.4.

Note that the method used for $m = 1$ can be used with suitable adjustments for the general case. It did not seem worth-while to make these adjustments since the result on totally umbilical manifolds is well-known. Moreover it is only the case $m = 1$ that is required for the later sections in this paper.

This completes the section on the general results we require. It is interesting to raise the question of what further results can be given on these lines. For instance we can prove that if $2m < n$ and there is a cone in $\mathbf{R}^n$ with cross-section an $(n-2)$-sphere which contains all the focal set of M, then M must lie on a round sphere. It would be interesting to know when $\eta(\Gamma)$ could be a manifold.

3. Totally focal spheres

In this section we consider some special properties of totally focal embeddings of spheres.

Theorem 3.1. *Let $M \subset \mathbf{R}^n$ be totally focal where M is homeomorphic to the sphere $\mathbf{S}^m$. Let $M' \subset N(M) \setminus \Gamma$ be any connected component of $\eta^{-1}(M)$. Then $\eta \mid M'\colon M' \to M$ is a homeomorphism, and $\pi \mid M'\colon M' \to M$ is of degree ± 1.*

Proof. The case when M' is the zero section is trivial, so we exclude this case from now on. If $m \neq 1$, then $M \cong \mathbf{S}^m$ is simply-connected, and M' is a connected covering space, so $\eta \mid M'$ must be a homeomorphism. If $m = 1$, so that M is a simple closed curve, then there will be a finite number of normals in $\eta^{-1}(p) \cap M'$ which we can call $(q_1, p), \cdots, (q_k, p)$ for any $p \in M$. Since $p, q_1, \cdots, q_k$ must be distinct we can order them by choosing a direction on the curve M, and this determines cross-sections $\sigma_i\colon M \to M'$, $i = 1, \cdots, k$ of the covering map $\eta \mid M'$ by putting $\sigma_i(p) = (q_i, p)$ where we suppose $p, q_1, \cdots, q_k$ is the given order. But if M' has a cross-section, it is trivial, so again $\phi \mid M'$ must be a homeomorphism.

Now consider $\pi \circ (\eta \mid M')^{-1}\colon M \to M$. This is a continuous map with no fixed point. Since any map of $\mathbf{S}^m$ into itself with no fixed points is homotopic to the antipodal map which has degree $(-1)^{m+1}$, we deduce that $\pi \circ (\eta \mid M')^{-1}$ and hence $\pi \mid M'$ have degree ± 1.

Theorem 3.2. *Let $M \subset \mathbf{R}^n$ be totally focal where M is homeomorphic to $\mathbf{S}^m$. Suppose that for some $x \in \mathbf{R}^n \setminus \eta(\Gamma)$, there are l points in $\eta^{-1}(x) \cap N^k(M)$. Then for any $p \in M$, $N_p^k \cap M$ contains at least l points.*

Proof. In [2] we showed that $\mathbf{R}^n \setminus \eta(\Gamma)$ is connected and contains M, and that each connected component of $N(M) \setminus \Gamma$ is a covering of $\mathbf{R}^n \setminus \eta(\Gamma)$. In particular the conditions imply that $N^k(M)$ is an l-fold covering of $\mathbf{R}^n \setminus \eta(\Gamma)$. Thus $\eta^{-1}(M) \cap N^k(M)$ must be an l-fold covering of M. But from Theorem 3.1 this covering must consist of l connected components $M_1, \cdots, M_l$ such that $\eta \mid M_i$ is a homeomorphism for $i = 1, \cdots, l$. Also $\pi \mid M_i$: $M_i \to M$ is of degree ± 1 and so, in particular, must be onto. Thus for any $p \in M$ there exist points $q_1, \cdots, q_l$ such that $(p, q_i) \in M_i \subset N^k(M)$. That is, $q_1, \cdots, q_l \in N_p^k \cap M$.

Note that this theorem can be interpreted in terms of distance functions. The conditions are equivalent to L_x having l critical points of index k, and the conclusion is that for any $p \in M$ there exist $q_1, \cdots, q_l \in M$ such that $L_{q_1}, \cdots, L_{q_l}$ all have p as a critical point of index k.

The next theorem can be interpreted in terms of height functions. It says that if one height function has a critical point of index k, then for every point $p \in M$ there is some height function which has p as a critical point of index k. With this interpretation it is seen to be similar to the above theorem. However there is some difficulty of defining the index in $\tilde{N}(M) \setminus \tilde{\Gamma}$, and our statement of Theorem 3.3 avoids this. It can be considered as a limiting case of Theorem 3.2, when x tends to a point at infinity.

Theorem 3.3. *Let $M \subset \mathbf{R}^n$ be totally focal where M is homeomorphic to $\mathbf{S}^m$. Let U be a component of $\tilde{N}(M) \setminus \tilde{\Gamma}$. Then for all $p \in M$ there is an infinite normal in $U \cap \pi^{-1}(p)$.*

Proof. By Theorem 3.2 we know that, for any $p \in M$, N_p^m is not empty since $N^m(M)$ cannot be empty. Also from Theorem 2.1 we know that it is an open convex subset of N_p. Since $N^m(M)$ is itself open in $N(M)$, this is sufficient to deduce that $N^m(M)$ is a fibre bundle over M with fibre which is solid. Hence there is a continuous cross-section μ: $M \to N^m(M)$; see [7, p. 55].

Now for each $p \in M$ all the intersections of the line in $\tilde{N}_p$ joining p to $\eta \circ \mu(p)$ with $\tilde{F}_p$ must lie on the segment, in N_p, between p and $\eta \circ \mu(p)$. So the segment of this line from p to the point at infinity which does not contain $\eta \circ \mu(p)$ lies entirely in N_p^0. This gives a map ϕ: $M \times I \to P^n \setminus \eta(\tilde{\Gamma})$ in which the image of (p, t) is $(p - tx)/(1 - t)$ if $\eta \circ \mu(p) = x$, $t \neq 1$ and $[x - p]$ if $\eta \circ \mu(p) = x$, $t = 1$.

Now consider $U \subset \tilde{N}(M) \setminus \tilde{\Gamma}$. This is a covering space of $\mathbf{P}^n \setminus \eta(\tilde{\Gamma})$ and so contains a connected covering space M' of M. By Theorem 3.1, $\eta \mid M'$ is a homeomorphism, and we can put $\psi_0 = (\eta \mid M')^{-1}$: $M \to M'$. Now we consider the covering map $\eta \mid U$ and apply the covering homotopy theorem to ψ_0 and ϕ. We obtain a lifting ψ: $M \times I \to U$ of ϕ with $\psi_0(p) = \psi(p, 0)$. The image $\psi(M \times \{1\})$ consists entirely of infinite normals, that is, $\psi(M \times \{1\}) \subset U \setminus N(M)$. Put $\psi_1(p) = \psi(p, 1)$ for all $p \in M$.

We now wish to show that for all $p \in M$, $\psi_1(M) \cap \pi^{-1}(p)$ is not empty. Now by Theorem 3.1, $\pi \circ \psi_0 \colon M \to M$ has degree ± 1, and by construction $\pi \circ \psi_0$ is homotopic to $\pi \circ \psi_1$. So $\pi \circ \psi_1$ also has degree ± 1 and thus must be onto. This is precisely what we required to prove the theorem.

4. Totally focal embeddings of $\mathbf{S}^2$ in $\mathbf{R}^4$

We now specialise the results in the previous section to the case when M is homeomorphic to $\mathbf{S}^2$ and $M \subset \mathbf{R}^4$.

Theorem 4.1. *Let $M \subset \mathbf{R}^4$ be totally focal where M is homeomorphic to $\mathbf{S}^2$. Then M is a round sphere lying in a hyperplane of $\mathbf{R}^4$.*

Proof. If $N^1(M)$ is empty, then no nondegenerate distance function can have a critical point of index 1. Hence they must each have just one maximum and one minimum. In other words the embedding is taut, and the result follows from [1].

We show that the above is the only case by assuming that $N^1(M)$ is not empty and getting a contradiction. By Theorem 3.2, if $N^1(M)$ is not empty, then N_p^1 is nonempty for all $p \in M$. Also $N^2(M)$ is not empty, and again by Theorem 3.2, N_p^2 is nonempty for all $p \in M$. So we have N_p^0, N_p^1 and N_p^2 are all nonempty. Thus since the closures of $N^1(M)$ and $N^0(M) \cup N^2(M)$ in $\tilde{N}(M) \setminus \tilde{\Gamma}$ are distinct components of $\tilde{N}(M) \setminus \tilde{\Gamma}$, Theorem 3.3 implies that these components intersect $\tilde{N}_p \setminus N_p$ for all $p \in M$. Thus $\tilde{N}_p^1$ and $\tilde{N}_p^0$, the closures of N_p^1 and $N_p^0 \cup N_p^2$ in $\tilde{N}_p \setminus \tilde{\Gamma}$, must intersect the line at infinity. This means that the conic $\tilde{F}_p$ in $\tilde{N}_p$ must intersect the line at infinity in distinct points. Also since every line through p intersects $\tilde{F}_p$ in real points, if $\tilde{F}_p$ is a proper conic, then p must lie inside it. Thus the polar line of p with respect to the conic $\tilde{F}_p$ intersects the line at infinity in $\tilde{N}_p^1$. Let $\tilde{L}_p$ be the projective line through p which intersects this polar line at infinity. Then the two points in which $\tilde{L}_p$ intersects $\tilde{F}_p$ are separated by p and the point at infinity on $\tilde{L}_p$, that is, there is one in each direction from p along $L_p = \tilde{L}_p \cap N_p$. Now $\bigcup \{L_p \colon p \in M\}$ is a vector line bundle over M and so must be trivial. Thus we can pick out a direction along L_p, and so find a cross-section of $N(M)$ which lies in Γ and has multiplicity 1. Therefore corresponding to this field of critical normals there is a tangent field of principal directions. But the tangent bundle of M has no 1-dimensional vector subbundle, and so we arrive at a contradiction. This proves the theorem.

Of course this theorem immediately raises the question as to whether there exists a totally focal embedding of $\mathbf{S}^2$ in $\mathbf{R}^5$ which does not have image a round sphere. Using Theorems 3.2 and 3.3 it is possible to show that the focal set F_p, which is a quadric, must be a cone in $\tilde{N}_p$. However, this does not seem to give enough information to construct a cross-section in $N^1(M)$, as in the proof of Theorem 4.1.

5. Totally focal embedding of S^1 in R^n

We need rather different techniques to deal with the case when M is a simple closed curve in $\mathbf{R}^n$. Thus the methods in this section tend to be rather ad hoc.

Theorem 5.1. *Let $M \subset \mathbf{R}^n$ be totally focal where M is a smooth simple closed curve. Then M is a round circle lying in a plane.*

Proof. The focal set in any normal projective $(n-1)$-plane $\tilde{N}_p$ consists of just a projective hyperplane. Thus $\tilde{F}_p$ is a projective $(n-2)$-plane in $\mathbf{P}^n$. But by Theorem 3.2, for all $p \in M$, N_p^1 is nonempty and therefore F_p is an $(n-2)$-plane in $\mathbf{R}^n$.

We divide the proof of this theorem into two cases. In the first case we suppose that M contains a semicircle A lying in a 2-plane. So there exist an $(n-2)$-plane Λ in $\mathbf{R}^n$ and $c \in \Lambda$ such that A is a semicircle centre c lying in the 2-plane through c orthogonal to Λ. So if p and q are the endpoints of A, then $N_p = N_q$. Also $F_r = \Lambda$ for all $r \in A$. The crucial observation is that the normal planes to A fill up $\mathbf{R}^n$, that is, $\mathbf{R}^n = \bigcup \{N_r : r \in A\}$. Thus if $x \in \eta(\Gamma)$ there exists some $r \in A$ such that $x \in N_r$ and hence $x \in F_r = \Lambda$. Thus $\eta(\Gamma) \subset \Lambda$, and so $F_r = \Lambda$ for all $r \in M$. Lemma 2.3 and Theorem 2.4 show that the curve M must be a plane circle with axis Λ, which must be the circle centre c since it contains the semicircle A.

The theorem will be proved by assuming next that M does not contain any plane semicircle and obtaining a contradiction. So suppose M contains no plane semicircle.

We have shown in Theorem 3.1 that $\eta^{-1}(M)$ consists of a finite number of connected components $M_0, M_1, \cdots, M_k$ say, such that $\eta \mid M_i$ is a homeomorphism for each i. Let $\sigma_i: M \to M$ be defined by $\sigma_i = \pi \circ (\eta \mid M_i)^{-1}$. Then $\sigma_0(p), \cdots, \sigma_k(p)$ are distinct points in M and are the feet of the normals through p. Since M is homeomorphic to $\mathbf{S}^1$, we can fix on an orientation and order these feet around the curve so that they are $\sigma_0(p), \sigma_1(p), \cdots, \sigma_k(p)$ in that order. Thus $\sigma_0(p) = p$ corresponds to the zero normal, $\sigma_1(p)$ must be a maximum point for the distance function L_p, $\sigma_2(p)$ will be a minimum point, and so on.

We now concentrate on σ_1 and observe that the length of the normal $\| p - \sigma_1(p) \|$ is a smooth function on M. We will show that it is constant. Let p be a point where this length is minimum or maximum, and put $q = \sigma_1(p)$.

Then clearly the line joining p to q is a double normal, that is, it is normal to M at both p and q. We let A be the arc of M from p to q in the agreed direction so that the distance function L_p has no critical points on A except for a minimum at p and a maximum at q. This means that F_p must intersect the segment pq at a point c which must be the only point in which $\eta(\tilde{\Gamma})$ intersects the projective line through these points. Let $\gamma: I \to \mathbf{P}^n \setminus \eta(\tilde{\Gamma})$ be a path mapping the unit interval into the segment of this projective line through p and q, which does not contain c. Given any starting point in $\eta^{-1}(p)$ this can be lifted to a unique path in $\tilde{N}(M) \setminus \tilde{\Gamma}$ with endpoint in $\eta^{-1}(q)$. Clearly if we start with $(p, p) \in \eta^{-1}(p)$, we obtain $\gamma_0: I \to \tilde{N}(M) \setminus \tilde{\Gamma}$ given by $\gamma_0(t) = (p, \gamma(t))$ so $\pi \circ \gamma_0(t) = p$ for all $t \in I$. Similarly if we start with $(q, p) \in \eta^{-1}(p)$, we obtain $\gamma_1: I \to \tilde{N}(M) \setminus \tilde{\Gamma}$ given by $\gamma_1(t) = (q, \gamma(t))$ so $\pi \circ \gamma_1(t) = q$ for all $t \in I$. Now suppose we start with some other point $(r, p) \in \eta^{-1}(p)$. We obtain $\gamma': I \to \tilde{N}(M) \setminus \tilde{\Gamma}$, and this path does not intersect γ_0 or γ_1 so $\pi \circ \gamma'(t) \neq p$ or q whatever $t \in I$. But since r is a critical point of the distance function L_p, it does not lie on A. Since $\pi \circ \gamma'(I)$ is connected, we conclude that $\pi \circ \gamma'(I) \subset M \setminus A$. In particular $\pi \circ \gamma'(I) \not\subset A$, and this means that the distance function L_q has no critical points on A except for p and q. This means that $p = \sigma_k(q)$.

Let $\Lambda = F_p \cap F_q$ so that Λ is a hyperplane in $N_p \cap N_q$ and passes through c. We now carry out a similar argument to the above. Suppose $x \in (N_p \cap N_q) \setminus \Lambda$ and lies on the same side of Λ as p. We join p to x by a path $\gamma: I \to (N_p \cap N_q) \setminus \Lambda$, and by the same argument as before deduce that L_x has no critical points on A apart from p and q. Further, since γ does not intersect the hyperplane at infinity, the lifted path must lie in $N^0(M)$ or $N^1(M)$. We deduce that p is a minimum for L_x and q is a maximum. Thus, for all $x \in (N_p \cap N_q) \setminus \Lambda$ lying on the same side of Λ as p, A must lie outside the open ball, centre x, radius $\|x - p\|$ and inside the closed ball, centre x, radius $\|x - q\|$. There is a similar argument with p and q interchanged.

The only way in which this is possible is that c is the mid-point of the segment pq in $\mathbf{R}^n$, Λ is perpendicular to pq, and A lies on the sphere, centre c, whose focal set is Λ. This means that $\Lambda \subset D_r$ for all $r \in A$. By Lemma 2.3, if $\dim \Lambda = n - 2$, then A is a plane semicircle, and we are assuming that M does not contain a plane semicircle so we assume $\dim \Lambda = n - 3$ and $F_p \neq F_q$. This means of course that $N_p \neq N_q$.

We next observe that if $r, s \in M$ are any two points, then $\tilde{N}_r \cap \tilde{N}_s$ is nonempty. In fact either $\tilde{N}_r = \tilde{N}_s$ or $\tilde{N}_r \cap \tilde{N}_s$ is a hyperplane of $\tilde{N}_r$. But $\tilde{F}_r$ is also a hyperplane of $\tilde{N}_r$ since $\dim M = 1$. Thus $\tilde{F}_r \cap \tilde{F}_s = \tilde{F}_r \cap (\tilde{N}_r \cap \tilde{N}_s)$ is nonempty and either $\tilde{F}_r = \tilde{F}_s$ or $\tilde{F}_r \cap \tilde{F}_s$ is a projective $(n - 3)$-plane which could lie in the hyperplane at infinity. Let us write $\tilde{\Lambda}$ for the projective

$(n-3)$-plane in $\mathbf{P}^n$ which contains Λ. Then $\tilde{\Lambda}$ is a hyperplane of $\tilde{F}_p$, and if $\tilde{F}_r \neq \tilde{F}_p$, then so is $\tilde{F}_p \cap \tilde{F}_r$. Thus either $\tilde{\Lambda} \subset \tilde{F}_r$ or $\tilde{\Lambda} \cap \tilde{F}_r$ is an $(n-4)$-plane.

Let us consider the case when for some $r \in M$, $\tilde{\Lambda} \cap \tilde{F}_r$ is an $(n-4)$-plane. Then for every $s \in A$, $\tilde{F}_s$ intersects $\tilde{F}_r$ in an $(n-3)$-plane and contains the $(n-3)$-plane $\tilde{\Lambda}$. Thus $\tilde{F}_s \cap \tilde{F}_r$ must contain both the $(n-4)$-plane $\Lambda \cap \tilde{F}_r$ and some point x in $\tilde{F}_r$ which is not in $\tilde{\Lambda}$. Thus $\tilde{F}_s$ contains both x and $\tilde{\Lambda}$, and must be the unique projective $(n-2)$-plane which contains both x and $\tilde{\Lambda}$. Thus $\tilde{F}_s$ must lie in the unique hyperplane of $\mathbf{P}^n$ which contains both $\tilde{\Lambda}$ and $\tilde{F}_r$. Call this $\tilde{\Pi}$.

So $\tilde{F}_s \subset \tilde{\Pi}$ for all $s \in A$, and therefore by Theorem 2.4, A lies on a circle or a line in $\mathbf{R}^n$. But A cannot lie on a line since pq is a double normal, so A is a plane semicircle, and this contradicts our assumption that M does not contain a semicircle.

This leaves the alternative that $\tilde{\Lambda} \subset \tilde{F}_r$ for all $r \in M$, and hence by Lemma 2.3, M lies on a 2-sphere or a 2-plane in $\mathbf{R}^n$. Since we are assuming that A is not a plane semicircle, M must lie on the 2-sphere with centre c and diameter $\|p-q\|$.

So we have shown that if p is a point such that $\|p-\sigma_1(p)\|$ is minimum or maximum, then $q=\sigma_1(p)$ and p are antipodal points on this sphere. Thus $\|r-\sigma_1(r)\|$ is constant for all $r \in M$, and $\|r-\sigma_1(r)\|$ is the diameter of the sphere.

Thus if $q=\sigma_1(p)$, then $p=\sigma_1(q)$. But we have shown that $p=\sigma_k(q)$. Hence $k=1$. This means that $\eta\colon N(M) \setminus \Gamma \to \mathbf{R}^n \setminus \eta(\Gamma)$ is a 2-fold covering. Hence every nondegenerate distance function L_x, $x \in \mathbf{R}^n \setminus \eta(\Gamma)$, has just two critical points, a minimum and a maximum. Thus M is taut and so must be a round circle in a 2-plane [1]. This contradicts the assumption that M contains no plane semicircle, and completes the proof of the theorem.

References

[1] S. Carter & A. West, *Tight and taut immersions*, Proc. London Math. Soc. **25** (1972) 701–720.
[2] ______, *Totally focal embeddings*, J. Differential Geometry **13** (1978) 251–261.
[3] ______, *A characterisation of isoparametric hypersurfaces in spheres*, to appear in J. London Math. Soc.
[4] T. E. Cecil & P. J. Ryan, *Tight spherical embeddings*, Lecture Notes in Math. Vol. 838, Springer, Berlin, 1979, 94–104.
[5] B. Y. Chen, *Geometry of submanifolds*, Marcel Dekker, New York, 1973.
[6] S. S. Chern & R. K. Lashof, *On the total curvature of immersed manifolds*, Amer. J. Math. **79** (1957) 306–318.
[7] N. Steenrod, *The topology of fibre bundles*, Princeton University Press, Princeton, 1951.
[8] R. J. Walker, *Algebraic curves*, Princeton University Press, Princeton, 1950.

UNIVERSITY OF LEEDS, ENGLAND