PACIFIC JOURNAL OF MATHEMATICS
Vol. 87, No. 2, 1980

UNIQUE BEST APPROXIMATION FROM A C^2-MANIFOLD IN HILBERT SPACE

THEAGENIS ABATZOGLOU

Given a C^2-manifold in a Hilbert space we examine whether a critical point of the distance function to the manifold is actually a global best approximation. We establish a criterium for the above in terms of the curvature of the manifold.

1. Introduction. In the last 10 years there have been papers, see [3], [4], [9], [10], [12], where the authors determine how close you have to be from a manifold to make sure that a critical point of the distance function to the manifold is a global best approximation. In their discussion the authors above use explicitly or implicitly the notions of curvature and radius of curvature and as long as the manifold does not bend back into itself too much, they use a global lower bound on the radius of curvature to conclude that if a point is within one third of the radius of curvature then it has a unique global approximation. In this paper we use different methods, mainly from differential geometry, to arrive at a sharp bound of one radius of curvature guaranteeing unique global best approximation.

2. Global best approximation to C^2-curves. We would like to establish some facts about C^2-curves which we will use later to obtain our results about global best approximation to C^2-manifolds. A version of Theorem 2.1b is known in n-dimensional Euclidean space and is due to Schwarz, see [7] page 38. Our proof holds for any Hilbert space and is different from the classical one.

We will need the following lemma to prove Theorem 2.1:

LEMMA 2.1. *Let x, y, z be in H, a Hilbert space, and assume that $\|y\| = 1$, $y \perp z$. Then:* $|(x, z)| \leqq \|z\| \sqrt{\|x\|^2 - (x, y)^2}$.

Proof. Write z as: $z = t(x - (x, y)y) + v$ where $v \perp x, y$. Then $\|z\|^2 = t^2(\|x\|^2 - (x, y)^2) + \|v\|^2$ so that: $|t| \leqq \|z\| / \sqrt{\|x\|^2 - (x, y)^2}$. We now estimate (x, z) and get

$$|(x, z)| = |t\,\|x\|^2 - t(x, y)^2| = |t(\|x\|^2 - (x, y)^2)|$$
$$\leqq \|z\| \sqrt{\|x\|^2 - (x, y)^2}\,.$$

THEOREM 2.1. *If γ is a C^2-curve embedded in H such that*

$\|\gamma(0)\| = R$, $\gamma(0) \perp \gamma'(0)$, $\|\gamma'(t)\| = 1$ *and* $\|\gamma''(t)\| \leqq 1/R$ *for all* t. *Then*

(a) $\|\gamma(t)\| \geqq R$ *for all* t, $|t| \leqq \pi R$

(b) $d/dt\,\|\gamma(t) - \gamma(0)\| \geqq 0$ *for* $0 < t \leqq \pi R$ *and* $\|\gamma(\pm\pi R) - \gamma(0)\| \geqq 2R$.

Proof. (a) We shall assume first that $\|\gamma(0)\| < R$. Let $F(t) = \|\gamma(t)\|^2$ and compute $F'(t)$ and $F''(t)$. We obtain

$$F'(t) = 2(\gamma'(t), \gamma(t))$$
$$F''(t) = 2 + 2(\gamma''(t), \gamma(t)) .$$

Since $\gamma'(0) \perp \gamma(0)$ and $\|\gamma''(t)\| \leqq 1/R$ we can assume there exists an interval $[0, t_1]$ where $F'(t) \geqq 0$ and $F'(t_1) = 0$ otherwise the proof is immediate. We now apply Lemma 2.1 with $x = \gamma(t)$, $y = \gamma'(t)$, $z = \gamma''(t)$ and obtain

$$\left|\frac{F''(t) - 2}{2}\right| \leqq \|\gamma''(t)\| \sqrt{F(t) - 1/4F'^2(t)}$$

or

$$2 - F'' \leqq 1/R\sqrt{4F - F'^2}$$

deleting the t's. Multiply the last inequality by F' in $[0, t_1]$ and get (1)

$$(1) \qquad F'(2 - F'') \leqq \frac{F'\sqrt{4F - F'^2}}{R}$$

which we rewrite as

$$\frac{F'(2 - F'')}{\sqrt{4F - F'^2}} \leqq \frac{F'}{R} .$$

Integrate from t to t_1 and get (2)

$$(2) \qquad \sqrt{4F(t_1)} - \sqrt{4F(t) - F'^2(t)} \leqq \frac{F(t_1) - F(t)}{R} .$$

Set $F(t_1) = a^2$ and $2a - a^2/R = b$, then (2) becomes (3)

$$(3) \qquad (b + F/R) \leqq \sqrt{4F - F'^2} .$$

We would like to show that $t_1 \geqq \pi R/2$. If we assume on the contrary that $t_1 < \pi R/2$ then it easily follows that $b + F/R \geqq 0$ for all t, t in $[0, t_1]$. We then square (3) and obtain

$$(4) \qquad F'^2 \leqq 4F - (b + F/R)^2 \text{ or } F' \leqq \sqrt{4F - (b + F/R)^2} ,$$

Integrate the last inequality from 0 to t_1 and get

$$\int_{R^2(1-\varepsilon)^2}^{a^2} \frac{dF}{\sqrt{4F-(b+F/R)^2}} \leqq \int_0^{t_1} dt$$

so that

$$R\sin^{-1}\left[\frac{2F(t)-(4R^2-2bR)}{\sqrt{(4R^2-2bR)^2-4b^2R^2}}\right]\Bigg|_{F(0)=(1-\varepsilon)^2R^2}^{F(t_1)=a^2} \leqq t_1 \,.$$

We evaluate the integral substituting $2a - a^2/R = b$, observing that when $t = 0$ equation (3) implies $F(t_1) = a^2 \geqq (1+\varepsilon)^2R^2$. So the above integral computes to:

$$R\left[\pi/2 - \sin^{-1}\left[\frac{2(1-\varepsilon)^2R^2 - 4R^2 + 4aR - 2a^2}{4R(a-R)}\right]\right] \leqq t_1$$

from which we obtain $R\pi/2 \leqq t_1$. This contradiction shows that $\|\gamma(t)\|$ is an increasing function on $[0, \pi R/2]$. In the case that $\|\gamma(0)\| = R$ we consider the curve $\delta(t) = \gamma(t) - \varepsilon\gamma(0)$ and let $\varepsilon \to 0$ to obtain once more that $\|\gamma(t)\|$ is an increasing function of t on $[0, \pi R/2]$. Now let $[0, t_2]$ be the largest interval where $\|\gamma(t)\| \geqq R$ with $\|\gamma(t_2)\| = R$. Suppose $F'(t) \leqq 0$ on $[t_3, t_2]$ and $F'(t_3) = 0$, then an entirely analogous argument shows that $t_2 - t_3 \geqq \pi R/2$. This concludes the proof of Theorem 2.1a.

(b) Now we shall prove Theorem 2.1b.

Define G by $G(t) = \|\gamma(t) - \gamma(0)\|^2$. Then

$G'(t) = 2(\gamma'(t), \gamma(t) - \gamma(0))$ and $G''(t) = 2 + 2(\gamma''(t), \gamma(t) - \gamma(0))$.

We apply Lemma 2.1 with $x = \gamma(t) - \gamma(0)$, $y = \gamma'(t)$, $z = \gamma''(t)$ and obtain,

$$|2 - G''(t)| \leqq 1/R\sqrt{4G(t) - G'^2(t)}\,.$$

Let us assume that $G'(t) \geqq 0$ on $[0, t_0]$; then we multiply the inequality by $G'(t)$ and get

$$(2 - G''(t))G'(t) \leqq 1/RG'(t)\sqrt{4G(t) - G'^2(t)}\,.$$

This last relation we rewrite as $(2 - G''(t))G'(t)/\sqrt{4G(t) - G'^2(t)} \leqq G'(t)/R$ and integrate from 0 to t, t in $[0, t_0]$ to obtain

$$\sqrt{4G(t) - G'^2(t)} \leqq G(t)/R$$

which we express as (5)

$$4G(t) - G'^2(t) \leqq G^2(t)/R^2\,. \tag{5}$$

Suppose there is t_0, $0 < t_0$, such that $0 < G'(t)$ for t in $(0, t_0)$ and $G'(t_0) = 0$. Then using (5) we get $4R^2 \leqq G(t_0)$. Set $c^2 = G(t_0)$ and consider the curve δ defined by

$$\delta(t) = \gamma(t) - \left(\frac{c - R}{c}\gamma(t_0) + \frac{R}{c}\gamma(0)\right).$$

Observe that $\delta(t_0) \perp \delta'(t_0)$ because $G'(t_0)=0$; also $\|\delta(t_0)\|=R$, $\|\delta'(t)\|=1$ and $\|\delta''(t)\| \leqq 1/R$ for all t. Then by Theorem 2.1a δ will stay outside the sphere of radius R centered at $((c-R)/c)\gamma(t_0) + (R/c)\gamma(0)$, on the interval $[t_0 - \pi R, t_0]$. We project δ radially into the above sphere; call the projected curve $\bar{\delta}$. Then the arc length of $\bar{\delta}$ from 0 to t_0 must be less or equal than t_0. However $\bar{\delta}(t_0)$ and $\bar{\delta}(0)$ are antipodal points of the sphere and by [11]: Theorem 5I, $t_0 \geqq \pi R$. We rewrite (5) as $\sqrt{4G(t) - G^2(t)/R^2} \leqq G'(t)$; this last inequality is equivalent to $2R\sqrt{1 - (G(t)/2R^2 - 1)^2} \leqq G'(t)$ which we integrate from 0 to t with $G(t) \leqq 4R^2$ we obtain:

$$t \leqq R[\sin^{-1}(G(t)/2R^2 - 1) - \sin^{-1}(-1)] \quad \text{i.e.,}$$
$$t \leqq R[\pi/2 + \sin^{-1}(G(t)/2R^2 - 1)] .$$

So then $G(t) = 4R^2$, t in $[0, t_0]$, t must be less or equal to πR. This implies now that $G(\pi R) \geqq 4R^2$.

We are almost ready now to apply the above theorem to best global approximation from curves. We need a definition: that of folding.

DEFINITION. Let γ be a curve in H. The folding of γ at $\gamma(t)$: $\phi(\gamma(t))$ is defined by

$$\phi(\gamma(t)) = \sup\{r \mid B(\gamma(t), s) \cap \gamma \text{ is connected for all } s \leqq r\}$$

where $B(x, r)$ is the ball of radius r centered at x. $\phi(\gamma(t))$ measures how much does γ twist back to $\gamma(t)$. We can now prove:

THEOREM 2.2. *Let γ be a C^2 curve embedded in a Hilbert space H; suppose $\|\gamma'(t)\| = 1$ and $\|\gamma''(t)\| \leqq 1/R$ for all t. Let x be in H and assume that $x - \gamma(0) \perp \gamma'(0)$. Then if $\|x - \gamma(0)\| < \min\{\phi(\gamma(0))/2, R\} = \mu$, $\gamma(0)$ is the best global approximation of x from γ.*

Proof. Suppose there is a $c \neq 0$ such that $\|x - \gamma(c)\| \leqq \|x - \gamma(0)\|$. Then it follows that $\|x - \gamma(c)\| < \mu$ and (1)

$$(1) \qquad \|\gamma(c) - \gamma(0)\| \leqq \|x - \gamma(c)\| + \|x - \gamma(0)\| < 2\mu .$$

It follows from Theorem 2.1a that $\|x - \gamma(0)\| < \|x - \gamma(t)\|$ for $0 < |t| < \pi R$; also by Theorem 2.1b, $\|\gamma(\pm\pi R) - \gamma(0)\| \geqq 2R$.

This forces $\pi R < |c|$ but since $B(\gamma(0), \|\gamma(c) - \gamma(0)\|) \cap \gamma$ is connected by hypothesis, we must have $\|\gamma(c) - \gamma(0)\| \geqq 2R$ which contradicts (1).

EXAMPLE. Let $\gamma(t) = 1/(1 - tx)$ in $L_2[-1, 1]$, $|t| < 1$.
Let us find the folding of γ at $\gamma(s) = 1/(1 - sx)$. Compute:

$$\begin{aligned}\frac{d}{dt}\|\gamma(t) - \gamma(s)\|^2 &= 2(\gamma(t) - \gamma(s), \gamma'(t))\\ &= 2\int_{-1}^{1}\left(\frac{1}{1 - tx} - \frac{1}{1 - sx}\right)\frac{x}{(1 - tx)^2}dx\\ &= 2(t - s)\int_{-1}^{1}\frac{x^2}{(1 - sx)(1 - tx)^3}dx\ .\end{aligned}$$

This shows that any point of γ has infinite folding. Now we need to compute the curvature. Define the arc length $s(t)$ by:

$$s(t) = \int_0^t \|\gamma'(t)\|\, dt = \int_0^t \sqrt{2/3\,\frac{1 + 3t^2}{(1 - t^2)^2}}dt\ ,$$

and use the chain rule to obtain

$$\frac{d^2\gamma}{ds^2} = \left(\frac{dt}{ds}\right)^2\frac{d^2\gamma}{dt^2} + \frac{d^2t}{ds^2}\,\frac{d\gamma}{dt}\ .$$

We also get the following formulas by straight forward computations:

$$\begin{aligned}\frac{dt}{ds} &= \sqrt{3/2\,\frac{(1 - t^2)^2}{1 + 3t^2}}\\ \frac{d^2\gamma}{dt^2} &= \frac{2x^2}{(1 - tx)^3}\\ \frac{dt^2}{ds^2} &= \frac{-9t(1 + t^2)(1 - t^2)^2}{(1 + 3t^2)^2}\end{aligned}$$

and

$$\frac{d\gamma}{dt} = \frac{x}{(1 - tx)^2}\ .$$

Next we compute $\|d^2\gamma/ds^2\|^2$ and after some work we find that

$$\left\|\frac{d^2\gamma}{ds^2}\right\|^2 = \frac{18}{5}\,\frac{(1 - t^2)}{(1 + 3t^2)^3}(5t^4 - 2t^2 + 1)$$

which has a maximum, for $|t| < 1$, equal to $18/5$ when $t = 0$.

By Theorem 2.2 all functions f in $L_2[-1, 1]$ whose distance from γ is less than $\sqrt{5/18} = 1/3\sqrt{5/2}$ have a unique best approximation from γ.

In conjunction to this example it is interesting to observe that if $f(x) = 1 + ax^2$ then $\gamma(0) = 1$ is the best global approximation to

f, for $|a| \leqq 5/6$ in which case $||f - 1|| = |a| \sqrt{2/5}$. Note that $||(1 + 5/6x^2) - 1|| = \sqrt{5/18}$. It is easy to see that if $a > 5/6$ then $1 + ax^2$ has at least 2 best approximations from γ. This suggests the sharpness of Theorem 2.2.

3. Metric curvature. Let M be a C^2 manifold embedded in a Hilbert space H and suppose $X(x_1, \cdots, x_n)$ is the inclusion map. Consider a point m in M and a vector v orthogonal to M at m, $||v||=1$.

DEFINITION. We define the directional curvature of M at m in the direction v to be:

$$\frac{1}{\rho(m, v)} = \max_{||w|| \leq 1} \frac{(Aw, w)}{(Bw, w)}$$

where

$$A = \left(\left(v, \frac{\partial^2 X}{\partial x_i \partial x_j}\right)\right)_{i,j=1}^n, \quad B = \left(\left(\frac{\partial X}{\partial x_i}, \frac{\partial X}{\partial x_j}\right)\right)_{i,j=1}^n.$$

$1/\rho(m, v)$ is also equal to the largest eigenvalue of $B^{-1}A$ and is called by Milnor in [8]: the maximum principal curvature of M at m in the direction v, $\rho(m, v)$ is called the radius of curvature at m in the direction v.

It was proved in [1] that if P is the metric projection of H into M and if P is continuous at x then $||P'(x)|| = \rho/(\rho - r)$ where $\rho = \rho(m, v)$ and $||r|| = ||x - P(x)||$.

We define the curvature, "metric curvature", of M at m by, $1/\rho(m) = \sup_v 1/\rho(m, v)$ where v is orthogonal to M at m and $||v||=1$.

LEMMA 3.1. *If γ is a geodesic on M then the curvature of γ at $\gamma(t)$ is less or equal to the metric curvature of M at $\gamma(t)$,* i.e,. $||\gamma''(t)|| \leqq 1/\rho(m)$.

Proof. Let γ be a geodesic on M, this means that $||\gamma'(t)|| = 1$ and $\gamma''(t) \perp M$ at $\gamma(t)$ for each t. Now the definition of directional curvature applied to γ would give us $1/\rho(\gamma(t), v) = (\gamma''(t), v)/||\gamma'(t)||^2$. It is easy to see now that the curvature (metric curvature) of γ at $\gamma(t)$ is equal to $||\gamma''(t)||$. On the other hand since γ lies on M

$$\gamma'(t) = \sum_{i=1}^n \frac{\partial X}{\partial x_i} \frac{dx_i}{dt}$$

and

$$\gamma''(t) = \sum_{j=1}^n \sum_{i=1}^n \frac{\partial^2 X}{\partial x_j \partial x_i} \frac{dx_i}{dt} \frac{dx_j}{dt} + \sum_{i=1}^n \frac{\partial X}{\partial x_i} \frac{d^2 x_i}{dt^2}.$$

From the last 2 equalities we obtain

$$1/\rho(\gamma(t), v) = \frac{(Aw, w)}{(Bw, w)} \quad \text{where} \quad w = \left(\frac{dx_1}{dt}, \cdots, \frac{dx_n}{dt}\right).$$

This means that,

$$1/\rho(\gamma(t), v) = \frac{(Aw, w)}{(Bw, w)} \leqq \max_{\|w\|=1} \frac{(Aw, w)}{(Bw, w)} = 1/\rho(m, v)$$

where

$$m = \gamma(t) .$$

We therefore conclude that if γ is a geodesic on M. Then the curvature of γ at $\gamma(t) = m$ is less or equal to the metric curvature of M at m.

4. The sectional curvature of M. In order to prove a sharper and generalized version of Theorem 2.2 we need to estimate the sectional curvature of M in terms of the metric curvature $1/\rho(m)$. Unfortunately we cannot use formulas for the sectional curvature which hold when M is embedded in R^m; so we have to rederive these formulas for n-dimensional manifolds embedded in a Hilbert space. The derivation follows [5], page 10.

We consider the set $\{\partial X/\partial x_i\}_{i=1}^n$ consisting of linearly independent vectors spanning the tangent space of M at each point. Using this set as a basis we let,

$$Y = \sum_{i=1}^n f_i \frac{\partial X}{\partial x_i} \text{ and } Z = \sum_{j=1}^n g_j \frac{\partial X}{\partial x_j}$$

where f_i and g_j are smooth real valued functions.

We now compute $\nabla'_Y Z$ where ∇' stands for the usual gradient.

$$\begin{aligned} \nabla'_Y Z &= \sum_{i=1}^n f_i \nabla'_{\partial X/\partial x_i}\left(\sum_{j=1}^n g_j \frac{\partial X}{\partial x_j}\right) = \sum_{i=1}^n f_i \sum_{j=1}^n \left(\frac{\partial g_j}{\partial x_i} \frac{\partial X}{\partial x_j} + g_j \frac{\partial^2 X}{\partial x_i \partial x_j}\right) \\ &= \sum_{i,j=1}^n \left(f_i \frac{\partial g_i}{\partial x_j} \frac{\partial X}{\partial x_j} + f_i g_j \frac{\partial^2 X}{\partial x_i \partial x_j}\right). \end{aligned}$$

Define $\nabla_Y Z$ by, $\nabla_Y Z = P(\nabla'_Y Z)$ where P is the orthogonal projection onto the tangent subspace spanned by $\{\partial X/\partial x_i\}_{i=1}^n$. Then we may write, $\nabla'_Y Z = \nabla_Y Z + \alpha(Y, Z)$ where $\alpha(Y, Z) \perp \partial X/\partial x_i$, $i=1, \cdots, n$. Through a straightforward calculation it can be verified that $\nabla_Y Z$ is covariant differentiation for the Riemannian connection of M; so that now we have the means of computing the sectional curvature of M.

Take any 2 orthonormal vectors on the tangent space of M at m; by an appropriate choice of normal coordinates we may express these vectors as $\partial X/\partial x_1$, $\partial X/\partial x_2$ with $\partial^2 X/\partial x_1^2$ and $\partial^2 X/\partial x_2^2 \perp M$ at m. We want to find out the sectional curvature of the section spanned by $\partial X/\partial x_1$ and $\partial X/\partial x_2$. That is we have to compute $(R(\partial X/\partial x_1, \partial X/\partial x_2)\partial X/\partial x_1, \partial X/\partial x_2)$ where $R(\partial X/\partial x_1, \partial X/\partial x_2)\partial X/\partial x_1$ is the curvature tensor of M. By definition:

$$R\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right)\frac{\partial X}{\partial x_1} = \nabla_{\partial X/\partial x_2}\left(\nabla_{\partial X/\partial x_1}\left(\frac{\partial X}{\partial x_1}\right)\right) - \nabla_{\partial X/\partial x_1}\left(\Delta_{\partial X/\partial x_2}\left(\frac{\partial X}{\partial x_1}\right)\right) + \nabla_{[\partial X/\partial x_1, \partial X/\partial x_2]}\frac{\partial X}{\partial x_1}$$

where $[Y, Z]$ is the usual Lie bracket. Let us remind the reader that when M is a surface in R^3 its sectional curvature at a point is usually called Gaussian curvature and it is equal to the product of the principal curvatures of M at the same point. Therefore the sectional curvature of a surface is less or equal to $(1/\rho(m))^2$.

LEMMA 4.1. *The sectional curvature of M at m is less or equal to* $(1/\rho(m))^2$.

Proof. First we perform a few calculations to obtain the curvature tensor.

By the definition of $\nabla'_Y Z$ we have

$$\frac{\partial^2 X}{\partial_{x_1^2}} = \nabla'_{\partial X/\partial x_1}\left(\frac{\partial X}{\partial x_1}\right) = \nabla_{\partial X/\partial x_1}\left(\frac{\partial X}{\partial x_1}\right) + \alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_1}\right).$$

Next we use this equality to get

$$\nabla_{\partial X/\partial x_2}\left(\nabla_{\partial X/\partial x_1}\left(\frac{\partial X}{\partial x_1}\right)\right) + \alpha\left(\frac{\partial X}{\partial x_2}, \nabla_{\partial X/\partial x_1}\left(\frac{\partial X}{\partial x_1}\right)\right)$$
$$= \nabla'_{\partial X/\partial x_2}\left(\nabla_{\partial X/\partial x_1}\left(\frac{\partial X}{\partial x_1}\right)\right) = \frac{\partial}{\partial x_2}\left(\frac{\partial^2 X}{\partial x_1^2} - \alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_1}\right)\right).$$

By an analogous computation we also obtain

$$\nabla_{\partial X/\partial x_1}\left(\nabla_{\partial X/\partial x_2}\left(\frac{\partial X}{\partial x_1}\right)\right) + \alpha\left(\frac{\partial X}{\partial x_1}, \nabla_{\partial X/\partial x_2}\left(\frac{\partial X}{\partial x_1}\right)\right)$$
$$= \frac{\partial}{\partial x_1}\left(\frac{\partial^2 X}{\partial x_2 \partial x_1} - \alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right)\right).$$

Furthermore we have $[\partial X/\partial x_1, \partial X/\partial x_2] = 0$.

The last 3 equalities imply now that

$$\left(R\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right)\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right) = \left(\frac{\partial}{\partial x_1}\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right), \frac{\partial X}{\partial x_2}\right)$$
$$- \left(\frac{\partial}{\partial x_2}\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_1}\right), \frac{\partial X}{\partial x_2}\right).$$

To simplify the last expression we observe that $\alpha(\partial X/\partial x_1, \partial X/\partial x_2) \perp \partial X/\partial x_2$ which we use to obtain $\partial/\partial x_1(\alpha(\partial X/\partial x_1, \partial X/\partial x_2), \partial X/\partial x_2) = 0$.

In turn this last equality yields

$$\left(\frac{\partial}{\partial x_1}\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right), \frac{\partial X}{\partial x_2}\right) = -\left(\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right), \frac{\partial^2 X}{\partial x_1 \partial x_2}\right)$$
$$= -\left\|\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right)\right\|^2.$$

Similar calculations lead to

$$-\left(\frac{\partial}{\partial x_2}\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_1}\right), \frac{\partial X}{\partial x_2}\right) = \left(\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_1}\right), \frac{\partial^2 X}{\partial x_2^2}\right)$$
$$= \left(\frac{\partial^2 X}{\partial x_1^2}, \frac{\partial^2 X}{\partial x_2^2}\right)$$

because both $\partial^2 X/\partial x_1^2$ and $\partial^2 X/\partial x_1^2$ are orthogonal to M at m.

By Lemma 3.1 we have

$$\left\|\frac{\partial^2 X}{\partial x_1^2}\right\| \leqq 1/\rho(m), \quad \left\|\frac{\partial^2 X}{\partial x_2^2}\right\| \leqq 1/\rho(m).$$

So finally we obtain the estimate,

$$\left(R\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right)\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right) = \left(\frac{\partial^2 X}{\partial x_1^2}, \frac{\partial^2 X}{\partial x_2^2}\right)$$
$$- \left\|\alpha\left(\frac{\partial X}{\partial x_1}, \frac{\partial X}{\partial x_2}\right)\right\|^2 \leqq (1/\rho(m))^2.$$

5. **Folding.** We extend the concept of folding to n-dimensional manifolds in Hilbert space. Define the folding $\phi(m)$ of M at m by

$$\phi(m) = \sup\{r \mid B(m, s) \cap M \text{ is connected for every } s \leqq r\}.$$

We proceed with a couple of properties of folding and an example.

LEMMA 5.1. *Suppose for every p in $B(m, r) \cap M$, $r > 0$, there exists a continuous curve γ_p in M such that $\gamma_p(0) = m$, $\gamma_p(1) = p$ and $\|\gamma_p(t) - m\|$ is increasing for $0 \leqq t \leqq 1$, then $\phi(m) \geqq r$.*

Proof. Clearly any such curve γ_p lies in $\overline{B(m, \|p - m\|)}$. For any $s \leqq r$ consider p, q in $B(m, s)$. Join both p, q to m by curves γ_p and γ_q which have the above properties. Then it is clear you can join p to q by a continuous curve within $B(m, s)$.

EXAMPLE. Let M be the manifold of fractional linear transformations: $(ax + b)/(cx + 1)$, $|c| < 1$, in $L_2[-1, 1]$. Fix a point $(\bar{a}x + \bar{b})/(\bar{c}x + 1) = m$ and take any $(ax + b)/(cx + 1)$ in M. Assume that $c \neq \bar{c}$.

Consider the curve $\gamma(t) = ([\bar{a} + (a - \bar{a})t]x + \bar{b} + (b - \bar{b})t)/([\bar{c} + (c - \bar{c})t]x + 1)$ in $L_2[-1, 1]$. Clearly $\gamma(0) = (\bar{a}x + \bar{b})/(\bar{c}x + 1)$ and $\gamma(1) = (ax + b)/(cx + 1)$. Now we compute the following,

$$\frac{d}{dt} ||\gamma(t) - \gamma(0)||^2 = 2(\gamma(t) - \gamma(0), \gamma(t))$$

$$= 2t\int_{-1}^{1} \frac{[(\bar{c}a - \bar{a}c)x^2 + [(a - \bar{a}) + \bar{c}b - \bar{b}c]x + b - \bar{b}]^2 dx}{(\bar{c}x + 1)[[\bar{c} + (c - \bar{c})t]x + 1]^3} > 0 .$$

By Lemma 5.1 we have now $\phi(m) = \infty$ for any m in M.

The next theorem gives a geometric illustration of $\phi(m)$.

THEOREM 5.1. *Suppose* $\overline{B(m, \phi(m) + \varepsilon)} \cap M$ *is compact for some* $\varepsilon > 0$. *Let* $X = X(x)$ *be the embedding of* M *into* H *where* $x = (x_1, \cdots, x_n)$. *Then* $\phi(m) \geqq \inf_{\bar{x}} ||m - X(\bar{x})||$ *where* $\bar{x}$ *is a critical point of* $F(x) = ||m - X(x)||$.

Proof. Assume F has no critical points in $\overline{B(m, \phi(m))}$. Applying the contrapositive of the definition of $\phi(m)$ and using pathwise connectedness we obtain a sequence $\{p_k\}$ with p_k in M, $||p_k - m|| \geqq \phi(m)$, and $\lim_{k\to\infty} ||p_k - m|| = \phi(m)$ such that p_k cannot be connected to m by a continuous curve within $\overline{B(m, ||p_k - m||)} \cap M$. Using the compactness hypothesis we select a subsequence of $\{p_k\}$ which we call again $\{p_k\}$, with the property that p_k converges to $p \in M$ and $||p - m|| = \phi(m)$. Set $p = X(x)$. By assumption there is a $y = (y_1, \cdots, y_n)$ such that $F'(x)(y) < 0$; also since X is locally a diffeomorphism there is a sequence $\{x_k\}$ in R^n with the property that $X(x_k) = p_k$ and $\lim_{k\to\infty} x_k = x$. By the continuity of F' we now have, $F'(x_k + ty)(y) < -\varepsilon$ for some fixed $\varepsilon > 0$, for all $k \geqq N$ and $0 \leqq t < \delta$, with appropriately chosen N, δ.

To complete the proof we need to estimate $F(x_k + ty) = ||m - X(x_k + ty)||$, we have $F(x_k + ty) = F(x_k + ty) - F(x_k) + F(x_k) - F(x) + F(x) = tF'(x_k + t_k y)(y) + F(x_k) - F(x) + F(x)$ with $0 < t < \delta, 0 \leqq t_k \leqq t$. If we now take k sufficiently large we see that $F(x_k + ty) < F(x)$ for a fixed $0 < t < \delta$. Also $F(x_k + sy) - F(x_k) = sF'(x_k + \bar{s}y)(y) < 0$ for $0 < s < t < \delta$.

This means that $p_k = X(x_k)$ can be joined to $X(x_k + ty)$, which lies in $B(m, \phi(m)) \cap M$, by a curve $\gamma(s) = X(x_k + sy)$, $0 \leqq s \leqq t$, which lies in $\overline{B(m, \|p_k - m\|)} \cap M$. Then we can connect $X(x_k + ty)$ to m within $B(m, \phi(m)) \cap M$ which shows that $X(x_k) = p_k$ is pathwise connected to m inside $\overline{B(m, \|p_k - m\|)} \cap M$, which is a contradiction.

6. Unique global best approximations to C^2-manifolds. In this section we prove our main theorem. We will use the exponential map, $\exp_m$, which is defined in the following way:

For m in M and v in the tangent space of M at m define $\exp_m(v) = \gamma(1)$ where γ is a geodesic on M such that $\gamma(0) = m$ and $\gamma'(0) = v$. Here $\|v\|$ is the arc length between $\gamma(0)$ and $\gamma(1)$.

We say M is geodesicially complete if every geodesic can be infinitely extended. This property of M implies that any 2 points of M can be joined by a geodesic.

While proving Theorem 6.1 we will make use of the following:

The Morse-Schoenberg Comparison Theorem. Let M be a Riemannan manifold of dimension n, and $\gamma: [0, L] \to M$ a geodesic parametrized by arc length. Let $R > 0$ be a constant. Then if all sectional curvatures of M along γ are less or equal to $1/R^2$ and γ has length $L < \pi R$, then γ has no conjugate points.

THEOREM 6.1. *Let M be a C^2, complete, connected, n-dimensional manifold embedded in a Hilbert space H. Suppose x is in H, m in M and $x - m \perp M$ at m. Assume $1/R \geqq \sup_{m' \in M} \{1/\rho(m') \mid \|m' - m\| \leqq 2R\}$. Then if $\|x - m\| < \min\{R, \phi(m)/2\} = \mu$, m is the unique best approximation from M to x.*

Proof. On the contrary suppose there is p in M, $p \neq m$ such that $\|x - p\| \leqq \|x - m\| < \mu$, then $\|p - m\| \leqq \|x - p\| + \|x - m\| < 2\mu$.

If we can now join m to p by a geodesic within $B(m, 2R)$ then an application of Theorem 2.1, together with Lemma 3.1, will contradict the assumption $\|x - p\| \leqq \|x - m\|$. By the definition of folding there exists a continuous curve c in $B(m, 2\mu)$ such that $c(0) = m$ and $c(1) = p$.

Let $s = \sup\{t \mid c(t)$ can be joined by a geodesic to m within $B(m, 2\mu)\}$. $s > 0$ because the exponential map is a local diffeomorphism at m. We will show that $s = 1$. Let $\{\gamma_k\}$ be a sequence of geodesics lying inside $B(m, 2\mu)$ and joining m to $c(t_k)$ where $t_k \nearrow s$. Express γ_k in terms of $\exp_m$, i.e.: $\gamma_k(t) = \exp_m(v_k t)$, with $c(t_k) = \exp_m(v_k)$. By Theorem 2.1b the arc lengths $\|v_k\|$ are bounded in norm, therefore there is a converging subsequence which we call again $\{v_k\}$ such that $v_k \to v$. Then by continuity of the $\exp_m$, see

[6] page 69, we have $\lim_{k\to\infty} \exp_m(v_k t) = \exp_m(vt)$ for all t in $[0, 1]$, and $\gamma(t) = \exp_m(vt)$ is a geodesic within $B(m, 2\mu)$ such that $\gamma(0)=m$ and $\gamma(1) = \exp_m(v) = c(s)$.

Recall that by Lemma 4.1 and by hypothesis, the sectional curvature of M in $B(m, 2R)$ M is less or equal to $1/R^2$. We also know from Theorem 2.1b that the arc length of γ from $\gamma(0)=m$ to $\gamma(1) = c(s)$ is less than πR. Now we apply the Morse-Schoenberg Comparison Theorem, see [12], page 344, to conclude that $\exp_m(v)$ is not a conjugate point of m along the geodesic $\exp_m(tv)$, $0 \leqq t \leqq 1$. Therefore by [8]: Theorem 18.1, $\exp_m$ is not critical at v, in other words $\exp_m$ is a diffeomorphism in a neighborhood of v. This now forces $s = 1$ and completes the proof of the theorem.

It is easy to verify the sharpness of this last theorem by considering a circle or a parabola in R^2.

References

1. Theagenis Abatzoglou, *The minimum norm projection on C^2-manifolds in R^n*, Trans. Amer. Math. Soc., **243** (1978), 115–122.
2. R. Bishop and R. Crittenden, *Geometry of manifolds*, Pure and Appl. Math., Vol. 15, Academic Press, New York and London, 1964.
3. C. K. Chui, E. R. Rozema, P. W. Smith and J. D. Ward, *Metric curvature, following and unique best approximation*, SIAM J. Math. Anal. Vol. F, No. **3**, (1976), 436–499.
4. C. K. Chui and P. W. Smith, *Unique best nonlinear approximation in Hilbert spaces*, Proc. Amer. Math. Soc., **49** (1975), 66–70.
5. S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. II, Interscience Publishers, New York-London-Sydney, 1969.
6. S. Lang, *Introduction to Differentiable Manifolds*, Interscience Publishers, John Wiley, New York.
7. MAA Studies in Mathematics, *Studies in Global Geometry and Analysis*, Vol. 4, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1967.
8. J. Milnor, *Morse theory*, Annals of Math. Studies, No. 51, Princeton University Press, 1963.
9. J. Rice, *Approximation of Functions*, Vol. II, Addison Wesley, Reading, Mass., 1969.
10. E. Rozema and P. W. Smith, *Nonlinear approximation in uniformly smooth Banach spaces*, Trans. Amer. Math. Soc., **188** (1974), 199–211.
11. J. J. Schaffer, *Geometry of Spheres in Normed Spaces*, Marcel Dekker, Inc., New York and Basel, 1976.
12. M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Vol. IV, Publish or Perish, Inc., Berkeley, 1975.
13. J. M. Wolfe, *On the unicity of nonlinear approximation in smooth spaces*, J. Approximation Theory, **12** (1974), 165–181.

Received December 20, 1978 and in revised from June 25, 1979.

IOWA STATE UNIVERSITY
AMES, IA 50011