ANNALS OF MATHEMATICS
Vol. 63, No. 1, January, 1956
Printed in U.S.A.

THE IMBEDDING PROBLEM FOR RIEMANNIAN MANIFOLDS

BY JOHN NASH

(Received October 29, 1954)

(Revised August 20, 1955)

Introduction and remarks

History. The abstract concept of a Riemannian manifold is the result of an evolution in mathematical attitudes [1, 2]. In an earlier period mathematicians thought more concretely of surfaces in 3-space, of algebraic varieties, and of the Lobatchevsky manifolds. As the more abstract view of manifolds came into favor a question naturally arose: To what extent are the abstract Riemannian manifolds a more general family than the sub-manifolds of euclidean spaces?

This question has been considered in various specializations and with assorted side conditions. In 1873 Schlaefli [3] discussed the local form of this imbedding problem. He conjectured that a neighborhood in an n-manifold would generally require an imbedding space of $(n/2)\ (n + 1)$ dimensions. In 1901 Hilbert [4] obtained a negative result, showing that the Lobatchevsky plane is not realizable as a smooth surface in E^3. Some contemporary negative theorems are due to Tompkins [5] and to Chern and Kuiper [6]. For example, a flat n-torus is not realizable in less than $2n$ dimensions.

Janet [7] solved the local problem for two-manifolds with analytic metric in 1926, and Cartan [8] immediately extended the result to n-manifolds, treating it as an application of his theory of Pfaffian forms. The dimensionality requirement was $(n/2)\ (n + 1)$, as conjectured by Schlaefli. This number is a plausible one, being the number of components of the metric tensor. The proof depended on power series development, so it was limited to local results and it required that the metric be analytic.

There are some theorems on the existence of isometric imbeddings in infinite dimensional spaces. This is a much simpler problem.

A recent discovery [9, 10] is that C^1 isometric imbeddings of Riemannian manifolds can be obtained in rather low dimensional spaces. At first glance some of these C^1 results seem inconsistent with the negative theorems, such as Hilbert's. Apparently C^1 imbeddings are very different from the smoother ones.

Until recently the only general results on imbeddings in the large were proved for the problem of Weyl. This problem is to realize in E^3 all two-manifolds with everywhere positive Gaussian curvature. Alexandrov [13] and Pogorelov [14] have been successful with a geometrical approach based on polyhedral approximations. H. Lewy [12] and L. Nirenberg [15] have treated the problem from the viewpoint of partial differential equations. These results can probably be sharpened with respect to differentiability, but dimension-wise they are clearly optimal.

Rigidity theory concerns the metric preserving perturbations of an imbedding.

A closed convex surface in E^3 is rigid, because it admits only trivial perturbations. But it becomes flexible if there is a hole in it. Apparently rigidity disappears completely when the imbedding space has enough dimensions.

Arrangement of this paper. There are four main divisions, called Parts A, B, C, and D. At the end of Part C the treatment of compact manifolds is complete and we state Theorem 2, which is essentially this: *Every compact Riemannian n-manifold is realizable as a sub-manifold of Euclidean* $(n/2)(3n + 11)$*-space.* In Part D this is made to apply to non-compact manifolds by means of a device which reduces the non-compact problem to the compact case. The device is extravagant with dimensions. Theorem 3 realizes non-compact n-manifolds in $(n/2)(n + 1)(3n + 11)$ dimensions.

The core of this paper is in Part B. There a perturbation process is developed and applied to construct a small finite perturbation of an imbedding such that the perturbed imbedding induces a metric that differs by a specified small amount from the metric induced by the original imbedding. This work is summarized at the end of Part B in Theorem 1. The interesting thing about the perturbation process is that it does not seem special to this imbedding problem. It may be an illustration of a general method applicable to a variety of problems involving partial differential equations.

Part A is devoted to the fairly straight-forward construction of a smoothing operator of the type required by the method of Part B. The operator's main properties are stated in equations (A15, 16, 17). In general, the four Parts are relatively independent in notation. Each depends only on the main results of the preceding part, not on the details.

Remarks. Some respects in which the results here should be improvable are these: The dimension bounds for the imbedding space should be lowered; the C^2 case should be included; and it should be proved that the process gives an analytic imbedding when the metric is analytic. The treatment of the C^2 and analytic cases would require new sets of estimates. A more unified approach to the problem which would not require the use of two separate sets of imbedding functions might reduce the dimension requirements substantially.

The methods used here may prove more fruitful than the results. Time will tell how much can be done with smoothing and "feed-back" methods like those applied in Part B. The device of Part D suggests an alternative way to imbed general two-manifolds by exploiting the results on Weyl's problem.

Acknowledgement. I am profoundly indebted to H. Federer, to whom may be traced most of the improvement over the first chaotic formulation of this work. N. Levinson also gave very helpful advice and information; and through constructive criticism, several others at M.I.T. helped me to improve the paper. This paper was supported in part by the Office of Naval Research.

Part A: A general smoothing operator

This part develops an analytical tool, the smoothing operator, which is essential to the perturbation process developed in Part B. A smoothing operator is first constructed to act on real functions of n-variables, that is functions on E^n. Then we define a smoothing operator for a manifold $\mathfrak{M}$ by imbedding $\mathfrak{M}$ in E^n and extending scalar functions on $\mathfrak{M}$ to functions on E^n, which can be smoothed by the E^n operator. Finally, we devise a canonical representation for tensors on $\mathfrak{M}$ in terms of sets of scalars and use this to smooth tensors. We also obtain three important general inequalities that describe the action of smoothing on functions.

Smoothing functions on E^n

Real functions of n-variables are smoothed by convolution with a certain kernel, K_θ, that we define below. Here θ is a parameter controlling the degree of smoothing. K_θ is defined by defining its Fourier transform $\bar{K}_\theta$.

Let $\psi(u)$ be a C^∞ function such that

$$\begin{aligned} &\text{for } u \leqq 1: && \psi(u) = 1, \\ &\text{for } 1 \leqq u \leqq 2: && \psi(u) \text{ is monotone decreasing,} \\ &\text{for } u \geqq 2: && \psi(u) = 0. \end{aligned}$$

To illustrate, we could take $\psi = e^{(e^{(1/1-u)}/u-2)}$ in the range $1 < u < 2$.

Suppose $x_1, x_2, \cdots n_n$ are the coordinates of E^n and $\xi_1, \xi_2, \cdots, \xi_n$ are the corresponding coordinates of the Fourier transform space. We define the transform $\bar{K}_\theta$ of the kernel as

$$\text{(A1)} \qquad \begin{aligned} \bar{K}_\theta &= \psi(\xi/\theta), \qquad \text{where} \\ \xi &= (\xi_1^2 + \xi_2^2 + \cdots \xi_n^2)^{\frac{1}{2}}. \end{aligned}$$

Thus $\bar{K}_\theta$ is a spherically symmetric non-negative C^∞ function which is 1 inside the sphere $\xi = \theta$, zero outside the sphere $\xi = 2\theta$, and smoothly decreasing with ξ in the annular region between the two spheres.

K_θ is the transform of $\bar{K}_\theta$; so K_θ is spherically symmetric; it is real because $\bar{K}_\theta$ is even; it is analytic because $\bar{K}_\theta$ vanishes, except for $\xi < 2\theta$; and $|K_\theta|$ decreases as rapidly as any negative power of the distance because all derivatives of $\bar{K}_\theta$ are continuous.

As θ varies K_θ will be more or less concentrated at the origin. But the integral over all of E^n of K_θ will always be the same. Since variation of θ changes $\bar{K}_\theta$ only in a way corresponding to a change of scale in the transform space, K_θ will change in a similar way, with normalization being preserved. Specifically, we can relate K_θ to K_1 (where $\theta = 1$) by the equation

$$\text{(A2)} \qquad K_\theta(x_1, x_2, \cdots x_n) = \theta^n K_1(\theta x_1, \theta x_2, \cdots \theta x_n).$$

Effect of convolution on derivatives

Convolution and differentiation commute under favorable conditions:

$$\text{(A3)} \qquad \frac{\partial}{\partial x_i}(K_\theta * f) = \left(\frac{\partial}{\partial x_i} K_\theta\right) * f.$$

In our applications of this f will vanish outside a compact region. Since K_θ (and also its derivatives, as we see below) decreases rapidly away from the origin, the favorable conditions will be more than met.

To consider a derivative of K_θ, observe that $\partial K_\theta / \partial x_i$, which we abbreviate to $K_{\theta,i}$ satisfies

$$\overline{K_{\theta,i}} = \xi_i (-1)^{\frac{1}{2}} \bar{K}_\theta .$$

Thus the transform of $K_{\theta,i}$ is a C^∞ function and vanishes outside the sphere $\xi = 2\theta$. So $| K_{\theta,i} |$ decreases rapidly in the same manner as $| K_\theta |$, and the same is true for the higher derivatives of K_θ.

From (A2) we see that

$$K_{\theta,i}(x_1, \cdots, x_n) = \theta^{n+1} K_{1,i}(\theta x_1, \cdots, \theta x_n),$$

so (A3) can be written as

$$\begin{aligned}(K_\theta * f)_{,i} &= \theta^{n+1} K_{1,i}(\theta x_i, \cdots, \theta x_n) * f \\ &= \theta^{n+1} \int \cdots \int K_{1,i}(\theta y_1, \cdots, \theta y_n) f(x_1 - y_1, \cdots, x_n - y_n)\, dy_1 \cdots dy_n \\ &= \theta \int \cdots \int K_{1,i}(z_1, \cdots, z_n) f((x_1 - z_1/\theta), \cdots, (x_n - z_n/\theta))\, dz_1 \cdots dz_n .\end{aligned}$$

Here we considered the integral formula for convolution, letting the kernel carry the dummy variables $y_1, y_2, \cdots, y_n$. Then we made a change of variables: $\theta y_i = z_i$. Now using the last equation we can say

$$\begin{aligned}| (K_\theta * f)_{,i} | &\leqq \theta (\max | f |) \int \cdots \int | K_{1,i} | \, dz_1 \cdots dz_n \\ &\leqq C\, \theta \max | f |,\end{aligned}$$

where C is the integral of the absolute value of $K_{1,i}$. Similarly we could obtain a bound for the size of a higher derivative of $K_\theta * f$. Each such bound would involve a constant, analogous to C, and θ to the power equal to the order of the derivative. A higher order derivative can be considered a derivative of a lower order derivative and we can bound it in terms of the maximum size of that derivative, before convolution with K_θ. To illustrate,

$$\begin{gathered}(K_\theta * f)_{,ji} = (K_\theta * f_{,j})_{,i} = (K_{\theta,i} * f_{,j}), \\ \therefore\ | (K_\theta * f)_{,ji} | \leqq C\, \theta \max | f_{,j} | .\end{gathered}$$

If we are not concerned with the precise sizes of the constants, such as C, which appear in these estimates we can put them all in one comprehensive statement. Let $\max^{(s)} f$ stand for the maximum of the values attained by the absolute values of all derivatives of f of order s at all points of E^n.

Then we can say generally:

$$\text{(A4)} \qquad \max^{(r)} (K_\theta * f) \leqq C_{rs} \theta^{r-s} \max^{(s)} f, \qquad \text{when } r \geqq s.$$

Here C_{rs} is to be a constant, independent of f, and in fact depending only on $r - s$.

Effect of varying θ

We need to know how rapidly $K_\theta * f$ and its derivatives change as θ varies. Of course we can say

$$\frac{\partial}{\partial\theta}(K_\theta * f) = \frac{\partial K_\theta}{\partial\theta} * f,$$

since we do not think of f as varying with θ. $\partial K_\theta/\partial\theta$ will also be a kernel with good properties, as we see below.

We can begin by considering

$$\text{(A5)} \qquad \overline{\frac{\partial}{\partial\theta} K_\theta} = \frac{\partial}{\partial\theta}\bar{K}_\theta = \frac{\partial}{\partial\theta}[\psi(\xi/\theta)] = -\xi/\theta^2\,\psi'(\xi/\theta) = \theta^{-1}\,\chi(\xi/\theta).$$

Here we introduce a new function, χ, defined by $\chi(u) = -u\,\psi'(u)$. Observe that $\chi(u) = 0$ for $u \leqq 1$ or $u \leqq 2$, and that $\chi(u) > 0$ for $1 < u < 2$. Also χ is C^∞, so we see that $\partial K_\theta/\partial\theta$ has the same general properties of analyticity, smallness at infinity, etc., that K_θ has.

Let L stand for the value of $\partial K_\theta/\partial\theta$ at $\theta = 1$. We can express $\partial K_\theta/\partial\theta$ in terms of L in a manner analogous to the way we obtained in (A2) an expression for K_θ as $\theta^n K_1(\theta x_1, \cdots)$. The only difference in the situations is the appearance of θ^{-1} in (A5). So we obtain

$$\text{(A6)} \qquad \frac{\partial}{\partial\theta} K_\theta = \theta^{n-1}\,L(\theta x_1, \theta x_2, \cdots, \theta x_n).$$

We want to express $L(x_1, x_2, \cdots, x_n)$ as the sum of n special functions L_i. We do this via the transform of L, which is $\chi(\xi)$. Thus we shall have

$$\text{(A7)} \qquad \bar{L} = \chi(\xi) = \textstyle\sum_i \bar{L}_i .$$

To define the $\bar{L}_i$ we construct a non-negative C^∞ function α_i for each transform variable ξ_i:

$$\alpha_i = 0, \quad \text{for} \quad |\xi_i| \leqq (2n)^{-\frac{1}{2}},$$

and

$$\alpha_i = e^{((2n)^{-\frac{1}{2}} - |\xi_i|)^{-1}}$$

for $|\xi_i| > (2n)^{-\frac{1}{2}}$.
Observe that when $\xi > (2)^{-\frac{1}{2}}$ then $\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2 > \frac{1}{2}$ and some $\xi_i > (2n)^{-\frac{1}{2}}$ so that some $\alpha_i > 0$. Therefore $\alpha_1 + \alpha_2 + \cdots + \alpha_n$ is positive and C^∞ in the region $1 \leqq \xi \leqq 2$ where $\chi(\xi)$ is non-vanishing. So we can define

$$\bar{L}_i = \frac{\alpha_i}{\alpha_1 + \alpha_2 + \cdots + \alpha_n}\,\chi(\xi), \quad \text{for} \quad \xi \geqq 1, \quad \text{and}$$

$$\bar{L}_i = 0, \quad \text{for} \quad \xi < 1.$$

These functions will be C^∞ everywhere and will satisfy (A7). Each $\bar{L}_i$ has the important properties:

(a) $\bar{L}_i = 0$ for $\xi \geqq 2$

(b) $\bar{L}_i = 0$ for $|\xi_i| \leqq (2n)^{-\frac{1}{2}}$.

The corresponding kernels L_i will clearly have all the good behavior properties at infinity, etc., that K_θ has. And we have

$$\sum_i L_i = L = \begin{Bmatrix} \text{the value at } \theta = 1 \\ \text{of } \dfrac{\partial}{\partial\theta} K_\theta \end{Bmatrix}.$$

Purpose of the L_i kernels

Each L_i is so constructed that when one forms the indefinite integral $\int_{-\infty}^{x_i} L_i\, dx_i$ the result is still a function that is small at infinity. We see this again via the transform $\bar{L}_i$. Let H_i^r stand for the rth member of a series of kernels developed from L_i and defined by

$$\bar{H}_i^r = (\xi_i(-1)^{\frac{1}{2}})^{-r}\bar{L}_i .$$

The properties of $\bar{L}_i$ that were emphasized above assure that $\bar{H}_i^r$ will be a C^∞ function of $\xi_1, \xi_2, \cdots, \xi_n$ which vanishes when $|\xi_i| \leqq (2n)^{-\frac{1}{2}}$ or $|\xi| \geqq 2$. Therefore the kernels H_i^r are analytic functions which decrease rapidly away from the origin.

The important property of the H_i^r is that

$$\frac{\partial^r H_i^r}{\partial x_i^r} = L_i . \quad \text{Also} \quad H_i^{r+1} = \int_{-\infty}^{x_i} H_i^r\, dx_i , \qquad \text{because}$$

$$\int_{-\infty}^{\infty} H_i^r\, dx_i = 0, \quad \text{since} \quad \bar{H}_i^r = 0 \quad \text{when} \quad \xi_i = 0.$$

The H_i^r help us to estimate the size of $\partial K_\theta/\partial\theta * f$ from data on the size of derivatives of f. By (A6) we have

$$\frac{\partial}{\partial\theta} K_\theta * f = \{\theta^{n-1} \textstyle\sum_i L_i(\theta x_1, \cdots \theta x_n)\} * f(x_1, \cdots, x_n).$$

Changing variables, $\theta x_i \to x_i$, as we did when estimating $(K_\theta * f)_{,i}$, we obtain

$$\begin{aligned} \partial K_\theta/\partial\theta * f &= \theta^{-1}[\{\textstyle\sum_i L_i(x_1, \cdots, x_n)\} * f(x_1/\theta, \cdots, x_n/\theta)] \\ &= \theta^{-1} \textstyle\sum_i \{L_i(x_1, \cdots, x_n) * f(x_1/\theta, \cdots, x_n/\theta)\} \\ &= \theta^{-1} \textstyle\sum_i \left\{H_i^r(x_1, \cdots, x_n) * \dfrac{\partial^r}{\partial x_i^r}[f(x_1/\theta, \cdots, x_n/\theta)]\right\} \\ &= \theta^{-1} \textstyle\sum_i \left\{H_i^r(x_1, \cdots, x_n) * \theta^{-r} \dfrac{\partial^r}{\partial(x_i/\theta)^r} f(x_1/\theta, \cdots x_n/\theta)\right\} \\ &= \theta^{-r-1} \textstyle\sum_i \left\{H_i^r * \dfrac{\partial^r}{\partial(x_i/\theta)^r} f(x_1/\theta, \cdots, x_n/\theta)\right\}. \end{aligned}$$

By the series of equations above we have expressed $(\partial K_\theta/\partial\theta) * f$ entirely in terms of rth order derivatives of f. These equations are, of course, invalid unless the rth order derivatives of f are continuous, which we assume. They lead to an inequality:

$$\left|\frac{\partial}{\partial\theta} K_\theta * f\right| \leqq \theta^{-r-1} \max\left\{\sum_i \left|\frac{\partial^r f}{\partial x_i^r}\right|\right\} \max_i \int \cdots \int |H_i^r| \, dx_1 \cdots dx_n$$
$$\leqq C_r\, \theta^{-r-1} \max{}^{(r)} f.$$

Since a derivative of $(\partial K_\theta/\partial\theta) * f$ is the same as $(\partial K_\theta/\partial\theta)$ convoluted with the corresponding derivative of f, we can also say

$$\left|\left(\frac{\partial}{\partial\theta} K_\theta * f\right)_{,i}\right| \leqq C_r\, \theta^{-r-1} \max{}^{(r)}(f_{,i})$$
$$\leqq C_r\, \theta^{-r-1} \max{}^{(r+1)} f.$$

Extending this principle, we obtain

$$\text{(A8)} \qquad \max{}^{(s)}\left[\frac{\partial}{\partial\theta} K_\theta * f\right] \leqq C_r\, \theta^{-r-1} \max{}^{(r+s)} f, \quad \text{or}$$
$$\leqq C_{t-s}\, \theta^{s-t-1} \max{}^{(t)} f.$$

This is valid only for $t \geqq s$, so far as we have shown.

Actually (A8) is also valid for $t < s$, if appropriate constants C_r are used for the negative r values. We don't need the L_i to see this. Beginning with (A6), we obtain

$$\left(\frac{\partial}{\partial\theta} K_\theta\right) * f = \theta^{n-1} L(\theta x_1, \cdots, \theta x_n) * f(x_1, \cdots, x_n),$$

$$\therefore \frac{\partial^r}{\partial x_i^r}\left[\frac{\partial}{\partial\theta} K_\theta * f\right] = \theta^{r+n-1} \frac{\partial r}{\partial(\theta x_i)^r} L(\theta x_1, \cdots, \theta x_n) * f(x_1, \cdots, x_n).$$

Making the change of variables $\theta x_i \to x_i$, the right hand term is

$$\theta^{r-1} \frac{\partial^r}{\partial x_i^r} L(x_1, \cdots, x_n) * f(x_1/\theta, \cdots, x_n/\theta)\,, \leqq C_{-r}\, \theta^{r-1} \max |f|.$$

Again we can generalize by replacing f by $(\partial^s/\partial x_i^s)f$. And we could deal with mixed partial derivatives. The general result would be

$$\max{}^{(s)}\left[\frac{\partial}{\partial\theta} K_\theta * f\right] \leqq C_{-r}\, \theta^{r-1} \max{}^{(s-r)} f.$$

This corresponds to (A8) with $-r$ in place of r. So we see that (A8) holds with positive or negative r or with $s \leqq t$ or $s > t$.

Smoothing on a manifold

Let the manifold $\mathfrak{M}$ be compact and analytic in the strong sense, so that it has an analytic imbedding $\mathfrak{R}$ in a euclidean space E^n. We can find a surrounding neighborhood, $\mathfrak{N}$, in E^n of $\mathfrak{R}$; and $\mathfrak{N}$ can be such that for any point x in $\mathfrak{N}$ there is a unique point $y(x)$ on $\mathfrak{R}$ which is the point of $\mathfrak{R}$ that is nearest to x. Also $y(x)$ can be an analytic function throughout $\mathfrak{N}$.[1]

Now let

$$\varphi(x) = \psi\left(\frac{\text{distance from } x \text{ to } y(x)}{\varepsilon}\right)$$

where ψ is the C^∞ function defined before above (A1) and ε is a small constant. If we assume ε is sufficiently small, $\varphi(x)$ will be a C^∞ function throughout $\mathfrak{N}$ and will vanish at all points near the boundary of $\mathfrak{N}$. Assume also that the definition of $\varphi(x)$ is extended by making it vanish identically outside $\mathfrak{N}$. Then $\varphi(x)$ is C^∞ everywhere.

Now if $f(y)$ is a function defined on $\mathfrak{R}$ we define an extension $f(x)$ to E^n by putting

$$f(x) = \varphi(x)\, f(y(x)), \quad \text{for} \quad x \in \mathfrak{N}, \qquad \text{and}$$

$$f(x) = 0, \quad \text{for} \quad x \notin \mathfrak{N}.$$

This extends $f(y)$ to a function $f(x)$ which agrees with the original function on $\mathfrak{R}$ and has the same degree of differentiability.

The method of smoothing is simple. Beginning with $f(y)$ on $\mathfrak{R}$ one extends to $f(x)$ in E^n. Then $K_\theta * f(x)$ is the smoothed function. The final (a logical formality) step is to restrict the definition of $K_\theta * f(x)$ to $\mathfrak{R}$ and again have a function defined only on $\mathfrak{R}$. But we must do more than simply present this definition of smoothing on $\mathfrak{R}$; we need to know how it affects derivatives of the function smoothed, etc. To do this we need

A standard size concept for derivatives

$\mathfrak{R}$ has an analytic metric induced by the imbedding in E^n. So at each point p of $\mathfrak{R}$ we can set up an internal system of geodesic normal coordinates. This system is not unique, but is unique up to orthogonal transformations. At p we can measure the size of the derivatives of order r of a function by considering all the systems of geodesic normal coordinates at p. We define $\text{size}_p^{(r)} f$ as the maximum over all these systems of the maximum of the absolute values of the various rth order derivatives of f with respect to the coordinates of that system. Then we call $\text{size}^{(r)} f$ the maximum of $\text{size}_p^{(r)} f$ over all points p of $\mathfrak{R}$.

We need to know how the measure $\text{size}^{(r)} f(y)$ of the sizes of the derivatives of f as a function on $\mathfrak{R}$ is related to the measure $\max^{(r)} f(x)$ of the sizes of the derivatives of the function extended to E^n. And when $f(y)$ is obtained by restricting the range of definition of a function defined throughout E^n we need to know how the sizes of the internal derivatives will be related to the sizes of the derivatives with respect to coordinates of E^n. In the first case it is fairly easy to see that there will be general inequalities of the form

$$\text{(A9)} \qquad \max{}^{(r)} f(x) \leqq \sum_{k=0}^{r} B_k^r \, \text{size}^{(k)} f(y)$$

[1] The existence of such a neighborhood $\mathfrak{N}$ is shown in Lemma 1 of [16].

where the coefficients are constants determined by the imbedding of $\mathfrak{R}$ in E^n and the function φ which was used in the extension of the function $f(y)$ defined only on $\mathfrak{R}$ to the function $f(x)$ defined on E^n.

Similarly, when a function $g(x)$ defined on E^n is specialized to a function, say $g(y)$, defined only on $\mathfrak{R}$ there is a conversion from bounds on derivatives with respect to the E^n coordinates ($\max^{(r)}$) to the internal measure of the size of derivatives. This has the form

$$\text{(A10)} \qquad \text{size}^{(r)}\ g(y) \ \leqq\ \sum_{k=0}^{r} D_k^r \max^{(k)}\ g(x).$$

Actually, $D_0^r = 0$, except for the trivial case $r = 0$, where $D_0^0 = 1$. The constants D_s^r depend only on the imbedding $\mathfrak{R}$.

Effect of smoothing on a manifold on derivatives

We are now ready to see how smoothing of a function on a manifold acts on derivatives, etc., and relate this action to the size of the original function and its derivatives. Suppose $f(y)$ was the original function. Then smoothing proceeds thus:

$$\text{(a)} \qquad f(y) \to f(x) \quad \text{by extension to } E^n,$$

$$\text{(b)} \qquad f(x) \to g(x) = K_\theta * f(x),$$

$$\text{(c)} \qquad g(x) \to g(y) \quad \text{by restriction to } \mathfrak{M}.$$

We call $g(y) = S_\theta f(y)$ so that here S_θ stands for the total operation of smoothing.

Corresponding to the two general inequalities, (A4) and (A8), for smoothing in E^n, we obtain two general inequalities for smoothing on $\mathfrak{R}$. For example, (A9) gives us bounds on $\max^{(r)} f(x)$ from $\text{size}^{(r)}$ data on $f(y)$. Then (A4) gives us $\max^{(r)}$ data on $g(x)$ from this data on $f(x)$. Finally (A10) gives us $\text{size}^{(r)}$ data on $g(y)$ from the $\max^{(r)}\ g(x)$ data. The outcome is a bound on $\text{size}^{(r)}\ g(y)$ in terms of $\text{size}^{(s)}\ f(y)$, $\text{size}^{(s-1)}\ f(y), \cdots, \text{size}^{(0)}\ f(y)$. If we use $S_\theta f$ for $g(y)$ and f for $f(y)$, and if we weaken the form of the bound by using the maximum constant involved, we get a bound of the form

$$\text{(A11)} \qquad \text{size}^{(r)}[S_\theta f] \ \leqq\ H_{rs}\theta^{r-s} \sum_{t=0}^{s} \text{size}^{(t)} f, \qquad \text{for } \theta \geqq 1,\ r \geqq s.$$

Exactly analogously we obtain from (A8):

$$\text{(A12)} \qquad \text{size}^{(r)} \left[\frac{\partial}{\partial\theta} S_\theta f\right] \leqq J_{rs}\, \theta^{r-s-1} \sum_{t=0}^{s} \text{size}^{(t)} f, \qquad \text{for } \theta \geqq 1.$$

We use the restriction $\theta \geqq 1$ so that we can majorize lower powers of θ by θ^{r-s} or θ^{r-s-1} in the two inequalities. The H_{rs} and J_{rs} coefficients depend only on the imbedding $\mathfrak{R}$ and φ.

Smoothing of tensors

Here the first step is to express each tensor in a normalized (non-tensorial) form in terms of a set of scalar functions defined over all of $\mathfrak{R}$. This corresponds to using a specific redundant coordinate system on $\mathfrak{R}$ in order to have a coordinate system without singularities. Let $x^1, x^2, \cdots, x^n$ be the coordinates of E^n and let $u^1, u^2, \cdots, u^\nu$ be any local system of coordinates in $\mathfrak{R}$, such as one of the family of geodesic normal coordinate systems used in defining the size$^{(r)}$ concept.

The imbedding defines a transformation

$$(u^1,\ u^2, \cdots,\ u^\nu) \rightarrow (x^1,\ x^2, \cdots,\ x^n) \tag{A13}$$

and the nearest point function $y(x)$ defines a transformation $x \rightarrow y(x)$ in the neighborhood $\mathfrak{R}$, which gives a transformation of coordinates

$$(x^1,\ x^2, \cdots,\ x^n) \rightarrow (u^1,\ u^2, \cdots,\ u^\nu). \tag{A14}$$

Both transformations are analytic and their composition in either order is the identity on $\mathfrak{R}$.

Now suppose $T^{\alpha\beta\cdots}_{\gamma\delta\cdots}$ is a tensor on $\mathfrak{R}$ referred to the coordinates $u^1, u^2, \cdots, u^\nu$. Then we define

$$\mathfrak{T}^{ij\cdots}_{kl\cdots} = T^{\alpha\beta\cdots}_{\gamma\delta\cdots} \frac{\partial x^i}{\partial u^\alpha} \frac{\partial x^j}{\partial u^\beta} \cdots \frac{\partial u^\gamma}{\partial x^k} \frac{\partial u^\delta}{\partial x^l} \cdots,$$

with the summation convention operating, and with $\partial x^i/\partial u^\alpha$, etc., taken from (A14). This definition has the correct invariance properties, so $\mathfrak{T}^{ij\cdots}_{kl\cdots}$ is defined globally on $\mathfrak{R}$ and is completely independent of the coordinates $u^1, u^2, \cdots, u^\nu$ through which it is obtained.

Because the composition of (A13) and (A14) is the identity, it is easy to see that the reverse conversion

$$T^{\alpha\beta\cdots}_{\gamma\delta\cdots} = \mathfrak{T}^{ij\cdots}_{kl\cdots} \frac{\partial u^\alpha}{\partial x^i} \frac{\partial u^\beta}{\partial x^j} \cdots \frac{\partial x^k}{\partial u^\gamma} \frac{\partial x^l}{\partial u^\delta} \cdots$$

yields the original tensor again from the normalized form $\mathfrak{T}^{ij\cdots}_{kl\cdots}$.

The smoothing operation for a tensor consists of three steps: (a) conversion to normalized form, (b) smoothing of each component of the normalized form by S_θ, and (c) converting the result back to a tensor on $\mathfrak{R}$ via the reverse conversion.

Derivative size concept for tensors

If T is a tensor we consider the standard local coordinate systems defined before and at the center of each system we consider all the derivatives of order r of each component of T. The maximum of the absolute values of these is then the local size$^{(r)}$ of T. Then the maximum of this over all the standard systems is size$^{(r)}$ T.

Associated with the conversion of tensors to normalized form and with the reverse conversion there will be conversions from size$^{(r)}$ T measures to size$^{(r)}$ measures on the components of $\mathfrak{T}$. These will give us inequalities quite analogous to (A9) and (A10). By combining these inequalities (we shall not bother to write them out) with (A11) and (A12) to estimate the effect of S_θ on the components of $\mathfrak{T}$, we can obtain the analogous bounds for smoothing of tensors via the process $T \to \mathfrak{T} \to S_\theta \mathfrak{T} \to S_\theta T$ (definition of $S_\theta T$). These bounds are

$$\text{(A15)} \qquad \operatorname{size}^{(r)}(S_\theta T) \leqq L_{rs}\theta^{r-s} \sum_{t=0}^{s} \operatorname{size}^{(t)} T \quad \text{for } r \geqq s,\ \theta \geqq 1$$

and

$$\text{(A16)} \qquad \operatorname{size}^{(r)} \left[\left(\frac{\partial}{\partial\theta} S_\theta \right) T \right] \leqq M_{rs}\, \theta^{r-s-1} \sum_{t=0}^{s} \operatorname{size}^{(t)} T \qquad \text{for } \theta \geqq 1.$$

Note that the process for smoothing tensors preserves certain simple properties that a tensor may have, such as symmetry or skew-symmetry, etc.

Concluding remarks

The $\operatorname{size}^{(r)}$ concept is suited to routine estimations concerning sums, differences, or products of functions or tensors. For example,

$$\operatorname{size}^{(1)}(fg) \leqq (\operatorname{size}^{(0)} f)(\operatorname{size}^{(1)} g) + (\operatorname{size}^{(1)} f)(\operatorname{size}^{(0)} g),$$

$$\operatorname{size}^{(0)}(T^j_i S^k_j) \leqq \nu \operatorname{size}^{(0)}(T^j_i) \operatorname{size}^{(0)}(S^k_i).$$

In the second estimate we invoke the summation convention and ν should be the dimensionality of $\mathfrak{R}$. These remarks are made to prepare for the frequent use of such elementary estimations in Part B, where we shall not specifically mention these properties of $\operatorname{size}^{(r)}$ in the instances where they are used.

Another general bound referring to the action of S_θ is derivable from (A15) and (A16). This is

$$\text{(A17)} \qquad \operatorname{size}^{(r)}(T - S_\theta T) \leqq N_{rs}\theta^{r-s} \sum_{t=0}^{s} \operatorname{size}^{(t)} T \quad \text{for } s \geqq r,\ \theta \geqq 1.$$

In the case $s = r$ this is derived trivially from (A15). When $s > r$ we use

$$T - S_{\theta_1} T = \int_{\theta_1}^{\infty} \left(\frac{\partial}{\partial\theta} S_\theta \right) T \, d\theta$$

and apply (A16).

Our more or less elaborate development of S_θ was undertaken to give us a smoothing operator for which (A15), (A16), and (A17) would hold. Without care in the definition one would obtain a weaker set of bounds. (A17) would probably have $\theta^{\max(r-s,-2)}$ instead of θ^{r-s}. (A15) would probably not be weakened.

There is a good heuristic interpretation for an operator such as S_θ. It is a low-pass filter which passes undiminished all frequencies below θ. Above 2θ it cuts off completely. Between θ and 2θ there is a variable attenuation, decreasing with increasing frequency so that the characteristic of the filter for all frequencies is a C^∞ function of frequency.

PART B: THE METRIC PERTURBATION THEOREM

The perturbation device

The perturbation process developed in this part of the paper is based on a method for finding an infinitesimal change in the imbedding of a manifold that will effect a specified infinitesimal change in the metric induced by that imbedding. The smoothing operator S_θ of Part A is used in connection with this method, the "perturbation device".

Let $\mathfrak{M}$ be a compact n-manifold smoothly imbedded in E^m. Let the Cartesian coordinates of E^m be $z_1, z_2, \cdots, z_m$. Referred to a set $x_1, x_2, \cdots, x_n$ of local coordinates in $\mathfrak{M}$, the metric tensor induced by the imbedding is

$$g_{ij} = \sum_\alpha \frac{\partial z_\alpha}{\partial x_i} \frac{\partial z_\alpha}{\partial x_j}. \tag{B1}$$

We can consider perturbations of the imbedding as rates of change, measured with respect to the change of a parameter. The parameter is unspecified and we indicate the rate of change of any quantity by placing a dot over it. Thus we have

$$\dot{g}_{ij} = \sum_\alpha \frac{\partial z_\alpha}{\partial x_i} \frac{\partial \dot{z}_\alpha}{\partial x_j} + \sum_\alpha \frac{\partial \dot{z}_\alpha}{\partial x_i} \frac{\partial z_\alpha}{\partial x_j}, \tag{B2}$$

which follows from (B1).

We want a method by which we can determine $\{\dot{z}_\alpha\}$ satisfying (B2) when $\{\dot{g}_{ij}\}$ is specified. We can make this problem simpler to solve by adding another condition to be satisfied by the rate $\{\dot{z}_\alpha\}$ of change of the imbedding. This is

$$\sum_\alpha \frac{\partial z_\alpha}{\partial x_i} \dot{z}_\alpha = 0, \qquad \text{for all } i, \tag{B3}$$

which requires the perturbation $\{\dot{z}_\alpha\}$ to be normal to the imbedding.

Consider the result of differentiating (B3) with respect to x_j:

$$\sum_\alpha \frac{\partial z_\alpha}{\partial x_i} \frac{\partial \dot{z}_\alpha}{\partial x_j} = -\sum_\alpha \frac{\partial^2 z_\alpha}{\partial x_j \, \partial x_i} \dot{z}_\alpha. \tag{B4}$$

The left member occurs in (B2) and the right member is symmetric in i and j (since the imbedding is to be reasonably smooth). Consequently we can use (B4) to modify (B2) and obtain

$$\dot{g}_{ij} = -2 \sum_\alpha \frac{\partial^2 z_\alpha}{\partial x_j \, \partial x_i} \dot{z}_\alpha. \tag{B5}$$

This is the condition the perturbation $\{\dot{z}_\alpha\}$ should satisfy to effect the metric perturbation $\{\dot{g}_{ij}\}$ when (B3) holds.

Now we have a much simpler type of requirements relating $\{\dot{z}_\alpha\}$ to $\{\dot{g}_{ij}\}$. Together (B3) and (B5) form a system of linear equations to be satisfied by the $\dot{z}_\alpha$, whereas before we had partial differential equations in the $\dot{z}_\alpha$.

How and when can we solve (B3)–(B5) for $\{\dot{z}_\alpha\}$ after $\{\dot{g}_{ij}\}$ has been specified? The number of variables $\dot{z}_\alpha$, which is m, should be at least as large as the number of linear equations, which is $\frac{1}{2}n^2 + 1\frac{1}{2}n$, taking the i, j symmetry into account. Probably m will have to be larger than $\frac{1}{2}n^2 + 1\frac{1}{2}n$ to insure that the equations are not singular at some points of $\mathfrak{M}$. So the equations will probably be underdetermined.

In Part C we construct an imbedding of $\mathfrak{M}$ such that (B3)–(B5) is everywhere non-singular. Here in Part B we assume that the imbedding has this property and simply make our results conditional upon this.

We must find a way to select a particular solution $\{\dot{z}_\alpha\}$, and do this in a smooth way, when the equations are underdetermined. A very simple requirement,

$$\text{(B6)} \qquad \begin{gathered} \textstyle\sum_\alpha (\dot{z}_\alpha)^2 = \text{minimum}, \\ \text{subject to satisfaction of (B3–B5),} \end{gathered}$$

selects a particular solution in a satisfactory manner. $\{\dot{z}_\alpha\}$ will have the same degree of differentiability as $\{\dot{g}_{ij}\}$.

The geometrical interpretation of the effect of (B6) is that it selects the nearest point to the origin in the plane of solutions $\{\dot{z}_\alpha\}$ of (B3)–(B5). We can also study the effect of (B6) from a more formal viewpoint. The system (B3)–(B5) is of the form:

$$\text{(a)} \qquad \textstyle\sum_{\alpha=1}^{m} C_{\mu\alpha}\dot{z}_\alpha = \varphi_\mu .$$

If we assume a solution in the form

$$\text{(b)} \qquad \dot{z}_\alpha = \textstyle\sum_{\nu=1}^{\nu=\frac{1}{2}n^2+1\frac{1}{2}n} C_{\nu\alpha}\, d_\nu ,$$

then the d_ν's must satisfy

$$\text{(c)} \qquad \textstyle\sum_{\alpha,\nu} C_{\mu\alpha}C_{\nu\alpha}\, d_\nu = \varphi_\mu .$$

Let

$$\text{(d)} \qquad E_{\mu\nu} = \textstyle\sum_\alpha C_{\mu\alpha}C_{\nu\alpha} ,$$

then (c) becomes

$$\text{(e)} \qquad \textstyle\sum_\nu E_{\mu\nu}\, d_\nu = \varphi_\mu .$$

This last equation (e), will be non-singular if det $\| E_{\mu\nu} \|$ is not zero. However this is Gram's determinant for the matrix $\| C_{\mu\alpha} \|$ and it cannot vanish unless $\| C_{\mu\alpha} \|$ has less than maximal rank $(\frac{1}{2}n^2 + 1\frac{1}{2}n)$. Because we are assuming that (a), which is (B3)–(B5) in a condensed notation, is non-singular, we know that rank $\| C_{\mu\alpha} \| = \frac{1}{2}n^2 + 1\frac{1}{2}n$. So (e) is non-singular. Because (e) is not underdetermined it has a solution in the form

$$\text{(f)} \qquad d_\nu = \| E_{\mu\nu} \|^{-1} \cdot \{\varphi_\mu\}.$$

Now from (b) we can express a special solution $\{\dot{z}^*_\alpha\}$ of (a) (or of (B3)–(B5), which is the same thing) in the form

(g) $$\dot{z}^*_\alpha = \| C_{\nu\alpha} \| \cdot \| E_{\mu\nu} \|^{-1} \cdot \{\varphi_\mu\} .$$

This special solution of (a) happens to be the one for which $\sum_\alpha (\dot{z}_\alpha)^2$ is minimized. Suppose $\{\dot{z}_\alpha\}$ is any other solution of (a). Then

(h) $$\sum_\alpha C_{\mu\alpha} (\dot{z}_\alpha - \dot{z}^*_\alpha) = 0.$$

We can write

(i) $$\sum_\alpha (\dot{z}_\alpha)^2 - \sum_\alpha (\dot{z}^*_\alpha)^2 = \sum_\alpha (\dot{z}_\alpha - \dot{z}^*_\alpha)^2 + 2\sum_\alpha \dot{z}^*_\alpha(\dot{z}_\alpha - \dot{z}^*_\alpha),$$

and the last term vanishes because

(j) $$\begin{aligned} 2\sum_\alpha \dot{z}^*_\alpha(\dot{z}_\alpha - \dot{z}^*_\alpha) &= 2\sum_\alpha [\sum_\nu d_\nu C_{\nu\alpha}] (\dot{z}_\alpha - \dot{z}^*_\alpha) \\ &= 2\sum_\nu d_\nu \sum_\alpha C_{\nu\alpha} (\dot{z}_\alpha - \dot{z}^*_\alpha) \\ &= 2\sum d_\nu \cdot 0 \\ &= 0. \end{aligned}$$

Here we employed the expression (b) for $\dot{z}^*_\alpha$.

Now that we know the last term of (i) is zero it follows that the right hand side is positive and $\sum_\alpha (\dot{z}_\alpha)^2 > \sum_\alpha (\dot{z}^*_\alpha)$. Thus the special solution of (B3)–(B5) that is given by (g) is the same as the one selected by (B6). This shows that (B6) determines a solution which is a well behaved function of $\{\dot{g}_{ij}\}$ and the derivatives of the imbedding functions, so long as (B3)–(B5) remains a non-singular system.

The solution of (B3)–(B5) determined by (B6), or equivalently by (g), has the following form of linear dependence on $\{\dot{g}_{ij}\}$ and analytic dependence on the imbedding derivatives:

(B7) $$\dot{z}_\alpha = \sum_{i \leqq j} \dot{g}_{ij} F_{\alpha ij} \left(\left\{\frac{\partial z}{\partial x_k}\right\}, \left\{\frac{\partial^2 z}{\partial x_k\, \partial x_l}\right\}\right).$$

$\{\dot{g}_{ij}\}$ is represented in (g) by $\{\varphi_\mu\}$, so this shows that $\{\dot{z}_\alpha\}$ depends linearly on $\{\dot{g}_{ij}\}$. The $F_{\alpha ij}$ are analytic functions of the first and second order derivatives of the imbedding functions, so long as these are such that (B3)–(B5) is non-singular.

To recapitulate, (B7) indicates the form and behavior of the solution of the system:

(B8a) $$\sum_\alpha \frac{\partial z_\alpha}{\partial x_i} \dot{z}_\alpha = 0$$

(B8b) $$-2 \sum_\alpha \frac{\partial^2 z_\alpha}{\partial x_i\, \partial x_j} \dot{z}_\alpha = \dot{g}_{ij}$$

(B8c) $$\sum_\alpha (\dot{z}_\alpha)^2 = \text{minimum, subject to (a) and (b).}$$

The solution is an imbedding perturbation rate $\{\dot{z}_\alpha\}$ which leads to the rate

$\{\dot{g}_{ij}\}$ of change of the metric induced by the imbedding. We call this method of determining a perturbation rate $\{\dot{z}_\alpha\}$ the "perturbation device".

Notational conventions

The work below becomes almost entirely the treatment of a problem in analysis, so a different (condensed) notation is appropriate. We shall drop the coordinate indices generally. Thus

$$\{z_\alpha\} \text{ becomes } z,\ \{\dot{z}_\alpha\} \text{ becomes } \dot{z},\left\{\frac{\partial z_\alpha}{\partial x_i}\right\} \text{ becomes } z',\left\{\frac{\partial^2 z_\alpha}{\partial x_i\,\partial x_j}\right\} \text{ becomes } z'', \text{ etc.}$$

We now write (B7) as

$$\text{(B9)} \qquad \dot{z} = F(z', z'') \boxtimes \dot{g}.$$

$\boxtimes$ indicates the (contracting) tensor product acting between F and $\dot{g}$. (B9) is the solution of

$$\text{(B10)} \quad \begin{array}{lrl} \text{(a)} & z' \circ \dot{z} & = 0 \\ \text{(b)} & -2z'' \circ \dot{z} & = \dot{g} \\ \text{(c)} & |\, \dot{z} \,| & = \text{minimum, subject to (a) and (b).} \end{array}$$

Here $\circ$ indicates the scalar product, summation over the index α of E^m. We also have

$$\text{(B11)} \qquad \dot{g} = 2z' \otimes \dot{z}'$$

as a modified condensation of (B2). Here $\otimes$ is a symmetrizing tensor product accompanied by a summation (as with $\circ$).

We shall deal with many inequalities on the sizes of functions and derivatives. Generally these will be tied to the parameter θ, which controls the smoothing operator (see Part A). θ will play a dual role, being both the parameter of smoothing and the parameter of the process. Dot will denote $\partial/\partial\theta$. [e.g.: $\dot{z} = \partial z/\partial\theta$]. The process will begin with a specific value of θ, called θ_0, and end with $\theta = \infty$.

Our canonical notation for bounds on the sizes of functions and their derivatives is explained by illustration:

$$T \lesssim K[\theta \mid {}^{-1}_{\ 2};\ {}^{2}_{4}]$$

indicates a whole system of bounds on the tensor T and its derivatives measured in terms of the size$^{(r)}$ concept of Part A (think of T as varying with θ), which are

$$\begin{aligned} \text{size}^{(0)}\ T &\lesssim K\theta^{-1} \\ \text{size}^{(1)}\ T &\lesssim K\theta^{-1} \\ \text{size}^{(2)}\ T &\lesssim K\theta^{-1} \\ \text{size}^{(3)}\ T &\lesssim K \\ \text{size}^{(4)}\ T &\lesssim K\theta \\ \text{size}^{(5)}\ T &\lesssim K\theta^2. \end{aligned}$$

In general, if the symbol is $[\theta \mid {}^{p,}_{r,} {}^{q}_{s}]$, where r and s are integers and $0 \leqq r \leqq s$, $p \leqq q$, the exponent of θ is p for size$^{(0)}$, size$^{(1)}$, $\cdots$, size$^{(r)}$. Then from size$^{(r)}$ to size$^{(s)}$ the exponent increases in arithmetic progression in such a way that we have θ^q for size$^{(s)}$. Almost always this increase will be one unit for each increase of the order of differentiation, as in the illustration. In other words, we usually have $q - p = s - r$, and also q and p are usually integers.

Most often we can abbreviate $[\theta \mid {}^{p,}_{r,} {}^{q}_{s}]$ to $[{}^{p,}_{r,} {}^{q}_{s}]$ and understand that θ is involved. Also we shorten symbols such as $[{}^{0,}_{3,} {}^{0}_{3}]$ to $[{}^{0}_{3}]$.

The perturbation process

As was remarked above, this process uses the perturbation device and the smoothing operator. It also uses "feed-back".

The equations defining the process are listed below:

$$\zeta = S_\theta z \tag{B12}$$

$$\dot{z} = F(\zeta', \zeta'') \boxtimes M \tag{B13}$$

$$\begin{aligned} E &= M - \dot{g} \\ &= 2(\zeta - z)' \otimes \dot{z}' \end{aligned} \tag{B14}$$

(equivalence shown below)

(B15) $u(p)$ = a special C^∞ function, nondecreasing everywhere and monotone increasing for $0 \leqq p \leqq 1$. Also, $u(p) = 0$ for $p \leqq 0$ and $u(p) = 1$ for $p \geqq 1$. To be specific think of $u(p) = \psi(2 - p)$, where ψ is the special function of Part A.

$$L(\theta) = \int_{\theta_0}^{\theta} E(\bar{\theta}) u(\theta - \bar{\theta})\, d\bar{\theta} \tag{B16}$$

(B17) G = the desired total change of the metric tensor (a symmetric covariant tensor).

(B18)

(a) $\displaystyle\int_{\theta_0}^{\theta} M(\bar{\theta})\, d\bar{\theta} = u(\theta - \theta_0) S_\theta G + S_\theta L(\theta)$, or

(b) $M = \dot{u}(\theta - \theta_0) S_\theta G + u(\theta - \theta_0) \dot{S}_\theta G + \dot{S}_\theta L + S_\theta \dot{L}$, or

(c) $$M = \dot{u}(\theta - \theta_0) S_\theta G + u(\theta - \theta_0) \dot{S}_\theta G + \dot{S}_\theta L + S_\theta \int_{\theta_0}^{\theta} E(\bar{\theta}) \dot{u}(\theta - \bar{\theta})\, d\bar{\theta}.$$

Interpretation of the quantities and equations

Each of the quantities has its interpretation, but our interpretation may seem unilluminating to many. ζ is a smooth approximation to the imbedding. M is the rate of metric change that is being "attempted." Since the actual rate, $\dot{z}$, of perturbation of the imbedding is computed from a formula which is like (B9) but has ζ in place of z, we cannot expect that the actual rate, $\dot{g}$, of metric change will be the same as M. If the imbedding function were ζ instead of z then $\dot{g}$ would be M. That is, (B13) gives the correct rate of imbedding perturbation to accomplish the metric change M, provided ζ is the imbedding. This implies that

$$M = 2\zeta' \otimes \dot{z}'. \tag{B19}$$

E is the error rate, or the excess of the attempted metric change rate over the actual rate, so $E = M - \dot{g}$. Since we have (B11),

$$\dot{g} = 2z' \otimes \dot{z}',$$

we can subtract this from (B19) to obtain the alternate E formula:

$$E = 2(\zeta - z)' \otimes \dot{z}'.$$

This formula is the more useful one for estimating the size of E.

L represents accumulated error. It is not the total accumulated error, but it includes all the error incurred up to $\theta - 1$. Between $\theta - 1$ and θ it includes only part of the error. Effectively, there is a lag in the inclusion of error in L. This effect is accomplished by the weighting function $u(\theta - \bar{\theta})$ in (B16). It is not really necessary to define L so elaborately. It could be the total error. The advantage of the more elaborate definition is that it ultimately makes it easier to view the process as controlled by a set of very tame integral equations. $\dot{L}$ has the simple integral expression

$$\dot{L} = \int_{\theta_0}^{\theta} E(\bar{\theta})\dot{u}(\theta - \bar{\theta})\, d\bar{\theta}.$$

If L were the total error $\dot{L}$ would be E. Then it would seem that M was defined from E and E from M. We avoid this complication with the definition we use.

The definition of M is based on the principle of "feeding in" the smoother parts of the desired metric correction first, saving the rougher parts for later. Referring to (B18a), we can consider $\int_{\theta_0}^{\theta} M$ the total "attempted" metric change from the beginning of the process at θ_0 to the current situation at θ. This is set equal to $S_\theta L + u(\theta - \theta_0)S_\theta G$. So it is the smooth part of L plus the smooth part of G weighted by the coefficient $u(\theta - \theta_0)$. The reason for attaching $u(\theta - \theta_0)$ to $S_\theta G$ in this formula is the simple one that for $\theta = \theta_0$ both sides of the equation (B18a) should be zero. At the beginning of the process a finite portion of G, specifically $S_{\theta_0}G$, is considered smooth enough to be "fed in." But it must be fed in gradually, so we use $u(\theta - \theta_0)$ to make the process start gradually. For $\theta \geqq \theta_0 + 1$ this factor $u(\theta - \theta_0)$ is just $+1$ and is irrelevant.

We can see how these definitions should work out if the process is convergent. The total metric change accomplished by the process from its start at θ_0 to the limit as $\theta \to \infty$ will be $\int_{\theta_0}^{\infty} \dot{g}\, d\theta$. From (B14),

$$\dot{g} = M - E, \tag{B20}$$ therefore

(B21)
$$\begin{aligned}\int_{\theta_0}^{\infty} \dot{g}\, d\theta &= \int_{\theta_0}^{\infty} M - \int_{\theta_0}^{\infty} E \\ &= u(\infty) S_\infty G + S_\infty L(\infty) - \int_{\theta_0}^{\infty} E \\ &= G + L(\infty) - L(\infty) \\ &= G.\end{aligned}$$

To put $L(\infty)$ for $\int_{\theta_0}^{\infty} E$ requires the assumption that $E \to 0$ as $\theta \to \infty$. (B21) verifies the general design of our "feed-back" process, but of course the main task is the proof of convergence. These remarks on interpretation should not be regarded as if presented as proofs. However, we shall use the equivalence of the two formulas of (B14).

To prove that the process works we first derive a set of appropriate *a priori* bounds on the quantities involved which would be satisfied if the equations defining the process have a solution up to some value θ_1 of θ. Second we prove a local continuation theorem about solutions of the equation system. The combination of the bounds and the local continuation theorem gives us the existence and uniqueness of the solution for all values of θ, but we assume that G is sufficiently small and that θ_0 is properly chosen in obtaining this result.

The estimates

These estimates, or bounds, form a self-interacting system because the size of each function tends to depend on the sizes of the others. We assume the system (B12) through (B18) has a solution for $\theta_0 \leqq \theta \leqq \theta_1$ and assume that the quantities E, M, ζ, z, L, etc., satisfy certain bounds in this range. Then we compute new bounds on the quantities from the defining equations. Finally we show that when G is sufficiently small and θ_0 properly chosen there exists a set of bounds which is satisfied by the initial values of the quantities and is such that the rederived bounds computed from the defining equations are all smaller than this original set of bounds.

The first bound is

(B22)
$$\zeta - z_0 \precsim \varepsilon \left[{0 \atop 2}\right].$$

z_0 stands for $z(\theta_0)$, the initial value of the imbedding function. We assume that z_0 is analytic and such that (B8a, b) is non-singular. $F(z', z'')$ will be well behaved when z, z', z'' are near z_0, z_0', z_0''. So (B22) is designed to insure the good behavior of $F(\zeta', \zeta'')$. ε must be a sufficiently small constant so that $F(\zeta', \zeta'')$ will be well behaved when (B22) holds.

There will be some value θ_a (assume $\theta_a \geqq 1$) such that for $\theta \geqq \theta_a$ we have $S_\theta z_0 - z_0 \leqq \varepsilon/2\left[{0 \atop 2}\right]$. So we make a requirement:

(B23)
$$\theta_0 \geqq \theta_a .$$

With this requirement we can satisfy (B22) by keeping $z - z_0$ small enough.

Let

$$z_0 \lesssim \alpha[{}^{0,1}_{3,4}]. \tag{B24}$$

This is for notational convenience only. z_0 is of course independent of θ. The bound α should be chosen so that (B24) holds for all $\theta \geqq 1$. For $z - z_0$ we assume a bound

$$z - z_0 \lesssim \beta[{}^{0,1}_{3,4}]. \tag{B25}$$

Note that if (B23) holds and β is small enough then (B22) holds. Adding (B24) and (B25),

$$\begin{aligned} z &\lesssim (\alpha + \beta)[{}^{0,1}_{3,4}] \\ &\lesssim \xi[{}^{0,1}_{3,4}]. \end{aligned} \tag{B26}$$

ξ is employed for notational advantages.

Other bounds used are:

$$L \lesssim \lambda[{}^{0}_{3}] \tag{B27}$$

$$M \lesssim \mu[{}^{-4,0}_{\ 0,4}] \tag{B28}$$

$$\dot{z} \lesssim \gamma[{}^{-4,0}_{\ 0,4}] \tag{B29}$$

$$E \lesssim \eta[{}^{-5,-2}_{\ 0,\ \ 3}] \tag{B30}$$

$$G \lesssim \delta[{}^{0}_{3}]. \tag{B31}$$

The bound on G is a handle by which we can refer to the size of G, which we shall assume to be as small as necessary to make the process converge. This smallness includes derivatives up to the third order.

The rederived bounds

Now we assume that (B22) through (B31) hold for a solution of the equations of the process for $\theta_0 \leqq \theta \leqq \theta_1$ and use the defining equations to compute new "rederived" bounds on the same quantities in terms of the original bounds. The new bounds hold in the same θ range and are distinguished by starred Greek letters.

First consider L. From (B16)

$$L(\theta) \lesssim \int_{\theta_0}^{\theta} \eta[\bar{\theta} \mid {}^{-5,-2}_{\ 0,\ \ 3}] u(\theta - \bar{\theta})\, d\bar{\theta}.$$

When $k \geqq 2$

$$\int_{\theta_0}^{\theta} (\bar{\theta})^{-k}\, d\bar{\theta} = \frac{\theta_0^{1-k} - \theta^{1-k}}{k - 1} \leqq \theta_0^{1-k}.$$

Since $|u| \leqq 1$ and $\theta_0 \geqq 1$ we can put

$$
\begin{aligned}
L &\lesssim \eta[\theta_0 \mid {}^{-4,-1}_{0,\ \ 3}] \\
&\lesssim \eta\theta_0^{-1}[{}^{0}_{3}] \qquad\qquad \text{(B32)}\\
&\lesssim \lambda^*[{}^{0}_{3}].
\end{aligned}
$$

To estimate M we use (B18c) and have

$$
\begin{aligned}
M &\lesssim \dot{u}'(\theta - \theta_0)C_1\delta[{}^{0,1}_{3,4}] + C_2\delta[{}^{-4,0}_{0,4}] \\
&\quad + C_2\lambda^*[{}^{-4,0}_{0,4}] + S_\theta \int_{\theta_0}^{\theta} \eta\, [\bar{\theta} \mid {}^{-5,-2}_{0,\ \ 3}]\, \dot{u}'(\theta - \bar{\theta})\, d\bar{\theta}. \qquad \text{(B33)}
\end{aligned}
$$

The first term is obtained from (A15) applied to (B31). The constant C_1 is simply the largest of the coefficients that would come in from (A15). In general, when we have unspecified constants appearing in estimates we shall simply number them $C_1, C_2, \cdots$ in order of occurrence. We shall not be concerned with the actual sizes of these constants. The second term comes from applying (A16) to (B31), remembering that $|u| \leqq 1$, and the third comes from (A16) and (B32).

For the first term we say

$$
\begin{aligned}
\dot{u}'(\theta - \theta_0)C_1\delta[{}^{0,1}_{3,4}] &\lesssim [\theta^4\dot{u}(\theta - \theta_0)]C_1\delta[{}^{-4,0}_{0,4}] \\
&\lesssim \max_{\theta_0 \leqq \theta \leqq \theta_0 + 1}\{\theta^4\} \max_\theta \{\dot{u}(\theta - \theta_0)\} C_1\delta[{}^{-4,0}_{0,4}] \\
&\lesssim C_3(\theta_0 + 1)^4\delta[{}^{-4,0}_{0,4}].
\end{aligned}
$$

This brings the first term into the $[{}^{-4,0}_{0,4}]$ form, leaving only the fourth term of (B33) to be treated.

We want the fourth term of (B33) majorized by a $[{}^{-4,0}_{0,4}]$ term. Let $\theta^* = \max(\theta_0, \theta - 1)$, then since $\dot{u}'(\theta - \bar{\theta}) = 0$ when $\bar{\theta} \leqq \theta - 1$, we can say

$$
\begin{aligned}
\text{the fourth term} &\lesssim S_\theta \int_{\theta^*}^{\theta} \eta\, [\bar{\theta} \mid {}^{-5,-2}_{0,\ \ 3}]\, \dot{u}\,(\theta - \bar{\theta})\, d\bar{\theta} \\
&\lesssim S_\theta \{\eta \max{}_p\, [\dot{u}(p)]\, [\theta^* \mid {}^{-5,-2}_{0,\ \ 3}]\},
\end{aligned}
$$

because the interval of integration is not more than one unit long, i.e., $\theta - \theta^* \leqq 1$. Continuing,

$$
\text{the fourth term} \lesssim C_4\eta[\theta^* \mid {}^{-5,-1}_{0,\ \ 4}],
$$

where C_4 is to take care of $\max_p \dot{u}(p)$ and the constant coefficients arising from (A15) (which tells us how S_θ acts). Then, since $\theta_0 \leqq \theta^* \leqq \theta$, we can say

$$
\begin{aligned}
\text{the fourth term} &\lesssim C_4\eta\theta_0^{-1}(\theta^*/\theta)^{-5}[\theta \mid {}^{-4,0}_{0,4}] \\
&\lesssim C_4\eta\theta_0^{-1}(\tfrac{1}{2})^{-5}[\theta \mid {}^{-4,0}_{0,4}] \\
&\qquad\qquad (\text{because } 1 \leqq \theta^* \leqq \theta, \text{ and } \theta^* \geqq \theta - 1) \\
&\lesssim C_5\eta\theta_0^{-1}[{}^{-4,0}_{0,4}].
\end{aligned}
$$

Combining these results we obtain

$$\text{(B34)}\qquad \begin{aligned} M &\lesssim \{C_3(\theta_0+1)^4\delta + C_2\delta + C_2\lambda^* + C_5\theta_0^{-1}\eta\}\left[{}^{-4,0}_{\ 0,4}\right] \\ &\lesssim \mu^*\left[{}^{-4,0}_{\ 0,4}\right]. \end{aligned}$$

To estimate $\dot z$ we need (B34) and an estimate on ζ. For ζ we write

$$\text{(B35)}\qquad \zeta \lesssim C_6\xi\left[{}^{0,3}_{3,6}\right],$$

by applying (A15) to (B26). The reason for extending this estimate to sixth derivatives is that $\dot z$ depends on ζ'', so that the fourth derivatives of $\dot z$ depend on the sixth derivatives of ζ.

A derivative of $\dot z$ will depend on derivatives of various orders of M and ζ. For example, we can write symbolically

$$\dot z' = F(\zeta', \zeta'') \boxtimes M' + (F_{\zeta'})\zeta'' \boxtimes M + (F_{\zeta''})\zeta''' \boxtimes M.$$

The function F and its partial derivatives $F_{\zeta'}$ and $F_{\zeta''}$ will be bounded when (B22) holds so we can say

$$\text{size } \dot z' \leqq \text{const. } \mu^*\theta^{-3} + \text{const. } C_6\xi\mu^*\theta^{-4} + \text{const. } C_6\xi\mu^*\theta^{-4}$$

by applying (B34) and (B35). Notice that the highest (least negative) power of θ comes from the term where M is differentiated. For a higher derivative of $\dot z$ there would be many terms and the highest power of θ would appear in the term with the maximum differentiation of M. If we majorize the lower powers of θ by the highest and observe that μ^* would appear once in each term we can put the general estimate for $\dot z$ in this form:

$$\text{(B36)}\qquad \begin{aligned} \dot z &\lesssim P_1(\xi)\mu^*\left[{}^{-4,0}_{\ 0,4}\right], &&\text{or} \\ &\lesssim \gamma^*\left[{}^{-4,0}_{\ 0,4}\right], \end{aligned}$$

where $P_1(\xi)$ is simply some fourth degree polynomial in ξ, the first of a series of such polynomials that we shall use, analogous to the series of numbered constants. Note that $\dot z$ has the same θ dependence as M.

We can use (B36) and an estimate

$$\zeta - z \lesssim C_7\xi\left[{}^{-3,1}_{\ 0,4}\right],$$

obtained via (A17) and (B26), to estimate E. Using $E = 2(\zeta - z)' \otimes \dot z'$ we obtain

$$\text{(B37)}\qquad \begin{aligned} E &\lesssim 2\{C_7\xi\left[{}^{-3,1}_{\ 0,4}\right]\}' \otimes \{\gamma^*\left[{}^{-4,0}_{\ 0,4}\right]\}', &&\text{or} \\ E &\lesssim 2C_7\xi\left[{}^{-2,1}_{\ 0,3}\right] \otimes \gamma^*\left[{}^{-3,0}_{\ 0,3}\right], &&\text{or} \\ E &\lesssim C_8\xi\gamma^*\left[{}^{-5,-2}_{\ 0,\ 3}\right] \\ &\lesssim \eta^*\left[{}^{-5,-2}_{\ 0,\ 3}\right]. \end{aligned}$$

This illustrates the pattern of the bound corresponding to the product of two bounds expressed in our notation.

The estimate of z is the most laborious. The simple estimate obtainable through $z = z_0 + \int_{\theta_2}^{\theta} \dot z$ is not good enough for the third derivatives of z. We need an estimate on integrals of M as an intermediate step. Here also the direct estimation is inadequate, but we can use (B18a) to say

$$\int_{\theta_2}^{\theta_3} M(\theta)\, d\theta = u(\theta_3 - \theta_0) S_{\theta_3} G + S_{\theta_3} L(\theta_3) - u(\theta_2 - \theta_0) S_{\theta_2} G - S_{\theta_2} L(\theta_2).$$

Applying (A15) to the G and L bounds, assuming $\theta_2 \leqq \theta_3$, and using $|u| \leqq 1$, we can say

$$\int_{\theta_2}^{\theta_3} M(\theta)\, d\theta \lesssim C_9(\delta + \lambda^*)\, [\theta_3 \mid {}^{0,1}_{3,4}].$$

The straightforward integration of the M estimate, (B34), yields

$$\int_{\theta_2}^{\theta_3} M(\theta)\, d\theta \lesssim \int_{\theta_2}^{\theta_3} \mu^* \, [{}^{-4,0}_{0,4}]\, d\theta$$

$$\lesssim C_{10}\, \mu^* \, [\theta_2 \mid {}^{-3,-1}_{0,\ 2}].$$

This estimate is stated for differentiation only up to the second order because for the third order $\int \theta^{-1}\, d\theta$ would lead to a logarithmic term; and for the fourth order θ_3^{+1} would majorize θ_2^{+1} rather than the θ_2 term majorizing the θ_3 term (which occurs up through the second order derivatives).

Now we have two estimates, one good for lower derivatives, and one good for higher ones. If we add the two estimates we can safely extend the range of the second and modify the first in the range covered by the second. This gives us

(B38) $$\int_{\theta_2}^{\theta_3} M(\theta)\, d\theta \lesssim C_{11}(\mu^* + \delta + \lambda^*)\, \{[\theta_2 \mid {}^{-3,1}_{0,4}] + [\theta_3 \mid {}^{-3,1}_{0,4}]\}.$$

We also need an $\dot F$ estimate in obtaining our z estimate. Symbolically written,

$$\dot F = F_{\zeta'} \dot\zeta' + F_{\zeta''} \dot\zeta''.$$

To use this we must estimate $\dot\zeta$, which is $(\dot{S_\theta z}) = \dot S_\theta z + S_\theta \dot z$. So from (A16) and (B26) and from (A15) and (B29) we derive

$$\dot\zeta \lesssim C_{12}\xi[{}^{-4,2}_{0,6}] + C_{13}\gamma[{}^{-4,2}_{0,6}]$$

$$\lesssim C_{14}(\xi + \gamma)[{}^{-4,2}_{0,6}].$$

Using this $\dot\zeta$ estimate we can now estimate $\dot F$ in a manner exactly analogous to the manner in which we estimated $\dot z = F \boxtimes M$. The result has the form

(B39) $$\dot F \lesssim P_2(\xi)(\xi + \gamma)[{}^{-3,2}_{0,5}] + P_3(\xi)(\xi + \gamma)[{}^{-2,2}_{0,4}]$$

$$\lesssim P_4(\xi)(\xi + \gamma)[{}^{-2,2}_{0,4}].$$

Finally, we need an estimate on F itself. Here the highest power of θ will come from the maximum differentiation of ζ''. So the estimate takes the form

$$F \lesssim P_5(\xi)\left[{}^{0,3}_{1,4}\right] \tag{B40}$$

and has the same θ dependence as ζ''.

The z estimate

We actually estimate $z - z_0$, then z is easily estimated from this result. First,

$$\begin{aligned} z(\theta_3) - z_0 &= \int_{\theta_0}^{\theta_3} \dot{z}\, d\theta \\ &= \int_{\theta_0}^{\theta_3} F \boxtimes M\, d\theta \\ &= \int_{\theta_0}^{\theta_3} F \boxtimes \left[-\frac{\partial}{\partial \theta_2}\int_{\theta_2}^{\theta_3} M\, d\theta\right] d\theta \\ &= -\int_{\theta_0}^{\theta_3} F \boxtimes \left(\frac{\partial}{\partial \theta_2}\int_{\theta_2}^{\theta_3} M\, d\theta\right) d\theta. \end{aligned}$$

Now we apply integration by parts and have

$$z(\theta_3) - z_0 = -\left[F \boxtimes \int_{\theta_2}^{\theta_3} M\, d\theta\right]_{\theta_2=\theta_0}^{\theta_2=\theta_3} + \int_{\theta_0}^{\theta_3} \dot{F} \boxtimes \left(\int_{\theta_2}^{\theta_3} M\, d\theta\right) d\theta_2 .$$

The first term can be evaluated and we obtain

$$\begin{aligned} &z(\theta_3) - z_0 \\ &\quad = F(\zeta'(\theta_0), \zeta''(\theta_0)) \boxtimes \int_{\theta_0}^{\theta_3} M\, d\theta + \int_{\theta_0}^{\theta_3} \left\{\dot{F}(\text{at } \theta_2) \boxtimes \int_{\theta_2}^{\theta_3} M\, d\theta\right\} d\theta_2 . \end{aligned} \tag{B41}$$

At this point we insert the estimates (B40) for F, (B39) for $\dot{F}$, and (B38) for $\int M\, d\theta$. This yields

(B42) $z - z_0$

$$\lesssim P_5(\xi)\left[\theta_0 \mid {}^{0,3}_{1,4}\right] \boxtimes C_{11}(\mu^* + \delta + \lambda^*)\left\{\left[\theta_0 \mid {}^{-3,1}_{0,4}\right] + \left[\theta_3 \mid {}^{-3,1}_{0,4}\right]\right\}$$

$$+ \int_{\theta_0}^{\theta_3} \left\{P_4(\xi)(\xi + \lambda)\left[\theta_2 \mid {}^{-2;2}_{0,4}\right]\right\} \boxtimes C_{11}(\mu^* + \delta + \lambda^*)\left\{\left[\theta_2 \mid {}^{-3,1}_{0,4}\right] + \left[\theta_3 \mid {}^{-3,1}_{0,4}\right]\right\} d\theta_2 .$$

Call the two terms on the right T_1 and T_2. Using $\theta_3 \geqq \theta_0 \geqq 1$ we can weaken and simplify T_1 and have

$$T_1 \lesssim P_6(\xi)(\mu^* + \delta + \lambda^*)\left[\theta_3 \mid {}^{0,1}_{3,4}\right].$$

T_2 must be handled with more care. Each derivative of T_2 would correspond to the sum of several integrals involving various powers of θ_2 and θ_3 in the integrand. For the r^{th} order derivatives of T_2 there are terms of the form

$$\int_{\theta_0}^{\theta_3} \text{constant} \cdot \theta_2^{s} \cdot \theta_3^{r-s-5}\, d\theta_2 ,$$

where $r = 0, 1, 2, 3, 4$ and either $s = r - 5$ (which makes $r - s - 5 = 0$) or s satisfies $-2 \leqq s \leqq r - 2$. These two alternatives correspond to the two expressions $[\theta_2 \mid {}^{-3,1}_{0,4}]$ and $[\theta_3 \mid {}^{-3,1}_{0,4}]$ which are added in T_2. These integrals give varied terms and we can handle the situation most clearly by simply listing all cases. (This is done in Figure 1.)

	$r = 0$	$r = 1$	$r = 2$	$r = 3$	$r = 4$
$s = r - 5$	$\frac{1}{4}\theta_0^{-4}$	$\frac{1}{3}\theta_0^{-3}$	$\frac{1}{2}\theta_0^{-2}$	θ_0^{-1}	$\log(\theta_3/\theta_0)$
$s = \quad -2$	$\theta_0^{-1}\,\theta_3^{-3}$	$\theta_0^{-1}\,\theta_3^{-2}$	$\theta_0^{-1}\,\theta_3^{-1}$	θ_0^{-1}	$\theta_0^{-1}\,\theta_3$
-1		$\log(\theta_3/\theta_0)\theta_3^{-3}$	$\log(\theta_3/\theta_0)\theta_3^{-2}$	$\log(\theta_3/\theta_0)\theta_3^{-1}$	$\log(\theta_3/\theta_0)$
0			θ_3^{-2}	θ_3^{-1}	1
1				$\frac{1}{2}\theta_3^{-1}$	$\frac{1}{2}$
2					$\frac{1}{3}$
majorizer	θ_0^{-4}	θ_0^{-3}	θ_0^{-2}	θ_0^{-1}	$\theta_0^{-1}\,\theta_3$

FIG. 1

By using the majorizing terms listed at the bottom of the chart we can say

$$T_2 \lesssim P_7(\xi)(\xi + \gamma)(\mu^* + \delta + \lambda^*)\theta_0^{-1}[\theta_3 \mid {}^{0,1}_{3,4}].$$

Because the pattern of powers of θ_0 does not fit into our notation scheme we have simply used the highest power (θ_0^{-1}) which occurs.

Now if we add the T_1 and T_2 estimates we get an estimate for $z(\theta_3) - z_0$:

$$\text{(B43)} \qquad \begin{aligned} z(\theta_3) - z_0 &\lesssim P_8(\xi)(1 + \xi + \gamma)(\mu^* + \delta + \lambda^*)[\theta_3 \mid {}^{0,1}_{3,4}], \quad \text{or} \\ z - z_0 &\lesssim \beta^*[{}^{0,1}_{3,4}]. \end{aligned}$$

Since $z = z_0 + (z - z_0)$ we can say

$$\text{(B44)} \qquad \begin{aligned} z &\lesssim (\alpha + \beta^*)[{}^{0,1}_{3,4}] \\ &\lesssim \xi^*[{}^{0,1}_{3,4}]. \end{aligned}$$

To conclude the rederivation of bounds we must consider the requirement (B22). We have stipulated $\theta_0 \geqq \theta_a$ so that

$$S_\theta z_0 - z_0 \leqq \varepsilon/2[{}^0_2].$$

Applying (A15) to (B43) we see that

$$\zeta - S_\theta z_0 = S_\theta(z - z_0) \leqq C_{15}\beta^*[{}^{-3,1}_{0,4}].$$

Adding inequalities,

$$\text{(B45)} \qquad \begin{aligned} \zeta - z_0 = (\zeta - S_\theta z_0) + (S_\theta z_0 - z_0) &\leqq (C_{15}\beta^* + \varepsilon/2)[{}^0_2] \\ &\leqq \varepsilon^*[{}^0_2]; \end{aligned}$$

we hope, of course, that $\varepsilon^* \leqq \varepsilon$.

Strong consistency of the bounds

We verify here that if δ is sufficiently small we can choose θ_0 so that $\theta_0 \geqq \theta_a$ as required and so that all the rederived bounds are smaller than the originals: $\lambda^* < \lambda$, $\mu^* < \mu$, $\varepsilon^* < \varepsilon$, etc. For clarity the formulas for all the rederived bounds are assembled here:

$$\text{(B46)}\qquad \begin{aligned} \lambda^* &= \eta\theta_0^{-1} && \text{(B32)} \\ \mu^* &= C_3(\theta_0 + 1)^4\delta + C_2\delta + C_2\lambda^* + C_5\theta_0^{-1}\eta && \text{(B34)} \\ \gamma^* &= P_1(\xi)\mu^* && \text{(B36)} \\ \eta^* &= C_8\xi\gamma^* && \text{(B37)} \\ \beta^* &= P_8(\xi)(1 + \xi + \gamma)(\mu^* + \delta + \lambda^*) && \text{(B43)} \\ \xi^* &= \alpha + \beta^* \qquad (\xi = \alpha + \beta) && \text{(B44)} \\ \varepsilon^* &= \varepsilon/2 + C_{15}\beta^* && \text{(B45).} \end{aligned}$$

Consider any set λ, μ, γ, η, β, ξ, ε (with $\xi > \alpha$) of positive original bounds and consider the behavior of the rederived bounds as $\theta_0 \to \infty$ and as $\delta \to 0$ for each θ_0 value. That is, consider

$$\lim_{\theta_0 \to \infty} \lim_{\delta \to 0}$$

of each rederived bound.

First we see that

$$\lim_{\theta_0 \to \infty} [\lim_{\delta \to 0} \lambda^*] = 0,$$

or $\lambda^* \to 0$. Since $\lambda^* \to 0$ we can see that $\mu^* \to 0$. Therefore $\gamma^* \to 0$, whence $\eta^* \to 0$. Because μ^*, δ, $\lambda^* \to 0$ we see that $\beta^* \to 0$. Therefore $\xi^* \to \alpha$ and $\varepsilon^* \to \varepsilon/2$. These observations make it clear that, regardless of the sizes of the bounds λ, μ, $\cdots$ assumed originally, if θ_0 is taken large enough and δ is sufficiently small the rederived bounds λ^*, μ^*, $\cdots$ will all be smaller than the original bounds.

Existence of the solution

To use the results above we need a general understanding of the equation system defining our perturbation process. We must know there will be a local continuation whenever we have a well-behaved solution of the system in an interval $\theta_0 \leqq \theta < \theta_1$. Actually, the equation system is a very innocuous one. The smoothing makes the solutions analytic functions and makes them very well behaved, at least over small ranges of θ.

The form which the system takes above, as (B12) through (B18), does not reveal its true character very directly. The extent of the taming effect of S_θ is not fully apparent. Another formulation, given below, makes this tameness apparent.

A removable aspect of the system as presented above is the differentiation with respect to spatial coordinates, which is indicated by priming, as with ζ'

or z'. This differentiation makes the system look like a partial differential equation system, a type of system where local continuations of solutions often do not generally exist.

If a quantity is defined by smoothing, such as $\zeta = S_\theta z$, then its space derivative can be expressed as the result of an appropriate operator's action on the original quantity. So let us say

$$\zeta' = S_\theta' z.$$

For higher derivatives we have analogous operators S_θ'', S_θ''', etc.; and for $\dot{S}_\theta$ there is a similar series $\dot{S}_\theta'$, $\dot{S}_\theta''$, $\cdots$. These derived operators also have smoothing properties. By using them we can recast the system as an innocuous functional-integral-equation system.

Regard z, z', L and $\dot{L}$ as basic quantities. Then the others are expressible as functions or functionals of these four, either direct or indirect:

$$\begin{aligned}
\zeta &= S_\theta z, \qquad \zeta' = S_\theta' z, \qquad \zeta'' = S_\theta'' z, \qquad \zeta''' = S_\theta''' z \\
M &= S_\theta(\dot{u}(\theta - \theta_0)G + \dot{L}) + \dot{S}_\theta(u(\theta - \theta_0)G + L) \\
M' &= S_\theta'(\dot{u}(\theta - \theta_0)G + \dot{L}) + \dot{S}_\theta'(u(\theta - \theta_0)G + L) \\
\dot{z} &= F(\zeta', \zeta'') \boxtimes M \\
\dot{z}' &= (F_{\zeta'} \cdot \zeta'' + F_{\zeta''} \cdot \zeta''') \boxtimes M + F \boxtimes M' \\
\dot{g} &= 2z' \otimes \dot{z}' \\
E &= M - \dot{g} \quad \text{or} \quad E = 2(\zeta' - z') \otimes \dot{z}'.
\end{aligned}$$

The basic quantities are to be equal to certain integrals:

$$z = z_0 + \int_{\theta_0}^{\theta} \dot{z}, \qquad z' = z_0' + \int_{\theta_0}^{\theta} \dot{z}'$$

$$L = \int_{\theta_0}^{\theta} u(\theta - \bar{\theta}) E(\bar{\theta})\, d(\bar{\theta}), \quad \dot{L} = \int_{\theta_0}^{\theta} \dot{u}(\theta - \bar{\theta}) E(\bar{\theta})\, d\bar{\theta}.$$

Now, since all the functions, functionals, and operations involved are well-behaved, in fact, analytic, we have a well-behaved integral equation system. Indeed, the system is especially tame. Since $\dot{z}$ and $\dot{z}'$ are defined through smoothed quantities exclusively and since z_0 and z_0' are analytic, therefore z and z' are smooth and analytic (as functions of the space variables). Consequently $\dot{g}$ and E are smooth. So L and $\dot{L}$ are also smooth. Thus all the quantities are kept smooth by the presence of S_θ in the equations.

Continuation

Suppose the system has a solution for $\theta_0 \leqq \theta < \theta_1$. We can show that this solution will satisfy the bounds λ, μ, etc. The bounds will certainly hold at the beginning of the solution, so they will hold in an interval $\theta_0 \leqq \theta \leqq \varphi$ where $\theta_0 < \varphi < \theta_1$, if they do not hold up to θ_1. But now the smaller bounds

λ^*, μ^*, $\cdots$ ε^* must hold in $\theta_0 \leqq \theta \leqq \varphi$. Therefore the larger bounds (from continuity) must hold in some interval past φ. This contradiction proves that the larger bounds, and therefore also the smaller rederived bounds, hold in the whole range $\theta_0 \leqq \theta < \theta_1$ of the solution.

Now suppose θ_1 is the limit of continuation of the solution. There will certainly be a solution in the closed interval $\theta_0 \leqq \theta \leqq \theta_1$ because of the bounds and the tameness of the system. But a standard argument, such as the Picard method or the functional fixed point approach, will show that the system has a local continuation past θ_1. Furthermore, the continuation of a solution will be unique. Therefore we have contradicted the hypothesis and we see that the solution will exist, will be unique, and will satisfy the bounds for all values $\theta \geqq \theta_0$.

Convergence to isometry

We must show that the imbeddings $z(\theta)$ occurring in the perturbation process tend to a limit imbedding which realizes the desired metric $G + g_0$, where g_0 is the metric of the initial imbedding.

The Cauchy criterion approach to the proof of convergence of the imbedding requires us to consider

$$\begin{aligned} z(\theta_2) - z(\theta_1)\,, \quad \text{which} \quad &= \int_{\theta_1}^{\theta_2} \dot{z}(\theta)\, d\theta\,, \quad \text{which} \\ &\lesssim \int_{\theta_1}^{\theta_2} \gamma\left[{}^{-4,\,0}_{\;\;0,\,4}\right] d\theta\,, \quad \text{or} \\ &\lesssim \gamma\left[\theta_1 \,\middle|\, {}^{-3,\,-1}_{\;\;0,\;\;\;2}\right]. \end{aligned}$$

This estimate on $z(\theta_2) - z(\theta_1)$ and its derivatives is good enough to show that the imbedding $z(\theta)$ and its first and second derivatives converge to a limit $z(\infty)$. The integration of the $\dot{z}$ bound is too crude an approach for the third derivatives.

To check the convergence of the metric, $g(\theta)$, induced by the imbedding to the desired limit, $g_0 + G$, observe that

$$\begin{aligned} g(\theta_1) &= g_0 + \int_{\theta_0}^{\theta_1} \dot{g}\, d\theta \\ &= g_0 + \int_{\theta_0}^{\theta_1} (M - E)\, d\theta \qquad \text{(B48)} \\ &= g_0 + \int_{\theta_0}^{\theta_1} M\, d\theta - \int_{\theta_0}^{\theta_1} E\, d\theta. \end{aligned}$$

Now apply (B18a):

$$g(\theta_1) = g_0 + u(\theta_1 - \theta_0) S_{\theta_1} G + S_{\theta_1} L(\theta_1) - \int_{\theta_0}^{\theta_1} E\, d\theta.$$

Assume $\theta_1 \geqq \theta_0 + 1$ so that $u(\theta_1 - \theta_0) = 1$. Then we can say

$$g(\theta_1) = g_0 + S_{\theta_1} G + S_{\theta_1} L(\theta_1) - \int_{\theta_0}^{\theta_1} [1 - u(\theta_1 - \theta)] E(\theta)\, d\theta - \int_{\theta_0}^{\theta_1} u(\theta_1 - \theta) E(\theta)\, d\theta.$$

The first integrand vanishes for $\theta \leqq \theta_1 - 1$, and the second integral is $L(\theta_1)$, so

$$\begin{aligned} g(\theta_1) = {} & g_0 + G + (S_{\theta_1} - 1)[G + L(\theta_1)] \\ & + \int_{\theta_1 - 1}^{\theta_1} [u(\theta_1 - \theta) - 1] E(\theta)\, d\theta. \end{aligned} \tag{B49}$$

Let us call the last two terms remainders R_1 and R_2. Then by (A17) and (B27) and (B31) we see that

$$R_1 \precsim C_{16}(\delta + \lambda)[\theta_1 \mid {}^{-3,0}_{\;0,3}].$$

From (B30),

$$R_2 \precsim \eta[\theta_1 - 1 \mid {}^{-5,-2}_{\;0,\;3}].$$

The bound on R_2 is good enough for third derivatives but the R_1 bound is not. So we see that the metric $g(\theta)$ converges to the desired limit metric and also see that the first and second derivatives converge to the corresponding derivatives of the limit metric. But we do not yet know that the third derivatives converge. The limit metric, $g_0 + G$, is C^3, of course.

More refined results

This section of Part B can be passed over without losing the continuity of the paper. Here we show that the limit imbedding is actually always C^3 and we treat the cases where G is $C^4, C^5, \cdots, C^\infty$.

The cases where G is C^k are treated by an inductive method. The C^k case is handled with the aid of the results of the C^{k-1} case. We can illustrate all the essential features of the induction (with minimum notational difficulty) by considering the step from C^3 to C^4. To begin, we can assume bounds

$$\begin{aligned} G &\precsim \delta'[{}^0_4] \\ z_0 &\precsim \alpha'[{}^0_4] \end{aligned} \tag{B50}$$

that have the same form as (B24) and (B31), with 4 replacing 3. However, we make a point of not requiring δ' to be small.

Assume that the perturbation process is applied just as if G were only C^3. We will have the bounds (B22) through (B31) on the quantities involved. Our result will be that we can deduce a new set of bounds that are analogous to these except for the systematic replacement of 3 by 4, 4 by 5, etc., in the indices referring to order of differentiation.

In the above work which computed the set of rederived bounds λ^*, μ^*, etc., wherever an estimate was derived for a quantity defined by smoothing the estimate could have been derived for any particular upper limit on the order of differentiation. Instead of (B34) we could have derived

$$M \lesssim \bar{\mu}\left[{}^{-4,\,1}_{\;0,\,5}\right].$$

This is not the kind of estimate we ultimately want for M. That would be an $\left[{}^{-5,\,0}_{\;0,\,5}\right]$ estimate. This is just a useful intermediate step. Similarly we can say

$$\zeta \lesssim \bar{C}_6\xi\left[{}^{0,\,4}_{3,\,7}\right] \qquad \text{and}$$

$$\dot{z} \lesssim \bar{\gamma}\left[{}^{-4,\,2}_{\;0,\,6}\right].$$

By integrating this and using (B50) to estimate fifth derivatives and by using (B26) for the lower derivatives we get a crude z bound:

$$z \lesssim \bar{\xi}\left[{}^{0,\,2}_{3,\,5}\right].$$

From this we have

$$\zeta - z \lesssim \bar{C}_7\bar{\xi}\left[{}^{-3,\,2}_{\;0,\,5}\right].$$

Since $E = 2(\zeta - z)' \otimes \dot{z}'$, it can be estimated from our above results and we find

$$E \lesssim \bar{\eta}\left[{}^{-5,\,-1}_{\;0,\;\;4}\right].$$

A weaker E estimate, more conveniently manipulated, is

$$E \lesssim \bar{\eta}\left[{}^{-4\frac{1}{2},\,-\frac{1}{2}}_{\;\;0,\;\;\;4}\right].$$

This weakened estimate behaves nicely in an integrand and we get new L and $\dot{L}$ estimates via the L and $\dot{}$ formulas. These are

$$\text{(B51)}\qquad \begin{aligned} L &\lesssim \bar{\lambda}\left[{}^{0,\,\frac{1}{2}}_{3,\,4}\right] \\ \dot{L} &\lesssim \bar{C}_5\bar{\eta}\left[{}^{-4\frac{1}{2},\,-\frac{1}{2}}_{\;\;0,\;\;\;4}\right]. \end{aligned}$$

Now reconsider the estimation of M and use the new L and $\dot{L}$ bounds and the new G bound, (B50). One gets

$$M \lesssim \mu^{\#}\left[{}^{-4\frac{1}{2},\,\frac{1}{2}}_{\;\;0,\;\,5}\right].$$

From this we obtain

$$\dot{z} \lesssim \gamma^{\#}\left[{}^{-4\frac{1}{2},\,\frac{1}{2}}_{\;\;0,\;\,5}\right], \qquad \text{then}$$

$$E \lesssim \eta^{\#}\left[{}^{-5\frac{1}{2},\,-1\frac{1}{2}}_{\;\;0,\;\;\;4}\right].$$

This sharper E bound yields improved bounds for L and $\dot{L}$:

$$\begin{aligned} L &\lesssim \lambda'\left[{}^{0}_{4}\right] \\ \dot{L} &\lesssim \bar{C}_5\eta^{\#}\left[{}^{-5\frac{1}{2},\,-1\frac{1}{2}}_{\;\;0,\;\;\;4}\right]. \end{aligned}$$

Now from these and the new G bounds we can derive

$$M \precsim \mu'\left[\begin{smallmatrix}-5 & ; & 0\\ 0 & ; & 5\end{smallmatrix}\right], \qquad \text{then}$$

$$\dot{z} \precsim \gamma'\left[\begin{smallmatrix}-5 & ; & 0\\ 0 & ; & 5\end{smallmatrix}\right], \qquad \text{then}$$

$$E \precsim \eta'\left[\begin{smallmatrix}-6 & ; & -2\\ 0 & ; & 4\end{smallmatrix}\right].$$

The z estimate depends on estimating $z - z_0$. This can be done exactly like it was done before in computing the rederived bounds. We can use a weak $\dot{\xi}$ estimate, improved only by extension to sixth derivatives. This gives us a weak $\dot{F}$ estimate, which is however adequate because it is used in combination with an improved estimate on $\int_{\theta_2}^{\theta_3} M\, d\theta$. This improved estimate comes from the strong M, L, and G bounds above (with μ', λ', δ'). The result is of the form

$$z - z_0 \precsim \beta'\left[\begin{smallmatrix}0 & ; & 1\\ 4 & ; & 5\end{smallmatrix}\right]$$

$$z = \precsim \xi'\left[\begin{smallmatrix}0 & ; & 1\\ 4 & ; & 5\end{smallmatrix}\right],$$

and this completes the set of new C^4 type estimates (indicated by primed greek letters).

Our result concerning the inductive extension of appropriate bounds to the cases where G is C^k also gives us a result for the C^∞ case. The (typical) induction step from the C^3 case to the C^4 case shown above, involved the use of a new G bound (B50). But we did not assume δ' was small. So our results on the C^k cases with $k > 3$ are really sharper than the exactly analogous results would be, because only G and its derivatives up to the third order need be small.

If G is C^∞ the result of each C^k case is valid, so the imbedding is C^∞.

We can show that when G is C^k the k^{th} derivatives of the imbedding converge so the limit imbedding is C^k. Also the k^{th} derivatives of the metric converge. The argument is the same for any value of k, and since we have not shown this for $k = 3$, we treat that case.

The basic fact is that the limits of the third derivatives of $S_\theta G$ as $\theta \to \infty$ are the third derivatives of G, and that this convergence is uniform. This follows from the uniform continuity of the third derivatives of G, a consequence of compactness. The fact can be symbolized by saying

$$\text{(B52)} \qquad G - S_\theta G \precsim \Delta_\theta\left[\begin{smallmatrix}0\\ 3\end{smallmatrix}\right],$$

where Δ_θ is a constant for each θ value and $\Delta_\theta \to 0$ as $\theta \to \infty$. Nothing can be said about the rapidity of the decrease of Δ_θ; the fact is purely qualitative.

The estimate of (B51) on L, although derived after we assumed G to be C^4, did not depend on this assumption. So we can use it here, where G is again only assumed to be C^3. By applying (A17) to it one can obtain

$$(1 - S_{\theta_1})L(\bar{\theta}) \precsim C_{17}\bar{\lambda}(\bar{\theta})^{\frac{1}{2}}\left[\theta_1 \,\middle|\, \begin{smallmatrix}-4 & ; & 0\\ 0 & ; & 4\end{smallmatrix}\right].$$

Assuming $\theta_2 \geqq \theta_1$, it follows that

$$S_{\theta_2}L(\bar{\theta}) - S_{\theta_1}L(\bar{\theta}) \precsim 2C_{17}\bar{\lambda}(\bar{\theta})^{\frac{1}{4}}[\theta_1 \mid {}^{-4,\,0}_{\;0,\,4}]. \tag{B53}$$

We shall need a bound on $L(\theta_2) - L(\theta_1)$. Assume $\theta_2 \geqq \theta_1 \geqq \theta_0 + 1$, then

$$\begin{aligned} L(\theta_2) - L(\theta_1) &= \int_{\theta_0}^{\theta_2} u(\theta_2 - \theta)E(\theta)\, d\theta - \int_{\theta_0}^{\theta_1} u(\theta_1 - \theta)E(\theta)\, d\theta \\ &= \int_{\theta_1 - 1}^{\theta_2} \{u(\theta_2 - \theta) - u(\theta_1 - \theta)\}E(\theta)\, d\theta, \end{aligned}$$

because $u(\theta_2 - \theta) = u(\theta_1 - \theta) = 1$ for $\theta \leqq \theta_1 - 1$ and because $u(\theta_1 - \theta) = 0$ for $\theta \geqq \theta_1$. Using (B30),

$$\begin{aligned} L(\theta_2) - L(\theta_1) &\precsim \int_{\theta_1 - 1}^{\theta_2} \eta[{}^{-5,\,-2}_{\;0,\;\;3}]\, d\theta \\ &\precsim \eta[\theta_1 - 1 \mid {}^{-4,\,-1}_{\;0,\;\;3}] \\ &\precsim 16\eta[\theta_1 \mid {}^{-4,\,-1}_{\;0,\;\;3}]. \end{aligned} \tag{B54}$$

We must have a more refined estimate on $\int_{\theta_1}^{\theta_2} M$. By (B18a), assuming $\theta_2 \geqq \theta_1 \geqq \theta_0 + 1$,

$$\begin{aligned} \int_{\theta_1}^{\theta_2} M\, d\theta &= S_{\theta_2} G + S_{\theta_2} L(\theta_2) - S_{\theta_1} G - S_{\theta_1} L(\theta_1) \\ &= S_{\theta_2} G - S_{\theta_1} G + S_{\theta_1}[L(\theta_2) - L(\theta_1)] + (S_{\theta_2} - S_{\theta_1})L(\theta_2) \\ &= T_a + T_b + T_c\,. \end{aligned} \tag{B55}$$

Our principal concern is the third derivatives, which our previous estimate, (B38), merely showed bounded, not decreasing. By (B52),

$$T_a \precsim (\Delta_{\theta_1} + \Delta_{\theta_2})[{}^0_3].$$

If $\theta_1, \theta_2 \to \infty$ the third derivatives of T_a will approach zero, which is what we want.

(B54) is strong enough to show that the third derivatives of T_b will approach zero as $\theta_1, \theta_2 \to \infty$. But T_c requires a more elaborate treatment. We can write

$$T_c = (S_{\theta_2} - S_{\theta_1})L(\theta_2) = (S_{\theta_2} - S_{\theta_1})[L(\bar{\theta}) + \{L(\theta_2) - L(\bar{\theta})\}]$$

where $\bar{\theta}$ should be regarded as smaller than θ_1. Then by (A15) applied to (B54) and by (B53) we obtain

$$T_c \precsim C_{18}\eta[\bar{\theta} \mid {}^{-4,\,-1}_{\;0,\;\;3}] + 2C_{17}\bar{\lambda}(\bar{\theta})^{\frac{1}{4}}[\theta_1 \mid {}^{-4,\,0}_{\;0,\,4}].$$

Observe now that T_c and its derivatives up to order 3 can be forced to be arbitrarily small if one first chooses $\bar{\theta}$ large enough to make the first term above small and then chooses θ_1 enough larger than $\bar{\theta}$ so that the second term is small. Effectively then, as θ_1, $\theta_2 \to \infty$ the term T_c and its derivatives up to the third order approach zero.

Since T_a''', T_b''', and T_c''' all approach zero as θ_1, $\theta_2 \to \infty$, we have shown that

$$\left[\int_{\theta_1}^{\theta_2} M \, d\theta\right]''' \to 0 \quad \text{as} \quad \theta_1, \theta_2 \to \infty.$$

For the lower derivatives of this intergal the simple estimate obtained by integrating (B28) is adequate. We can combine the two approaches in a single estimate:

$$\text{(B56)} \qquad \int_{\theta_1}^{\theta_2} M \, d\theta \lesssim \mu[\theta_1 \mid {}^{-3,\,-1}_{0,\;\;2}] + \mu_{\theta_1}[\theta_1 \mid {}^{-3,0}_{0,3}].$$

μ_θ is to be a constant for each θ which approaches zero as $\theta \to \infty$.

We show that the third derivatives of the imbedding $z(\theta)$ converge by a Cauchy criterion argument. The difference $z(\theta_3) - z(\theta_1)$ can be expressed in a form exactly analogous to (B41), where θ_1 appears in place of θ_0:

$$z(\theta_3) - z(\theta_1) = F(\zeta'(\theta_1), \zeta''(\theta_1)) \boxtimes \int_{\theta_1}^{\theta_3} M \, d\theta \qquad (= T_1)$$

$$+ \int_{\theta_1}^{\theta_3} \left\{ \dot{F}(\text{at } \theta_2) \boxtimes \int_{\theta_2}^{\theta_3} M \, d\theta \right\} d\theta_2 \quad (= T_2).$$

The analogue of the estimate we obtained, below (B42), for T_1 is not adequate for our needs here. But the analogous T_2 estimate

$$T_2 \leqq P_7(\xi)(\xi + \gamma)(\mu^* + \delta + \lambda^*)\theta_1^{-1}[\theta_3 \mid {}^{0,1}_{3,4}]$$

is quite good because θ_1^{-1} becomes small as θ_1, $\theta_2 \to \infty$.

The estimate (B56) was obtained so that we could handle T_1. We use this with (B40) for F and obtain

$$T_1 \leqq P_9(\xi)\mu[\theta_1 \mid {}^{-3,\,-1}_{0,\;\;2}] + P_{10}(\xi)\mu[\theta_1 \mid {}^{-4,\,-1}_{0,\;\;3}] + P_{11}(\xi)\mu_{\theta_1}[\theta_1 \mid {}^{-3,0}_{0,3}].$$

This shows that $T_1''' \to 0$ as θ_1, $\theta_2 \to \infty$ and that is what we need to show that

$$[z(\theta_3) - z(\theta_1)]''' \to 0 \quad \text{as} \quad \theta_1, \theta_2 \to \infty.$$

So we have verified that the third derivatives of the imbedding converge uniformly to the third derivatives of the limit imbedding (which therefore must exist and be continuous).

To see that the third derivatives of the metric converge consider

$$g(\theta_2) - g(\theta_1) = \int_{\theta_1}^{\theta_2} \dot{g} \, d\theta = \int_{\theta_1}^{\theta_2} (M - E) \, d\theta$$

$$= \int_{\theta_1}^{\theta_2} M \, d\theta - \int_{\theta_1}^{\theta_2} E \, d\theta$$

$$\lesssim \mu[\theta_1 \mid {}^{-3,\,-1}_{0,\;\;2}] + \mu_{\theta_1}[\theta_1 \mid {}^{-3,0}_{0,3}] + \eta[\theta_1 \mid {}^{-4,\,-1}_{0,\;\;3}],$$

from (B56) and (B30). Since $\mu_{\theta_1} \to 0$ as $\theta_1 \to \infty$, this gives us the convergence of the metric and its derivatives up to the third order.

Summary of results

The main results of this part of the paper can be summarized in one theorem:

THEOREM 1:

HYPOTHESES:

(1) *$\mathfrak{M}$ is a compact manifold analytically imbedded in a euclidean space.*

(2) *The system* (B8a, b) *of linear equations in the $\dot{z}_\alpha$ is non-singular at all points of the imbedding.*

(3) *G is a symmetric covariant tensor on $\mathfrak{M}$ representing the change we want to make in the metric induced by the imbedding of $\mathfrak{M}$. We want to accomplish this change by modifying the imbedding.*

(4) *G is C^k where $3 \leqq k \leqq \infty$.*

(5) *θ_0 is the parameter determining the initial amount of smoothing in our pertubation process.*

CONCLUSIONS:

If θ_0 is taken sufficiently large and if G and its derivatives up to the third order are sufficiently small, then the perturbation process will produce a perturbed imbedding of $\mathfrak{M}$ that is C^k and induces a metric tensor on $\mathfrak{M}$ which differs by the amount G from the metric induced by the original imbedding.

PART C: PREPARATORY METRIC APPROXIMATION

In Theorem 1 of Part B we have the means for making small changes in the metric induced by an imbedding. Here in Part C we learn how to arrange that only a small change is needed. This solves the imbedding problem for compact Riemannian manifolds.

The addition property

Suppose a manifold $\mathfrak{M}$ has two imbeddings, one by functions z^α into E^m and the other by functions y^β into E^p. Let

$$\text{(C1)} \qquad g_z = \sum_\alpha \frac{\partial z_\alpha}{\partial x_i}\frac{\partial z_\alpha}{\partial x_j} \quad \text{and} \qquad g_y = \sum_\beta \frac{\partial y_\beta}{\partial x_i}\frac{\partial y_\beta}{\partial x_j}$$

be the two metric tensors defined by the two imbeddings, both referred to a system $x_1, x_2, \cdots x_n$ of local coordinates in $\mathfrak{M}$. The total set of functions z_1, $z_2, \cdots z_m, y_1, y_2, \cdots y_p$ defines an imbedding of $\mathfrak{M}$ into the product space $E^m \times E^p$. The metric tensor induced on $\mathfrak{M}$ by this product imbedding is $g_z + g_y$. This is the addition property of metric tensors.

This property enables us to separate into two parts the problem of constructing an imbedding of $\mathfrak{M}$ such that we can successfully apply Theorem 1 and obtain an isometric imbedding. Suppose g is the intrinsic metric of $\mathfrak{M}$ which we want to realize by an imbedding. We first find a z-imbedding which is "perturbable", that is, one such that the equation system (B8a, b) is everywhere non-singular. Then we find a y-imbedding such that $g_z + g_y$ is close to g. The perturbation process is now applied only to the z-imbedding, directed towards effecting the change $G = g - (g_z + g_y)$ in the metric induced by the z-imbedding.

Actually, the form of Theorem 1 forces us to proceed somewhat like this. This theorem tells us that for any perturbable imbedding a sufficiently small change G in the metric can be accomplished. To use it we must be able to make G small without changing the imbedding which is to be perturbed. And derivatives of G up to the third order must be small. Thus it is rather important that we make $G(= g - g_z - g_y)$ small by adjusting the y-imbedding, leaving the z-imbedding fixed.

The metric g_y must be a positive metric, so it is clear that it would be impossible to use the approach we outlined above unless $g - g_z$ were a positive metric, which g_y would approximate to. If we have a z-imbedding that is perturbable (i.e. where (B8a, b) is everywhere a non-singular system) we can always simply make a change of scale, if necessary, to make g_z as small as desired so that $g - g_z$ will be positive. This does not affect the quite qualitative question of the singularity or non-singularity of (B8a, b).

A simplified approach

The method we actually use here for constructing the z and y imbeddings has a certain intricacy occasioned by our desire to bound the number of dimensions needed for the final imbedding and to get a relatively good bound. But if one throws out all concern for the number of dimensions to be used the problem can be handled rather simply. Therefore we indicate here the simpler approach for the benefit of those who may not want to bother with the more complicated details of our method.

One could get by with two lemmas:

L_1 : *Every compact differentiable manifold may be imbedded as an analytic submanifold of some euclidean space in such a way that* (B8a, b) *is everywhere non-singular.*

L_2 : *Any compact Riemannian manifold with C^k positive metric, where* $3 \leqq k \leqq \infty$, *can be represented as an analytic sub-manifold of euclidean space so that the induced metric and its derivatives up to the third order approximate the given metric as closely as desired.*

These lemmas would not give us any indication of the number of dimensions necessary.

L_1 can be proved very easily by a direct construction. Take an analytic imbedding of $\mathfrak{M}$ in E^{2n}, where $n = \dim \mathfrak{M}$. Let $v_1, v_2, \cdots, v_{2n}$ be the coordinates of E^{2n}. Then the $2n^2 + 3n$ functions $v_1, v_2, \cdots, v_{2n}$; $v_1^2, v_2^2, \cdots, v_{2n}^2$; v_1v_2,

$v_1v_3, \cdots, v_2v_3, \cdots$ define an imbedding of $\mathfrak{M}$ in $(2n^2 + 3n)$-space if they are used as imbedding functions. This imbedding has the property we want, that (B8a, b) is everywhere non-singular. To verify this, consider n of the v's as local coordinates at a point of $\mathfrak{M}$. [See (C7).]

L_2 could be proved by considering finite dimensional approximations to a C^k isometric imbedding of $\mathfrak{M}$ in Hilbert space. Or one could use a result of J. Schwartz that implies L_2. It is relatively easy to patch up a good approximation to a metric when there is no limit on the number of imbedding functions to be used. Each function might vanish except on a small neighborhood of $\mathfrak{M}$.

Outline of the method

Our method for constructing y-imbeddings can use a minimum number of dimensions if the z-imbedding is chosen with some care. It turns out that if $g - g_z$ is close to a favorable metric γ the y-imbedding can be made in $n^2 + 3n$ dimensions.

Our first step is to find a favorable metric γ. Then we construct a z-imbedding such that $g - g_z$ approximates γ (derivatives are not involved here). Finally we construct the y-imbedding so that g_y approximates $g - g_z$ and its derivatives.

Determination of γ

The special mechanism we use in constructing the y-imbedding requires a set $\{\psi^r\}$ of functions on $\mathfrak{M}$ such that the set of symmetric tensors

$$M^r_{ij} = \frac{\partial \psi^r}{\partial x_i} \frac{\partial \psi^r}{\partial x_j} \tag{C2}$$

has at every point of $\mathfrak{M}$ a subset of $\frac{1}{2}n^2 + \frac{1}{2}n$ tensors that are linearly independent. The x_i's are local coordinates in $\mathfrak{M}$. The metric γ will be the sum over r of the M^r_{ij}, or in other words, it will be the metric induced by the imbedding defined by using the ψ^r as imbedding functions.

Our construction of the ψ^r is based on a dimensionality argument. If we arrange to deal exclusively with algebraic functions and conditions we can think in terms of the precise dimensionality concepts of algebraic geometry (such as those based on transcendance degree). Therefore let $\mathfrak{A}$ be an algebraic representation[2] of $\mathfrak{M}$ in E^a. The mapping between $\mathfrak{M}$ and $\mathfrak{A}$ can be as differentiable as the differentiability structure on $\mathfrak{M}$, and analytic if $\mathfrak{M}$ has a *global* analytic structure (*not* mere overlapping local coordinates analytically related). Then a C^k metric on $\mathfrak{M}$ becomes a C^k metric on $\mathfrak{A}$.

Let us first see how we can find a set of functions which has the independence or "(C2) property" in the neighborhood of a point of $\mathfrak{A}$. Suppose $x_1, x_2, \cdots, x_n$ are local coordinates. Then the $\frac{1}{2}n^2 + \frac{1}{2}n$ functions

$$f^{ij} = x_i + x_j, \qquad \text{where } i \leqq j, \tag{C3}$$

[2] Reference [16] proves that there is an algebraic representation of any closed differentiable manifold.

can easily be seen to suffice. At any point of $\mathfrak{A}$ some n of the coordinates u_1, $u_2, \cdots, u_a$ of E^a will be suitable as local coordinates. Therefore the $\frac{1}{2}a^2 + \frac{1}{2}a$ functions

$$f^{\beta\delta} = u_\beta + u_\delta, \qquad \text{where } \beta \leqq \delta, \tag{C4}$$

will have the "(C2) property" everywhere. But this is more functions than we want to use.

A plausibility argument for the number of functions ψ^r that should be necessary so that there are always $\frac{1}{2}n^2 + \frac{1}{2}n$ linearly independent M^r_{ij} at a point of $\mathfrak{A}$ goes as follows: $\frac{1}{2}n^2 + \frac{1}{2}n$ functions are good locally, so they would fail on a sub-manifold of $n - 1$ dimensions. Adding a total of $n - 1$ functions should reduce it step by step to a zero-dimensional set of singular points. Then one more function should eliminate these. Thus we should need $\frac{1}{2}n^2 + 1\frac{1}{2}n$ functions in all.

$\frac{1}{2}n^2 + 1\frac{1}{2}n$ is the correct number, although that argument is not rigorous. Define $\frac{1}{2}n^2 + 1\frac{1}{2}n$ functions

$$\psi^r = \textstyle\sum_\beta C^r_\beta u_\beta \qquad \begin{bmatrix} r = 1, 2, \cdots, \frac{1}{2}n^2 + 1\frac{1}{2}n \\ \beta = 1, 2, \cdots, a \end{bmatrix} \tag{C5}$$

as linear combinations of the coordinates of E^a. We shall show that a generic choice of the coefficients C^r_β automatically gives a set of ψ^r with the desired property. Let $s = \frac{1}{2}n^2 + \frac{1}{2}n$. Then there are $s + n$ of the ψ^r and $(s + n)a$ of the coefficients C^r_β.

Our dimensionality argument is based on analyzing the family of ways in which a set of the ψ^r can fail to define independent M^r_{ij}. If the M^r_{ij} are not linearly independent at a point p of $\mathfrak{A}$ they lie in some linear sub-space H_p of the space L_p of all values of symmetric tensors (with two subscripts) at p. We can consider only sub-spaces H_p which have one less dimension than the whole linear space L_p. Since $\dim L_p = s$, $\dim H_p = s - 1$ and the dimension of the family of sub-spaces H_p of L_p is $s - 1$. The dimension of the family of all H_p for all points p of $\mathfrak{A}$ is $n + s - 1$.

For any r the coefficients $(C^r_1, C^r_2, \cdots, C^r_a)$ can be chosen so that ψ^r is any one of the functions $f^{\beta\delta}$ of (C4). Therefore for any particular H_p, one can select $(C^r_1, C^r_2, \cdots, C^r_a)$ so that M^r_{ij} (which they determine, via ψ^r) does not lie in H_p. If this were not so the $f^{\beta\delta}$ would not have the "(C2) property," but we saw that they did have this property. Since not all selections of the constants determining ψ^r make M^r_{ij} lie in H_p, the dimension of the family of selections of $(C^r_1, C^r_2, \cdots, C^r_a)$ that do make $M^r_{ij} \epsilon H_p$ is not more than $a - 1$.

The dimension of the family of selections of all the C^r_β which make all the M^r_{ij} lie in H_p is clearly not more than $(s + n)(a - 1)$ since there are $s + n$ sets of a coefficients determining the $s + n$ functions ψ^r which determine the M^r_{ij}. Now since the family of all H_p has dimension $n + s - 1$, the dimension of the family of selections of the C^r_β for which there is some H_p such that all the M^r_{ij} lie in H_p at p can be at most $[(s + n)(a - 1)] + [n + s - 1]$. But this number is $(s + n)a - 1$, one less than the dimension of all selections of the C^r_β. Consequently with a generic selection of the C^r_β there will be no point p and tensor sub-space H_p where all the M^r_{ij} lie in H_p. In other terms, the ψ^r will have the "(C2)", or independence, property we want.

One can be quite explicit about how the C^r_β can be chosen. The n-dimensional variety associated with $\mathfrak{A}$ (which is the smallest variety containing $\mathfrak{A}$) can be defined by a set of polynomial equations in the coordinates of E^a. Adjoin all the coefficients involved in these equations to the field of rationals to produce an extension F. If the C^r_β are algebraically independent over F the ψ^r will have the desired property.[3] Obviously this is a sufficient but not necessary condition for the proper selection of the C^r_β.

We have said that the "favorable metric" γ would be the sum of the M^r_{ij}, so that

$$\text{(C6)} \qquad \gamma_{ij} = \sum\nolimits_r M^r_{ij} = \sum\nolimits_r \frac{\partial \psi^r}{\partial x_i}\frac{\partial \psi^r}{\partial x_j}.$$

We also said that our procedure was to make $g - g_z \approx \gamma$. This means $g_z \approx g - \gamma$, so we must certainly have $g - \gamma$ a positive metric. To take care of this let us assume that a definite choice of the C^r_β is made in such a way that they are small enough to make $g - \gamma$ be a positive metric.

The Z-imbedding

There will be a C^1 imbedding of $\mathfrak{M}$ in E^{2n} which realizes the metric $g - \gamma$ exactly [9]. We can think of this as a mapping from $\mathfrak{A}$ into E^{2n} and approximate it in the C^1 sense by an algebraic imbedding $\mathfrak{B}$. Let g_b be the metric induced by this imbedding $\mathfrak{B}$. The approximation between g_b and $g - \gamma$ can be as close as desired, but it does not extend to derivatives.

Let $v_1, v_2, \cdots, v_{2n}$ be the coordinates of E^{2n}. At any point of $\mathfrak{B}$ some n of these will be suitable as local coordinates; let $x_1, x_2, \cdots, x_n$ be this subset.

Now consider the functions

$$\text{(C7)} \qquad \begin{aligned} z_i &= x_i \\ z_{ij} &= x_i x_j \end{aligned} \qquad i \leqq j.$$

If these are the functions z_α of (B8a, b) that system takes a very simple form:

$$\sum\nolimits_\alpha \frac{\partial z_\alpha}{\partial x_i}\dot{z}_\alpha = \dot{z}_i = 0$$

$$-2\sum\nolimits_\alpha \frac{\partial^2 z_\alpha}{\partial x_i\,\partial x_j}\dot{z}_\alpha = -2\dot{z}_{ij} = \dot{g}_{ij}.$$

[3] By using rational coefficients in the approximating polynomials used in [16] one would obtain an algebraic imbedding $\mathfrak{A}$ such that the equations defining the corresponding variety would have rational coefficients. Then it would suffice to simply select the C^r_β as independent transcendentals over the field of rationals, without reference to $\mathfrak{A}$.

The solution is apparent for any $\dot{g}_{ij}$. The system is non-singular in the most obvious way. Each linear equation has just one variable (remember the $\dot{z}_\alpha$ are the variables) with non-vanishing coefficient and this is a different variable in each equation.

The system (B8a, b) has the property that once one has a set of functions z_α that makes the system non-singular then the introduction of new functions z_α (which also introduces new variables $\dot{z}_\alpha$) can only improve the situation. The coefficients of each variable $\dot{z}_\alpha$ in the equations can be regarded as a vector V_α with $\frac{1}{2}n^2 + \frac{1}{2}n + n$ or $\frac{1}{2}n^2 + 1\frac{1}{2}n$ components. The system is non-singular if there are $\frac{1}{2}n^2 + 1\frac{1}{2}n$, which is $s + n$, linearly independent V_α's. Additional functions z_α simply introduce more V_α's.

Define $s + 2n$ (or $\frac{1}{2}n^2 + 2\frac{1}{2}n$) quadratic functions of the coordinates of E^{2n} by

$$z_\alpha = \sum_{1 \leqq \beta \leqq 2n} C_\alpha^\beta v_\beta + \sum_{1 \leqq \beta \leqq \delta \leqq 2n} D_\alpha^{\beta\delta} v_\beta v_\delta , \tag{C8}$$

where $1 \leqq \alpha \leqq s + 2n$. At any point of $\mathfrak{B}$ we can make the z_α contain the functions defined by (C7) which are suitable at that point by an appropriate choice of the C's and the D's. An argument exactly analogous to the one we used above in finding the ψ^r shows that a generic choice of the C's and D's gives functions z_α that make (B8a, b) non-singular. If the C's and D's are algebraically independent over the field of definition for $\mathfrak{B}$ they define satisfactory functions z_α.

Now how do we arrange that $g_z \approx g - \gamma$? We select all $D_\alpha^{\beta\delta} \approx 0$ and $C_\alpha^\beta \approx 0$ unless $\beta = \alpha$, when we select $C_\alpha^\alpha \approx 1$. This makes the z-imbedding, which is in E^{s+2n}, be approximately the same as $\mathfrak{B}$, because all the z^α are quite small, except for the first $2n$, and these are approximately the coordinates $v_1, v_2, \cdots, v_{2n}$ of E^{2n}. Thus $g_z \approx g_b$ and hence $g_z \approx g - \gamma$ since $g_b \approx g - \gamma$.

The Y-imbedding

This is constructed, with the aid of the ψ^r of (C2), by means of a special device. The device will produce an imbedding with metric g_y approximating any metric of the form

$$g(a_1, a_2, \cdots, a_{s+n}) = \sum_r a_r M_{ij}^r , \tag{C9}$$

where the M_{ij}^r are those of (C2) and the a_r are positive analytic functions on $\mathfrak{M}$. (Regard $\mathfrak{M}$ as having an analytic structure corresponding to those of its imbeddings $\mathfrak{A}$ and $\mathfrak{B}$.) This approximation will apply to derivatives also.

Note that $g(1, 1, \cdots 1) = \gamma$. Since the M_{ij}^r are linearly independent at each point and $g - g_z \approx \gamma$, we can certainly represent $g - g_z$ in the form

$$g - g_z = \sum_r \alpha_r M_{ij}^r \tag{C10}$$

at any point of $\mathfrak{M}$, and we can have $| \alpha_r - 1 | \leqq \varepsilon$ where ε is a small uniform bound that depends on the closeness of the approximation of $g - g_z$ to γ.

But can we do this in a uniform manner so that the α_r become continuous functions on $\mathfrak{M}$? The solution for the α_r of (C10) will be non-unique at every point because there are $s + n$ of the α_r and only s components of $g - g_z$, thus only s equations. This redundancy can be removed by the same device which removed the redundancy of (B8a, b) in Part B. Here we desire positive α_r, so we specify:

$$\sum_r (\alpha_r - 1)^2 = \text{minimum, subject to (C10).} \tag{C11}$$

This determines the α_r uniquely and makes them have the same differentiability as $g - g_z$. Since (C10) has a solution with $|\alpha_r - 1| \leqq \varepsilon$, the solution of the modified system (C10, 11) will have $|\alpha_r - 1| \leqq \epsilon (s + n)^{\frac{1}{2}}$. Assume ε is small enough so that this makes the α_r necessarily positive.

We define $g(a_1, a_2, \cdots, a_{n+s})$ by approximating the α_r by positive analytic functions a_r. The a_r must also approximate the α_r for derivatives up to the third order. Hence $g - g_z$ must be C^3 so that the α_r are C^3. Since $a_r \approx_3 \alpha_r$ we shall have $g(a_1, a_2, \cdots, a_{n+s}) \approx_3 g - g_z$.

The device

This device is reminiscent of one used for constructing C^1 imbeddings.[4] Suppose λ is a large constant. We define $2(s + n)$, which is $n^2 + 3n$, y-imbedding functions as

$$\text{(C12)} \qquad \begin{aligned} y_r &= \frac{(a_r)^{\frac{1}{2}}}{\lambda} \sin(\lambda \psi^r) \\ \bar{y}_r &= \frac{(a_r)^{\frac{1}{2}}}{\lambda} \cos(\lambda \psi^r). \end{aligned}$$

When these are used as imbedding functions they induce the metric

$$g_y = \sum_r \frac{\partial y_r}{\partial x_i} \frac{\partial y_r}{\partial x_j} + \sum_r \frac{\partial \bar{y}_r}{\partial x_i} \frac{\partial \bar{y}_r}{\partial x_j}.$$

When this is expanded by substituting the formulas for y_r and $\bar{y}_r$ many terms cancel. All terms which contain λ^{-1} occur as pairs which differ only by containing either $[\sin(\lambda\psi^r)]\cdot[\cos(\lambda\psi^r)]$ or $[\cos(\lambda\psi^r)]\cdot[-\sin(\lambda\psi^r)]$ and cancel together. The remaining terms can be combined with the identity $\sin^2 + \cos^2 = 1$. This finally gives

$$\begin{aligned} g_y &= \sum_r a_r \frac{\partial \psi^r}{\partial x_i} \frac{\partial \psi^r}{\partial x_j} + \lambda^{-2} \sum_r \frac{\partial(a_r)}{\partial x_i} \frac{\partial(a_r)}{\partial x_j}, \text{ or} \\ g_y &= \sum_r a_r M^r_{ij} + \lambda^{-2} \bar{g} \\ &= g(a_1, \cdots, a_{s+n}) + \lambda^{-2} \bar{g}. \end{aligned}$$

Now $\bar{g}$ is an analytic tensor independent of λ. By choosing λ very large the error $\lambda^{-2}\bar{g}$ and any number of its derivatives can be made as small as desired. So we have

$$g_y \approx_3 g(a_1, a_2, \cdots, a_{s+n}) \approx_3 g - g_z, \tag{C14}$$

[4] See equation (13) page 387 of reference [9].

where $\approx_3$ indicates approximation up to third derivatives. Thus we have $g_y + g_z \approx_3 g$, which is what we need to apply Theorem 1. This requires, of course, that g be C^3.

Summary and applications

We used $s + 2n$, which is $\frac{1}{2}n^2 + 2\frac{1}{2}n$, z-imbedding functions and $2(s + n)$, or $n^2 + 3n$, y-imbedding functions. This is $3s + 4n$ or $1\frac{1}{2}n^2 + 5\frac{1}{2}n$ functions altogether. The z-imbedding was analytic and made (B8a, b) everywhere non-singular, so that Theorem 1 could be applied to it. The y-imbedding was also analytic and was adjustable so that $g_z + g_y$ could approximate g as closely as desired (this approximation including derivatives up to the third order). For this, g had to be C^3.

The z- and y-imbeddings can be arranged to take up arbitrarily little space. The z-imbedding approximates a C^1 imbedding of $\mathfrak{M}$ which realizes the metric $g - \gamma$. Since a C^1 isometric imbedding can be made arbitrarily small (and highly twisted) so can the z-imbedding. If the parameter λ is very large the y-imbedding is very small.

The fact that the z-imbedding approximates a C^1 imbedding serves to prevent self-intersections in the final imbedding of the manifold in $E^{1\frac{1}{2}n^2+5\frac{1}{2}n}$. Since the amount of perturbation needed can be made arbitrarily small by adjusting the y-imbedding, the application of the perturbation process to the z-imbedding need not produce self-intersections.

Now we state the result obtained by combining the work of Part C with Theorem 1 of Part B. This is our "main theorem."

Imbedding of compact manifolds

THEOREM 2. *A compact n-manifold with a C^k positive metric has a C^k isometric imbedding in any small volume of euclidean* $(n/2)(3n + 11)$*-space, provided* $3 \leqq k \leqq \infty$.

PART D: NON-COMPACT MANIFOLDS

Our treatment here is not a direct attack. It exploits a device by which the imbedding problem for non-compact manifolds is reduced to the problem for compact manifolds. This approach gives a poor upper bound on the number of dimensions needed for the imbedding space; but that is the price of taking a shortcut to the non-compact case.

A special mapping

Our basic tool is a C^∞ mapping of E^n onto the n-sphere S^n. Most of E^n is mapped into the "north pole" of S^n. The interior of the unit disk of E^n covers the remaining portion of S^n in a one-to-one manner. Any mapping of E^n on S^n with these properties will serve our purpose, but to illustrate we construct one.

Take the case of the (x, y) plane, or E^2. This can be mapped on the 2-sphere $\xi^2 + \eta^2 + \zeta^2 = \frac{1}{4}$ as follows:

For $x^2 + y^2 < 1$, let

$$Q = \exp (x^2 + y^2 - 1)^{-1}$$

$$\xi = \frac{xQ}{Q^2 + x^2 + y^2}$$

$$\eta = \frac{yQ}{Q^2 + x^2 + y^2}$$

$$\zeta = \frac{1}{2} - \frac{Q^2}{Q^2 + x^2 + y^2};$$

$$\text{for} \quad x^2 + y^2 \geqq 1, \qquad \xi = \eta = 0, \qquad \zeta = \tfrac{1}{2}.$$

It is easy to see that ξ, η, and ζ are C^∞ functions because Q is a C^∞ function if it is assigned the value zero for $x^2 + y^2 \geqq 1$. A direct check verifies that

$$\xi^2 + \eta^2 + \zeta^2 = \tfrac{1}{4}.$$

The equations define a non-singular one-to-one mapping of the open disk $x^2 + y^2 < 1$ onto the sphere minus the "north pole" ($\xi = \eta = 0, \zeta = \frac{1}{2}$). This mapping is obtainable by taking a mapping of the open disk onto the whole plane,

$$\bar{x} = x/Q$$

$$\bar{y} = y/Q$$

and following this by the classical conformal mapping of the plane onto the sphere (minus the "north pole"). The effect is to give a mapping which has a C^∞ extension to the rest of the plane where all points not interior to the disk map into the "north pole".

Patch mappings

Consider a C^∞ Riemannian n-manifold $\mathfrak{M}$ (the metric need not be C^∞, but $\mathfrak{M}$ has a C^∞ structure). A local coordinate system or neighborhood N in $\mathfrak{M}$ can be regarded as the image of the unit disk D of E^n under a C^∞ mapping of D into $\mathfrak{M}$. This mapping should be one-to-one, non-singular, and extensible to an open set of E^n containing D, as a C^∞ non-singular mapping. Then an open set containing N is mapped on the open set containing D by the C^∞ inverse mapping.

For brevity call S the n-sphere. Our special mapping will map any open set containing D onto S in a C^∞ manner. This mapping, with the inverse mapping $N \to D$, gives a mapping

$$N \xrightarrow{\varphi} S$$

which is C^∞ and has a C^∞ extension to an open set containing N. φ maps all points on the boundary of N or outside N into the "north pole" of S. Clearly we can extend the definition of φ to all points of $\mathfrak{M}$ by mapping all other points into this "north pole". φ will remain C^∞. This mapping φ has non-vanishing Jacobian in the interior of N, so φ^{-1} is C^∞ there. φ is called a patch mapping.

Appropriate coverings for $\mathfrak{M}$

$\mathfrak{M}$ can be covered by a family of disk neighborhoods N_i in such a way that we can divide the N_i among $n + 1$ classes where: No two N_i of the same class overlap. Each N_i overlaps only a finite number of other N_i.

How is such a covering constructed? First obtain a regular star-finite cellular sub-division of $\mathfrak{M}$. Then form a disk neighborhood corresponding to each vertex, edge, $\cdots$, face, or cell of the cellular sub-division. Each disk neighborhood that corresponds (for example) to an edge covers the middle section of the edge but not the end points. These are covered by the neighborhoods which correspond to them in their role of vertices. In this way no two edge neighborhoods are allowed to meet. The same principle applies up the series of dimensions. The $n + 1$ dimensions from 0 through n give rise to the $n + 1$ classes of neighborhoods.

Within each N_i we can select a slightly smaller disk neighborhood $\bar{N}_i$ such that the $\bar{N}_i$ also cover $\mathfrak{M}$. These $\bar{N}_i$ should correspond to sub-disks of D (through the mappings between D and the N_i). Then we can select a C^∞ function u_i for each $\bar{N}_i$ which is positive interior to $\bar{N}_i$ and zero on the boundary and outside $\bar{N}_i$. Each u_i can be regarded as defined and C^∞ over all of $\mathfrak{M}$.

Now if we define

$$v_i = u_i / \textstyle\sum_i u_i$$

the v_i form a partition of unity by C^∞ functions, each of which is positive interior to the corresponding sub-neighborhood $\bar{N}_i$ and zero everywhere else.

Assignment of metrics

Each N_i has an associated patch mapping

$$N_i \xrightarrow{\varphi_i} S_i .$$

We write S_i to distinguish different n-spheres for different N_i. The mapping φ_i has a C^∞ non-singular inverse φ_i^{-1} on $\bar{N}_i$.

Consider a metric γ_{i0} on S_i. This gives a corresponding metric g_{i0} on N_i. Actually g_{i0} can be regarded as defined for all $\mathfrak{M}$ because it will be zero at the boundary of N_i and can be extended by defining it as zero on the rest of $\mathfrak{M}$. If it is C^k on S_i it will be C^k on $\mathfrak{M}$ also. If we select a metric γ_{i0} on each S_i which is positive, C^∞, and sufficiently small the corresponding metrics g_{i0} will add on $\mathfrak{M}$ to a metric

$$g_0 = \textstyle\sum_i g_{i0}$$

which is C^∞ and everywhere shorter than g, the metric we are given on $\mathfrak{M}$, and which we want to realize by an imbedding of $\mathfrak{M}$. For example, each γ_{i0} could be the metric induced by an imbedding of S_i in E^{n+1} as a small geometrical sphere.

Now, since $g - g_0$ is a positive metric,

$$g_i = g_{i0} + v_i(g - g_0)$$

will be positive and will be as differentiable as g (say C^k). g_i differs from g_{i0} only within $\bar{N}_i$ where the mapping φ_i has a non-singular inverse φ_i^{-1}. Therefore φ_i^{-1} carries $g_i - g_{i0}$ over to S_i as a C^k non-negative metric $\gamma_i - \gamma_{i0}$. That is, there is a positive C^k metric γ_i on S_i which corresponds, via φ_i, to the metric g_i on N_i.

Consider the sum

$$\begin{aligned}\textstyle\sum_i g_i &= \textstyle\sum_i g_{i0} + (\sum_i v_i)(g - g_0)\\ &= g_0 + (g - g_0)\\ &= g.\end{aligned}$$

Realization of metrics

We have defined a C^k metric γ_i on each S_i. Assuming $3 \leqq k \leqq \infty$, γ_i can be realized by a C^k imbedding of S_i in $E^{(n/2)(3n+11)}$ by our Theorem 2. We can always let the "north pole" of S_i be at the origin.

Now consider all the N_i of one of the $n + 1$ classes, let us say class C. The S_i corresponding to each has an imbedding in $E^{(n/2)(3n+11)}$ that realizes γ_i and maps the "north pole" into the origin. The corresponding patch mappings φ_i, together with these imbeddings, define a mapping ψ_C of $\mathfrak{M}$ into $E^{(n/2)(3n+11)}$. This ψ_C is C^∞ and maps all points of $\mathfrak{M}$, except those interior to any neighborhood N_i of class C, into the origin. Each of these other points can be in only one N_i, so the mapping is unambiguously defined.

This mapping ψ_C induces a metric on $\mathfrak{M}$ which is the sum,

$$g_c = \sum_{N_i \in C} g_i ,$$

of the metrics g_i associated with neighborhoods N_i of class C. The product mapping

$$\psi = \psi_1 \times \psi_2 \times \cdots \times \psi_{n+1}$$

maps $\mathfrak{M}$ into $(n + 1)(n/2)(3n + 11)$ dimensional space and is also C^k. It induces the metric

$$g = \sum_C g_C,$$

as we desired. If the image of $\mathfrak{M}$ under ψ has no self-intersections it is an isometric imbedding. In any case it is a C^k isometric immersion.

Avoidance of self-intersections

Self-intersections can be avoided by using the fact that the isometric imbedding of S_i can be made as small as desired. Let α_i be the minimum distance of a point of $\bar{N}_i$ from the origin after it is mapped into $E^{1\frac{1}{2}n^2+5\frac{1}{2}n}$ by φ_i and the imbedding of S_i. Let β_i be the maximum distance from the origin of points of N_i after being mapped into $E^{1\frac{1}{2}n^2+5\frac{1}{2}n}$. Think of i as a serial index running 1, 2, 3 $\cdots$. Now all we need do to avoid self-intersections is to arrange the imbeddings of the S_i in order of increasing i so that for all i

$$\beta_i < \min_{j<i} \alpha_j .$$

Why is this sufficient? First, any two points of $\mathfrak{M}$ that are interior to a common N_i are distinguished through the imbedding of S_i. So we need only consider pairs of points which lie in completely different sets of neighborhoods. In this case one of the points will be in a sub-neighborhood $\bar{N}_i$ with lower index than the other. Then this point will be further from the origin with respect to the set of $1\frac{1}{2}n^2 + 5\frac{1}{2}n$ coordinates associated with N_i than the other point can be.

The result

Our theorem is:

THEOREM 3. *Any Riemannian n-manifold with C^k positive metric, where $3 \leqq k \leqq \infty$, has a C^k isometric imbedding in $(1\frac{1}{2}n^3 + 7n^2 + 5\frac{1}{2}n)$-space, in fact. in any small portion of this space.*

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

BIBLIOGRAPHY

1. C. F. GAUSS, Disquisitiones generales circa superfices curvas, Werke IV, 1827.
2. B. RIEMANN, *Über die hypotesen welche der geometrie zugrunde liegen*, Gött. Abh., 13 (1868), pp. 1–20.
3. L. SCHLAEFLI, *Nota alla memoria del. Sig. Beltrami, Sugli spazii di curvatura constante*, Ann. di mat., 2e serie, 5 (1871–1873), pp. 170–193.
4. D. HILBERT, *Ueber flächen von constanter Gausscher krümmung*, Trans. Amer. Math. Soc., 2 (1901), pp. 87–99.
5. C. TOMPKINS, *Isometric embedding of flat manifolds in Euclidean space*, Duke Math. J., 5 (1939), pp. 58–61.
6. S. S. CHERN and N. H. KUIPER, *Some theorems on the isometric imbedding of compact Riemann manifolds in Euclidean space*, Ann. of Math., 56 (1952), pp. 422–430.
7. M. JANET, *Sur la possibilité de plonger un espace riemannien donné dans un espace euclidien*, Annales de la société polonaise de mathématique, 5 (1926), pp. 38–43.
8. E. CARTAN, *Sur la possibilité de plonger un espace riemannien donné dans un espace euclidien*, Annals de la société polonaise de mathématique, 6 (1927), pp. 1–7.
9. J. NASH, *C^1 isometric imbeddings*, Ann. of Math., 60 (1954), pp. 383–396.
10. N. H. KUIPER, *On C^1 isometric imbeddings*, Proc. Kon. Ac. Wet. Amsterdam A 58 (Indigationes Mathematicae) no. 4 (1955), pp. 545–556.
11. H. WEYL, *Über die bestimmung einer geschlossenen konvexen flache durch ihr linienelement*, Vierteljahrsschrift der naturforschender Gesellschaft, Zurich, 61 (1916), pp. 40–72.
12. H. LEWY, *On the existence of a closed convex surface realizing a given Riemannian metric*, Proc. Nat. Acad. Sci. U.S.A., 24, No. 2 (1938), pp. 104–106.
13. A. D. ALEXANDROV, Intrinsic geometry of convex surfaces, OGIZ, Moscow-Leningrad, 1948.
14. A. V. POGORELOV, Deformation of convex surfaces, Gosudarstv. Izdat. Tehn.—Teor. Lit., Moscow-Leningrad, 1951. (Also see later papers of Alexandrov and Pogorelov.)
15. LOUIS NIRENBERG, *The Weyl and Minkowski problems in differential geometry in the large*, Comm. Pure Appl. Math., 6 (1953), pp. 337–394.
16. J. NASH, *Real algebraic manifolds*, Ann. of Math., 56 (1952), pp. 405–421.
17. DANILO BLANUŠA, *Über die Einbettung hyperbolischer Räume in euklidische Räume*, Monatschefte für Mathematik, 59 Band, 3 Heft (1955), pp. 217–229.