CHARACTERISTICS AND EXISTENCE OF ISOMETRIC EMBEDDINGS

ROBERT L. BRYANT, PHILLIP A. GRIFFITHS AND DEANE YANG

Let (M^n, ds^2) be an n-dimensional Riemannian manifold. A well-known problem is to prove the existence of a local C^∞ isometric embedding

$$(M^n, ds^2) \hookrightarrow E^{n(n+1)/2}. \tag{1}$$

By this we mean that there is a smooth isometric embedding of a neighborhood of a given point $x_0 \in M$; to simplify notation, we shall refer to this neighborhood also as M.

When (M, ds^2) is real analytic, the Burstin–Cartan–Janet–Schafly theorem (cf. the references given in [3, 11]) shows that such local isometric embeddings exist.

When $n = 2$ it is also known that local C^∞ isometric embeddings exist in a neighborhood of a point x_0 where the Gaussian curvature $K(x_0) \neq 0$.

When $n \geqslant 2$ it has been proved by R. Greene [7] that local C^∞ isometric embeddings

$$(M^n, ds^2) \rightarrow E^{(n(n+1)/2)+n}$$

always exist.

In general we may consider the exterior differential system (I, ω) whose integrals give local isometric embeddings

$$(M^n, ds^2) \rightarrow E^{(n(n+1)/2)+s}. \tag{2}$$

The basic invariant of (I, ω) is its characteristic sheaf $\mathcal{M}$. We may think of $\mathcal{M}$ as a family of vector spaces $\mathcal{M}_{(x,\xi)}$ of varying dimension whose support

$$\text{supp} \, \mathcal{M} = \{ (x, \xi) : \dim \mathcal{M}_{(x,\xi)} > 0 \}$$

is the characteristic variety Ξ of (I, ω). However, $\mathcal{M}$ contains much more information, both locally and globally, than Ξ alone. The system (I, ω) is determined when $s = 0$, underdetermined when $s > 0$, and overdetermined when $s < 0$. This is reflected in the properties of $\mathcal{M}$ in a precise way (cf. the appendix to §II(c)).

In particular, let us consider the case $s = 0$. Although the system (I, ω) is only invariant under the group $E(n)$ of Euclidean motions, it turns out that both $\mathcal{M}$

and Ξ are invariant under the full group $A(n)$ of affine linear transformations. For example, when $n = 2$ the system is elliptic when $K > 0$ and hyperbolic when $K < 0$; the sign of K, but not its value, is invariant under the action of $A(2)$. Therefore, it is not surprising that understanding $\mathcal{M}$ is basically an algebro-geometric problem, one dealing with the interplay between the Gauss equations and the linear systems of quadrics that constitute the second fundamental forms of embeddings (1).

Our first main result, Theorem A in §I(b), shows that a rational involution j is canonically defined on Ξ, and that $\mathcal{M}$ is uniquely determined by the pair (Ξ, j). It is a general fact that the characteristic sheaf $\mathcal{M}$ uniquely determines the symbol map of an involutive exterior differential system (I, ω). We find as a corollary that the second fundamental form of a general $M^n \subset E^{n(n+1)/2}$ is uniquely determined up to general linear transformations by (Ξ, j).

The characteristic variety is a union $\Xi = \bigcup_{x \in M} \Xi_x$ of projective algebraic varieties $\Xi_x \subset P^{n-1}$; Theorem B, as stated in §II(c), describes the local and global structure of Ξ_x. In summary, we find that for the general $M^n \subset E^{n(n+1)/2}$,

$$
\begin{cases}
\Xi_x \text{ is smooth for } n \leqslant 4, \\
\Xi_x \text{ is non-empty for } n \geqslant 3, \\
\deg \Xi_{C,x} = n, \text{ where } \Xi_{C,x} \text{ is the complexification of } \Xi_x, \\
\Xi_{x,sg} \neq \varnothing \text{ for } n = 4m + 2, m \geqslant 1, \text{ where } \Xi_{x,sg} \text{ is the singular locus of } \Xi_x.
\end{cases}
$$

$\{\Xi_{C,x}, x \in M\}$ is a family of beautiful and remarkable hypersurfaces. For example, when $n = 4$ the quotient $\Xi_{C,x}/j$ is an Enriques surface in $G(4, 10)$ first considered by Cayley and Sylvester.

It is a general fact that the characteristic sheaf and variety of an exterior differential system (I, ω) induce the usual (at least for the characteristic variety) objects for the linear variational equations of an integral manifold N of (I, ω). For the isometric embedding system the symbol map of these linear variational equations are fiberwise of the form

$$\gamma_H : W \otimes S^2 V^* \to K \tag{3}$$

where

$$
\begin{cases}
W \cong R^{n(n-1)/2} & \text{is the normal space,} \\
V \cong R^n & \text{is the cotangent space,} \\
K \cong R^{n^2(n^2-1)/12} & \text{is the space of curvature-like tensors.}
\end{cases}
$$

In fact, (3) is just the map given by polarizing the (quadratic) Gauss equations. When $n \geqq 3$, $\dim(W \otimes S^2 V^*) < \dim K$; therefore, the linearized isometric embedding equations "appear" to be overdetermined. Intuitively, this is because the naïve system

$$(dx, dx) = ds^2, \tag{4}$$

which is determined, has been prolonged so as to uncover the geometry as expressed in terms of the 2nd fundamental form and curvature tensor.

In concrete terms, we may try to represent a local isometric embedding (1) as a graph

$$x \rightarrow (x, z(x))$$

where

$$x = (x^1, \dots, x^n) = (x^i) \in \mathsf{R}^n$$

$$z = (z^1, \dots, z^{n(n-1)/2}) = (z^\mu) \in \mathsf{R}^{n(n-1)/2}.$$

The Gauss equations are then expressed by relations of the (approximate) form

$$R_{ijkl}(x) = \sum_\mu \det \left\| \begin{matrix} \dfrac{\partial^2 z^\mu}{\partial x^i \partial x^k} & \dfrac{\partial^2 z^\mu}{\partial x^i \partial x^l} \\[2mm] \dfrac{\partial^2 z^\mu}{\partial z^j \partial z^k} & \dfrac{\partial^2 z^\mu}{\partial x^j \partial x^l} \end{matrix} \right\| \tag{5}$$

When $n = 2$ this is a Monge–Ampere equation, but when $n = 3$ it gives 6 equations in the 3 unknowns $z^\mu(x)$. For $n \geqq 4$ the situation only worsens.

The general theory of exterior differential systems suggests "deprolonging" the system (I, ω) by replacing $\mathcal{M}$ by $\mathcal{M} \otimes \mathcal{O}(-1)$. When this is done, the characteristic variety Ξ remains the same, but the symbol mapping now becomes a square matrix. (This "deprolonging" of (I, ω) does not lead to (4), for which the characteristic variety is wrong.) In fact, this deprolonged system may be concretely realized by a certain uncoupling of the linearization of the naïve system (4) into an $n \times n$ system plus an algebraic one. All of this is explained in §II, the upshot being that the linearized isometric embedding system reduces to an $n \times n$ system with characteristic sheaf $\mathcal{M}(-1)$ whose microlocal structure can be explicitly determined (cf. Theorem C in §II(d)).

When $n = 3$ the situation is especially beautiful. The characteristic varieties Ξ_x are cubic curves in P^2; and the pair (Ξ_x, j) determines a "parent cubic" $\Psi_x \subset \mathsf{P}^2$. By putting the equation of Ψ_x in standard form, the 2nd fundamental form of an $M^3 \subset \mathsf{E}^6$ and characteristic cubic Ξ_x can be made very explicit and simple (cf. Theorem D in §III(a)).

It is well known that a smooth real cubic curve $\Xi_x \subset \mathsf{P}^2$ has either two or one components. The linearized isometric embedding system is accordingly strictly hyperbolic or real principal type. Moreover, for a certain range of the modulus σ of the parent cubic, the system is symmetric hyperbolic in the sense of Friederichs [6]. This is all worked out in Theorem E, §III(b).

In §III(c) we investigate the relationship between the modulus σ and the solutions to the Gauss equations. On a 3-dimensional Riemannian manifold it is well-known that the curvature tensor is determined by the Ricci tensor R_{ij}. For our purposes it is more natural to use the (equivalent) Einstein tensor $\mathcal{G}$. We

show in Theorem F that the range of values of σ for which the Gauss equations may be solved depends on the signature of $\mathscr{G}$.

The last three sections of this paper combine the geometric information obtained thus far with PDE theory to prove our

MAIN THEOREM. *Let* (M, ds^2) *be a 3-dimensional* C^∞ *Riemannian manifold and* $x_0 \in M$ *a point such that the Einstein tensor* $\mathscr{G}(x_0)$ *is not* (L^2) *where* $L \in T^*_{x_0}(M)$. *Then there exists a local* C^∞ *isometric embedding of a neighborhood of* x_0 *into* E^6.

Before discussing our proof, we remark our general results on characteristic varieties reduce the question of local C^∞ isometric embeddings (1) for any n to a local existence theorem, with suitable bounds, for a type of linear PDE's that we call "generic." To explain what this means, we denote by $\mathscr{L}_n$ the space of $n \times n$ matrices and by

$$\mathscr{L}_{n,k} \subset \mathscr{L}_n$$

the subvariety of matrices of rank $\leqq n - k$ (thus $\mathscr{L}_{n,1}$ is given by det $= 0$). There is a canonical stratification

$$\mathscr{L}_n \supset \mathscr{L}_{n,1} \supset \mathscr{L}_{n,2} \supset \cdots \supset \mathscr{L}_{n,n-1} \supset (0) \tag{6}$$

with the properties

$$\begin{cases} (\mathscr{L}_{n,k})_{sg} = \mathscr{L}_{n,k+1} \\ \text{codim} \, \mathscr{L}_{n,k} = k^2. \end{cases}$$

Consider now a linear PDE system

$$P(x, D)u = f, \qquad x \in \Omega, \tag{7}$$

where u, f are n-vectors of functions and $P(x, D)$ is an $n \times n$ matrix of 1st order linear differential operators. We view the symbol of (7) as a map

$$P(x, \xi) : T^*(\Omega) \backslash \{0\} \to \mathscr{L}_n, \tag{8}$$

and shall say that (7) is *generic* in case the mapping (8) is transverse to the stratification (6) (transverse here has a precise meaning in terms of ideals of defining functions). In particular, if we set

$$\Xi_k = P(x, \xi)^{-1}(\mathscr{L}_{n,k}),$$

then

$$\begin{cases} \Xi = \Xi_1 \text{ is the characteristic variety of (7);} \\ (\Xi)_{sg} = \Xi_2 \text{ is the singular locus of } \Xi; \text{ and} \\ \Xi_k \neq \varnothing \Rightarrow \text{codim} \, \Xi_k = k^2. \end{cases}$$

The linearized isometric embedding equations turn out to be generic in this sense.

§§IV–V are a compilation of various results needed to prove the Main Theorem. Although most are not new, they are presented in a form particularly suited to our application. Then in §VI the proofs of Theorem H and the Main Theorem are given.

First, the Nash–Moser–Schwartz–Sergeraert theorem is stated. This is a generalized implicit function theorem using a simple version of the Nash–Moser iteration scheme. The iteration method was first devised by Nash [16] to prove his famous C^∞ global isometric embedding theorem, and was simplified by J. Moser [15]. The theorem given here is due to J. T. Schwartz [17] with a refinement due to M. F. Sergeraert [19].

To use the Nash–Moser iteration scheme, a family of smoothing operators is needed. The existence of such operators on R^n and compact manifolds is well known (cf. [17, 18]). However, the domains used here are manifolds with boundary. In §IV(c), the construction of smoothing operators on R^n and T^n is recalled; and then to define smoothing operators on compact manifolds with boundary, an extension operator defined by E. Stein [21] is used. The extension operator will also be an important tool in §V.

§IV(d) is essentially a review of some results obtained in §§II–III that will be needed for the proof of Theorem H. The only difference is that the equations are given explicitly in local co-ordinates, with all the invariant, abstract definitions unraveled. In particular, the $n \times n$ system of PDE's which is equivalent to the linearized isometric embedding system is described in detail.

When trying to embed the general $M^3 \subset E^6$, we find that the linearized isometric embedding equations reduce to a 3×3 linear hyperbolic system. §V contains a detailed discussion on how to solve such a system with the bounds necessary for the Nash–Moser–Schwartz–Sergeraert theorem. There are essentially two types of hyperbolic systems. The easier of the two to study is called "symmetric hyperbolic." Such systems were first considered by K. O. Friederichs [6] who showed that the Cauchy problem is well-posed for such systems. S. Klainerman [13] used the Nash–Moser iteration scheme to prove the global existence of decaying solutions to certain nonlinear hyperbolic PDE's. We reproduce here a clever argument he used, combining some calculus lemmas with Gronwall's inequality to prove that symmetric hyperbolic systems can be solved with the desired bounds. Since we use a different set of Banach spaces than he did, our results extend his slightly.* (See *Corrections* A and B at end of paper.)

The second type of hyperbolic system is called "strictly hyperbolic." These were first studied by Leray, Petrowsky, and Gårding. Estimates for such a system are now usually proved using a "symmetrizing" pseudodifferential operator (cf. [23]). Although the proof of the needed estimates is essentially the same as for a symmetric hyperbolic system, the details are greatly complicated by the use of pseudodifferential operators. Therefore, the proofs for such systems have been reserved for the appendix to §V.

Finally all the results given in §IV–V are brought together to prove Theorem

---

*Hamilton has proved the same estimates using a different technique in his paper *The Inverse Function Theorem of Nash–Moser*, Bull. Amer. Math. Soc. **17** (1982), 65–222.

H which shows that all metrics sufficiently close to a metric which is isometrically embeddable are also isometrically embeddable. A short argument exploiting the Burstin–Cartan–Janet–Schafly theorem then completes the proof of the Main Theorem.

We would like to thank Eric Berger, S. S. Chern, Robert Greene, Sergiu Klainerman, Richard Melrose, Louis Nirenberg, Donald Spencer, and S. T. Yau for their help and encouragement.

I. The characteristic variety of the isometric embedding system.

(a)   We consider a submanifold $M^n \subset \mathsf{E}^N$. Since we are working locally we shall fix a point $x_0 \in M$ and work in a small neighborhood, still denoted by M, of x_0. We shall denote by (M, ds^2) the abstract Riemannian manifold associated to $M \subset \mathsf{E}^N$, and we shall consider the embedded manifold as the image of an isometric embedding

$$x : (M, ds^2) \to \mathsf{E}^N. \tag{1.a.1}$$

By $\mathscr{F}(M)$ we shall denote the manifold of Darboux frames $(x; e_1, \ldots, e_n; e_{n+1}, \ldots, e_N)$ associated to $M \subset \mathsf{E}^N$. Employing summation convention and using the ranges of indices $1 \leq i, j \leq n$, $n+1 \leq \mu, \nu \leq N$, we have the structure equations (cf. (1.27) and (1.31) in [4] for notations)[1]

$$\begin{cases} dx = \omega^i e_i \\ de_i = \psi_i^j e_j \\ d\omega^i = -\psi_j^i \wedge \omega^j \\ d\psi_j^i + \psi_k^i \wedge \psi_j^k = h_i^\mu \wedge h_j^\mu \\ h_i^\mu = H_{ij}^\mu \omega^j, \qquad H_{ij}^\mu = H_{ji}^\mu \\ h_i^\mu \wedge h_j^\mu = \tfrac{1}{2} R_{ijkl} \omega^k \wedge \omega^l \end{cases}$$

Here $\psi = \|\psi_j^i\|$ is (the pullback to $\mathscr{F}(M)$ of) the connection form for (M, ds^2) and

$$H = H_{ij}^\mu e_\mu \otimes \omega^i \omega^j$$

is the 2nd fundamental form of $M \subset \mathsf{E}^N$. The intrinsic and extrinsic geometry are related by the last of these equations, the Gauss equations, which following the notation in (1.36)–(1.38) of [4] we write as

$$\gamma(H, H) = R.$$

We will denote by $\mathscr{F}(M)$ the principal $O(n)$-bundle of orthonormal frames on

---

[1]Generally speaking we shall follow the terminology, notations, and ranges of indices from [4]. One exception is that the abstract in Riemannian manifold is here denoted by (M, ds^2) instead of by $\tilde{M}$ as in [4].

(M, ds^2) and by $\mathcal{F}(\mathsf{E}^N)$ the frame manifold for E^N. With the notations of §II(a) of [4] (cf. 2.25) we denote by (I, ω) the exterior differential system on the manifold $X \subset \mathcal{F}(M) \times \mathcal{F}(\mathsf{E}^N) \times (W \otimes S^2 V^*)$ which gives the isometric embeddings of (M, ds^2) into E^N. More precisely, the admissible integral manifolds $N \subset X$ of (I, ω) are in one-to-one correspondence with the local isometric embeddings. Since (I, ω) is a quasi-linear Pfaffian system (loc. cit., (2.20)), on a suitable bundle of projective spaces lying over X there is defined the characteristic sheaf $\mathcal{M}$ whose support is the characteristic variety Ξ of (I, ω) (cf. [5] for these definitions). Our object in this section is to study $\mathcal{M}$ and Ξ in detail.

For this we let $V = T_{x_0}(M)$ be the tangent space to M, $W = N_{x_0}(M)$ be the normal space, and

$$H \in W \otimes S^2 V^*$$

the 2nd fundamental form evaluated at x_0. Below we shall define a proper algebraic subvariety $D \subset W \otimes S^2 V^*$ of non-general 2nd fundamental forms, and we let $U = W \otimes S^2 V^* \backslash D$ be the dense Zariski open set of *general* 2nd fundamental forms (the definition of U is given in §I(c)).

Definition. We shall say that the embedding (1.a.1) is *general* in case $H \in U$.

Remarks. Since we are working locally we may assume that all 2nd fundamental forms $H_x \in N_x(M) \otimes S^2 T_x^*(M)$ are general.

For the remainder of this paper we shall only consider embeddings (1.a.1) *that are general.* Non-general embeddings satisfy a PDE system strictly larger than the isometric embedding system.

With this assumption it will, at least for the time being, suffice to study the characteristic sheaf and characteristic variety over the point $x_0 \in M$. Consequently, for the remainder of §I we adopt the following:

Notations. (i) $\mathcal{M}$ will denote the restriction of the characteristic sheaf to $\mathsf{P}V^* \cong \mathsf{RP}^{n-1}$, and $\Xi \subset \mathsf{P}V^*$ will be the (real) characteristic variety.

(ii) $V_\mathsf{C}, W_\mathsf{C}$ will denote the complexifications of V, W, and $\mathcal{M}_\mathsf{C}, \Xi_\mathsf{C}$ will denote the natural extensions of $\mathcal{M}, \Xi$ to $\mathsf{P}V_\mathsf{C}^* \cong \mathsf{CP}^{n-1}$.

We shall call $\mathcal{M}_\mathsf{C}$ and Ξ_C the *complex characteristic sheaf* and *complex characteristic variety*, respectively (in other words, if we just say characteristic variety then we are referring to the real characteristic variety).

(b) The following study is purely algebro-geometric. We will study the characteristic sheaf and characteristic variety of the isometric embedding system for a general $M^n \subset \mathsf{E}^{n(n+1)/2}$ over a fixed point $x_0 \in M$.

Let V, W be real vector spaces where W (but not V) has a non-degenerate symmetric bilinear form $w \otimes w' \to w \cdot w'$ (which in practice will be positive definite). We denote by

$$K \subset \Lambda^2 V^* \otimes \Lambda^2 V^* \subset \mathrm{Hom}(\otimes^4 V, \mathsf{R})$$

the space of curvature like tensors (cf. §I(a) of [4]). Given

$$H \in W \otimes S^2 V^* \cong \mathrm{Hom}(S^2 V, W)$$

we define

$$\gamma_H : W \otimes S^2 V^* \to K \tag{1.b.1}$$

by the polarized Gauss equations (loc. cit., equation (1.36))

$$\gamma_H(G)(v_1, v_2, v_3, v_4) = \tfrac{1}{2} \{ H(v_1, v_3) \cdot G(v_2, v_4) + H(v_2, v_4) \cdot G(v_1, v_3)$$
$$- H(v_1, v_4) \cdot G(v_2, v_3) - H(v_2, v_3) \cdot G(v_1, v_4) \}$$

where $v_1, v_2, v_3, v_4 \in V$ (here, we are viewing $\gamma_H(G)$ as an element of $\mathrm{Hom}(\otimes^4 V, \mathbf{R})$). Using the bilinear form to identify W with W^*, the transpose of (1.b.1) is

$$\gamma_H^* : K^* \to W \otimes S^2 V. \tag{1.b.2}$$

As will now be explained, the symbol of the isometric embedding system is obtained by localizing (1.b.2) at points $\xi \in \mathbf{P} V^*$.

For this, we let $L_\xi \subset V^*$ denote the line corresponding to $\xi \in \mathbf{P} V^*$. Then

$$(L_\xi)^* = \mathscr{O}_{\mathbf{P} V^*}(1)_\xi$$

is the fibre of the standard line bundle $\mathscr{O}_{\mathbf{P} V^*}(1) = \mathscr{O}(1)$ at $\xi \in \mathbf{P} V^*$. Composing the evaluation map

$$S^2 V \to S^2(L_\xi)^*$$

with (1.b.2) gives the dual symbol map

$$\gamma_{H,\xi}^* : K^* \to W \otimes \mathscr{O}(2)_\xi \tag{1.b.3}$$

of the isometric embedding system (cf. §III(a) in [4]). If we denote by $\mathscr{K}^*, \mathscr{W}$ the sheaves of sections of the trivial vector bundles with respective fibres K^*, W, then with the standard notations $\mathscr{O}(k) = \otimes^k \mathscr{O}(1)$ and $\mathscr{W}^*(k) = \mathscr{W}^* \otimes_{\mathscr{O}} \mathscr{O}(k)$, the dual symbol maps (1.b.3) collectively give a sheaf map

$$\gamma_H^* : \mathscr{K}^* \to \mathscr{W}(2).$$

We denote by $\mathscr{M}$ the cokernel of γ_H^*, so that we have

$$\mathscr{K}^* \xrightarrow{\;\gamma_H^*\;} \mathscr{W}(2) \longrightarrow \mathscr{M} \longrightarrow 0. \tag{1.b.4}$$

Definitions. (i) $\mathscr{M}$ is the *characteristic sheaf* (associated to $H \in W \otimes S^2 V^*$); (ii) the *characteristic variety* $\Xi \subset \mathbf{P} V^*$ is the support of $\mathscr{M}$.

Remark. As indicated by our terminology, the dual

$$\gamma_{H,\xi} : W \otimes S^2 L_\xi \to K$$

of (1.b.3) is the *symbol map* at $\xi \in \mathsf{P}V^*$ of the isometric embedding system. Since

$$\Xi = \{\xi \in \mathsf{P}V^* : \gamma_{H,\xi}^* \text{ fails to be surjective}\}$$

$$= \{\xi \in \mathsf{P}V^* : \gamma_{H,\xi} \text{ fails to be injective}\},$$

we see that our definition is the usual one for characteristic varieties.

All of these constructions make sense when we extend scalars by complexifying. For example, the analogue of (1.b.4) over $\mathsf{P}V_{\mathbb{C}}^*$ will be denoted by

$$\mathcal{K}_{\mathbb{C}}^* \to \mathcal{W}_{\mathbb{C}}^*(2) \to \mathcal{M}_{\mathbb{C}} \to 0$$

where $\mathcal{M}_{\mathbb{C}}$ is the complex characteristic sheaf with support the complex characteristic variety

$$\Xi_{\mathbb{C}} = \{\xi \in \mathsf{P}V_{\mathbb{C}}^* : \dim(\ker \gamma_{H,\xi} : W_{\mathbb{C}} \otimes S^2 L_\xi \to K_{\mathbb{C}}) \geqq 1\}.$$

Any coherent sheaf $\mathcal{M}$ is, at least after tensoring with some $\mathcal{O}(k)$, the characteristic sheaf of an exterior differential system. However, for the isometric embedding system the characteristic sheaf and characteristic variety have an especially beautiful description that we now explain.

Choose bases $\{w_\mu\}, \{v_i\}$ for W, V and denote the dual bases by $\{w^\mu\}, \{\omega^i\}$. Then

$$H = H_{ij}^\mu w_\mu \otimes \omega^i \omega^j.$$

Elements of $\mathsf{P}V^*$ will be written as

$$\xi = \xi_i \omega^i$$

where $[\xi_1, \ldots, \xi_n]$ are homogeneous coordinates; similarly

$$\lambda = \lambda_\mu w^\mu$$

will be a point of $\mathsf{P}W^*$. With the notations

$$\begin{cases} H^\mu = H_{ij}^\mu \omega^i \omega^j \in S^2 V^* \\ H_\lambda = \lambda_\mu H^\mu, \end{cases}$$

in the language of classical algebraic geometry $|H_\lambda|_{\lambda \in \mathsf{P}W^*}$ *is a linear system of quadrics on* $\mathsf{P}V^*$.

Definition. We define

$$\Sigma \subset \mathsf{P}W^* \times \mathsf{P}V^*$$

by

$$\Sigma = \{ (\lambda, \xi) : H_\lambda = \xi \circ \eta \text{ for some } \eta \in \mathbf{P}V^* \}.$$

The equation $H_\lambda = \xi \circ \eta$ means that

$$\lambda_\mu H_{ij}^\mu = \xi_i \eta_j + \xi_j \eta_i .\tag{1.b.5}$$

The projections onto the two factors of $\mathbf{P}W^* \times \mathbf{P}V^*$ induce a diagram

$$\begin{array}{ccc}
 & \Sigma & \\
{\scriptstyle \pi_1} \swarrow & & \searrow {\scriptstyle \pi_2} \\
\Lambda \subset \mathbf{P}W^* & & \Xi \subset \mathbf{P}V^*
\end{array}\tag{1.b.6}$$

where

$$\begin{cases} \Lambda = \{ \lambda \in \mathbf{P}W^* : \operatorname{rank} H_\lambda \leqq 2 \} \\ \Xi = \{ \xi \in \mathbf{P}V^* : H_\lambda = \xi \circ \eta \text{ for some } \lambda \in \mathbf{P}W^*, \eta \in \mathbf{P}V^* \}. \end{cases}\tag{1.b.7}$$

To justify the notation we have the

(1.b.8) PROPOSITION. *The subvariety Ξ in (1.b.7) coincides with the previously defined characteristic variety.*

Proof. This is a variant of the computation in the proof of proposition 3.10 in §III(a) of [4]. The point is that the symbol map

$$\gamma_{H,\xi} : W \to K$$

has a simple prescription that we now review. Choose coordinates so that $\xi = \omega^n$ and write elements $R \in K$ as $R = R_{ijkl}\omega^i \wedge \omega^j \otimes \omega^k \wedge \omega^l$. Then, with the additional index range $1 \leqq \alpha, \beta \leqq n - 1$ it turns out (loc. cit. (3.11)) that

$$\gamma_{H,\xi}(G) = G^\mu H_{\alpha\beta}^\mu \omega^\alpha \wedge \omega^n \otimes \omega^\beta \wedge \omega^n$$

where $G = G^\mu w_\mu \in W$. Setting $\lambda_\mu = G^\mu$ the condition

$$\gamma_{H,\xi}(G) = 0\tag{1.b.9}$$

is equivalent to

$$\lambda_\mu H_{\alpha\beta}^\mu = 0, \qquad 1 \leqq \alpha, \beta \leqq n - 1.$$

Clearly this is the same as

$$\lambda_\mu H_{ij}^\mu = \omega^n \circ \eta$$

where $\eta = \lambda_\mu H_{ni}^\mu \omega^i$. Finally, (1.b.9) is equivalent to the relation defining Ξ in (1.b.7).   Q.E.D.

Remarks. (i) The proof shows that the following conditions are equivalent:

$$\begin{cases} H = \xi \circ \eta & \text{for some} \quad \eta \in V^* \\ \gamma_{H,\xi}(\lambda_\mu w^\mu) = 0 \\ H_\lambda|_{\xi^\perp} = 0. \end{cases} \tag{1.b.10}$$

Hyperplanes $\xi^\perp \subset V$ satisfying either of these conditions for some $\lambda \in \mathbf{P}W^*$ are classically called *asymptotic hyperplanes*. (ii) Of course, everything we have said is also valid over **C**.

The following will be proved in §I(c) below.

(1.b.11) PROPOSITION. *If H is chosen generally, then Σ is smooth and the projection*

$$\pi_2 : \Sigma \to \Xi \tag{1.b.12}$$

is birational.

Equipped with these propositions we may describe the characteristic sheaf. For this, we define an involution

$$j_\Sigma : \Sigma \to \Sigma$$

by

$$j_\Sigma(\xi) = \eta$$

where the condition

$$H_\lambda = \xi \circ \eta \tag{1.b.13}$$

is satisfied. Using the birational map (1.b.12) we then define a *rational* involution

$$j : \Xi \to \Xi$$

by

$$j(\xi) = \eta$$

where (1.b.13) is satisfied for some $\lambda \in \mathbf{P}W^*$. We note that j fails to be defined over the locus

$$\Xi_{sg} = \left\{ \xi \in \Xi : \dim \pi_2^{-1}(\xi) \geqq 1 \right\}.$$

As will be proved in §I(c), Ξ_{sg} is the singular locus of Ξ and (1.b.12) turns out to be a *canonical* resolution of the singularities of Ξ on which the rational involution j becomes well-defined. With this understood, we define the rational map

$$\pi : \Xi \to \Lambda$$

by $\pi(\xi) = \lambda$ where (1.b.13) is satisfied, and then we define the coherent sheaf on Ξ

$$\pi^*\mathcal{O}_\Lambda(k) = (\pi_2)_*(\pi_1^*\mathcal{O}_\Lambda(k)) \qquad (1.b.14)$$

where, for $\mathcal{F}$ a coherent sheaf on Σ, $(\pi_2)_*\mathcal{F}$ is the direct image sheaf on Ξ.

We note that all of this discussion extends in the obvious way to $\Sigma_\mathbf{C}$, $\Xi_\mathbf{C}$, $\Lambda_\mathbf{C}$, $j_\mathbf{C}$, $\pi_\mathbf{C}$, etc.
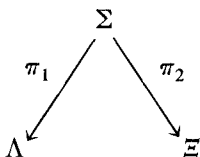
THEOREM A.   (i) *The characteristic sheaf is given by*

$$\mathcal{M} = \pi^*\mathcal{O}_\Lambda(1) \otimes \mathcal{O}_\Xi(2). \qquad (1.b.15)$$

(ii) *In case* $\dim W = n(n-1)/2$ *is the embedding codimension, the characteristic sheaf uniquely determines* $H \in W \otimes S^2 V^*$, *up to* $\mathrm{GL}(W) \times \mathrm{GL}(V)$.

(1.b.16) COROLLARY.   *The 2nd fundamental form of a general* $M^n \subset \mathbf{E}^{n(n+1)/2}$ *is uniquely determined, up to* $\mathrm{GL}(W) \times \mathrm{GL}(V)$, *by the pair* (Ξ, j) *consisting of the characteristic variety* Ξ *together with the involution* j.

Proof of the corollary.   As remarked above, given Ξ we may construct a canonical resolution of singularities $\Sigma \to \Xi$ on which j becomes a holomorphic involution. If $\Lambda = \Xi/j$ is the quotient variety, then we obtain a diagram

constructed from (Ξ, j). We then have

$$\mathcal{M} = (\pi_2)_*(\pi_1^*\mathcal{O}_\Lambda(1)) \otimes \mathcal{O}_\Xi(2),$$

which shows that the characteristic sheaf is uniquely determined by (Ξ, j). Now apply (ii) in Theorem A.   Q.E.D.

When $n = 3$ the corollary is (a slightly corrected version of) a theorem of Tennenblatt [24].

Proof of (i) *in Theorem* A.   For $\xi \in \mathbf{P}V^*$ we denote by $m_\xi \subset \mathcal{O}_{\mathbf{P}V^*,\xi}$ the maximal ideal in the local ring $\mathcal{O}_{\mathbf{P}V^*,\xi}$ and then we denote by

$$M_\xi = \mathcal{M}_\xi / m_\xi \cdot \mathcal{M}_\xi$$

the fibre of $\mathcal{M}$ at ξ. Over $\mathbf{P}V^*$ we define a vector bundle $\mathcal{F}$ with fibres

$$\mathcal{F}_\xi = S^2\xi^\perp \subset S^2 V$$

where $\xi^\perp \subset V$ is the hyperplane corresponding to $\xi \in \mathbf{P}V$. Then the dual vector bundle has fibres

$$\mathcal{F}_\xi^* \cong S^2(V^*/L_\xi)$$

where $L_\xi \subset V^*$ is the line corresponding to ξ.

According to (3.4)–(3.6) in [4], for $\omega \in L_\xi$ and $\varphi, \eta \in V^*/L_\xi$ the map

$$(\varphi \circ \eta) \otimes \omega^2 \to \tfrac{1}{2}\{\varphi \wedge \omega \otimes \eta \wedge \omega + \eta \wedge \omega \otimes \varphi \wedge \omega\}$$

induces an inclusion

$$j_\xi : \mathscr{F}_\xi^* \otimes S^2 L_\xi \hookrightarrow K.$$

The duals of the j_ξ then induce a surjective sheaf mapping

$$\mathscr{K}^* \xrightarrow{\;\;j^*\;\;} \mathscr{F}(2) \longrightarrow 0.$$

Now, and this is the main point, the localized and dualized Gauss mappings

$$\gamma_{H,\xi}^* : K^* \to W \otimes S^2 L_\xi^*$$

factor through j_ξ^* to induce a commutative diagram

$$
\begin{array}{c}
0 \\
\uparrow \\
\\
\\
\mathscr{F}(2) \longrightarrow \mathscr{W}(2) \longrightarrow \mathscr{M} \\
j^* \uparrow \quad \swarrow\; \gamma_H^* \\
\mathscr{K}^*
\end{array}
\qquad (1.b.17)
$$

(cf. (3.9) and proposition (3.10) in [4]). We will use the exact sheaf sequence

$$\mathscr{F} \xrightarrow{\;\gamma_H^*\;} \mathscr{W} \longrightarrow \mathscr{M}(-2) \longrightarrow 0 \qquad (1.b.18)$$

obtained by tensoring the top row of the exact sequence (1.b.17) with $\mathcal{O}(-2)$ to describe the fibres $M_\xi \otimes L_\xi^2$ of $\mathscr{M}(-2)$.

For this we have the fibre sequence of (1.b.18)

$$S^2 \xi^\perp \xrightarrow{\;\gamma_{H,\xi}^*\;} W \longrightarrow M_\xi \otimes L_\xi^2 \longrightarrow 0. \qquad (1.b.19)$$

If we consider $H \in W \otimes S^2 V^*$ as an element $H \in \operatorname{Hom}(S^2 V, W)$, then $\gamma_{H,\xi}^*$ in (1.b.19) is simply the map obtained by restricting

$$H : S^2 V \to W$$

to the subspace

$$S^2 \xi^\perp \subset S^2 V.$$

Thus, denoting by

$$U_\xi = \left(H(S^2 \xi^\perp)\right)^\perp \subset W^*$$

the annihilator of $H(S^2\xi^\perp) \subset W$, we have

$$M_\xi \otimes S^2 L_\xi \cong W/H(S^2\xi^\perp)$$

$$\cong U_\xi^*.$$

Using coordinates as above where $\xi = \omega^n$

$$U_\xi = \{\lambda : \lambda_\mu H_{\alpha\beta}^\mu = 0 \text{ for } 1 \leqq \alpha, \beta \leqq n-1\}$$

$$= \{\lambda \in W^* : H_\lambda = \xi \circ \eta \text{ for some } \eta \in V^*\}.$$

Summarizing, *the fibres of*

$$\pi_2 : \Sigma \to \Xi$$

are projective linear subspaces $\mathsf{P} U_\xi \subset \mathsf{P} W^*$ *where*

$$M_\xi \otimes S^2 L_\xi \cong U_\xi^*. \tag{1.b.20}$$

Since for $\lambda \in \Lambda$

$$(\pi_1)^{-1}(\lambda) = \{\xi \in \mathsf{P} V^* : \lambda \in \mathsf{P} U_\xi\}$$

we infer that the fibre over $\xi \in \mathsf{P} V^*$ of the coherent sheaf $(\pi_2)_*((\pi_1)^* \mathscr{O}_\Lambda(1))$ is

$$H^0(\mathsf{P} U_\xi, \mathscr{O}_{\mathsf{P} U_\xi}(1)) \cong U_\xi^*. \tag{1.b.21}$$

Comparing (1.b.20) and (1.b.21) gives (i) in Theorem A.

Proof of (ii) *in Theorem* A. In case $\dim W = n(n-1)/2$ is the embedding codimension, both $\mathscr{F}$ and W are vector bundles of the same rank r over $\mathsf{P} V^*$. The restrictions of $H \in \mathrm{Hom}(S^2 V, W)$ to the subspaces $S^2\xi^\perp \subset S^2 V$ induce

$$H_\xi : \mathscr{F}_\xi \to W.$$

The condition that H_ξ be an isomorphism for a general point $\xi \in \mathsf{P} V^*$ is equivalent to the condition that H be *ordinary* in the sense of §2(b) of [4]. In this case the isometric embedding system is involutive near $x_0 \in M$, and it is a general result (cf. [5]) that, for an involutive Pfaffian differential system in good form, the characteristic sheaf uniquely determines the symbol. Since the symbol of the isometric embedding system is H, part (ii) of Theorem A follows from this general result.

However, it is worthwhile to prove the result directly in the case at hand. For ordinary $H \in \mathrm{Hom}(S^2 V, W)$ the top row in (1.b.17) gives the exact sheaf sequence

$$0 \to \mathscr{F}_{\mathsf{C}}(2) \to \mathscr{W}_{\mathsf{C}}(2) \to \mathscr{M}_{\mathsf{C}} \to 0 \tag{1.b.22}$$

over $PV_{\mathbf{C}}^*$. Roughly speaking, given only $\mathscr{M}$ we shall reconstruct this sequence, and this will lead to the desired result.

For $\xi \in PV_{\mathbf{C}}^*$ there are exact sequences

$$0 \to \xi^\perp \to V_{\mathbf{C}} \to L_\xi^* \to 0$$

$$0 \to S^2(\xi^\perp) \to S^2(V_{\mathbf{C}}) \to V_{\mathbf{C}} \otimes L_\xi \to 0.$$

The second of these gives the exact bundle sequence

$$0 \to \mathscr{F}_{\mathbf{C}} \to S^2\mathscr{V}_{\mathbf{C}} \to \mathscr{V}_{\mathbf{C}}(1) \to 0 \tag{1.b.23}$$

over $PV_{\mathbf{C}}^*$ (here $\mathscr{V}_{\mathbf{C}}$ is the trivial bundle $PV_{\mathbf{C}}^* \times V_{\mathbf{C}}$).

Remark. Taking 1st Chern classes in (1.b.23) gives

$$c_1(\mathscr{F}_{\mathbf{C}}) = -n\omega$$

where $\omega \in H^2(PV_{\mathbf{C}}^*, \mathbf{Z})$ is the standard generator. Using (1.b.22) this gives

$$c_1(\mathscr{M}_{\mathbf{C}}) = n\omega.$$

Since $\operatorname{supp}\mathscr{M}_{\mathbf{C}} = \Xi_{\mathbf{C}}$ we obtain (cf. Theorem B below)

$$\deg \Xi_{\mathbf{C}} = n.$$

Returning to the proof of (ii) in Theorem A, from the standard results ([20])

$$H^0(PV_{\mathbf{C}}^*, S^2\mathscr{V}_{\mathbf{C}}) = S^2V_{\mathbf{C}}$$

$$H^0(PV_{\mathbf{C}}^*, \mathscr{V}_{\mathbf{C}}(1)) = V_{\mathbf{C}} \otimes V_{\mathbf{C}}$$

$$H^i(PV_{\mathbf{C}}^*, S^2\mathscr{V}_{\mathbf{C}}(-1)) = 0 \qquad i = 1, 2$$

and the exact cohomology sequences of $(1.b.23) \otimes \mathscr{O}(-1)$ and (1.b.23), we obtain respectively

$$H^1(PV_{\mathbf{C}}^*, \mathscr{F}_{\mathbf{C}}(-1)) \cong V_{\mathbf{C}}$$

$$H^1(PV_{\mathbf{C}}^*, \mathscr{F}_{\mathbf{C}}) \cong \Lambda^2 V_{\mathbf{C}}.$$

The exact cohomology sequences of $(1.b.22) \otimes \mathscr{O}(-3)$ and $(1.b.22) \otimes \mathscr{O}(-2)$ then give respectively

$$H^0(PV_{\mathbf{C}}^*, \mathscr{M}_{\mathbf{C}}(-3)) \cong V_{\mathbf{C}}$$

$$0 \to W_{\mathbf{C}} \to H^0(PV_{\mathbf{C}}^*, \mathscr{M}_{\mathbf{C}}(-2)) \to \Lambda^2 V_{\mathbf{C}} \to 0. \tag{1.b.24}$$

We will now complete proof of (ii) in Theorem A. Given $\mathscr{M}_{\mathbf{C}}$ and $PV_{\mathbf{C}}^*$ we use

the natural identification (1.b.24) to define the sequence

$$0 \longrightarrow Z \longrightarrow H^0(\mathbf{P}V_{\mathbf{C}}^*, \mathscr{M}_{\mathbf{C}}(-3)) \otimes V_{\mathbf{C}} \xrightarrow{\ \mu\ } \Lambda^2 V_{\mathbf{C}} \longrightarrow 0$$
$$\|$$
$$V_{\mathbf{C}} \otimes V_{\mathbf{C}}$$

(1.b.25)

where μ is exterior product and Z is the kernel of μ. Using the vertical equality in (1.b.25) and product pairing $\nu : \mathscr{M}_{\mathbf{C}}(-3) \otimes \mathscr{O}(1) \to \mathscr{M}_{\mathbf{C}}(-2)$ we obtain a diagram

$$
\begin{array}{ccc}
0 \longrightarrow Z \longrightarrow & & H^0(\mathbf{P}V_{\mathbf{C}}^*, \mathscr{M}_{\mathbf{C}}(-3)) \otimes V_{\mathbf{C}} \\
\quad\ \nu \downarrow & & \qquad\qquad \downarrow \nu \\
0 \longrightarrow W_{\mathbf{C}} \xrightarrow{\ i\ } & & H^0(\mathbf{P}V_{\mathbf{C}}^*, \mathscr{M}_{\mathbf{C}}(-2))
\end{array}
$$

where, *by definition*, $W_{\mathbf{C}} = \nu(Z)$. By localizing the inclusion i we obtain

$$\rho : \mathscr{W}_{\mathbf{C}} \to \mathscr{M}_{\mathbf{C}}(-2).$$

Tensoring with $\mathscr{O}(2)$ and taking global sections gives

$$H^0(\mathbf{P}V_{\mathbf{C}}^*, \mathscr{W}_{\mathbf{C}}(2)) \xrightarrow{\ \rho\ } H^0(\mathbf{P}V_{\mathbf{C}}^*, \mathscr{M}_{\mathbf{C}})$$
$$\|$$
$$W_{\mathbf{C}} \otimes S^2 V_{\mathbf{C}}$$

Using the metric to have an identification $W_{\mathbf{C}}^* \cong W_{\mathbf{C}}$, we have

$$(\ker \rho)^{\perp} \subset W_{\mathbf{C}} \otimes S^2 V_{\mathbf{C}}^*. \tag{1.b.26}$$

Construction. Given the characteristic sheaf $\mathscr{M}$ over $\mathbf{P}V^*$ we define the subspace

$$II_{\mathbf{C}} \subset S^2 V_{\mathbf{C}}^*$$

to be the image of $(\ker \rho)^{\perp}$ in (1.b.26) under the natural map

$$\left(W_{\mathbf{C}} \otimes S^2 V_{\mathbf{C}}^*\right) \otimes W_{\mathbf{C}}^* \to S^2 V_{\mathbf{C}}^*.$$

Thus, $II_{\mathbf{C}}$ is a linear system of quadrics on $\mathbf{P}V_{\mathbf{C}}^*$, and is in fact the linear system of quadrics given by the 2nd fundamental form H. In this way we reconstruct H, up to $GL(W_{\mathbf{C}}) \times GL(V_{\mathbf{C}})$ acting on $W_{\mathbf{C}} \otimes S^2 V_{\mathbf{C}}^*$, from $\mathscr{M}$.

To complete the proof we must show that H is determined up to the real group $GL(W) \times GL(V)$.

When $n = 2$ the characteristic variety is either empty (elliptic case), two points (hyperbolic case), or one point counted twice (parabolic case). In suitable coordinates H is respectively $(\omega^1)^2 + (\omega^2)^2, \omega^1 \omega^2, (\omega^1)^2$.

Thus we may assume that $n \geqq 3$. The crucial point is the following

(1.b.27) LEMMA.   *If $H \in W \otimes S^2 V^*$ is general, then the conditions*

$$\begin{cases} S \in \mathfrak{gl}(W) \times \mathfrak{gl}(V) \\ S \cdot H = 0 \end{cases}$$

imply that $S = 0$.

Proof.   It will suffice to prove the same result over C. The 1-parameter group $g(t) = \exp(tS)$ acts on $\mathbf{P}\,W_{\mathbf{C}}^* \times \mathbf{P}\,S^2 V_{\mathbf{C}}^*$ and leaves $\Sigma_{\mathbf{C}}$ invariant. Thus $g(t)$ induces a holomorphic vector field $\theta_S \in H^0(\Sigma_S, \Theta)$ with the special property that it is induced by a linear vector field on $\mathbf{P}^N$ under a suitable embedding $\Sigma_{\mathbf{C}} \subset \mathbf{P}^N$ $(= \mathbf{P}(W_{\mathbf{C}}^* \otimes S^2 V_{\mathbf{C}}^*)$ in our case).

When $n = 3$, $\Sigma_{\mathbf{C}} \cong \Xi_{\mathbf{C}}$ is a smooth curve of genus one (cf. Theorem B) and it is well known that $\theta_S = 0$.

When $n \geqq 4$ we even have that

$$H^0(\Sigma_{\mathbf{C}}, \Theta) = 0. \tag{1.b.28}$$

For example, when $n = 4$ we will also see in Theorem B that $\Sigma_{\mathbf{C}} \cong \Xi_{\mathbf{C}}$ is a smooth K3 surface and in this case (1.b.28) is well known.

In general, the proof of Theorem B will imply that the dualizing sheaf $\omega_{\Xi_{\mathbf{C}}}$ is invertible, and hence

$$\omega_{\Xi_{\mathbf{C}}} \cong \mathscr{O}_{\Xi_{\mathbf{C}}}$$

since $\deg \Xi_{\mathbf{C}} = n$. By Grothendieck–Serre duality and the fact that $\Sigma_{\mathbf{C}} \to \Xi_{\mathbf{C}}$ is a rational resolution of singularities

$$\dim H^0(\Sigma_{\mathbf{C}}, \Theta) = \dim H^n\left(\Xi_{\mathbf{C}}, \omega_{\Sigma_{\mathbf{C}}} \otimes \Omega^1_{\Sigma_{\mathbf{C}}}\right)$$

$$= \dim H^n\left(\Xi_{\mathbf{C}}, \omega_{\Xi_{\mathbf{C}}} \otimes \Omega^1_{\Xi_{\mathbf{C}}}\right)$$

$$= \dim H^0\left(\Xi_{\mathbf{C}}, \mathrm{Hom}\left(\Omega^1_{\Xi_{\mathbf{C}}}, \mathscr{O}\right)\right)$$

$$= 0.$$

Q.E.D. for lemma (1.b.27).

To complete the proof of (ii) in Theorem B we suppose given general $H, H' \in W \otimes S^2 V^*$ with

$$H' = TH$$

where $T \in \mathrm{GL}(W_{\mathbf{C}}) \times \mathrm{GL}(S^2 V_{\mathbf{C}}^*)$. Then

$$H' = \overline{T} H$$

$$0 = (T - \overline{T})H,$$

so that lemma 1.b.27 applies to give that $S = \sqrt{-1}\,(T - \overline{T}) = 0$.   Q.E.D.

(c) We continue our study of the characteristic sheaf $\mathscr{M}$ and characteristic variety Ξ over $x_0 \in M$ where $M^n \subset \mathbf{E}^N$ is a general submanifold. In this section we shall restrict to the case of the embedding dimension

$$N = n(n + 1)/2 = n + n(n - 1)/2.$$

Keeping the notations from §I(b), for $\xi \in \mathbf{P}V^*$ the symbol map for the isometric embedding system is

$$\gamma_{H,\xi} : W \to K.$$

Using bases so that $\xi = \omega^n$ and w_μ is an orthonormal basis for W, we have seen in the proof of proposition (1.b.8) that for $G = G^\mu w_\mu \in W$

$$\gamma_{H,\xi}(G) = G^\mu H_{\alpha\beta}^\mu \omega^\alpha \wedge \omega^n \otimes \omega^\beta \wedge \omega^n.$$

As noted there (cf. (1.b.10)) there is a natural isomorphism

$$\ker \gamma_{H,\xi} \cong \{\lambda = \lambda_\mu w^\mu \in W^* : H_\lambda = \xi \circ \eta \text{ for some } \eta \in V^*\}. \qquad (1.c.1)$$

Definition. We define $\Xi_k \subset \mathbf{P}V^*$ by

$$\Xi_k = \{\xi \in \mathbf{P}V^* : \dim \ker \gamma_{H,\xi} \geqq k\}.$$

The same prescription also defines $\Xi_{\mathbf{C},k} \subset \mathbf{P}V_\mathbf{C}^*$.

Since, by generality, the map $\xi \mapsto \gamma_{H,\xi}$ will be injective we have $\Xi_n = \varnothing$ and $\Xi_1 = \Xi$. In fact, there is a filtration

$$\varnothing = \Xi_n \subset \Xi_{n-1} \subset \cdots \subset \Xi_2 \subset \Xi_1 = \Xi; \qquad (1.c.2)$$

as will be seen below, this is a stratification with very remarkable properties (also over $\mathbf{C}$).

THEOREM B. *For H general we have*
  (i) $\operatorname{codim} \Xi_\mathbf{C} = 1$ *and* $\deg \Xi_\mathbf{C} = n$,
  (ii) *for any k*

$$\operatorname{codim} \Xi_{\mathbf{C},k} = k^2$$

$$\deg \Xi_{\mathbf{C},k} = \det \left\| \begin{matrix} \binom{n}{k} & \cdots & \binom{n}{2k-1} \\ \vdots & & \vdots \\ \binom{n}{1} & \cdots & \binom{n}{k} \end{matrix} \right\| \neq 0;$$

  (iii) $(\Xi_{\mathbf{C},k})_{sg} = \Xi_{\mathbf{C},k+1}$;
  (iv) $\pi_2 : \Sigma_\mathbf{C} \to \Xi_\mathbf{C}$ *is a canonical rational resolution of singularities; and*

(v) *for the real characteristic variety*

$$\begin{cases} (\Xi_k)_{sg} = \Xi_{k+1} \\ \Xi \neq \varnothing & \text{if} \quad n \geq 3 \\ \Xi_{sg} \neq \varnothing & \text{if} \quad n = 4m + 2, \quad m \geq 1. \end{cases}$$

Remarks. Here, Ψ_{sg} denotes the singular locus of a variety Ψ (over either $\mathbb{R}$ or $\mathbb{C}$). To say that $\Sigma_{\mathbb{C}} \xrightarrow{\pi_2} \Xi_{\mathbb{C}}$ is a rational resolution of singularities means in particular that: (a) $\Sigma_{\mathbb{C}}$ is smooth; (b) the fibres $\Sigma_{\mathbb{C}\xi} = \pi_2^{-1}(\xi)$ are rational varieties; and (c) the direct image sheaves satisfy

$$R^q_{\pi_2}(\mathcal{O}_{\Sigma_{\mathbb{C}}}) = \begin{cases} \mathcal{O}_{\Xi_{\mathbb{C}}} & q = 0 \\ 0 & q \geq 1 \end{cases} \tag{1.c.3}$$

In our situation it will also be the case that: (d) the dualizing sheaf $\omega_{\Xi_{\mathbb{C}}}$ of $\Xi_{\mathbb{C}}$ is invertible and satisfies

$$R^q_{\pi_2}(\Omega^n_{\Sigma_{\mathbb{C}}}) = \begin{cases} \omega_{\Xi_{\mathbb{C}}} & q = 0 \\ 0 & q \geq 1 \end{cases} \tag{1.c.4}$$

$$j_*(\Omega^n_{\Xi_{\mathbb{C}} \setminus \Xi_{\mathbb{C},sg}}) = \begin{cases} \omega_{\Xi_{\mathbb{C}}} & q = 0 \\ 0 & q \geq 1 \end{cases} \tag{1.c.5}$$

where $j : \Xi_{\mathbb{C}} \setminus \Xi_{\mathbb{C},sg} \hookrightarrow \Xi_{\mathbb{C}}$ is the inclusion. Although rather technical to state, this last condition is very strong. Effectively, it says that from an algebro-geometric viewpoint the singularities of $\Xi_{\mathbb{C}}$ are negligible.

(1.c.6) COROLLARY. *For the characteristic variety* Ξ *of the isometric embedding system* (I, ω) *of* $M^n \subset E^{n(n+1)/2}$ *we have*

(i) $\Xi \neq \varnothing$ *if* $n \geq 3$ (*in particular*, (I, ω) *cannot be elliptic for* $n \geq 3$);

(ii) Ξ *is smooth if* $M \subset E^{n(n+1)/2}$ *is general and* $n \leq 4$;

(iii) Ξ *is not smooth for any* $M \subset E^{n(n+1)/2}$ *if* $n = 4m + 2$, $m \geq 1$.

Remarks. (i) In perhaps more standard PDE terminology, the isometric embedding system is of principal type if $n \leq 4$ but has multiple characteristics if $n = 4m + 2$, $m \geq 1$.

(ii) Essentially the same proof will give the following for the isometric embedding system of a general $M^n \subset E^{n+r}$

(i) $$\dim \Xi_{\mathbb{C}} = \max(-1, r - (n-1)(n-2)/2 - 1)$$

(ii) $$\dim \Xi_{\mathbb{C},k} = \max\left(-1, r - (n-1)(n-2)2 - k^2\right)$$

(iii) $$\deg \Xi_{\mathbb{C}} = \binom{n(n+1)/2 - r}{n-1}$$

(iv) $$(\Xi_{\mathbb{C},k})_{sg} = \Xi_{\mathbb{C},k+1}$$

(1.c.7) COROLLARY. $\Xi_C = \emptyset$ if $r \leqq (n-1)(n-2)/2$.

This corollary was the starting point of the rigidity results in [4].

The idea behind the proof of Theorem B is based on the following three considerations:

(1.c.8) Let A, B be vector spaces (real or complex) and define the standard determinantal subvarieties

$$\Psi_l \subset P \operatorname{Hom}(A, B)$$

by

$$\Psi_l = \{ T \in \operatorname{Hom}(A, B) : \operatorname{rank} T \leqq l \}.$$

(To be precise, Ψ_l is defined by the ideal given by all $(l+1) \times (l+1)$ minors of T.) Then, if $\dim A = a$ and $\dim B = b$,

$$\operatorname{codim} \Psi_l = (a-l)(b-l).$$

In particular, if $a = b = n$ and we set $k = n - l$, then $\operatorname{codim} \Psi_{n-k} = k^2$.

(1.c.9) If we let $G = G(a-l, A)$ be the Grassmannian of $(a-l)$ planes $\Lambda \subset A$ and define

$$\Sigma_l \subset G \times P \operatorname{Hom}(A, B)$$

by

$$\Sigma_l = \{ (\Lambda, T) : \Lambda \subseteq \ker T \},$$

then

$$\pi : \Sigma_l \to \Psi_l$$

is a rational resolution of singularities having the properties (a)–(d) listed above (over C). Note that for $T \in \operatorname{Hom}(A, B)$ with $\dim(\ker T) \geqq a - l$,

$$\pi^{-1}(T) = G(a-l, \ker T).$$

In particular, it is reasonable to expect (a)–(d) to hold for any variety that is locally a subvariety of Ψ_l situated transverse to the stratification $\Psi_1 \subset \Psi_2 \subset \cdots \subset \Psi_l$.

(1.c.10) Let

$$\varphi : E \to F$$

be a holomorphic vector bundle map over a smooth complex algebraic variety Σ and define

$$\Sigma_l = \{ x \in \Sigma : \varphi(x) : E_x \to F_x \text{ has rank} \leqq l \}.$$

If $r(E), r(F)$ are the respective ranks of E, F, then under the assumption that

$$\rho = \operatorname{codim} \Sigma_l = (r(E) - l)(r(F) - l),$$

the fundamental class $\sigma_l \in H^{2\rho}(\Sigma, \mathbf{Z})$ of Σ_l is given by Porteous' formula (cf. [2]):

$$\sigma_l = \Delta_l(c(E), c(F)).$$

Here, $c(E)$ and $c(F)$ are the Chern classes of E and F, and the notation is that of [2] (where a proof of Porteous' formula is given).

   The idea behind the proof is to apply the considerations (1.c.8)–(1.c.10) to the symbol map of the isometric embedding system (thus, $\Sigma = \mathbf{P}V^*$ or $\mathbf{P}V_{\mathbf{C}}^*$ in (1.c.10)). To implement this we will need to establish strong transversality properties of the symbol map for a general $M^n \subset \mathbf{E}^N$. Although it seems quite reasonable that "general for the 2nd fundamental form of $M^n \subset \mathbf{E}^N \Rightarrow$ transverse for the symbol map" this is by no means obvious, and in fact in algebraic geometry one knows all too well that such "obvious" general position properties are frequently false. Fortunately, we are able to adapt our situation to be able to apply known transversality results for sufficiently twisted subvarieties of Grassmannians meeting general special Schubert cycles. Our proof of the qualitative part of Theorem B (over $\mathbf{C}$) will be largely based on the paper [14] by Kleiman–Landolfi, and so our first step will be to briefly review [14] in a form suitable for our use.

   Review of [14]. Let $\mathscr{U}$ be a complex vector space and $G_n(\mathscr{U})$ the Grassmann manifold of *co*dimension-n planes $E \subset \mathscr{U}$. We denote the universal bundle sequence by

$$0 \longrightarrow \mathscr{S} \longrightarrow \mathscr{U} \overset{\pi}{\longrightarrow} \mathscr{D} \longrightarrow 0 \qquad (1.\text{c}.11)$$

where $\mathscr{S}, \mathscr{U}, \mathscr{D}$ have respective fibres

$$\begin{cases} \mathscr{S}_E = E \\ \mathscr{U}_E = \mathscr{U} \\ \mathscr{D}_E = \mathscr{U}/E. \end{cases}$$

Given an n-plane $F \subset \mathscr{U}$ we set

$$\sigma_k(F) = \left\{ E \in G_n(\mathscr{U}) : \dim(E \cap F) \geqq k \right\}$$

$$= \left\{ E \in G_n(\mathscr{U}) : \operatorname{rank}(\pi : F \to \mathscr{U}/E) \leqq n - k \right\}.$$

Since any two n-planes $F, F' \subset \mathscr{U}$ are related by $\mathrm{GL}(\mathscr{U})$, any two special Schubert cycles $\sigma_k(F), \sigma_k(F')$ are related by $\mathrm{GL}(\mathscr{U})$ acting on $G_n(\mathscr{U})$. Choosing a basis $f_1, \ldots, f_n$ for F, we may think of the vectors

$$f_j(E) = \pi(f_j) \in \mathscr{U}/E$$

as giving n holomorphic sections $\pi(f_j) \in H^0(G_n(\mathcal{U}), \mathcal{D})$. Then

$$\sigma_k(F) = \{ E \subset G_n(E) : \dim(\operatorname{span} f_1(E), \ldots, f_n(E)) \leqq n - k \}.$$

In particular, $\sigma_1(F)$ is the locus

$$\pi(f_1) \wedge \cdots \wedge \pi(f_n) = 0.$$

The $\sigma_k(F)$ are among the *special Schubert cycles* discussed in Kleiman–Landolfi [14]. In particular they have properties

$$\begin{cases} \operatorname{codim} \sigma_k(F) = k^2; \\ (\sigma_k(F))_{sg} = \sigma_{k+1}(F); \\ \sigma_k(F) \text{ is Cohen–Macaulay and normal; and} \\ \sigma_k(F) \text{ has a canonical rational resolution of singularities} \end{cases} \qquad (1.c.12)$$

The first three properties are given in §4 of [14] (cf. also (6.3)), and the last one is explained in §§5, 6. The term canonical refers to a succession of ordinary monoidal transformations, with non-singular centers, that are defined purely in terms of the scheme structure of $\sigma_k(F)$.

The canonical resolution of singularities of these special Schubert cycles also has the property (d) (cf. (1.c.4) and (1.c.5)) listed in the remarks following the statement of Theorem B. Since we will only marginally use this result here (cf. the proof of (1.b.28) for $n \geqq 5$) we will not give the standard but somewhat technical proof.

The mapping π in (1.c.11) induces a holomorphic mapping between rank n bundles over $G_n(\mathcal{U})$

$$\pi : \mathcal{F} \to \mathcal{D} \qquad (1.c.13)$$

(here, $\mathcal{F}$ is the trivial bundle $F \times G_n(\mathcal{U})$). Since

$$\operatorname{codim}\left\{ E \in G_n(\mathcal{U}) : \operatorname{rank}\left(F \xrightarrow{\pi} \mathcal{U}/E \right) \leqq n - k \right\} = k^2,$$

Porteous' formula applies to give the fundamental class σ_k of $\sigma_k(F)$ as

$$\sigma_k = \det \begin{Vmatrix} c_k & \cdots & c_{2k-1} \\ \vdots & & \vdots \\ c_1 & \cdots & c_k \end{Vmatrix}, \qquad (1.c.14)$$

where $c_j = c_j(\mathcal{D})$ are the universal Chern classes.

Remark. Actually, (1.c.14) is a consequence of *Giambelli's formula* (cf. page 205 in [8]) from classical Schubert calculus. We have emphasized the more general Porteous' formula since it may be useful in other problems of computing degrees of characteristic varieties.

Review of [14] (*continued*). Next we consider a smooth subvariety X $\subset G_n(\mathscr{U})$. Setting

$$\sigma_k(F, X) = \sigma_k(F) \cdot X, \qquad (1.c.15)$$

it is natural to ask how many of the properties (1.c.12) are inherited by the $\sigma_k(F, X)$. Clearly, this requires a transversality condition on the intersection (1.c.15). If this condition is satisfied for one n-plane, then it will be satisfied for a general n-plane $F \subset \mathscr{U}$ (but usually not for all F), so what is required is a condition on X that a general intersection (1.c.15) be suitably transversal. One such condition is given in [14].

Definition. A smooth, irreducible subvariety $X \subset G_n(\mathscr{U})$ is said to be *twisted* in case $\mathscr{D}_X(-1) = \mathscr{D}(-1) \otimes \mathscr{O}_X$ is generated by its global sections.

Here, $\mathscr{O}_X(1)$ refers to the standard line bundle under the composite embedding

$$X \hookrightarrow G_n(\mathscr{U}) \xrightarrow{\quad p \quad} \mathbf{P}(\Lambda^{N-n}\mathscr{U})$$

where $\dim \mathscr{U} = N$ and p is the Plücker embedding.

Remark. In particular it follows that, for a twisted $X \subset G_n(\mathscr{U})$, the universal quotient bundle $\mathscr{D}_X \cong \mathscr{D}_X(-1) \otimes \mathscr{O}(1)$ is very ample, whereas over the whole Grassmannian $\mathscr{D} \to G_n(\mathscr{U})$ is not even ample (unless $n = 1$ or $n = N - 1$). A slight additional argument shows that the normal bundle to X in $G_n(\mathscr{U})$ is also very ample, so the above is a reasonable notion of "twisted."

In §7 of [14] it is proved that the properties (1.c.12) are valid for a general $\sigma_k(F, X)$ provided that $X \subset G_n(\mathscr{U})$ is twisted. We shall not write out these properties here, but rather refer to Theorems 7.1, 7.3, 7.5–7.8 in [14] for the statements to be used below.

It may also be shown that (1.c.4), (1.c.5) follow from the analogous statements for the $\sigma_k(F)$. As mentioned above, the proof of this technical and for our present purposes not really essential result will be omitted.

By Porteous' formula applied to the restriction of (1.c.13) to a general $X \subset G_n(\mathscr{U})$ the fundamental class $\sigma_k \in H^{2k^2}(X, \mathbf{Z})$ of $\sigma_k(F, X)$ is given by the same formula

$$\sigma_k = \det \begin{Vmatrix} c_k & \cdots & c_{2k-1} \\ \vdots & & \vdots \\ c_1 & \cdots & c_k \end{Vmatrix} \qquad (1.c.16)$$

as (1.c.14), where now $c_k = c_k(Q_X)$.

Proof of (i)–(iv) *in Theorem* B. We now consider an algebraic 2nd fundamental form

$$H \in W \otimes S^2 V^* \cong \mathrm{Hom}(W^*, S^2 V^*)$$

where $\dim W = n(n-1)/2$. We set

$$\mathscr{U} = S^2 V_{\mathbb{C}}$$

$$F_H = (H(W^*))^\perp.$$

In coordinates, if $H = H_{ij}^\mu w_\mu \otimes \omega^i \omega^j$ then for $\lambda = \lambda_\mu w^\mu \in W_{\mathbb{C}}^*$

$$H(\lambda) = H_\lambda = \lambda_\mu H_{ij}^\mu \omega^i \omega^j \in S^2 V_{\mathbb{C}}^*$$

$$F_H = \left\{ q^{ij} e_i e_j \in S^2 V_{\mathbb{C}} : q^{ij} H_{ij}^\mu = 0 \text{ for all } \mu \right\}.$$

Note that for $H \in \mathrm{Hom}(W^*, S^2 V^*)$ injective (in particular, for general H)

$$\dim F_H = n(n+1)/2 - n(n-1)/2$$

$$= n.$$

With the identification (projective duality)

$$\mathsf{P} V_{\mathbb{C}}^* \xrightarrow{\ \sim\ } G_1(V_{\mathbb{C}})$$

given by

$$\xi \rightarrow \xi^\perp,$$

we define a mapping

$$\varphi : \mathsf{P} V_{\mathbb{C}}^* \rightarrow G_n(\mathscr{U})$$

by

$$\varphi(\xi) = S^2(\xi^\perp).$$

Setting $X = \varphi(\mathsf{P} V_{\mathbb{C}}^*)$ it is easy to see that $\mathsf{P} V_{\mathbb{C}}^* \rightarrow X$ is biholomorphic, and we will identify $\mathsf{P} V_{\mathbb{C}}^*$ with X. The fibres of the universal bundle sequence (1.c.11) over $\xi \in X$ are now

$$\mathscr{S}_\xi = S^2(\xi^\perp)$$

$$\mathscr{U}_\xi = S^2 V_{\mathbb{C}}$$

$$\mathscr{D}_\xi = S^2 V_{\mathbb{C}} / S^2(\xi^\perp).$$

In coordinates, if $\xi = \omega^n$ and we use the additional index range $1 \leqq \alpha, \beta \leqq n-1$,

$$\mathscr{S}_\xi = \left\{ q^{\alpha\beta} e_\alpha e_\beta \right\}$$

$$\mathscr{U}_\xi = \left\{ q^{ij} e_i e_j \right\}$$

$$\mathscr{D}_\xi \cong \left\{ q^{\alpha n} e_\alpha e_n \right\}.$$

In other words, if $L_\xi \subset V_C^*$ is the line corresponding to $\xi \in P_C^*$, then there is a natural isomorphism

$$\mathscr{D}_\xi \cong V_C \otimes L_\xi^*$$

and consequently

$$\mathscr{D}_X \cong \mathscr{V}_C \otimes \mathscr{O}_X(1) \tag{1.c.17}$$

where $\mathscr{V}_C \to X$ is the trivial bundle with fibre V_C. It follows (just barely) that $X \subset G_n(\mathscr{U})$ is twisted, and moreover

$$c(\mathscr{D}_X) = (1 + \omega)^n$$

$$c_q(\mathscr{D}_X) = \binom{n}{q}\omega^q \tag{1.c.18}$$

where $\omega \in H^2(PV_C^*, Z)$ is the standard generator.

Now, and this is the point of the above construction, (i)–(iv) in Theorem B will follow from the fact that $\varphi(PV_C^*) \subset G_n(S^2V_C)$ is twisted and (1.c.18) once we show that

$$\sigma_k(F_H, X) = \Xi_{C,k}. \tag{1.c.19}$$

For this, let $\xi \in PV_C^*$ and choose coordinates so that $\xi = \omega^n$. Then

$$\varphi(\xi) = S^2(\xi^\perp) = \{ q^{\alpha\beta}e_\alpha e_\beta \} \subset S^2V_C$$

$$\varphi(\xi)^\perp = S^2(\xi^\perp)^\perp = \{ p_{in}\omega^i\omega^n \} \subset S^2V_C^*$$

$$F_H = \{ q^{ij}e_i e_j : q^{ij}H_{ij}^\mu = 0 \text{ for all } \mu \} \subset S^2V_C$$

$$(F_H)^\perp = \{ H_\lambda = \lambda_\mu H_{ij}^\mu \omega^i\omega^j \} \in S^2V_C^*.$$

Since

$$\dim(\varphi(\xi) \cap F_H) = \dim(\varphi(\xi)^\perp \cap F_H^\perp),$$

and from the above descriptions

$$\varphi(\xi)^\perp \cap F_H^\perp = \{\lambda \in W_C^* : H_\lambda = \xi \circ \eta \text{ for some } \eta\},$$

we see that the conditions

$$\begin{cases} \dim(S^2(\xi^\perp) \cap F_H) \geqq k \\ \dim\{\lambda \in W_C^* : H_\lambda = \xi \circ \eta \text{ for some } \eta\} \geqq k \end{cases}$$

are equivalent. Since the first of these conditions gives $\sigma_k(F_H, X)$ and, by (1.c.1), the second gives $\Xi_{C,k}$, we conclude (1.c.19) and with its parts (i)–(iv) of Theorem B.

Definition. We will say that a 2nd fundamental form $H \in \mathrm{Hom}(W_C^*, S^2 V_C^*)$ is *general* in case: (i) H is injective, so that $F_H = (H(W^*))^\perp \subset S^2 V_C$ is an n-plane, and (ii) F_H meets the subvariety $X \subset G_n(S^2 V_C)$ transversely in the scheme-theoretic sense. A real 2nd fundamental form is *general* in case the corresponding complex 2nd fundamental form is general.

For H general we have now established Proposition (I.b.11), Theorem A (condition i in the definition of general $\Rightarrow$ ordinary as in the proof of Theorem A), and Theorem B where all statements are over C. The corresponding real assertions will be a consequence of the following argument.

Proof of (v) *in Theorem* B.   The first statement follows immediately from (iii) and the Jacobian criterion. (The general result is that

$$(\Psi)_{sg} \cap \mathsf{R}P^{n-1} = (\Psi_{sg}) \cap \mathsf{R}P^{n-1}$$

for any algebraic variety $\Psi \subset \mathsf{C}P^{n-1}$ whose idea is generated by *real* homogeneous polynomials.)

To prove the second part of (v) we recall our notational convention that $\Xi \subset \mathsf{P}V^* \cong \mathsf{R}P^{n-1}$ is the real characteristic variety and $\Xi_C \subset \mathsf{P}V_C^* \cong \mathsf{C}P^{n-1}$ the complex variety having the same real generators for its defining ideal as Ξ. In the embedding dimension, Ξ_C is a hypersurface and we have proved that

$$\deg \Xi_C = n.$$

Consequently, for $n \equiv 1 \bmod 2$ the real characteristic variety Ξ is a non-empty hypersurface because of the following well-known

(1.c.20) *Remark.*   Let $I \subset \mathsf{R}[\xi_1, \ldots, \xi_n]$ be a homogeneous ideal with complexification $I_C \subset \mathsf{C}[\xi_1, \ldots, \xi_n]$. Suppose that the variety Ψ_C defined by I_C has no multiple components and $\deg \Psi_C$ is odd. Then the real variety Ψ defined by I is non-empty.

Proof.   Suppose that $\dim \Psi_C = m$ and let $G_m(\mathsf{R}^n)$ (respectively $G_m(\mathsf{C}^n)$) denote the *co*dimension m planes in $\mathsf{R}P^{n-1}$ (respectively $\mathsf{C}P^{n-1}$). If $\deg \Psi_C = d$ then a Zariski open set $\mathscr{U} \subset G_m(\mathsf{C}^n)$ of planes $\Lambda \cong \mathsf{C}P^{n-m-1}$ all meet Ψ_C in d distinct smooth points. Since $G_m(\mathsf{R}^n) \subset G_m(\mathsf{C}^n)$ is not contained in any complex algebraic subvariety, there must be real planes Λ meeting Ψ_C in d distinct points. Complex conjugation then induces an involution on the finite set $\Lambda \cdot \Psi_C$, and since $\#(\Lambda \cdot \Psi_C) = d$ is odd there must be a fixed point.   Q.E.D. for (1.c.20).

To prove that $\Xi \neq \varnothing$ for *all* $n \geqq 3$ we proceed somewhat differently. The following argument is motivated by the proof of the similar result in [12].

We consider $H \in \mathrm{Hom}(W^*, S^2 V^*)$ and make the identification

$$S^2 V^* = \mathrm{Hom}^{(s)}(V, V^*),$$

where the right-hand side is the symmetric homomorphisms of V to V^*. We set

$$F_H = (H(W^*))^\perp \subset \mathrm{Hom}^{(s)}(V^*, V).$$

Then since $\dim H(W^*) = n(n-1)/2$,

$$\dim F_H = n.$$

To show that $\Xi \neq \emptyset$ we must show that there is $\xi \in V^*\backslash\{0\}$ such that

$$\langle A, \xi \circ \eta \rangle = 0 \tag{1.c.21}$$

for *some* $\eta \in V^*\backslash\{0\}$ and *all* $A \in F_H$. Writing (1.c.21) as

$$\langle A(\xi), \eta \rangle = 0,$$

the failure of (1.c.21) to hold means that the map

$$\alpha_\xi : F_H \to V$$

given by

$$\alpha_\xi(A) = A(\xi)$$

is *surjective* for *all* $\xi \in V^*\backslash\{0\}$. Since $\dim F_H = n$, α_ξ must be an isomorphism. In other words, if (1.c.21) fails then

$$A \in F_H\backslash\{0\}, \qquad \xi \in V^*\backslash\{0\} \Rightarrow A(\xi) \neq 0.$$

Equivalently,

$$A \in F_H\backslash\{0\} \Rightarrow \det A \neq 0.$$

According to the paper [1] of Adams–Lax–Phillips, the existence of a subspace

$$F \subset \mathrm{Hom}^{(s)}(\mathbf{R}^n, \mathbf{R}^n)$$

satisfying the two conditions

$$\begin{cases} \dim F = n \\ \det A \neq 0 \quad \text{for all} \quad A \in F \end{cases}$$

implies that $n \leqq 2$. This proves the second part of (v).

To prove that $\Xi_{sg} \neq \emptyset$ if $n = 4m + 2$, $m \geqq 1$, we shall use the remark (1.c.20) together with (iii) and the first part of (v) from Theorem B. Thus

$$\deg \Xi_{C,sg} = \det \left\| \begin{matrix} \binom{n}{2} & \binom{n}{3} \\ \binom{n}{1} & \binom{n}{2} \end{matrix} \right\| = \frac{n^2(n^2-1)}{2 \cdot 2 \cdot 3} .$$

If $n = 2(2m+1)$ this is

$$\deg \Xi_{C,sg} = \frac{(2m+1)^2(16m^2 + 16m + 3)}{3} \equiv 1 \qquad \mathrm{mod}\, 2.$$

Remark.   Naturally we would conjecture that

$$\Xi_{sg} \neq \varnothing \quad \text{for} \quad n \geqq 5.$$

II. The linearized isometric embedding system.

(a)   Let (M, ds^2) be a Riemannian manifold with metric $g = g_{ij}(x)dx^i dx^j$ in local coordinates. Fixing a point $p_0 \in M$ and replacing M by a small neighborhood of p_0, still denoted by M, we want to begin discussing the problem of finding a local C^∞ isometric embedding

$$x : (M, ds^2) \rightarrow \mathsf{E}^{n(n+1)/2} \tag{2.a.1}$$

Thus we want to solve the 1st order, non-linear P.D.E. system

$$(dx, dx) = g \tag{2.a.2}$$

in a neighborhood of p_0. Now (2.a.2) is a determined (i.e., the same number of unknowns as equations) system, but the "geometry", which relates the 2nd fundamental form of an embedding (2.a.1) with the curvature of g via the Gauss equations, can only be uncovered by differentiating (2.a.2).

This suggests that we approach the problem more intrinsically via the exterior differential system of isometric embeddings set up in §II(a) of [4]. This is especially true since the differential system formulation leads quite naturally to the solution of the embedding problem in the real-analytic case (the BCJS-theorem, cf. the references cited in [3] and §II(b) of [4]). Finding admissible integral manifolds of an exterior differential system is also a 1st order, non-linear P.D.E. system (cf. any proof of the Cartan–Kähler theorem). What is suggested is that we try to solve this system by an iteration scheme using strong solvability problems of the linearized equations, which of course must be established in the particular case at hand.

Following some further general remarks, we will in the next section give the general formulation for the variational equations of an exterior differential system. These will turn out to be a 1st order linear P.D.E. system whose symbol map and characteristic variety are in a very natural way induced from that of the exterior differential system (cf. [5] for the definitions and terminology we are using here). When applied to the isometric embedding system, we obtain a linear P.D.E. system, the essential part of whose symbol is given over $p_0 \in M$ by

$$\gamma_H : W \otimes S^2 V^* \rightarrow K$$

where H is the 2nd fundamental form at $x_0 = x(p_0)$ of an embedding (2.a.1), and where the remaining notations are given below (1.b.1). Since

$$\dim K = \frac{n^2(n^2 - 1)}{12}$$

$$\dim W \otimes S^2 V^* = \frac{n^2(n-1)^2}{4}$$

the linearized equation appears to be overdetermined (i.e., there are more equations than unknowns) when $n \geqq 3$. This is because the isometric embedding system (I, ω) has been obtained by prolongation of the non-involutive system (I_0, ω) of §II(a) in [4] (cf. (2.7) there), and it is a general phenomenon that the prolongation of a determined system appears to be overdetermined.

However, when the Cauchy problem for the linearized equation is well posed, this difficulty can be sometimes overcome. For example, in $\mathbb{R}^{n+1}$ with variables $(x, t) = (x_1, \ldots, x_n, t)$ we consider a 2nd order hyperbolic system

$$u_{tt} - a^{ij} u_{x_i x_j} = 0 \tag{2.a.3}$$

for one function $u(x, t)$ with initial data

$$u(x, 0) = f(x), \qquad u_t(x, 0) = g(x).$$

In the general theory of differential systems, (2.a.3) would be written as the 1st order system in $n + 1$ unknown functions $(u_1, \ldots, u_n, v)$

$$\begin{cases} v_{x_i} - (u_i)_t = 0 & i = 1, \ldots, n \\ (u_i)_{x_j} - (u_j)_{x_i} = 0 & 1 \leqq i < j \leqq n \\ v_t - a^{ij} (u_i)_{x_j} = 0 \end{cases} \tag{2.a.4}$$

with the initial data

$$\begin{cases} u_i(x, 0) = f_{x_i}(x) \\ v(x, 0) = g(x) \end{cases}$$

Now (2.a.4) seems to be overdetermined when $n \geqq 2$. However, by differentiating the first equations in (2.a.4) the functions

$$w_{ij} = (u_i)_{x_j} - (u_j)_{x_i}$$

satisfy

$$\begin{cases} (w_{ij})_t = 0 \\ w_{ij}(x, 0) = 0 \end{cases}$$

and hence $w_{ij} \equiv 0$. In other words, we may omit the middle equations in (2.a.4) to have a determined system.

This procedure will not work for a general 2nd order equation

$$b^{\alpha\beta} u_{y_\alpha y_\beta} = 0$$

in $\mathbb{R}^{n+1}$ with coordinates $(y_1, \ldots, y_{n+1})$ unless we can single out a non-characteristic hypersurface so that, after a change of variables, it can be written in the form (2.a.3).

Similarly, for a vector-valued function $U = (u^1, \ldots, u^k)$ the 1st prolongation of the determined system

$$\begin{cases} U_t - A^i U_{x_i} = 0 \\ U(x, 0) = F(x), \end{cases}$$

where the A^i are $k \times k$ matrices, is the seemingly overdetermined system

$$\begin{cases} V_{x_i} - (U_i)_t = 0 \\ (U_i)_{x_j} - (U_j)_{x_i} = 0 \\ V_t - A^i (U_i)_t - A_i^i U_i = 0 \\ U_i(x, 0) = F_{x_i}(0) \end{cases}$$

As before, the 2nd equations may be omitted to obtain a determined system.

Now even if this procedure of finding determined subsystems could be made to work in general, to be applicable to C^∞ problems it requires at the very least that Holmgren's uniqueness theorem be valid, and most likely the whole Cauchy problem would have to be well posed. Fortunately, for the isometric embedding problem it turns out that the linearized version of the naïve system (2.a.2) *uncouples* along M in a very special way and may be effectively used without prolongation. This procedure, which in some sense already occurs in Nash's original paper [16], will be given in §II(c) following our discussion in §II(b) of the general linearization of exterior differential systems.

(b)   Let (I, ω) be an exterior differential system on a manifold X, and let

$$f : N \to X \tag{2.b.1}$$

be an admissible integral manifold. We will derive the equations for the first order variations of (2.b.1) as an integral manifold of (I, ω).

Step one.   It is well known that the first order variations of any map (2.b.1) are given by sections $\nu \in C^\infty(E)$ of the normal bundle $E = f^* T(X)/T(N)$ to $f(N)$ in X. In suitable coordinates (x^i, x^α) on X and (y^i) on N, (2.b.1) is given by

$$y \to (y^i, x^\alpha(y))$$

A variation is

$$y \to (y^i, x^\alpha(y, t)) \tag{2.b.2}$$

Setting $\dot{} = \partial/\partial t\big|_{t=0}$ the tangent to (2.b.2) is

$$\nu = \dot{x}^\alpha \frac{\partial}{\partial x^\alpha} \tag{2.b.3}$$

Step two. Suppose that $\dim N = n$, $\dim X = N$, and use the ranges of indices $1 \leq i, j \leq n$, $n + 1 \leq \alpha, \beta \leq N$. Denote by $G_n(X) \to X$ the Grassmann bundle whose fiber

$$G_n(X)_p = G_n(T_p(X))$$

over $p \in X$ consists of all n-planes in the tangent space $T_p(X)$. A local coordinate system (x^i, x^α) on X induces local coordinates $(x^i, x^\alpha, l_i^\alpha)$ on $G_n(X)$, where $(x^i, x^\alpha, l_i^\alpha)$ corresponds to the n-plane with basis

$$e_i = \frac{\partial}{\partial x^i} + l_i^\alpha \frac{\partial}{\partial x^\alpha}$$

Any maximal rank mapping (2.b.1) has a *canonical lift*

$$
\begin{array}{ccc}
 & & G_n(X) \\
 & \overset{f_*}{\nearrow} & \downarrow \\
N & \overset{f}{\longrightarrow} & X
\end{array}
\quad , \tag{2.b.4}
$$

given by $f_*(y) = f_*(T_y(N))$, or in the above local coordinates by

$$f_*(y) = \left(y^i, x^\alpha(y), \frac{\partial x^\alpha}{\partial y^i}(y) \right) \tag{2.b.5}$$

On $G_n(X)$ there is a canonical exterior differential system (J, ω) whose admissible integral manifolds are the canonical lifts $f_* : N \to G_n(X)$. Locally, (J, ω) is the Pfaffian system generated by the 1-forms

$$\theta^\alpha = dx^\alpha - l_i^\alpha dx^i$$

with independence condition $dx^1 \wedge \cdots \wedge dx^n \neq 0$.

To see this, let $g : N \to G_n(X)$ be any mapping with $g^*(dx^1 \wedge \cdots \wedge dx^n) \neq 0$. Then by the implicit function theorem locally g may be given by

$$y \to \left(y^i, x^\alpha(y), l_i^\alpha(y) \right)$$

The condition

$$g^*(\theta^\alpha) = 0$$

is

$$l_i^\alpha(y) = \frac{\partial x^\alpha}{\partial x^i}(y),$$

which is (2.b.5).

Step three. Now let

$$g_t : N \to G_n(X) \tag{2.b.6}$$

be a 1-parameter family of maps given locally by

$$y \to \left(y^i, x^\alpha(y,t), l_i^\alpha(y,t) \right)$$

The associated infinitesimal variation is given by (2.b.3) where we replace X by $G_n(X)$, and it is

$$\dot{x}^\alpha \frac{\partial}{\partial x^\alpha} + \dot{l}_i^\alpha \frac{\partial}{\partial l_i^\alpha} \tag{2.b.7}$$

The condition that for each t (2.b.6) be an integral manifold of (J,ω) is

$$l_i^\alpha(y,t) = \frac{\partial x^\alpha}{\partial y^i}(y,t). \tag{2.b.8}$$

Setting $g = g_0$, the corresponding condition on a general normal vector

$$\tilde{\nu} = \nu^\alpha \frac{\partial}{\partial x^\alpha} + \nu_i^\alpha \frac{\partial}{\partial l_i^\alpha} \tag{2.b.9}$$

to $g(N) \subset G_n(X)$ is, by (2.b.7) and (2.b.8),

$$\frac{\partial \nu^\alpha}{\partial y^i} = \nu_i^\alpha \tag{2.b.10}$$

In summary:
Given an admissible integral manifold $g(= f_*)$ of (J,ω) the infinitesimal variations of g as integral manifolds of (J,ω) are given by normal vectors (2.b.9) that satisfy the 1st order linear P.D.E. system (2.b.10).

Step four. Now suppose that (I,ω) is an exterior differential system on X whose integral elements $V_n(I,\omega) \subset G_n(X)$ are locally given by

$$F^r\left(x^i, x^\alpha, l_i^\alpha \right) = 0.$$

Differentiating

$$F^r\left(y^i, x^\alpha(y,t), l_i^\alpha(y,t) \right) \equiv 0$$

gives

$$\frac{\partial F^r}{\partial l_i^\alpha} \dot{l}_i^\alpha + \frac{\partial F^r}{\partial x^\alpha} \dot{x}^\alpha = 0 \tag{2.b.11}$$

Let

$$g : N \to V_n(I,\omega) \subset G_n(X)$$

be the canonical lift of an integral manifold $f: N \to X$ of (I,ω). Then the conditions that a normal vector $\nu = \nu^\alpha(\partial/\partial x^\alpha)$ to $f(N) \subset X$ be the infinitesimal

variation of a 1-parameter of integral manifolds of (I, ω) are given by combining (2.b.10) and (2.b.11), and are

$$\frac{\partial F^r}{\partial l_i^\alpha} \frac{\partial \nu^\alpha}{\partial y^i} + \frac{\partial F^r}{\partial x^\alpha} \nu^\alpha = 0 \qquad (2.b.12)$$

If we recall that the first prolongation $(I^{(1)}, \omega)$ of (I, ω) is given by the restriction of the canonical system (J, ω) to $V_n(I, \omega)$, and that $(I^{(1)}, \omega)$ is the quasi-linear Pfaffian system

$$\begin{cases} \theta^\alpha = 0 \\ d\theta^\alpha \equiv -\pi_i^\alpha \wedge \omega^i \qquad \mod\{\theta^\alpha\} \end{cases}$$

with symbol relations

$$b_\alpha^{ri}\pi_i^\alpha \equiv 0 \qquad \mod\{\omega^i, \theta^\alpha\}$$

given by

$$b_\alpha^{ri} = \frac{\partial F^r}{\partial l_i^\alpha},$$

then we may write (2.b.12) as

$$\begin{cases} \nu = \nu^\alpha \dfrac{\partial}{\partial x^\alpha} \text{ is normal to } f(N) \subset X \\ b_\alpha^{ri} \dfrac{\partial \nu^\alpha}{\partial y^i} + c_\alpha^r \nu^\alpha = 0 \end{cases} \qquad (2.b.13)$$

(2.b.14) *Conclusion. The equations of variation of an integral manifold* $f: N \to X$ *of* (I, ω) *are given by a linear 1st order differential operator*

$$D : C^\infty(E) \to C^\infty(F) \qquad (2.b.15)$$

between vector bundles E, F over N. In coordinates, D is given by (2.b.13), and in intrinsic terms

$$\begin{cases} E = \textit{normal bundle to } f(N) \textit{ in } X \\ F = \textit{co-normal bundle to } V_n(I, \omega) \textit{ in } G_n(X) \end{cases}$$

In the last sentence it is understood that we pull back the co-normal bundle of $V_n(I, \omega)$ in $G_n(X)$ to N by the canonical lift f_* of f.

Remarks. (i) It is interesting to note that the linearized version (2.b.15) already involves the 1st prolongation of (I, ω).

(ii) Especially important is the observation that the symbol of (2.b.15), in the usual P.D.E. sense, is induced by the symbol of the exterior differential system $(I^{(1)}, \omega)$ as defined in §I of [5].

(iii) Suppose that (I_t, ω_t) is a 1-parameter of exterior differential systems on X with $(I_0, \omega_0) = (I, \omega)$, and let us ask for the variational equations of a 1-parameter family of integral manifolds $f_t : N \to X$ of (I_t, ω_t). Then essentially the same procedure shows that (2.b.15) is replaced by an inhomogeneous system

$$b_\alpha^{ri} \frac{\partial \nu^\alpha}{\partial y^i} + c_\alpha^r \nu^\alpha = a^r \tag{2.b.16}$$

where, if $F^r(x^i, x^\alpha, l_i^\alpha, t) = 0$ defines $V_n(I_t, \omega_t) \subset G_n(X)$, then

$$a^r = \frac{\partial F^r}{\partial t} \Big|_{t=0}$$

(c)   We shall now deal directly with the variational equations of the naïve system (2.a.2). These are obtained by differentiating

$$(dx_t, dx_t) = g_t \tag{2.c.1}$$

at $t = 0$. Setting $\dot{} = \partial/\partial t|_{t=0}$ and using the notations

$$\begin{cases} x = x_0 \\ y = \dot{x}_t \\ k = \frac{1}{2}(\dot{g}_t) \end{cases}$$

the variational equations of (2.a.2) are

$$(dx, dy) + (dy, dx) = 2k \tag{2.c.2}$$

Both sides are sections of $S^2 T^*(M)$; consequently (1.c.2) is an $n(n+1)/2 \times n(n+1)/2$ linear inhomogeneous system, and the main point is to *uncouple* the system along M by breaking y into tangential and normal components.

Thus, let $\{e_i, e_\mu\}$ be a Darboux frame field along $x(M) \subset \mathbf{E}^{n(n+1)/2}$ with $\{\omega^i\}$ the dual co-frame to $\{e_i\}$. Recall the structure equations (cf. §I(a))

$$\begin{cases} dx = \omega^i e_i \\ d\omega^i = -\psi_j^i \wedge \omega^j \\ de_i = \psi_i^j e_j \\ de_\mu \equiv H_{ij}^\mu e_i \omega^j \quad \mod\{e_\mu\} \end{cases}$$

where $\|\psi_j^i\|$ is the connection matrix and $H = H_{ij}^\mu e_\mu \otimes \omega^i \omega^j$ is the second fundamental form. If $W \to M$ is the normal bundle to $x(M)$, then H gives a map

$$H : W \to S^2 T^*(M) \tag{2.c.3}$$

by

$$\lambda_\mu e_\mu \to \lambda_\mu H_{ij}^\mu \omega^i \omega^j.$$

We shall assume that the embedding $x : M \to \mathbf{E}^{n(n+1)/2}$ is non-degenerate in the sense that (2.c.3) is injective, and we denote by $II \subset S^2 T^*(M)$ the image. Then $II = \bigcup_{p \in M} II_p$ where $II_p \subset S^2 T_p^*(M)$ is a subspace of co-dimension n. We define $II^\perp \subset S^2 T(M)$ by $II^\perp = \bigcup_{p \in M} II_p^\perp$ where $II_p^\perp$ is the annihilator of II_p. We may think of $II^\perp$ as the relations (depending on $p \in M$)

$$b^{ij} H_{ij}^\mu = 0, \qquad \mu = n + 1, \ldots, n(n + 1)/2. \tag{2.c.4}$$

To uncouple the equations (2.c.2) we write

$$y = y^i e_i + y^\mu e_\mu$$

$$dy \equiv \left(dy^i + y^{\,j} \psi_j^{\,i} - y^\mu H_{ij}^\mu \omega^{\,j} \right) e_i \qquad \bmod\{e_\mu\}.$$

Then (2.c.2) becomes

$$\left(dy^i + y^{\,j} \psi_j^{\,i} \right) \circ \omega^i - y^\mu H_{ij}^\mu \omega^i \circ \omega^j = k_{ij} \omega^i \circ \omega^j \tag{2.c.5}$$

To put these in intrinsic form we consider the 1-form

$$\varphi = y^i \omega^i \in C^\infty(T^*(M))$$

Its covariant differential is

$$D\varphi = \left(dy^i - y^{\,j} \psi_i^{\,j} \right) \otimes \omega^i$$

$$= \left(dy^i + y^{\,j} \psi_j^{\,i} \right) \otimes \omega^i \in C^\infty(T^*(M) \otimes T^*(M))$$

If we denote by

$$D^{(s)}\varphi = \left(dy^i + y^{\,j} \psi_j^{\,i} \right) \circ \omega^i \in C^\infty(S^2 T^*(M))$$

the symmetrized covariant differential and set

$$H^\mu = H_{ij}^\mu \omega^i \circ \omega^j \in C^\infty(S^2 T^*(M)),$$

then (2.c.5) is

$$D^{(s)}\varphi = y^\mu H^\mu + k$$

This is equivalent to

$$D^{(s)}\varphi \equiv k \qquad \bmod II \tag{2.c.6}$$

The formula

$$\varphi \to D^{(s)}\varphi \qquad \bmod II$$

defines a 1st order linear differential operator

$$D : C^\infty(T^*(M)) \to C^\infty(S^2 T^*(M)/II) \tag{2.c.7}$$

and (2.c.6) is the same as

$$D\varphi = l \qquad\qquad (2.c.8)$$

where l is the projection of k to $S^2 T^*(M)/II$. In summary:

(2.c.9)   *The solutions y to the variational equations* (2.c.2) *are in one-to-one correspondence with the solutions φ to* (2.c.8).

Explicitly, if $b_r = \{b_r^{ij} e_i e_j\}$ give a basis for $II^{\perp}$ and if we write

$$D^{(s)}\varphi = (D^{(s)}\varphi)_{ij} \omega^i \circ \omega^j,$$

then the equations (2.c.8) are

$$b_r^{ij}\big((D^{(s)}\varphi)_{ij} - k_{ij}\big) = 0$$

If $\varphi = y^i \omega^i$ is a solution to these equations, then by our nondegenerary assumption we have

$$D^{(s)}\varphi = y^\mu H^\mu + k$$

for unique y^μ's. Finally, $y = y^i e_i + y^\mu e_\mu$ is a solution to (2.c.2).

We will give some observations on the system (2.c.8).

(2.c.10)   *The bundles $T^*(M)$ and $S^2 T^*(M)/II$ both have rank n, so that* (2.c.8) *is an $n \times n$ determined system.*

(2.c.11)   *The characteristic variety of the operator* (2.c.8) *is equal to the characteristic variety Ξ of the isometric embedding system.*

Proof.   The symbol of (2.c.8) is given by

$$\sigma_\xi(D)(\eta) = \big\{ b_r^{ij}(\xi_i \eta_j + \xi_j \eta_i) \big\}$$

where $\xi, \eta \in T^*(M)$. The condition $\sigma_\xi(D)(\eta) = 0$ is therefore equivalent to

$$\lambda_\mu H_{ij}^\mu = \xi_i \eta_j + \xi_j \eta_i$$

for some $\lambda = \lambda_\mu e^{*\mu} \in W^*$. Writing this as

$$H_\lambda = \xi \circ \eta$$

we obtain exactly the condition that $\xi \in \Xi$ (cf. (1.b.10)).   Q.E.D.

Remark.   The characteristic variety of (2.c.8) coincides with Ξ *as a scheme*. However, the characteristic sheaf of (2.c.8) is not equal to the characteristic sheaf of the isometric embedding system (cf. Appendix to §II(c)).

We shall conclude this section by considering briefly the variational equations (2.c.2) for a general

$$M^n \subset \mathsf{E}^{n(n+1)/2+s}$$

If we define W to be the bundle of 2nd osculating spaces to $x(M)$ modulo the tangent bundle, then we may still assume that (2.c.3) is injective. In this case exactly the same derivation goes through, so that the variational equations (2.c.2) are still equivalent to (2.c.8) where D is given by (2.c.7). We now separate into over- and underdetermined cases.

$s < 0$.   There the system (2.c.8) is *over*determined. The dimension of the complex characteristic variety is given by the 2nd remark following (1.c.6), and we are in the situation of [4].

$s > 0$.   There the system (2.c.8) is *under*determined. In particular, where $s \geq n$ the operator (2.c.7) is zero and the variational equation (2.c.2) may be solved *algebraically* by taking the y^i arbitrary and determining the y^μ by the conditions

$$D^{(s)}(y^i\omega^i) = y^\mu H^\mu$$

This is the situation considered by Nash [16] and Greene [7].

Appendix to §II(c).   We shall attempt to clarify how the *same* characteristic variety Ξ turns up in the two *quite different* linearized isometric embedding systems. For this we shall use the notations and terminology of [5]. Let (I, ω) be a quasi-linear exterior differential system with characteristic sheaf $\mathcal{M}$. Then

$$\text{Supp}\,\mathcal{M} = \Xi_{\mathbf{C}} \subset \mathbf{P}V_{\mathbf{C}}^* \qquad (2.c.12)$$

is the complex characteristic variety of (I, ω). The condition that (I, ω) be involutive is given by the vanishing of the torsion of (I, ω) and by homological conditions on $\mathcal{M}$ (loc. cit.).

Suppose now that (I, ω) is involutive and denote by $\mathcal{M}^{(1)}$ the characteristic sheaf of the 1st prolongation $(I^{(1)}, \omega)$ of (I, ω). Then (loc. cit.)

$$\mathcal{M}^{(1)} = \mathcal{M} \otimes \mathcal{O}(1) \qquad (2.c.13)$$

Moreover, $(I^{(1)}, \omega)$ is involutive with characteristic variety $\Xi^{(1)}$ given by

$$\Xi^{(1)} = \Xi$$

according to (2.c.12) and (2.c.13). Finally, if (I, ω) is "determined" in the sense that its symbol matrices are square, then $(I^{(1)}, \omega)$ is "seemingly overdetermined" in the sense that its symbol matrices are $k \times m$ with $k < m$.

Conversely, suppose that (I, ω) is involutive and is "seemingly overdetermined" in the above sense. For example, we may consider the isometric embedding system for a general $M^n \subset \mathbf{E}^{n(n+1)/2}$ with symbol matrices

$$B_\xi : W \to K.$$

Suppose we also suspect that (I, ω) is, at least formally, the prolongation of a determined system. For example, the isometric embedding system is formally the

1st prolongation of the determined system (2.a.2). Then to find this determined system we should try to "deprolong" (I, ω) by considering

$$\mathcal{M}(-1) = \mathcal{M} \otimes \mathcal{O}(-1).$$

If $\mathcal{M}(-1)$ satisfies the homological conditions for involutivity, then it appears to be the case that the linearized equations for (I, ω) microlocally uncouple into an algebraic system plus a system with characteristic sheaf $\mathcal{M}(-1)$. This suggests the following

Definition. An involutive exterior differential system (I, ω) is said to be *determined* if for some k the sheaf $\mathcal{M}(-k) = \mathcal{M} \otimes \mathcal{O}(-k)$ satisfies the homological conditions for involutivity and has square symbol matrices.

(2.c.14) PROPOSITION. *The isometric embedding system of a general $M^n \subset \mathbf{E}^{n(n+1)/2}$ is determined.*

We will not give a formal proof of this result, but simply remark that the symbol mapping corresponding to $\mathcal{M}(-1)$ turns out to be the symbol map of the linearized equation (2.c.2).

It is instructive to conclude this discussion by explaining directly how the same characteristic variety Ξ turns up in the two *quite different* linearized isometric embedding systems. Keeping our above notations we recall that (cf. (1.b.10))

$$\Xi = \{ [\xi] \in \mathbf{P}V^* : \text{there exist } \lambda \in W^* \text{ and } \eta \in V^* \text{ with } H_\lambda = \xi \circ \eta \} \quad (2.c.15)$$

In indices the relation $H_\lambda = \xi \circ \eta$ is

$$\lambda_\mu H_{ij}^\mu = \xi_i \eta_j + \xi_j \eta_i . \quad (2.c.16)$$

The key observation is that the symbol map $\sigma_1(\xi)$ of the linearized isometric embedding system (I, ω) has $\lambda \in \ker \sigma_1(\xi)$, while the symbol map $\sigma_2(\xi)$ of the linearized equation (2.c.2) has $\eta \in \ker \sigma_2(\xi)$.

More precisely, recalling from the proof of Theorem A our notation

$$\mathscr{F}_\xi^* = S^2(V^*/L_\xi)$$

and the canonical inclusion

$$j_\xi : \mathscr{F}_\xi^* \otimes S^2 L_\xi \hookrightarrow K,$$

the symbol map

$$\sigma_1(\xi) : W^* \to K$$

of the linearization of (I, ω) is given by

$$\sigma_1(\xi)(\lambda) = j_\xi(\tilde{H}_\lambda) \quad (2.c.17)$$

where $\tilde{H}_\lambda \in S^2(V^*/L_\xi)$ is the projection of $H_\lambda = \lambda_\mu H_{ij}^\mu \omega^i \circ \omega^j \in S^2 V^*$ (cf. (1.b.17)).

On the other hand, the symbol map

$$\sigma_2(\xi) : V^* \to S^2 V^*/II^\perp$$

of the linearized and uncoupled equation (2.c.2) is given by

$$\sigma_2(\xi) = [\xi \circ \eta] \tag{2.c.18}$$

where $[\xi \circ \eta]$ is the projection of $\xi \circ \eta$ to $S^2 V^*/II^\perp$.

It is clear that the conditions

$$\sigma_1(\xi)\lambda = 0$$

$$\sigma_2(\xi)\eta = 0$$

are each equivalent to (2.c.16). In fact, the map

$$\ker \sigma_1(\xi) \to \ker \sigma_2(\xi)$$

given by

$$\lambda \to \eta,$$

where (2.c.16) is satisfied, is an isomorphism. It is for this reason that the two characteristic varieties are equal *as schemes.*

(d)   Theorem B and its consequences, especially corollary (1.c.6), give us fairly strong information on the microlocal behavior of the linearized isometric embedding system (2.c.8).

THEOREM C.   *Let $M^n \subset E^{n(n+1)/2}$ be a general submanifold. Then at a point $(x_0, \xi_0) \in \Xi_k \backslash \Xi_{k+1}$ the linearized isometric embedding system (2.c.8) is microlocally equivalent to a system of pseudo-differential equations whose leading symbol is*

$$\sigma(x,\xi) = \left\| \begin{array}{ccc|c} L_{11}(x,\xi) & \cdots & L_{1k}(x,\xi) & \\ \vdots & & \vdots & 0 \\ L_{k1}(x,\xi) & \cdots & L_{kk}(x,\xi) & \\ \hline & 0 & & \|\xi\| I_{n-k} \end{array} \right\| \tag{2.d.1}$$

where the $L_{\alpha\beta}(x,\xi)$ are homogeneous of degree one in ξ with $L_{\alpha\beta}(x_0,\xi_0) = 0$ and are generic in the sense that

$$\bigwedge_{1 \leq \alpha, \beta \leq k} d_\xi L_{\alpha\beta}(x,\xi) \neq 0$$

in a conical neighborhood of (x_0, ξ_0).

(2.d.2) COROLLARY.   *When $n = 2, 3, 4$ the variational equations of the isometric embedding of a general $M^n \subset E^{n(n+1)/2}$ are locally solvable.*

Proof of Theorem C.   This is pretty much equivalent to a more general result whose formulation may shed additional light on the situation. For this we use the

notations

$$\Sigma = \mathsf{P}(\mathrm{Hom}(\mathsf{R}^m, \mathsf{R}^m))$$

$$\cup$$

$$\Sigma_k = \left\{ A \in \mathrm{Hom}(\mathsf{R}^m, \mathsf{R}^m) : \mathrm{rank}\, A \leqq m - k \right\}$$

Thus, Σ is the projectivized space of $m \times m$ matrices and $\Sigma_k \subset \Sigma$ is the algebraic subvariety whose ideal is generated by all $(m - k + 1) \times (m - k + 1)$ minors of a variable matrix. It is well known that

$$\begin{cases} \mathrm{codim}\, \Sigma_k = k^2 \\ (\Sigma_k)_{sg} = \Sigma_{k+1} \end{cases}$$

We shall refer to

$$\Sigma_1 \subset \Sigma_1 \subset \cdots \subset \Sigma_{m-1} \subset \Sigma \tag{2.d.3}$$

as the *canonical stratification* on the space of $m \times m$ matrices.

We now consider a symbol

$$p(x, \xi) : T^*(M) \backslash \{0\} \to \mathrm{Hom}(\mathsf{R}^m, \mathsf{R}^m)$$

given by an asymptotic sum

$$p(x, \xi) \sim p_1(x, \xi) + p_0(x, \xi) + \cdots$$

whose leading term is homogeneous of degree one. This leading term induces a map

$$p_1(x, \xi) : \mathsf{P}T^*(M) \to \Sigma \tag{2.d.4}$$

for each fixed $x \in M$.

Definition. We shall say that the symbol $p(x, \xi)$ is *fibrewise transverse* in case, for each fixed $x \in M$, the mapping (2.d.4) is transverse to the canonical stratification (2.d.3).

Remarks. (i) If we set

$$\Xi_k = (p_1)^{-1} \Sigma_k$$

then it follows that $\Xi = \Xi_{m-1}$ is the characteristic variety and that either Ξ_k is empty or else

$$\mathrm{codim}\, \Xi_k = k^2.$$

It follows moreover that

$$(\Xi_k)_{sg} = \Xi_{k+1}.$$

In particular, $\Xi_{sg} = \Xi_2$.

(ii) If we write $p_1(x,\xi) = \| p_{1\alpha\beta}(x,\xi)\|$, then this definition has only to do with the principal symbol $\| p_{1\alpha\beta}(x,\xi)\|$ for each fixed x. It does not say anything about the Poisson brackets $\{ p_{1\alpha\beta}(x,\xi), p_{1\gamma\delta}(x,\xi)\}$ and whether or not the Ξ_k are involutive for $k \geqq 2$. It also says nothing about the sub-principal symbol of $p(x,\xi)$.

Theorem C is a consequence of

THEOREM C'. *Let $p(x,\xi)$ be a fibrewise transverse symbol. Then microlocally, $p(x,\xi)$ is equivalent to a symbol whose leading term has the form* (2.d.1) *with*

$$\bigwedge_{1 \leqq \alpha, \beta \leqq k} d_\xi L_{\alpha\beta}(x,\xi) \neq 0.$$

The assertion "Theorem C' $\Rightarrow$ Theorem C" is a consequence of Theorem B. To prove Theorem C' we consider $p(x,\xi)$ as an operator

$$P(x,D) : C_0^\infty(E) \to C_0^\infty(F)$$

between compactly supported sections of vector bundles E, F over the open set Ω in question. (We may also assume that P is properly supported.) Let $(x_0, \xi_0) \in \Xi_k \backslash \Xi_{k+1}$ and choose a frame $\{e_i(x)\}$ for E so that

$$\ker p_1(x_0,\xi_0) = \operatorname{span}\{e_1(x_0), \ldots, e_k(x_0)\}$$

Using the index ranges $1 \leqq \alpha, \beta \leqq k$ and $k+1 \leqq \mu, \nu \leqq m$ write

$$p_1(x,\xi) = \left\| \begin{matrix} L_{\alpha\beta}(x,\xi) & L_{\alpha\mu}(x,\xi) \\ L_{\mu\beta}(x,\xi) & L_{\mu\nu}(x,\xi) \end{matrix} \right\|$$

where

$$\begin{cases} L_{\alpha\beta}(x_0,\xi_0) = L_{\mu\beta}(x_0,\xi_0) = 0 \\ \det\| L_{\mu\nu}(x_0,\xi_0)\| \neq 0 \end{cases} \tag{2.d.5}$$

We may assume that $\det\| L_{\mu\nu}(x,\xi)\| \neq 0$ throughout a conical neighborhood $V \subset T^*(\Omega)\backslash\{0\}$. For each α, β set

$$M_{\alpha\beta} = \det \left\| \begin{matrix} L_{\alpha\beta} & L_{\alpha,k+1} & \cdots & L_{\alpha,m} \\ L_{k+1,\beta} & L_{k+1,k+1} & \cdots & L_{k+1,m} \\ \vdots & \vdots & & \vdots \\ L_{m,\beta} & L_{m,k+1} & \cdots & L_{m,m} \end{matrix} \right\|$$

$$= \det\| L_{\alpha\beta}\|\det\| L_{\mu\nu}\| + \{\text{terms containing products } L_{\mu\beta} \cdot L_{\alpha\nu}\}$$

By our assumption, the k^2 functions $M_{\alpha\beta}$ generate the ideal of $\Xi_k \cap V$.

We have

$$p_1(x,\xi)e_\mu(x) = \sum_i L_{i\mu}(x,\xi)f_i(x)$$

In our conical neighborhood V we choose a new frame $\tilde{f}_i(x,\xi)$ where

$$\tilde{f}_\alpha(x,\xi) = f_\alpha(x)$$

$$\tilde{f}_\mu(x,\xi) = \left\{ \frac{1}{\|\xi\|} \sum_i L_{i\mu}(x,\xi)f_i(x) \right\} = \frac{1}{\|\xi\|} p_1(x,\xi)e_\mu(x)$$

Then $\{\tilde{f}_i(x,\xi)\}$ is a frame for $\pi^*F \to V$ that is homogeneous of degree zero in ξ. Using this frame the matrix of $p_1(x,\xi)$ is

$$\left\| \begin{matrix} \tilde{L}_{\alpha\beta}(x,\xi) & 0 \\ \tilde{L}_{\mu\beta}(x,\xi) & \|\xi\|\delta_{\mu\nu} \end{matrix} \right\|$$

where the conditions (2.d.5) are still fulfilled. We next define a new frame for $\pi^*E \to V$ by

$$\begin{cases} \tilde{e}_\alpha(x,\xi) = e_\alpha(x) - \dfrac{1}{\|\xi\|}\left\{ \sum_\mu \tilde{L}_{\mu\alpha}(x,\xi)e_\mu(x) \right\} \\ \tilde{e}_\mu(x,\xi) = e_\mu(x) \end{cases}$$

The matrix of $p_1(x,\xi)$ is then

$$\left\| \begin{matrix} \tilde{L}_{\alpha\beta}(x,\xi) & 0 \\ 0 & \|\xi\|\delta_{\mu\nu} \end{matrix} \right\|$$

where the conditions (2.d.5) are also still fulfilled. Since, with obvious notation,

$$\tilde{M}_{\alpha\beta} = \tilde{L}_{\alpha\beta}$$

and these functions generate the ideal of $\Xi_k \cap V$, we have

$$\bigwedge_{a,\beta} d_\xi \tilde{L}_{\alpha\beta}(x,\xi) \neq 0 \tag{2.d.6}$$

by our fibrewise transversality assumption.

In summary, given $(x_0,\xi_0) \in \Xi_k \setminus \Xi_{k+1}$, there exists a conical neighborhood V of (x_0,ξ_0) and $m \times m$ matrices $a(x,\xi)$ and $b(x,\xi)$ defined and homogeneous of degree zero in $T^*(\Omega) \setminus \{0\}$ and invertible in V, such that in V

$$a(x,\xi)p_1(x,\xi)b(x,\xi) = \left\| \begin{matrix} \tilde{L}_{\alpha\beta}(x,\xi) & 0 \\ 0 & \|\xi\|\delta_{\mu\nu} \end{matrix} \right\| \tag{2.d.7}$$

where (2.d.6) is satisfied. If we let $A(x,D)$ and $B(x,D)$ be the pseudo-differential

operators (made properly supported) corresponding to $a(x, \xi)$ and $b(x, \xi)$, then

$$A(x, D)P(x, D)B(x, D) = \tilde{P}(x, D)$$

where

$$\tilde{p}(x, \xi) \sim \tilde{p}_1(x, \xi) + \tilde{p}_0(x, \xi) + \cdots$$

where $\tilde{p}_1(x, \xi)$ is given in V by the right-hand side of (2.d.7). This is exactly what is required in Theorem C'.   Q.E.D.

Proof of Corollary (2.d.2).   Since either $\Xi_{sg} = \emptyset$ or

$$\mathrm{codim}\ \Xi_{sg} = 4,$$

the operator (2.c.8) is of principal type when $n = 2, 3, 4$ (but *not* when $n = 6, 10, 14, \dots$), and the corollary follows from a theorem of Hormander [10], [22].

However, for later use we want to describe how (2.d.1) may be used to obtain local solutions with bounds. For this we write (2.c.8) as a P.D.E. system

$$Pu = f \tag{2.d.8}$$

where

$$P(x, D) : C^\infty(\Omega, D) \to C^\infty(\Omega, F)$$

is a linear 1st order operator between sections of vector bundles E, F over an open set $\Omega \subset \mathbf{R}^n$. With $x_0 \in \Omega$ a reference point we set $\Omega_\delta = \{ |x - x_0| < \delta \}$ and define

$$\begin{cases} \|u\|_{s,\delta}^2 = \int_{\mathbf{R}^n} (\delta^{-2} + |\xi|^2)^s |\hat{u}(\xi)|^2 \, d\xi \\ H_s(\Omega_\delta, E) = \text{completion of } C_0^\infty(\Omega_\delta, E) \text{ in } \|\ \|_{s,\delta} \end{cases}$$

Then it can be shown that (cf. [10]):

(2.d.9)   *There are constants* $C = C_s$ *and* δ_0 *such that for* $\delta \leqq \delta_0$ *and* $f \in H_s(\Omega_\delta, F)$ *the equation* (2.d.8) *has a solution* $u \in H_s(\Omega_\delta, E)$ *with*

$$\|u\|_{\delta,s} \leqq \delta C \|f\|_{\delta,s} \tag{2.d.10}$$

III. The characteristic variety and Gauss equations in the 3-dimensional case.

(a)   For a 3-dimensional vector space V we consider a 3-dimensional subspace $II \subset S^2 V^*$. Note that the annihilator $II^\perp \subset S^2 V$ is also 3-dimensional. Elements of II will be denoted by $P, Q, \dots$. We recall that the *characteristic variety*

$$\Xi \subset \mathbf{P} V^*$$

is given by

$$\Xi = \{\chi \in \mathsf{P}V^* : Q = \xi \circ \eta \text{ for some } Q \in II \text{ and } \eta \in V^*\}.$$

In this case, we say that $\xi \in \mathsf{P}V^*$ is *characteristic*. The following is a strengthening of (i) in Theorem B when $n = 3$.

(3.a.1) PROPOSITION. *For any II, either $\Xi = \mathsf{P}V^*$ in which case II consists of all multiples of a fixed $\xi \in V^*$, or else Ξ is a cubic curve in $\mathsf{P}V^*$.*

Proof. Choose a basis x^1, x^2, x^3 of V^* and let $Q^i = H^i_{jk} x^j x^k$ be a basis of II. If $\xi = \xi_i x^i \in \mathsf{P}V^*$ (i.e., $[\xi_1, \xi_2, \xi_3]$ are the homogeneous coordinates of ξ), then the condition $\xi \in \Xi$ is equivalent to

$$\lambda_i Q^i = \xi \circ \eta$$

for some λ, η. Explicitly, this is

$$\lambda_i H^i_{jk} = \tfrac{1}{2}(\xi_j \eta_k + \xi_k \eta_j). \tag{3.a.2}$$

This is 6 homogeneous equations for the 6 unknowns $\{\lambda_i, \eta_i\}$. Since the Q^i are linearly independent, if (3.a.2) is satisfied, then we cannot have either $\lambda = 0$ or $\eta = 0$. Thus ξ is characteristic if, and only if, the 6 equations (3.a.2) are dependent; i.e.,

$$\begin{vmatrix} \xi_1 & 0 & 0 & 0 & \xi_3 & \xi_2 \\ 0 & \xi_2 & 0 & \xi_3 & 0 & \xi_1 \\ 0 & 0 & \xi_3 & \xi_2 & \xi_1 & 0 \\ H^1_{11} & H^1_{22} & H^1_{33} & 2H^1_{23} & 2H^1_{31} & 2H^1_{12} \\ H^2_{11} & H^2_{22} & H^2_{33} & 2H^2_{23} & 2H^2_{31} & 2H^2_{12} \\ H^3_{11} & H^3_{22} & H^3_{33} & 2H^3_{23} & 2H^3_{31} & 2H^3_{12} \end{vmatrix} = 0 \tag{3.a.3}$$

This is the equation given in Cartan [5] and Tenenblat [24]. We write it as

$$F(\xi) = 0, \tag{3.a.4}$$

where clearly $F(\xi)$ is a cubic polynomial or else $F \equiv 0$. In the latter case, every $Q \in II$ is decomposable and we may see that every Q is a multiple of a fixed $\xi \in \mathsf{P}V^*$, as follows: We picture $|II| = \mathsf{P}II$ as a linear system of quadrics $Q \subset \mathsf{P}V \cong \mathsf{P}^2$. By assumption, each Q is a pair of lines, one of which must then be fixed as Q varies. Q.E.D.

In addition to $\Xi \subset \mathsf{P}V^*$ we shall have occasion to use the complex curve $\Xi_{\mathbb{C}} \subset \mathsf{P}V^*_{\mathbb{C}}$, defined by the same equation (3.a.4) but where $\xi \in \mathsf{P}V^*_{\mathbb{C}}$. Proposition (3.a.1) is also valid for $\Xi_{\mathbb{C}}$.

By Theorem B, we see that II is a *general* subspace in case $\Xi_{\mathbb{C}}$ is smooth.
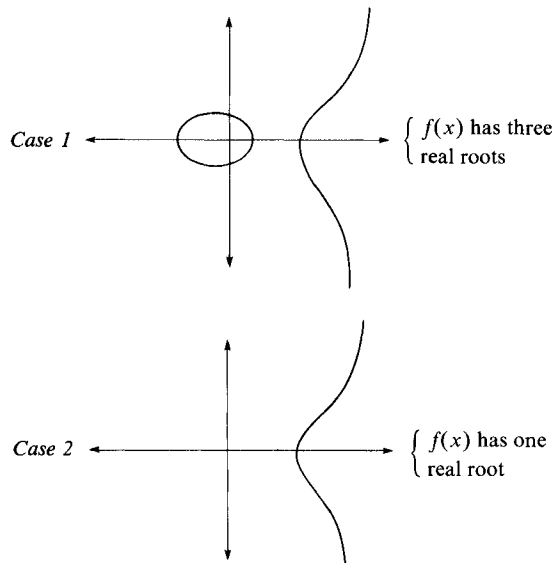
Remarks.   (i) It is easy to see that

$$\{\Xi_{\mathbf{C}} \text{ is smooth}\} \Leftrightarrow \{\Xi \text{ is smooth and irreducible}\}. \qquad (3.a.5)$$

The implication $\Rightarrow$ is clear (if Ξ is reducible, then any two components of $\Xi_{\mathbf{C}}$ must intersect in a singular point). Conversely, if Ξ is irreducible, then $\Xi_{\mathbf{C}}$ is irreducible (if $F = QL$ where L is linear, then either $L = \overline{L}$ or else $Q = M\overline{L}$ where $M = \overline{M}$ is linear). Then any singular point ξ of the irreducible complex cubic curve $\Xi_{\mathbf{C}}$ must be real (otherwise, ξ and $\overline{\xi}$ would be distinct singular points and by Bezout's theorem $\Xi_{\mathbf{C}}$ would contain the line $\overline{\xi}$).

(ii) Since Ξ has odd degree, it is non-empty (cf. (1.c.20)). In fact, it is a well-known consequence of Hanarck's theorem that $\Xi \subset \mathbf{R}P^2$ has either one or two components, each of which is a differentiable circle. Using real affine coordinates (x, y) the Weierstrass normal form of F is

$$y^2 = f(x) = x^3 + ax + b$$

and the possible pictures are

Case 1 $\left\{ \begin{array}{l} f(x) \text{ has three} \\ \text{real roots} \end{array} \right.$

Case 2 $\left\{ \begin{array}{l} f(x) \text{ has one} \\ \text{real root} \end{array} \right.$

As we shall see, the distinction between these two cases is of great analytic significance.

(3.a.6) PROPOSITION.   *If II is general, then II contains no perfect squares and no two-dimensional subspaces $S \subset II$ which depend on only two variables.*

Proof.   Suppose that we could choose bases x^i of V^* and Q^i of II so that Q^1, Q^2 depend only on x^1, x^2. From (3.a.3) we see that $F(\xi_1, \xi_2, 0) \equiv 0$, so that F contains ξ_3 as a factor.

Now suppose we could choose bases so that $Q^3 = (x^3)^2$. Then (3.a.3) reduces to

$$\begin{vmatrix} \xi_1 & 0 & 0 & \xi_3 & \xi_2 \\ 0 & \xi_2 & \xi_3 & 0 & \xi_1 \\ 0 & 0 & \xi_2 & \xi_1 & 0 \\ H^1_{11} & H^1_{22} & 2H^1_{23} & 2H^1_{31} & 2H^1_{12} \\ H^2_{11} & H^2_{22} & 2H^2_{23} & 2H^2_{31} & 2H^2_{12} \end{vmatrix} = 0.$$

Expanding along the middle row, we see that $F(\xi)$ is linear in ξ_3. It is then well-known (and easy to verify) that this implies that either ξ_3 is a factor of F or that $[0, 0, 1]$ is a singular point of Ξ.   Q.E.D.

A remarkable feature of 3-dimensions is that there are *three* cubic curves in the story. Two of these are $\Xi \subset \mathbf{P}V^*$ and $\Lambda \subset \mathbf{P}W$ that we have already encountered in general (cf. (1.b.6)); the remaining one is what we shall call the *parent cubic*

$$\Psi \subset \mathbf{P}V,$$

and it will now be described.

Definition. Let $G \in S^3V^*$ be a cubic form. We define the subspace $\partial G \subset S^2V^*$ by

$$\partial G = \{\partial G/\partial e \in S^2V^* : e \in V\}.$$

Note that ∂G is a subspace of dimension 0 (if $G = 0$), 1 (if G is a perfect cube), 2 (if G depends only on two variables), or 3 (if G depends on all variables).

(3.a.7) PROPOSITION. *If II is general, then there exists $G \in S^3V^*$ such that*

$$\partial G = II.$$

Moreover, G is unique up to scalar multiples.

Proof. Consider the exact sequence (polynomial Poincaré lemma)

$$0 \longrightarrow S^3V^* \xrightarrow{d} S^2V^* \otimes V^* \xrightarrow{d} V^* \otimes \Lambda^2V^* \xrightarrow{d} \Lambda^3V^* \longrightarrow 0$$
$$\quad\quad\quad (10) \quad\quad\quad (18) \quad\quad\quad\quad (9) \quad\quad\quad\quad (1)$$

of forms of degree 3 on V (the maps d are given by exterior differentiation). The number in parentheses under each vector space is its dimension. Consider the subspace $II \otimes V^* \subseteq S^2V^* \otimes V^*$; it has dimension 9. Since $d(S^3V^*)$ has dimension 10 and $S^2V^* \otimes V^*$ only has dimension 18, it follows that

$$d(S^3V^*) \cap (II \otimes V^*) \neq (0).$$

In other words, there exists $G \in S^3V^*$ so that $G \neq 0$ and $dG \in II \otimes V^*$. This clearly is equivalent to $\partial G \subset II$. If G depends on $k < 3$ variables, then

$\dim(\partial G) = k < 3$ and ∂G would depend on only k variables, contradicting Proposition (3.a.6) and our hypothesis of generality. Thus G depends on all three variables and we must have $\partial G = II$.

Now we must prove uniqueness. Suppose that $\tilde{G} \in S^3V^*$ also satisfies $\partial \tilde{G} = II$. Then $\partial \tilde{G} = \partial G$. Let e_1, e_2, e_3 be a basis of V. Then $Q_i = \partial G / \partial e_i$ and $\tilde{Q}_i = \partial \tilde{G} / \partial e_i$ are both bases of II. Therefore, there exists a matrix $A = (A_j^i)$ satisfying $\tilde{Q}_j = A_j^i Q_i$. Since A is 3×3, it has a real eigenvalue, say λ, and a corresponding eigenvector $a = (a^i)$ so that $a^i A_i^j = \lambda a^j$. This gives

$$\frac{\partial \tilde{G}}{\partial e} = a^i \tilde{Q}_i = a^i A_i^j Q_j = \lambda a^j \frac{\partial G}{\partial e_j} = \lambda \frac{\partial G}{\partial e}$$

where $e = a^i e_i$. It follows that

$$\frac{\partial}{\partial e}(\tilde{G} - \lambda G) = 0.$$

Thus $\tilde{G} - \lambda G$ depends on at most two variables, so $\partial(\tilde{G} - \lambda G) \subset II$ has dimension at most 2. By Proposition (3.a.6) (see the above argument constructing G), we see that $\partial(\tilde{G} - \lambda G) = 0$. This can only happen if $\tilde{G} = \lambda G$.   Q.E.D.

(3.a.8) PROPOSITION. *Let $II \subset S^2V^*$ be general and let $G \in S^3V^*$ be such that $\partial G = II$. Then the locus $G = 0$ defines a nonsingular cubic curve $\Psi_C \subset \mathbf{P}V_C$.*

Proof. Referring to (3.a.5), we shall show that Ψ is smooth and irreducible.

To begin, suppose $G = Qx^3$ where $Q \in S^2V^*$ is a quadric. Let $e_1, e_2 \in V$ be such that $x^3(e_1) = x^3(e_2) = 0$. Then

$$\partial G = \{ Q, x^3 \partial Q / \partial e_1, x^3 \partial Q / \partial e_2 \} = II.$$

Since $\dim II = 3$, we see that the linear forms $\partial Q / \partial e_1$ and $\partial Q / \partial e_2$ are independent. But then, by the very definition, the covector $\xi = \xi_1 \partial Q / \partial e_1 + \xi_2 \partial Q / \partial e_2$ is characteristic for all $(\xi_1, \xi_2) \neq 0$. Thus Ξ contains a line, contradicting the generality of II.

Now, suppose that Ψ_C is irreducible but singular. We may suppose coordinates x^i chosen so that the (real!) node or cusp of $G = 0$ is at $[0, 0, 1]$. It follows that G has the form

$$G = G_3(x^1, x^2) + G_2(x^1, x^2)x^3.$$

This implies that the quadrics in $\partial G = II$ are all linear in x^3. Consulting equation (3.a.3), this clearly implies that ξ_3 divides F, so II could not be general.   Q.E.D.

We may now use the classical normal form for a non-singular cubic. We record this lemma without proof, see [9]. It is valid over either $\mathbf{R}$ or $\mathbf{C}$.

(3.a.9) LEMMA. *If $G \in S^3V^*$ is a non-singular cubic form, then there exists a unique number σ and a basis x^1, x^2, x^3 of V^* unique up to permutation and*

simultaneous scaling (by roots of unity, if any) so that

$$G = (x^1)^3 + (x^2)^3 + (x^3)^3 + 6\sigma x^1 x^2 x^3.$$

Moreover, $\sigma \neq -1/2$ or ∞ since Ψ is singular in these cases.

Let us take G in this standard form and compute II and Ξ. A basis for II can be written

$$\partial G = II = \begin{cases} (x^1)^2 + 2\sigma x^2 x^3 \\ (x^2)^2 + 2\sigma x^3 x^1 \cdot \\ (x^3)^2 + 2\sigma x^1 x^2 \end{cases}$$

A covector $\xi = \xi_i x^i$ is characteristic iff the determinant

$$\begin{vmatrix} \xi_1 & 0 & 0 & 0 & \xi_3 & \xi_2 \\ 0 & \xi_2 & 0 & \xi_3 & 0 & \xi_1 \\ 0 & 0 & \xi_3 & \xi_2 & \xi_1 & 0 \\ 1 & 0 & 0 & 2\sigma & 0 & 0 \\ 0 & 1 & 0 & 0 & 2\sigma & 0 \\ 0 & 0 & 1 & 0 & 0 & 2\sigma \end{vmatrix} = 0.$$

Multiplying the first, second, and third columns by 2σ and subtracting them from the fourth, fifth, and sixth columns respectively, we see that this is equivalent to

$$\begin{vmatrix} -2\sigma\xi_1 & \xi_3 & \xi_2 \\ \xi_3 & -2\sigma\xi_2 & \xi_1 \\ \xi_2 & \xi_1 & -2\sigma\xi_3 \end{vmatrix} = 0.$$

Expanding this, we get the result

$$\sigma(\xi_1^3 + \xi_2^3 + \xi_3^3) + (1 - 4\sigma^3)\xi_1\xi_2\xi_3 = 0.$$

It follows that we cannot have $\sigma = 0$ since Ξ is assumed to be non-singular. We may now write the equation in the form

$$F = \xi_1^3 + \xi_2^3 + \xi_3^3 + 6\mu\xi_1\xi_2\xi_3 = 0$$

where

$$\mu = \frac{1 - 4\sigma^3}{6\sigma} \cdot$$

In order to avoid singularity of Ξ we cannot have $\mu = -1/2$. This would give

$$\frac{1 - 4\sigma^3}{6\sigma} = -1/2$$

or

$$4\sigma^3 - 1 = 3\sigma$$

or

$$4\sigma^3 - 3\sigma - 1 = (\sigma - 1)(4\sigma^2 + 4\sigma + 1) = (\sigma - 1)(2\sigma + 1)^2 = 0.$$

Thus, we cannot allow $\sigma = 1$ either. We summarize this discussion by

THEOREM D.   *Let* $II \subset S^2V^*$ *be a general subspace, and let* $G \in S^3V^*$ *satisfy* $\partial G = II$. *Write* G *in its normal form*

$$G = \left(x^1\right)^3 + \left(x^2\right)^3 + \left(x^3\right)^3 + 6\sigma x^1 x^2 x^3.$$

Then we have $\sigma \neq -1/2, 0, 1, \infty$ *and*

$$II = \begin{cases} \left(x^1\right)^2 + 2\sigma x^2 x^3 \\ \left(x^2\right)^2 + 2\sigma x^3 x^1 \\ \left(x^3\right)^2 + 2\sigma x^1 x^2 \end{cases} \qquad (3.a.10)$$

$$F = \xi_1^3 + \xi_2^3 + \xi_3^3 + 6\mu\xi_1\xi_2\xi_3$$
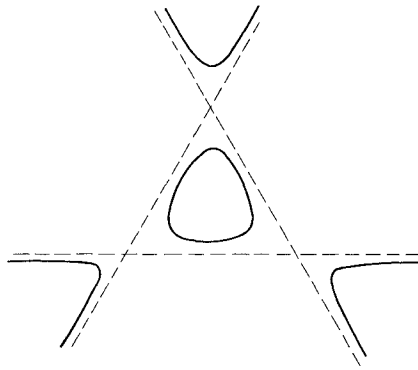
where

$$\mu = \frac{1 - 4\sigma^3}{6\sigma}.$$

Remark.   We graph μ as a function of σ:
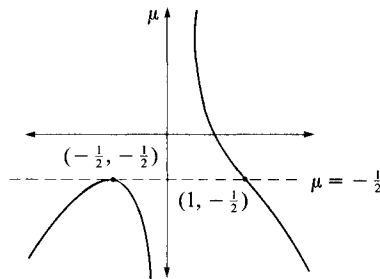
It follows that when $\mu < -1/2$, there are three distinct real values of σ which give the same value of μ. We can now draw pictures of the cubic Ψ for the interesting intervals of σ. The dashed lines are the co-ordinate axes $x^i = 0$, the line at infinity is $x^1 + x^2 + x^3 = 0$ (it contains the three real flexes), and the solid lines constitute the curve. (See next page for figures.)

Graphs of

$$G = (x^1)^3 + (x^2)^3 + (x^3)^3 + 6\sigma x^1 x^2 x^3 = 0.$$

$$-\infty < \sigma < -1/2$$
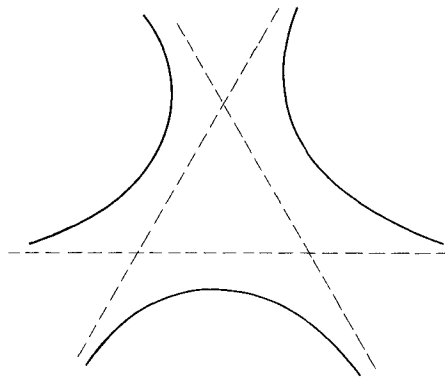
$$-1/2 < \sigma < \infty$$

(b) We now consider a submanifold $M^3 \subset \mathbf{E}^6$. For each $x \in M$ the 2nd fundamental form is given by

$$H_x \in N_x(M) \otimes S^2 T_x^*(M),$$

and it defines a subspace

$$II_x \subset S^2 T_x^*(M)$$

where $II_x = \{ H_x \cdot \nu : \nu \in N_x(M) \}$.

We recall that $M^3 \subset \mathbf{E}^6$ is *general* in case each II_x is a 3-dimensional subspace of $S^2 T_x^*(M)$ that is general.

For a general $M^3 \subset \mathbf{E}^6$, we consider the linearized isometric embedding

equations (cf. (2.c.7) and (2.c.8)). Setting

$$E = T^*(M)$$

$$F = S^2T^*(M)/II \cong (II^\perp)^*$$

we consider (2.c.7) as a 1st order linear differential operator

$$D : C^\infty(E) \to C^\infty(F). \tag{3.b.1}$$

In terms of local coordinates x^i on M and local frames for E, F we will have

$$Du = A^i(x)\frac{\partial u}{\partial x^i} + B(x)u$$

where $u = {}^t(u^1, u^2, u^3)$ and the $A^i(x)$, $B(x)$ are 3×3 matrices.

Definitions. (i) We shall say that (3.b.1) is *symmetric* in case we may choose frames for E, F so that the $A^i(x)$ are symmetric.

(ii) If, in addition, some linear combination $\xi_i(x)A^i(x)$ is positive definite, then we say that (3.b.1) is *symmetric hyperbolic*.

(iii) Finally, (3.b.1) is *hyperbolic* if for each $x \in M$ the characteristic cubic $\Xi_x \subset \mathbf{P}T_x^*(M)$ consists of two components.

Remarks. The notion of symmetric hyperbolic coincides with the classical notion of Friedrichs [6].

Referring to the pictures in remark (ii), following the definition of general in §III(a), (3.b.1) is hyperbolic if each Ξ_x falls in Case 1. For ξ inside the oval, we then have a picture where each line through ξ meets Ξ_x in 3 *distinct real* points.
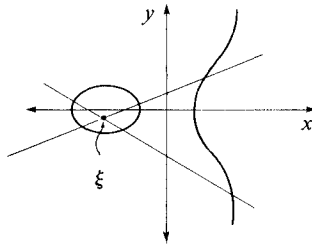
Choose coordinates (x^1, x^2, t) on M so that dt lies inside the oval, and for any choice of frames for E, F write

$$D = \tilde{A}^0(x,t)\partial/\partial t + \tilde{A}^1(x,t)\partial/\partial x^1 + \tilde{A}^2(x,t)\partial/\partial x^2 + B(x,t).$$

For $\xi = (\xi_1, \xi_2, \eta) \in T_x^*(M)$ the symbol is

$$\sigma_D(\xi) = \eta\tilde{A}^0 + \xi_1\tilde{A}^1 + \xi_2\tilde{A}^2.$$

Since $dt = [0, 0, 1] \notin \Xi_x$ we have $\det \tilde{A}^0 \neq 0$. We may then change frames so that

$$D = I \partial/\partial t + A^1(x, t) \partial/\partial x^1 + A^2(x, t) \partial/\partial x^2 + B(x, t). \qquad (3.b.2)$$

The line joining $[0, 0, 1]$ to the point $[\xi_1, \xi_2, 0]$ on the line $\eta = 0$ at infinity is given parametrically by

$$\xi_\lambda = \left[\lambda^{-1} \xi_1, \lambda^{-1} \xi_2, 1 \right]$$

where $\lambda = 0$ corresponds to $[0, 0, 1]$. Thus

$$\sigma_D(\xi_\lambda) = I + \lambda^{-1} \xi_1 A^1 + \lambda^{-1} \xi_2 A^2 = \lambda^{-1}(\lambda I + \xi_1 A^1 + \xi_2 A^2).$$

It follows that the equation

$$\det\|\lambda I + \xi_1 A^1 + \xi_2 A^2\| = 0$$

has, for each $(\xi_1, \xi_2) \neq (0, 0)$, three distinct real roots $\lambda_i(\xi_1, \xi_2)$ $(i = 1, 2, 3)$. Consequently, (3.b.1) is hyperbolic in the sense of Petrovsky–Gårding–Leray (cf. [10]).

Referring to the normal forms for G and F given in Theorem D, we define the following open intervals in R:

$$\begin{cases} I_1 = \{\sigma < -1/2\} \\ I_2 = \{-1/2 < \sigma < 0\} \\ I_3 = \{0 < \sigma < 1\} \\ I_4 = \{1 < \sigma\} \end{cases}. \qquad (3.b.3)$$

THEOREM E. (i) *The linearized isometric embedding system is always symmetric.*

  (ii) *It is symmetric hyperbolic if, and only if, $\sigma \in I_1 \cup I_4$.*

  (iii) *It is hyperbolic if, and only if $\sigma \in I_1 \cup I_2 \cup I_4$.*

  (iv) *In each I_k, it is of principal type.*

Proof. The statement only concerns the symbol mappings

$$\sigma_D : T_x^*(M) \otimes T_x^*(M) \to S^2 T_x^*(M)/II_x.$$

For fixed $x \in M$ we set $V = T(M)$ and let $II \subset S^2 V^*$ correspond to $II_x \subset S^2 T_x^*(M)$. Then, for $\xi, \eta \in V^*$

$$\sigma_D(\xi)(\eta) = \left[\xi \circ \eta \right] \qquad (3.b.4)$$

where $[\xi \circ \eta]$ is the equivalence class of $\xi \circ \eta$ in $S^2 V^*/II$. If $\{x^i\} \in V^*$ is a basis relative to which we have the normal form of Theorem D (note that there is no orthonormality requirement), then the basic observation is that by (3.a.10)

$$x^i \circ x^i \equiv -2\sigma x^j \circ x^k, \qquad i, j, k \text{ distinct}. \qquad (3.b.5)$$

The vectors $x^j \circ x^k$ (j, k distinct) therefore project to a basis of S^2V^*/II, and by (3.b.4) and (3.b.5)

$$\sigma_D(\xi)(\eta) = \sum{}' \left(\tfrac{1}{2}(\xi_j \eta_k + \xi_k \eta_j) - 2\sigma \xi_i \eta_i \right) \left[x^j \circ x^k \right] \qquad (3.b.6)$$

where $\sum'$ means the sum over distinct i, j, k. In terms of this basis

$$\sigma_D(\xi)(\eta) = \xi_1 A^1 \eta + \xi_2 A^2 \eta + \xi_3 A^3 \eta$$

where the A^i are the following 3×3 matrices

$$A^1 = \begin{bmatrix} -2\sigma & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -2\sigma & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$A^3 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -2\sigma \end{bmatrix}$$

From this it is clear that (3.b.1) is symmetric.

In order that the system be symmetric hyperbolic, it is necessary that the characteristic cubic have two components, and that for some $\xi = \xi_i$ the matrix $A_\xi = A^i \xi_i$ be positive definite. Since

$$A_\xi = \begin{bmatrix} -2\xi_1 \sigma & \xi_3 & \xi_2 \\ \xi_3 & -2\xi_2 \sigma & \xi_1 \\ \xi_2 & \xi_1 & -2\xi_3 \sigma \end{bmatrix}$$

we see that the characteristic variety is given by

$$\sigma(\xi_1^3 + \xi_2^3 + \xi_3^3) + (1 - 4\sigma^3)\xi_1 \xi_2 \xi_3 = 0$$

as expected. We have already seen that this has two components exactly when $\sigma < -1/2$, $-1/2 < \sigma < 0$, or $\sigma > 1$.

The condition of positivity is that we have, for some ξ,

$$-2\xi_i \sigma > 0 \qquad \text{for all } i \tag{1}$$

$$4\sigma^2 \xi_i \xi_j > \xi_k^2 \qquad \text{for } i, j, k \text{ distinct} \tag{2}$$

$$\sigma(\xi_1^3 + \xi_2^3 + \xi_3^3) + (1 - 4\sigma^3)\xi_1 \xi_2 \xi_3 > 0. \tag{3}$$

Equation (2) clearly implies that $4\sigma^2 > 1$, so $|\sigma| > 1/2$.

When $\sigma > 1$, conditions (1) and (3) imply (2) as follows: If $-2\xi_i \sigma > 0$, we must

have all $\xi_i < 0$. Thus we get, for example,

$$0 > \sigma(\xi_1^3 + \xi_2^3) + \xi_1\xi_2\xi_3.$$

Since (3) is equivalent to $\sigma(\xi_1^3 + \xi_2^3) + \xi_1\xi_2\xi_3 > \sigma\xi_3(4\xi_1\xi_2 - \xi_3^2)$, we see that we must have $\sigma\xi_3(4\xi_1\xi_2 - \xi_3^2) < 0$. Since $\sigma\xi_3 < 0$, we get $4\xi_1\xi_2 - \xi_3^2 > 0$. The other two inequalities in (2) follow similarly. Thus A_ξ is positive definite if, and only if, ξ is in the interior of the even circuit of Ξ.

If $\sigma < -1/2$ (the only other possibility), the situation is more delicate. The conditions (1) and (3) imply that each $\xi_i > 0$ and that ξ must lie inside the even circuit of Ξ. We must show that every point in the even circuit actually satisfies (2) as well. To do this, we note that the point $[1, 1, 1]$ satisfies (1), (2), and (3); thus, *some* point of the interior of the even circuit satisfies (2). Now consider the pair of equations

$$4\sigma^2\xi_1\xi_2 = \xi_3^2$$

$$\sigma(\xi_1^3 + \xi_2^3 + \xi_3^3) + (1 - 4\sigma^3)\xi_1\xi_2\xi_3 = 0.$$

Multiplying the first by $\sigma\xi_3$ and adding it to the second, we get $\sigma(\xi_1^3 + \xi_2^3) = -\xi_1\xi_2\xi_3$. Squaring both sides of this equation, we get

$$\sigma^2(\xi_1^3 + \xi_2^3)^2 = \xi_1^2\xi_2^2\xi_3^2 = 4\sigma^2\xi_1^3\xi_2^3$$

so

$$(\xi_1^3 - \xi_2^3)^2 = 0$$

so the only *real* points of intersection of the curves satisfy $\xi_1 = \xi_2$. From this we deduce $\xi_3 = -2\sigma\xi_1$. Thus the unique *real* point of intersection of Ξ and the quadric $4\sigma^2\xi_1\xi_2 - \xi_3^2 = 0$ is $[1, 1, -2\sigma]$. Since the regions $4\sigma^2\xi_1\xi_2 - \xi_3^2 > 0$ and $\det(A_\xi) > 0$ have $[1, 1, 1]$ in common, it follows that Ξ and the quadric must be tangent at $[1, 1, -2\sigma]$ and that the interior of the even circuit of Ξ lies inside the quadric $4\sigma^2\xi_1\xi_2 - \xi_3^2 = 0$. Clearly, the same argument holds for the other two equations of (2).

This completes the proof of Theorem E.

(c)   In the previous section, we saw that the modulus σ of the second fundamental form II of a general submanifold $M^3 \subseteq E^6$ completely determines the type (e.g., symmetric hyperbolic, etc.) of the linearized isometric embedding equations. In this section, we want to study the relationship between the curvature of the induced metric ds^2 on M^3 and the modulus σ.

For comparison, consider the case of a surface $M^2 \subseteq E^3$. For $x \in M^2$, we set $V = T_xM$ and note that II_x consists of a single quadric. By choosing an appropriate basis x^1, x^2 of V^*, we may display the relationship between the possible normal forms of II_x and the Gauss curvature $K(x)$ of the induced metric

ds^2 at x:

$$II_x = \begin{cases} \left(x^1\right)^2 + \left(x^2\right)^2 & K(x) > 0 \\ \left(x^1\right)^2 & K(x) = 0 \\ \left(x^1\right)^2 - \left(x^2\right)^2 & K(x) < 0 \end{cases}$$

This is, of course, well known.

In the case of a general $M^3 \subseteq \mathsf{E}^6$, the relationship between the modulo and the Riemann curvature tensor R is not so direct.

The Gauss equations are written

$$\gamma(H,H) = R \tag{3.c.1}$$

where, now, we think of $R \in K \subseteq S^2(\Lambda^2 V^*)$ as given and $H \in W \otimes S^2 V^*$ as unknown. In the case $n = 3$, we know that $K = S^2(\Lambda^2 V^*)$. If we set $\tilde{\gamma}(H) = \gamma(H,H)$, the quadratic map

$$\tilde{\gamma} : W \otimes S^2(V^*) \to S^2(\Lambda^2 V^*) \tag{3.c.2}$$

is $O(W)$-invariant and $GL(V)$-equivariant. It follows that $\tilde{\gamma}$ maps $O(W) \times GL(V)$-orbits to $GL(V)$-orbits.

The orbits of $GL(V)$ acting on $S^2(\Lambda^2 V^*)$ are well understood. If we regard $S^2(\Lambda^2 V^*)$ as the space of quadratic forms on $\Lambda^2 V$, then each element $R \in S^2(\Lambda^2 V^*)$ has a well-defined signature, (p,q), where p is the dimension of the largest subspace of $\Lambda^2 V$ on which R is positive definite and q is the dimension of the largest subspace on which R is negative definite. A $GL(V)$-orbit in $S^2(\Lambda^2 V^*)$ then consists of all R with a given signature (p,q):

$$\mathcal{O}_{p,q} = \left\{ R \in S^2(\Lambda^2 V^*) \,\middle|\, \mathrm{sign}(R) = (p,q) \right\}.$$

The following topological features of the orbits will be important in our discussion: The orbit $\mathcal{O}_{p,q}$ is open iff $p + q = 3$. The orbits $\mathcal{O}_{3,0}$ and $\mathcal{O}_{0,3}$ are convex cones in $S^2(\Lambda^2 V^*)$. The closure of an orbit $\mathcal{O}_{p,q}$ consists of all the orbits $\mathcal{O}_{p',q'}$ with $p' \leqslant p$ and $q' \leqslant q$.

The geometric meaning of R as a quadratic form on $\Lambda^2 V$ is well known: If e_1, e_2 are orthonormal with respect to the metric ds^2, then $R(e_1 \wedge e_2)$ is the sectional curvature of the two-plane spanned by e_1 and e_2. Henceforth, when we speak of the *signature* of R we mean the signature as defined above. We remind the reader that this signature is *not* the same as the signature of the Ricci tensor, $\mathrm{Ric} \in S^2(V^*)$, which may be formed from R by use of the metric $ds^2 \in S^2(V^*)$.

The orbit structure of $O(W) \times GL(V)$ acting on $W \otimes S^2(V^*)$ is somewhat more complicated. However, we need only a description of the orbits in the open set $U \subseteq W \otimes S^2(V^*)$ consisting of *general* second fundamental forms. For each $H \in U$, let $\sigma(H)$ denote the modulus of the parent cubic of the corresponding

subspace $II \subset S^2 V^*$. The function σ is constant on the orbits of $O(W) \times GL(V)$ acting on U and can be seen to be smooth there. It follows, by Theorem D, that we may write U as the disjoint union of open sets $\{U_l \mid 1 \leqslant l \leqslant 4\}$ where $H \in U_l$ if and only if $\sigma(H) \in I_l$ (see (3.b.3)). Referring to Theorem E, we shall say that $H \in U$ is *symmetric hyperbolic* (resp. *hyperbolic, principal type*) if $H \in U_1 \cup U_4$ (resp. $H \in U_1 \cup U_2 \cup U_4$, $H \in U_1 \cup U_2 \cup U_3 \cup U_4$).

THEOREM F. *Let $R \in S^2(\Lambda^2 V^*)$ be given and suppose the signature of R is (p,q).*

  (i) *If $q \geqslant 2$, the Gauss equations (3.c.1) have a symmetric hyperbolic solution.*
  (ii) *If $(p,q) \neq (0,0)$ or $(0,1)$ then (3.c.1) has a hyperbolic solution.*
  (iii) *If $(p,q) \neq (0,0)$, then (3.c.1) has a principal type solution.*

Remark. What we shall actually prove is that the range of σ on $\tilde{\gamma}^{-1}(R) \cap U$ is described by the following Table, in which (k,l) refers to the signature of R and the union of intervals underneath is the range of σ on $\tilde{\gamma}^{-1}(R) \cap U$, except for the two starred entries, where the range of σ may be larger. That this Table implies Theorem F is an immediate consequence of Theorem E and (3.b.3).

TABLE

(3,0)	(2,1)*		(1,2)		(0,3)
$I_2 \cup I_3$	$I_2 \cup I_3$		$I_1 \cup I_2 \cup I_3 \cup I_4$		$I_1 \cup I_2 \cup I_3 \cup I_4$
(2,0)		(1,1)*		(0,2)	
$I_2 \cup I_3$		$I_2 \cup I_3$		$I_1 \cup I_2 \cup I_3 \cup_4$	
	(1,0)		(0,1)		
	I_2		I_3		
		(0,0)			
		$\varnothing$			

Proof of Table. Let

$$U(\lambda) = \{ H \in U \mid \sigma(H) = \lambda \}$$

where $\lambda \notin \{-1/2, 0, 1, \infty\}$. $U(\lambda)$ is a union of $O(W) \times GL(V)$-orbits. Because of the equivariance of $\tilde{\gamma}$, it follows that $\tilde{\gamma}(U(\lambda)) \subseteq S^2(\Lambda^2 V^*)$ is a union of $GL(V)$-orbits. It follows from our previous discussion that the signature of $R \in S^2(\Lambda^2 V^*)$ completely determines whether or not $R \in \tilde{\gamma}(U(\lambda))$. Obviously, this is the same as saying that the range of σ on $\gamma^{-1}(R) \cap U$ depends only on the signature of R.

Let $\mathscr{F}(W)$ denote the space of bases $w = (w_1, w_2, w_3)$ of W and let $\mathscr{F}(V^*)$ denote the space of bases $y = (y^1, y^2, y^3)$ of V^*. For each $\lambda \notin \{-1/2, 0, 1, \infty\}$ we define

$$Q^i(y, \lambda) = (y^i)^2 + 2\lambda y^j \circ y^k$$

$$H(w, y, \lambda) = w_i \otimes Q^i(y, \lambda).$$

Clearly, this gives a map for each $l = 1, 2, 3, 4$

$$\mathcal{F}(W) \times \mathcal{F}(V^*) \times I_l \to U_l.$$

By Theorem D, this map is surjective and satisfies $H(w, y, \lambda) \in U(\lambda)$ for all λ. The fibers of this map are given by the twelve curves

$$t \mapsto \left(t^{-2}(w \cdot \epsilon), \pm t(y \cdot \epsilon)\right) \qquad (t > 0)$$

where $\epsilon \in S_3$ and S_3 is the sub-group of permutation matrices in $GL(3, \mathbf{R})$. Thus, in order to determine $\tilde{\gamma}(U(\lambda))$ as a union of $GL(V)$-orbits, it suffices to determine the possible signatures of $R = \tilde{\gamma}(H(w, y, \lambda))$ as $w \in \mathcal{F}(W)$ and $y \in \mathcal{F}(V^*)$ vary. Now, we may write

$$H(w, y, \lambda) = w_i \otimes Q^i(y, \lambda) = H_{ij}(w, \lambda) \otimes y^i \circ y^j$$

where

$$\|H_{ij}(w, \lambda)\| = \begin{vmatrix} w_1 & \lambda w_3 & \lambda w_2 \\ \lambda w_3 & w_2 & \lambda w_1 \\ \lambda w_2 & \lambda w_1 & w_3 \end{vmatrix}.$$

Writing

$$R = R_{ijkl}(y^i \wedge y^j) \circ (y^k \wedge y^l),$$

the Gauss equations $\tilde{\gamma}(H(w, y, \lambda)) = R$ become (cf. (1.b.1))

$$R_{ijkl} = H_{ik}(w, \lambda) \cdot H_{jl}(w, \lambda) - H_{il}(w, \lambda) \cdot H_{jk}(w, \lambda). \tag{3.c.3}$$

This equation may be written in terms of symmetric 3×3 matrices as follows: Let $\mathcal{S}$ denote the vector space of 3×3 symmetric matrices. For each $\lambda \notin \{-1/2, 0, 1, \infty\}$ we define a linear transformation $L_\lambda : \mathcal{S} \to \mathcal{S}$ by

$$L_\lambda \begin{bmatrix} a_1 & b_3 & b_2 \\ b_3 & a_2 & b_1 \\ b_2 & b_1 & a_3 \end{bmatrix} = \begin{bmatrix} b_1 - \lambda^2 a_1 & -\lambda a_3 + \lambda^2 b_3 & -\lambda a_2 + \lambda^2 b_2 \\ -\lambda a_3 + \lambda^2 b_3 & b_2 - \lambda^2 a_2 & -\lambda a_1 + \lambda^2 b_1 \\ -\lambda a_2 + \lambda^2 b_2 & -\lambda a_1 + \lambda^2 b_1 & b_3 - \lambda^2 a_3 \end{bmatrix}$$

If we define $h(w) \in \mathcal{S}$ and $G(R) \in \mathcal{S}$ by

$$h_{ij}(w) = w_i \cdot w_j \tag{3.c.5}$$

$$G_{ij}(R) = R_{klmn} \tag{3.c.6}$$

where (i, k, l) and (j, m, n) are even permutations of $(1, 2, 3)$, then the equations (3.c.3) may be written

$$G(R) = L_\lambda(h(w)). \tag{3.c.7}$$

Now it is easy to see that the signature of R as defined above is the same as the usual signature of $G(R)$ as a symmetric matrix. On the other hand, $h(w)$ is necessarily positive definite, being the matrix of inner products of a basis of W.

Let $\mathscr{P} \subseteq \mathscr{S}$ denote the convex cone of positive definite matrices. Our discussion has shown that

(3.c.8) PROPOSITION. *The following are equivalent*

(i) *For a given* $R \in S^2(\Lambda^2 V^*)$ *of signature* (k, l), *there exists an* $H \in W \otimes S^2 V^*$ *satisfying* $\sigma(H) = \lambda$ *and*

$$\gamma(H, H) = R.$$

(ii) $\mathscr{O}_{(k, l)} \subseteq \tilde{\gamma}(U(\lambda))$.

(iii) *There exists a matrix of signature* (k, l) *in*

$$L_\lambda(\mathscr{P}) \subseteq \mathscr{S}.$$

Using (3.c.8), the proof of the Table follows from the case analysis below. In what follows, we always assume that $\lambda \notin \{-1/2, 0, 1, \infty\}$.

(3.c.9) PROPOSITION. L_λ *is invertible. In fact*

$$L_\lambda^{-1} = \frac{L_\lambda}{\lambda^4 - \lambda}.$$

Proof. Compute.

(3.c.10) PROPOSITION. $L_\lambda(\mathscr{P})$ *does not contain* 0 *(and, hence, no matrix of signature* $(0, 0)$).

Proof. $0 \notin \mathscr{P}$ and L_λ is invertible. Q.E.D.

(3.c.11) PROPOSITION. *Consider the matrix*

$$S = \begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}$$

(i) *If* $b = 0$, *then* S *has a triple eigenvalue at* $\mu = a$.

(ii) *If* $b \neq 0$, *then* S *has a double eigenvalue at* $\mu = a - b$ *and a single eigenvalue at* $\mu = a + 2b$.

Proof. Compute.

(3.c.12) PROPOSITION. *If* $L_\lambda(\mathscr{P})$ *contains matrices with signature* $(3, 0)$, *then* $\lambda \in I_2 \cup I_3$.

Proof. Suppose that $L_\lambda(\mathscr{P}) \cup \mathscr{P} \neq \emptyset$, and let $M \in \mathscr{P}$ be such that $L_\lambda(M)$ $\in \mathscr{P}$. Define $T: \mathscr{S} \to \mathscr{S}$ by

$$T\begin{bmatrix} a_1 & b_3 & b_2 \\ b_3 & a_2 & b_1 \\ b_2 & b_1 & a_3 \end{bmatrix} = \begin{bmatrix} a_2 & b_1 & b_3 \\ b_1 & a_3 & b_2 \\ b_3 & b_2 & a_1 \end{bmatrix}.$$

Clearly T preserves $\mathscr{P}$ and commutes with L_λ. Thus $L_\lambda(TM), L_\lambda(T^2M) \in \mathscr{P}$. Since $\mathscr{P} \cap L_\lambda^{-1}(\mathscr{P})$ is a convex cone, we see that $\overline{M} = 1/3(M + TM + T^2M)$ belongs to $\mathscr{P} \cap L_\lambda^{-1}(\mathscr{P})$. We write

$$\overline{M} = \begin{bmatrix} a & b & b \\ b & a & b \\ b & b & a \end{bmatrix}.$$

Now $\overline{M} > 0$ so we may divide by a so as to reduce to the case $a = 1$. We then have

$$L_\lambda(\overline{M}) = \begin{bmatrix} b - \lambda^2 & \lambda^2 b - \lambda & \lambda^2 b - \lambda \\ \lambda^2 b - \lambda & b - \lambda^2 & \lambda^2 b - \lambda \\ \lambda^2 b - \lambda & \lambda^2 b - \lambda & b - \lambda^2 \end{bmatrix}.$$

Referring to (3.c.11) we see that the conditions

$$\overline{M} > 0$$

$$L_\lambda(\overline{M}) > 0$$

are incompatible unless $-1/2 < \lambda < 1$. Q.E.D.

(3.c.13) PROPOSITION. *$L_\lambda(\mathscr{P})$ contains matrices of signature $(0,1)$ if and only if $\lambda \in I_3$.*

Proof. Any matrix with signature $(0,1)$ can be written

$$Q = -\begin{bmatrix} r_1^2 & r_1 r_2 & r_1 r_3 \\ r_1 r_2 & r_2^2 & r_2 r_3 \\ r_1 r_3 & r_2 r_3 & r_3^2 \end{bmatrix}.$$

Setting

$$\Delta = (2\lambda^3 + 1)r_1 r_2 r_3 - \lambda^2(r_1^3 + r_2^3 + r_3^3)$$

we may compute

$$\det L_\lambda^{-1}(Q) = \frac{\Delta^2}{(\lambda - \lambda^4)^3}.$$

Hence, if $L_\lambda^{-1}(Q) > 0$ we must have $\lambda - \lambda^4 > 0$ so $0 < \lambda < 1$. Conversely, if $0 < \lambda < 1$, we may set $r_1 = r_2 = r_3 = 1$ and see that $L_\lambda^{-1}(Q) > 0$.   Q.E.D.

(3.c.14) PROPOSITION. $L_\lambda(\mathscr{P})$ contains matrices of signature $(1,0)$ if and only if $\lambda \in I_2$.

Proof. Keeping the same notations as in (3.c.13), we see that $-Q$ has signature $(1,0)$. We also compute

$$\det L_\lambda^{-1}(-Q) = \frac{\Delta^2}{(\lambda^4 - \lambda)^3} .$$

Thus if $L_\lambda^{-1}(-Q) > 0$ we must have $\lambda^4 - \lambda > 0$, i.e., $\lambda > 1$ or $\lambda < 0$. Now $L_\lambda(\mathscr{P})$ is an open set, and any open neighborhood of such a $-Q$ must contain positive definite matrices. By (3.c.12), $L_\lambda(\mathscr{P})$ contains positive definite matrices only if $-1/2 < \lambda < 0$ or $0 < \lambda < 1$. Hence, we conclude that the only possibility for λ is $-1/2 < \lambda < 0$. Conversely, if $\lambda \in I_2$, then setting $r_1 = r_2 = r_3 = 1$, we get $L_\lambda^{-1}(Q) > 0$.   Q.E.D.

(3.c.15) PROPOSITION. $L_\lambda(\mathscr{P})$ contains matrices of signature $(0,2)$, $(1,2)$, and $(0,3)$ for all λ.

Proof. If $L_\lambda(\mathscr{P})$ contains a matrix of signature $(0,2)$, then it necessarily contains matrices of signature $(1,2)$ and $(0,3)$. This is because $L_\lambda(\mathscr{P})$ is open and any neighborhood of a matrix with signature $(0,2)$ contains matrices of signature $(1,2)$ and $(0,3)$.

If $\lambda \in I_3$, then $L_\lambda(\mathscr{P})$ contains a matrix of signature $(0,1)$, hence we are done by the above argument.

If $\lambda \in I_1 \cup I_2 \cup I_4$, then the matrix

$$M = \begin{bmatrix} 1 & b & b \\ b & 1 & b \\ b & b & 1 \end{bmatrix}$$

with $b = \lambda(\lambda + 2)/(2\lambda^2 + 1)$ is positive definite and $L_\lambda(M)$ has signature $(0,2)$.
Q.E.D.

(3.c.16) PROPOSITION. $L_\lambda(\mathscr{P})$ contains matrices of signature $(2,0)$ if and only if $\lambda \in I_2 \cup I_3$.

Proof. If $L_\lambda(\mathscr{P})$ contains a matrix of signature $(2,0)$ it must contain a matrix of signature $(3,0)$. By (3.c.12) this forces $\lambda \in I_2 \cup I_3$.
If $\lambda \in I_3$, then the matrix

$$M = \begin{bmatrix} 1 & b & b \\ b & 1 & b \\ b & b & 1 \end{bmatrix}$$

where $b = \lambda(\lambda + 2)/(2\lambda^2 + 1)$ is positive definite and $L_\lambda(M)$ has signature $(2,0)$.

If $\lambda \in I_2$, then $L_\lambda(\mathscr{P})$ contains matrices of signature $(1,0)$ by (3.c.14) and hence must contain matrices of signature $(2,0)$. Q.E.D.

(3.c.17) PROPOSITION. *If $\lambda \in I_2 \cup I_3$ then $L(\mathscr{P})$ contains matrices of signature* $(2,1)$ *and* $(1,1)$.

Proof. If $\lambda \in I_2$, then $L_\lambda(\mathscr{P})$ contains matrices of signature $(1,0)$. Hence, by openness, it contains matrices of signature $(2,1)$ and $(1,1)$.

If $\lambda \in I_3$, then $L_\lambda(\mathscr{P})$ contains matrices of signature $(0,1)$. Again, openness implies that $L_\lambda(\mathscr{P})$ contains matrices of signature $(2,1)$ and $(1,1)$. Q.E.D.

IV. The set-up for the Nash–Moser–Schwartz–Sergeraert theorem.

(a) We begin by defining the manifolds and Banach spaces which will be used in §§IV–VI.

Let $\Omega \subset (0,1)^{n-1}$ be an open convex set. The manifold which will be embedded into Euclidean space is

$$M = \overline{\Omega} \times \left[\tfrac{1}{4}, \tfrac{3}{4} \right], \qquad \dim M = n.$$

Although this will only be done for $n = 3$, the discussion in §§IV–V is valid for all $n \geqslant 1$. Therefore, we will restrict to $n = 3$ in §VI only.

The co-ordinates on M will be $(x^1, \ldots, x^{n-1}, t)$, $x = (x^1, \ldots, x^{n-1}) \in \Omega$, $t \in [1/4, 3/4]$, in §V, but in §VI we will change "t" to "x^n". We do this because in §V, it is useful to distinguish a "time co-ordinate" t in contrast to the "space co-ordinates" $x^1, \ldots, x^{n-1}$. On the other hand, in §VI, when the proofs of the main theorems are being given, it will be more convenient to "forget" the distinguished co-ordinate.

In §V a linear hyperbolic system of partial differential equations on M will be solved by extending the system to one on a larger manifold

$$X = T^{n-1} \times [0,1], \qquad T^{n-1} = \mathsf{R}^{n-1} / \mathsf{Z}^{n-1},$$

where hyperbolic systems are more easily solved. Co-ordinates on X will be $(x^1, \ldots, x^{n-1}, t)$, $0 \leqslant x^i$, $t < 1$. Another domain we will use is the "spacelike hypersurface"

$$T_t = T \times \{t\} \subset X, \qquad t \in [0,1].$$

Given a multi-index $\alpha \in (\mathsf{Z}^+)^{n-1}$, we will denote

$$D_x^\alpha = \left(\frac{\partial}{\partial x^1} \right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x^{n-1}} \right)^{\alpha_{n-1}}.$$

However, if $\alpha \in (\mathsf{Z}^+)^n$, then

$$D_x^\alpha = \left(\frac{\partial}{\partial x^1} \right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x^n} \right)^{\alpha_n} = \left(\frac{\partial}{\partial x^1} \right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x^{n-1}} \right)^{\alpha_{n-1}} \left(\frac{\partial}{\partial t} \right)^{\alpha_n}.$$

Let $\mathscr{D}$ be M, X, or T_t; V a normed vector space, and E a vector bundle over $\mathscr{D}$ with a smooth norm on the fibers. We denote by $C^\infty(\mathscr{D}, V)$ $(C^\infty(E))$ the

space of V-valued functions (sections of E) which are smooth up to the boundary. $L^2(\mathscr{D}, V)$ $(L^2(E))$ is the space of square-integrable V-valued functions (sections of E) with respect to the measure $dx = dx^1 \ldots dx^{n-1}$ or $dx\,dt = dx^1 \ldots dx^n$, depending on $\mathscr{D}$. The L^2-norm will be denoted

$$|u|_{L^2(\mathscr{D})} = \left(\int_{\mathscr{D}} |u|^2 \, dx\,(dt) \right)^{1/2}.$$

The Sobolev spaces $L_k^2(\mathscr{D}, V)$ and $L_k^2(E)$ are defined to be the completions of $C^\infty(\mathscr{D}, V)$ and $C^\infty(E)$ respectively, using the following norms:

$$|u|_{L_k^2(\mathscr{D})} = \left(\sum_{|\alpha| \leqslant k} |D_x^\alpha u|_{L^2(\mathscr{D})}^2 \right)^{1/2}, \qquad k \geqslant 0.$$

Here, α is a multi-index running over $(\mathbf{Z}^+)^{\dim \mathscr{D}}$.

Given $u : X \to V$, we also define the following norms:

$$|u|_{L_{k,l}^2(T_l)} = \left(\sum_{j=0}^l \sum_{|\alpha| \leqslant k-j} \left| \frac{\partial^j}{\partial t^j} D_x^\alpha u \right|_{L^2(T_l)}^2 \right)^{1/2}, \qquad 0 \leqslant t \leqslant 1;$$

$$|u|_{L_{k,l}^2(X)} = \left(\sum_{j=0}^l \sum_{|\alpha| \leqslant k-j} \left| \frac{\partial^j}{\partial t^j} D_x^\alpha u \right|_{L^2(X)}^2 \right)^{1/2}.$$

Here, $\alpha = (\alpha_1, \ldots, \alpha_{n-1})$, $D_x^\alpha = (\partial / \partial x^1)^{\alpha_1} \ldots (\partial / \partial x^{n-1})^{\alpha_{n-1}}$.

The vector bundle $S^2 T^* M$ has a canonical trivialization with respect to the co-ordinates $(x^1, \ldots, x^n)$. Using this and the flat metric, we can define, as above, the spaces $C^\infty(S^2 T^* M), L_k^2(S^2 T^* M)$ of sections of the bundle. In local co-ordinates, given any section $g = g_{ij} \, dx^i \, dx^j$,

$$|g|_{L^2(M)} = \left(\int_M \sum_{i,j=1}^n (g_{ij})^z \, dx \, dt \right)^{1/2}$$

(b) The main tool we shall use to prove theorem G is the Nash–Moser–Schwartz–Sergeraert theorem. The heart of this theorem is a simplified version of the Nash–Moser iteration scheme. Just as the Picard iteration scheme leads to the usual implicit function theorem, the Nash–Moser scheme leads to a generalized implicit function theorem.

Let $\Phi : E \to F$ be a Fréchet differentiable map between two Banach spaces; usually it will be a nonlinear differential operator mapping between the appropriate Sobolev or Hölder spaces. Given $u_0 \in E$ and $f_0 \in F$ satisfying $\Phi(u_0) = f_0$, the usual implicit function theorem says that if the linearized operator $\Phi'(u_0)$ has a bounded right inverse, then for f sufficiently close to f_0, there exists $u \in E$ satisfying

$$\Phi(u) = f. \tag{1}$$

One way to prove this is to construct a sequence $u_n \to u$ using the Picard iteration scheme, i.e.,

$$u_{n+1} = u_n + \Phi'(u_0)^{-1}(f - \Phi(u_n)), \qquad n \geqslant 0$$

However, when solving a linear differential equation

$$Pu = f,$$

where we assume P is a nth order operator with smooth coefficients, it is not always possible to prove that if $f \in L_k^2$ or C^k, then $u \in L_{k+n}^2$ or C^{k+n}. In other words, in solving for u, we have suffered a "loss of derivatives." Then viewing $P: L_{k+n}^2 \to L_k^2$ or $C^{k+n} \to C^k$, the inverse to P is unbounded. Therefore, the Picard iteration scheme diverges badly since more derivatives are lost for each iteration.

If there is to be any hope of solving for u satisfying (1), the loss of derivatives must be controlled. Nash's fundamental idea was to modify the correction term by smoothing it, and to use Newton's iteration scheme, a much more rapidly convergent method than Picard's. The iteration scheme he used was a fairly complicated one and several people have given simplified schemes which usually yield weaker results than Nash's original formulation.

The iteration scheme used here is of the form

$$u_{n+1} = u_n + S_n\Phi'(u_n)^{-1}(f - \Phi(u_n)), \qquad n \geqslant 0, \tag{2}$$

where S_n is a sequence of smoothing operators, i.e., $S_n u$ is smooth and for n large, $S_n u$ approximates u. Observe that we need to have $\Phi'(u)$ be invertible for all u near u_0; this is a stronger assumption than needed for the standard implicit function theorem. Further assumptions on Φ are needed to prove that an iteration sequence like (2) converges. The key requirement is that Φ satisfy estimates of the form

$$|\Phi'(u)v|_k \leqslant c(|u|_\alpha)(|v|_{k+\beta} + |u|_{k+\gamma}|v|_\beta)$$

$$|\Phi'(u)^{-1}h|_k \leqslant c'(|u|_{\alpha'})(|h|_{k+\beta'} + |u|_{k+\gamma'}|v|_{\beta'}), \tag{3}$$

where $\alpha, \beta, \gamma, \alpha', \beta', \gamma'$ are fixed constants and $|\ |_s$ denotes a Hölder or Sobolev norm. The essential feature to notice is that there are no terms of the form $|u|_{k+\alpha}|v|_{k+\beta}$ or $|u|_{k+\alpha'}|h|_{k+\beta'}$. Such "quadratic" terms cause the iteration scheme to diverge. However, it turns out that estimates like (3) are not unreasonable to expect. The basic example to which almost everything finally reduces is given by lemma (5.6.1) which says roughly that for any two functions f and g,

$$|fg|_k \leqslant C_k(|f|_k|g|_0 + |f|_0|g|_k).$$

As can be seen in §§V–VI, this lemma seems to be the key fact when applying the Nash–Moser–Schwartz–Sergeraert theorem.

The theorem we give here is exactly the same as the one stated in [17], except that we have put the assumptions into the form of [3]. However, in [17], only a C^k-existence theorem is proved, one apparently far weaker than Nash's original theorem which gave smooth solutions. M. F. Sergeraert, in [19], was able to make clever use of Schwartz's own estimates and a standard interpolation inequality (cf. (4.b.1)) to prove a C^∞-existence theorem.

We will not use here the specific Banach spaces defined in IV(a). Instead the Nash–Moser–Schwartz–Sergeraert theorem will be stated for a general scale of Banach spaces.

Definition. $\{ E_k, k = 0, 1, 2, \ldots \}$ is a *scale of Banach spaces* if
1. Each E_k is a Banach space with norm $|\ |_k$.
2. For all $k \leqslant l$, $E_k \supset E_l$; and

$$|u|_k \leqslant |u|_l \qquad u \in E_l$$

We set

$$E_\infty = \bigcap_{k \geqslant 0} E_k .$$

The standard examples of such a scale are the scale of Sobolev spaces defined earlier and the scale of C^k-spaces.

In either case, E_∞ is a space of smooth functions. We say that a scale $\{ E_k \}$ has *smoothing operators* if there exists a family of operators:

$$S_\theta : E_0 \to E_\infty , \qquad \theta \geqslant 1$$

which satisfy the following estimates for any $0 \leqslant k \leqslant l$:

$$|S_\theta u|_l \ \leqslant \ M_{l,k} \theta^{l-k} |u|_k , \qquad u \in E_k \qquad (S1)$$

$$|(I - S_\theta)u|_k \ \leqslant \ M_{l,k} \theta^{k-l} |u|_l , \qquad u \in E_l \qquad (S2)$$

$$\left| \frac{d}{d\theta} S_\theta u \right|_k \ \leqslant \ M_{l,k} \theta^{k-l-1} |u|_l , \qquad u \in E_l \qquad (S3)$$

Here, $M_{l,k}$ is a constant independent of u.

$$\lim_{\theta \to +\infty} |(I - S_\theta)u|_k = 0, \qquad u \in E_k \qquad (S4)$$

Before discussing the theorem, we state and prove an interpolation inequality which will be quite essential both in proving the Nash–Moser–Schwartz–Sergeraert theorem and in applying it.

(4.6.1) LEMMA. *Let* $\{ E_k \}$ *be a scale of Banach spaces with smoothing operators* S_θ, $\theta \geqslant 1$. *Then the following estimate holds*:

$$|u|_j \ \leqslant \ \hat{M}_{i,j,k} |u|_i^{(k-j)/(k-i)} |u|_k^{(j-i)/(k-i)}, \qquad u \in E_k , \quad i \leqslant j \leqslant k$$

Here, $\hat{M}_{i,j,k}$ *is a constant independent of* u.

Proof.

$$|u|_j \leqslant |(I - S_\theta)u|_j + |S_\theta u|_j \leqslant M_{j,i}\theta^{j-i}|u|_j + M_{k,j}\theta^{k-j}|u|_k$$

This estimate holds for any $\theta \geqslant 1$. In particular, set $\theta = (|u|_i/|u|_k)^{1/k-i}$. We then have:

$$|u|_j \leqslant (M_{j,i} + M_{k,j})|u|_i^{k-j/k-i}|u|_k^{j-i/k-i} \quad \text{Q.E.D.}$$

Let $\{E_k\}, \{F_k\}$ be scales of Banach spaces with smoothing operators; we will denote the norms of both scales by $|\ |_k$. D_k will denote the closed unit ball in E_k, and for any $r \geqslant 0$, $D_k'(r)$ will be the ball of radius r centered about 0 in F_k.

NASH–MOSER–SCHWARTZ–SERGERAERT THEOREM. *Let* $\Phi : D_0 \to F_0$ *be a map satisfying the following assumptions:*
(1) $\Phi(0) = 0$
(2) *For any* $k \geqslant 0$, $\Phi : D_0 \cap E_k \to F_k$ *is a twice Fréchet differentiable map*
(3) *The Fréchet derivatives* Φ', Φ'' *satisfy the following estimates:*
For any $k \geqslant 0$; $u \in D_0 \cap E_k$; $v, w \in E_k$

$$|\Phi'(u)v|_k \leqslant C_k'(|v|_k + |u|_k|v|_0) \tag{4.b.3}$$

$$|(\Phi''(u)v, w)|_k \leqslant C_k''(1 + |u|_k)|v|_k|w|_k \tag{4.b.4}$$

(4) *There is an integer* $\alpha \geqslant 0$ *such that for any* $u \in D_\alpha$, $\Phi'(u)$ *has a right inverse* $Q(u) : F_\alpha \to E_0 \cap D_\alpha$ *satisfying the following estimate:*

$$|Q(u)v|_{k-\alpha} \leqslant K_k(|v|_k + |u|_k|v|_\alpha) \qquad k \geqslant \alpha, \quad u \in D_\alpha \cap E_k, \quad v \in E_k. \tag{4.b.5}$$

Then there exists $\delta > 0$ *such that for any integer* $\eta \geqslant 11$, *we can find a continuous map* $\Psi : D'_{\eta\alpha}(\delta) \to D_\alpha$ *satisfying:*

$$\Phi \circ \Psi(u) = u \qquad u \in D'_{\eta\alpha}(\delta).$$

Moreover, for any $k > \alpha$ *and* $\lambda > 50(k - \alpha)/33k$, $\Psi : D'_{\eta\alpha}(\delta) \cap F_{k+\lambda} \to D_\alpha \cap E_k$. *In particular,* $\Psi : D'_{\eta\alpha}(\delta) \cap F_\infty \to D_\alpha \cap E_\infty$.

A proof of this result can be found in [17] and [19].[2]

(c)   In order to use the interpolation inequality (4.b.1) and the Nash–Moser–Schwartz–Sergeraert theorem on the scales defined in IV(a), smoothing operators must exist. First, we will construct smoothing operators for $\{L_k^2(\mathbf{R}^n)\}$. These operators are easily adapted to any compact manifold, (cf. [18]). However, our domains are manifolds with boundary. To define smoothing operators on these domains, we will use extension operators. The extension operators will also be useful in studying hyperbolic systems on manifolds with boundary.

---

[2]See also HAMILTON, *The Inverse Function Theorem of Nash–Moser*, Bull. Amer. Math. Soc. **7** (1982), 65–222.

The smoothing operators we construct for $\{L_k^2(\mathbf{R}^n)\}$ are exactly the same as those constructed for $\{C^k(T^n)\}$ in [17]. The proof of the inequalities S1–S4 are the same as well; a different Hölder inequality is used. The proofs are elementary and are omitted here.

Let $\hat{a} : \mathbf{R}^n \to \mathbf{R}$ be a compactly supported C^∞ function with $\hat{a} \equiv 1$ in a neighborhood of 0. Let $a : \mathbf{R}^n \to \mathbf{R}$ be the inverse Fourier transform of $\hat{a}$;

$$a(x) = \int_{\mathbf{R}^n} \hat{a}(\xi) e^{2\pi i x \cdot \xi} d\xi, \qquad d\xi = d\xi_1 \ldots d\xi_n \,.$$

$a(x)$ is a Schwartz function and satisfies

$$\int x^\alpha a(x) \, dx = \begin{cases} 1 & \text{if} \quad \alpha = 0 \\ 0 & \text{if} \quad \alpha \neq 0 \end{cases},$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$, $x^\alpha = x_1 \ldots x_n$, $dx = dx_1 \ldots dx_n$. Given $u \in L^2(\mathbf{R}^n)$ and $\theta \geq 1$, we define the smoothing operator $S_\theta : L^2(\mathbf{R}^n) \to L_\infty^2(\mathbf{R}^n)$ to be

$$(S_\theta u)(x) = \theta^n \int_{\mathbf{R}^n} a(\theta(x - y)) u(y) \, dy.$$

Remark. The set of operators $\{S_\theta, \theta > 1\}$ defined here are also smoothing operators on $L_k^2(T^n)$. To do this, we write a function on T^n as a periodic function on $\mathbf{R}^n$. It is easy to check that S_θ preserves whatever periodic properties a function has.

We now turn to extension operators. The existence of such operators is given by the following result of Stein, [21]:

STEIN EXTENSION THEOREM. *Let D be either a bounded convex domain or a domain with smooth boundary in $\mathbf{R}^n$. Then there exists an extension operator $\mathscr{E} : L^2(D) \to L^2(\mathbf{R}^n)$ satisfying:*
  (1) *for any $x \in D$ and $f \in C^\infty(D)$, $(\mathscr{E}f)(x) = f(x)$;*
  (2) *$\mathscr{E}$ is a bounded linear operator,*

$$\mathscr{E} : L_k^p(D) \to L_k^p(D), \qquad 1 \leq p \leq \infty, \quad 0 \leq k < \infty.$$

We will only need this statement for $p = 2$.

It is now clear that given $D = X$ or M, we can define a family of smoothing operators $\{S_\theta'\}$ simply as

$$S_\theta' u = (S_\theta \mathscr{E} u)\big|_D \,.$$

Therefore, we can use both lemma (4.b.1) and the Nash–Moser–Schwartz–Sergeraert theorem on the manifolds M and X.

(d) We now recall the results of §II but here using co-ordinates. This is necessary because we do not know how to prove estimates using moving frames. Although the equations will only be needed for embeddings of a three-dimensional manifold in $\mathbf{E}^6$, we will derive the results for an n-dimensional manifold in $\mathbf{E}^N$.

Suspending the definition given in §II(a), let M be any smooth n-dimensional manifold, possibly with boundary. We will, however, use the Sobolev spaces defined in §II(a).

We define the functional Φ to be the left hand side of (2.a.2),

$$\Phi(u) = (du, du) = \sum_{i,j=1}^{n} \left(\frac{\partial u}{\partial x^i}, \frac{\partial u}{\partial x^j} \right) dx^i dx^j \qquad \text{in local co-ordinates.} \quad (4.d.1)$$

This defines a map

$$\Phi : L_{k+1}^2(M, \mathsf{E}^N) \rightarrow L_k^2(S^2 T^* M), \qquad k \geqslant 0.$$

The Fréchet derivative Φ is given by

$$\Phi'(u) : L_{k+1}^2(M, \mathsf{E}^N) \rightarrow L_k^2(S^2 T^* M), \qquad u \in L_{k+1}^2(M, \mathsf{E}^N),$$

$$\Phi'(u)y = \sum_{i,j=1}^{n} \left[\left(\frac{\partial u}{\partial x^i}, \frac{\partial y}{\partial x^j} \right) + \left(\frac{\partial y}{\partial x^i}, \frac{\partial u}{\partial x^j} \right) \right] dx^i dx^j. \tag{4.d.2}$$

This corresponds to the linearized or variational equation (2.c.2).

Fix a smooth embedding $u_0 : M \rightarrow \mathsf{E}^N$. We want to apply the Nash–Moser–Schwartz–Sergeraert theorem to the functional Φ. The main requirement is to be able to solve the linearized equation (2.c.2) which we write as

$$\Phi'(u)y = h, \qquad u \text{ "near" } u_0, \tag{4.d.3}$$

with the estimate given by (4.b.5).

We may assume that u and h are smooth, and u is a general embedding of M in E^N. Given $y : M \rightarrow \mathsf{E}^N$, we can write

$$y = y' + y''$$

where y' is the component of y tangent to M and y'' the normal component with respect to the embedding u.

Let φ be as in §II(c), the 1-form dual to y' with respect to the metric $\Phi(u)$. In local co-ordinates,

$$\varphi = \varphi_i dx^i$$

where

$$\varphi_i = \left(\frac{\partial u}{\partial x^i}, y' \right) = \left(\frac{\partial u}{\partial x^i}, y \right).$$

We also split the second partials of u into their tangent and normal components, obtaining in local co-ordinates

$$\frac{\partial^2 u}{\partial x^i \partial x^j} = \Gamma_{ij}^m \frac{\partial u}{\partial x^m} + H_{ij}, \qquad 1 \leqslant i, j < n.$$

The Γ_{ij}^m are the Christoffel symbols of the Levi–Civita connection for $\Phi(u)$ with respect to the local co-ordinates (x^i). Set

$$H = H_{ij} dx^i dx^j \in C^\infty(S^2 T^* M).$$

H is the second fundamental form of the embedding u.

Then in local co-ordinates, (2.c.5) looks like:

$$\frac{1}{2}\left(\frac{\partial \varphi_i}{\partial x^j} + \frac{\partial \varphi_j}{\partial x^i}\right) - \Gamma_{ij}^m \varphi_m - (H_{ij}, y'') = h_{ij} \qquad 1 \leqslant i, j \leqslant n \qquad (4.d.4)$$

Clearly, solving for y is equivalent to finding φ and y'' satisfying (4.d.4). If we let $N \geqslant (n(n+1)/2) + n$, then by setting $\varphi = 0$, we get at each point $x \in M$, a determined or undetermined system of linear equations for the unknown $y''(x)$ which has $N - n$ components. This is easily solved with the estimates needed for the Nash–Moser theorem, leading to a local C^∞ isometric embedding theorem of any M^n in E^N, $N \geqslant (n(n+1)/2) + n$, (cf. [7]). However, if $N < (n(n+1)/2) + n$, the intrinsic piece with the "symmetric covariant differential" cannot be suppressed. A "decoupling procedure" was described in §II(c) and we review it here.

Recall that $II \subset S^2 T^* M$ is the subbundle defined by the second fundamental form. In local co-ordinates, for any $p \in M$,

$$II_p = \left\{ \left(\nu, \frac{\partial^2 u}{\partial x^i \partial x^j}(p)\right) dx^i dx^j \mid \nu \in N_p M \right\}$$

We also defined the "annihilator bundle" of II, $II^\perp \subset S^2 T^* M$. Each fiber is given by:

$$II_p^\perp = \left\{ B^{ij} \frac{\partial}{\partial x^i \partial x^j} \mid B^{ij} H_{ij}(p) = 0 \right\}$$

Now restricting to a contractible piece of M, we can choose a set of smooth sections

$$b_r = b_r^{ij} \frac{\partial^2}{\partial x^i \partial x^j} \in C^\infty(II^\perp)$$

such that at each point p, $\{b_r(p)\}$ span $II_p^\perp$. Applying each of these sections to the system (4.d.4), we obtain

$$b_r^{ij}\left(\frac{\partial \varphi_i}{\partial x^j} - \Gamma_{ij}^k \varphi_k\right) = b_r^{ij} h_{ij} \qquad (4.d.5)$$

We restate (2.c.9) in the context here:

(4.d.6) The solutions y to the linearized equation (4.d.3) are in one-to-one correspondence with the solutions $\varphi = \varphi_i dx^i$ to (4.d.5).

Thus, to solve the linearized isometric embedding system, we must solve a system of first order partial differential equations. As described in §II(c), the system is underdetermined if $N > n(n + 1)/2$, determined if $N = n(n + 1)/2$, and overdetermined if $N < n(n + 1)/2$. A study of the overdetermined case is done in [4]. We will investigate only the determined system, when $N = n(n + 1)/2$.

Given a system of PDE's, the first step to determining the existence of solutions is to look at the characteristic variety or equivalently the principal symbol of the differential operator. One hopes to find something that is elliptic or hyperbolic because these are the best understood PDE's. For such equations, existence theorems are classical and good estimates are not hard to find.

In our case, the characteristic variety of the operator defined by the left hand side of (4.d.5) is the one for the operator in (2.c.11). It is therefore the one studied in great detail in §I and §III. We find that for $n \geqslant 5$, the characteristic variety has singularities; in linear PDE theory such an operator is said to have "multiple characteristics." Very little is known about the existence of solutions to such equations. For $n = 3, 4$, the characteristic variety is generally nonsingular; such an operator is called "real principal type" and as stated in corollary (2.d.2), local solutions are known to exist. However, we are not yet able to prove the estimates needed for the Nash–Moser theorem.

Finally, when $n = 3$, §III describes in detail when the linearized equation is hyperbolic. In §V we will describe in detail the estimates which will lead to a local C^∞ isometric embedding theorem for the general M^3 in E^6.

V. Estimates for linear hyperbolic systems.

(a) We now define what a linear hyperbolic partial differential operator is; and state the theorems which give the existence of a right inverse satisfying the estimates necessary for the Nash–Moser–Schwartz–Sergeraert theorem. The remaining sections of §V contain a proof of Theorem G, when the operator is symmetric hyperbolic. The proof of Theorem G' for a strictly hyperbolic operator has been relegated to the appendix. (See *Corrections* A at end of paper.)

A first order determined hyperbolic system on M or X is of the form

$$Pu = f$$

where u and f are R^m-valued functions; and

$$P = A^n \frac{\partial}{\partial t} + A^i \frac{\partial}{\partial x^i} + B;$$

where A^i, A^n, B are smooth $m \times m$ matrix-valued functions satisfying one of the following:

(5.a.1) P is *symmetric hyperbolic* if A^i, A^n are symmetric and A^n is positive definite.

(5.a.2) P is *strictly hyperbolic* if A^n is nonsingular and for any $(x, t) \in M$ or X, $\xi \in \mathsf{R}^{n-1} \backslash \{0\}$, the polynomial $p(\tau) = \det(\tau A^n(x, t) + \xi_i A^i(x, t))$ has n distinct real roots.

Given a hyperbolic operator P on M, define the following quantities:

$$|P|_{L^2_k(M)} = \sum_{j=1} |A^j|_{L^2_k(M)} + |B|_{L^2_k(M)},$$

(5.a.3)

$$\lambda(P) = \inf_{\substack{(x,t) \in M \\ v \in \mathbb{R}^m \setminus \{0\}}} \frac{|v \cdot A^n(x,t)v|}{|v|^2}.$$

$\lambda(P)$ gives a lower bound for the magnitudes of the eigenvalues of $A^n(x,t)$.

THEOREM G. *Fix* $0 < a < b$ *and an integer* $\alpha > [n/2]$. *Let* P *be a symmetric hyperbolic operator on* M *satisfying*

$$|P|_{L^2_\alpha(M)} \leqslant b$$

(5.a.5)

$$\lambda(P) \geqslant a$$

(5.a.6)

Then P *has a right inverse*

$$Q : C^\infty(M, \mathbb{R}^m) \to C^\infty(M, \mathbb{R}^m)$$

which satisfies the following estimate:

$$|Qf|_{L^2_k(M)} \leqslant C_k(|P|_{L^2_k(M)}|f|_{L^2_1(M)} + |f|_{L^2_k(M)}), \qquad k \geqslant \alpha, \quad f \in C^\infty(M, \mathbb{R}^m).$$

(5.a.7)

where C_k *is a constant independent of* u, f, *and* P.

The corresponding statement for a strictly hyperbolic operator requires further assumptions.

Let $\mathscr{H} \subset (\mathbb{R}^m \otimes \mathbb{R}^m) \times \cdots \times (\mathbb{R}^m \otimes \mathbb{R}^m)$ denote the set of all $(n-1)$-tuples of $m \times m$ matrices, $(A^1, \ldots, A^{n-1})$, such that for any $\xi \in \mathbb{R}^{n-1} \setminus \{0\}$, $\sum_{j=1}^{n-1} A^j \xi_j$ has m distinct real eigenvalues. $\mathscr{H}$ is an open set in $\mathbb{R}^{(n-1)m^2}$.

THEOREM G'. *Fix a set* $\mathscr{A} \subset \mathscr{H}$, *diffeomorphic to the closed ball in* $\mathbb{R}^{(n-1)m^2}$, *a real number* $b > 0$, *and a positive integer* $\alpha \geqslant n + 3$.

Given any strictly hyperbolic operator $P = A^n(\partial/\partial t) + A^i(\partial/\partial x^i) + B$ *on* M *such that*

$$\left((A^n)^{-1}A^1, (A^n)^{-1}A^2, \ldots, (A^n)^{-1}A^{n-1}\right) \in C^\infty(M, \mathscr{A}),$$

(5.a.8)

$$|P|_{L^2_\alpha(M)} < b,$$

(5.a.9)

there exists a right inverse $Q : C^\infty(M, \mathbb{R}^m) \to C^\infty(M, \mathbb{R}^m)$ *satisfying*

$$|Qf|_{L^2_k(M)} \leqslant C_k(|P|_{L^2_{k+\alpha}(M)}|f|_{L^2_1(M)} + |f|_{L^2_k(M)}), \qquad k \geqslant \alpha, \quad f \in C^\infty(M, \mathbb{R}^m).$$

(5.a.10)

Here, C_k *is a constant depending on* $\mathscr{A}$, b, *and* α.

The basic idea involved in constructing the right inverse Q is to extend P to a hyperbolic operator on X, solve a fixed initial value problem there, and restrict

the solution back to M. The Sobolev estimates (5.a.8) and (5.a.9) are obtained by a technique used by S. Klainerman in [13]. The main ingredients are the standard energy integral estimates, Gronwall's inequality, and some elementary calculus lemmas. The proof for a symmetric hyperbolic system is fairly straightforward. However, the basic estimates for a strictly hyperbolic system are proved using pseudodifferential operators. A discussion of this plus the other modifications of the proof to Theorem G to make it work for a strictly hyperbolic operator are contained in the appendix §AV.

(b) Before embarking on our excursion into hyperbolic systems, we need some calculus lemmas which will be quite essential in proving (5.a.8).

In §IV(c), we observed that the scales $\{L_k^2(M)\}$, $\{L_k^2(X)\}$, and $\{L_k^2(T_t)\}$ all have families of smoothing operators. Therefore, on each scale, the interpolation inequality (4.b.1) holds. We will use this lemma to obtain estimates on the product of two functions and on the composition of two functions.

(5.b.1) LEMMA [13].   *For any $f, g \in C^\infty(\mathcal{D})$; $\mathcal{D} = M, X$, or T; the following estimates hold for any multi-index α, $|\alpha| = k \geqslant 2$:*

(a) $|D^\alpha(fg)|_{L^2(\mathcal{D})} \leqslant G_k\big(|f|_{L_k^2(\mathcal{D})}|g|_{L^2(\mathcal{D})} + |f|_{L^2(\mathcal{D})}|g|_{L_k^2(\mathcal{D})}\big)$

(b) $|D^\alpha(fg) - fD^\alpha g|_{L^2(\mathcal{D})} \leqslant G_k\big(|f|_{L_k^2(\mathcal{D})}|g|_{L^2(\mathcal{D})} + |f|_{L_1^2(\mathcal{D})}|g|_{L_{k-1}^2(\mathcal{D})}\big)$

G_k *is a constant depending only on $\mathcal{D}$ and k.* (See *Corrections* B at end of paper.)

Proof.   We use the following elementary fact: given any $a, b \geqslant 0$ and integers $k \geqslant i \geqslant 0$,

$$a^{i/k}b^{(k-i)/k} \leqslant a + b.$$

Proof of (a):

$|D^\alpha(fg)|_{L_k^2(\mathcal{D})} \leqslant G_k'' \sum_{l=0}^{k} |f|_{L_l^2(\mathcal{D})}|g|_{L_{k-l}^2(\mathcal{D})}$ by Leibniz's rule

$\leqslant G_k' \sum_{l=0}^{k} |f|_{L^2(\mathcal{D})}^{(k-l)/k}|f|_{L_k^2(\mathcal{D})}^{l/k}|g|_{L^2(\mathcal{D})}^{l/k}|g|_{L_k^2(\mathcal{D})}^{(k-l)/k}$ by (4.b.1)

$\leqslant G_k\big(|f|_{L^2(\mathcal{D})}|g|_{L_k^2(\mathcal{D})} + |f|_{L_k^2(\mathcal{D})}|g|_{L^2(\mathcal{D})}\big).$

Proof of (b):

$|D^\alpha(fg) - fD^\alpha g|_{L^2(\mathcal{D})}$

$\leqslant G_k'' \sum_{l=1}^{k} |f|_{L_l^2(\mathcal{D})}|g|_{L_{k-l}^2(\mathcal{D})}$

$\leqslant G_k' \sum_{l=1}^{k} |f|_{L_1^2(\mathcal{D})}^{(k-l)/(k-1)}|f|_{L_k^2(\mathcal{D})}^{(l-1)/(k-1)}|g|_{L^2(\mathcal{D})}^{(k-l)/(k-1)}|g|_{L_{k-1}^2(\mathcal{D})}^{(l-1)/(k-1)}$

$\leqslant G_k\big(|f|_{L_1^2(\mathcal{D})}|g|_{L_{k-1}^2(\mathcal{D})} + |f|_{L_k^2(\mathcal{D})}|g|_{L^2(\mathcal{D})}\big).$ Q.E.D.

Remark. The lemma holds also for f and g vector-valued; f matrix-valued and g vector-valued; or f and g both matrix-valued.

(5.b.2) LEMMA. *Let $\mathscr{D} = M$, X, or T_t. Let $\varphi \in C^{\infty}(B)$, B a compact domain in R^N. Then for any $u \in C^{\infty}(\mathscr{D}, B)$, we have the following:*

$$\left| \varphi \circ u \right|_{L^2_k(\mathscr{D})} \leqslant M_k |u|_{L^2_k(\mathscr{D})}, \qquad k \geqslant 1$$

(See *Corrections* C at end of paper.)

Proof. Let α be any multi-index, $|\alpha| = k$. Using the chain rule, we have:

$$D^{\alpha}(\varphi \circ u) = \sum_{\substack{e_{\mu_1} + \cdots + e_{\mu_l} = \beta \\ l \leqslant k \\ \alpha_1 + \cdots \alpha_l = \alpha}} \left[\left(D_y^{\beta} \varphi \right) \circ u \right] (D_x^{\alpha_1} u^{\mu_1}) \cdots (D_x^{\alpha_l} u^{\mu_l})$$

where $u = (u^1, u^2, \dots)$ and $\{e_{\mu}\}$ is the standard basis for R^n. Therefore,

$$\left| D^{\alpha}(\varphi \circ u) \right|_{L^2(\mathscr{D})}$$

$$\leqslant M_k'' \left(\sup_{y \in B} \sum_{|\beta| \leqslant k} |D_y^{\beta} \varphi(y)| \right) \sum_{\substack{1 \leqslant l \leqslant k \\ k_1 + \cdots + k_l = k}} |u|_{L^2_{k_1}(\mathscr{D})} \cdots |u|_{L^2_{k_l}(\mathscr{D})}.$$

Apply the interpolation inequality (4.b.1) to each factor in the terms of the summation. We obtain

$$\left| D^{\alpha}(\varphi \circ u) \right|_{L^2(\mathscr{D})} \leqslant M_k' \left(\sum_{l=0}^{k} |u|_{L^2(\mathscr{D})}^l \right) |u|_{L^2_k(\mathscr{D})}.$$

However since u is a map into a bounded domain,

$$|u|_{L^2(\mathscr{D})} \leqslant (\mathrm{vol}(\mathscr{D})) \left(\sup_{y \in B} |y| \right).$$

Therefore, if we sum the estimate over all $|\alpha| \leqslant k$, the lemma is proved. Q.E.D.

(c)   The Cauchy problem for a linear hyperbolic system is known to be well posed. Proofs for both the symmetric hyperbolic and the strictly hyperbolic cases may be found in [23]. The result we shall need is:

(5.c.1) PROPOSITION. *Let P be a hyperbolic operator on $X \cong T^{n-1} \times [0, 1]$; i.e., it satisfies either (5.a.1) or (5.a.2). Then given smooth functions $\phi : T^n \to \mathsf{R}^m$, $f : X \to \mathsf{R}^m$, there is a unique solution $u : X \to \mathsf{R}^m$ to the system:*

$$Pu = f$$

$$u(x, 0) = \phi(x)$$

Moreover, u is smooth.

We will not prove the existence of the solution u. However, the regularity and uniqueness of u will be implied by the estimates we prove for symmetric hyperbolic systems in this section and for strictly hyperbolic systems in the appendix.

For the rest of this section, we will consider only symmetric hyperbolic systems, and only those of the form:

$$\frac{\partial u}{\partial t} + A^i \frac{\partial u}{\partial x^i} + Bu = f. \tag{5.c.2}$$

The A^i, B, and f will always be smooth functions on X.

The first estimate is the standard energy integral inequality whose proof requires the A^i's to be symmetric.

(5.c.3) PROPOSITION. *Let u be a solution to (5.c.2) with any L^2 initial value. Then for any $t \in [0, 1]$, the following estimate holds*:

$$\frac{d}{dt} |u|_{L^2(T_t)} \leqslant c(t)|u|_{L^2(T_t)} + |f|_{L^2(T_t)}$$

where $c(t) = \sum_{i=1}^{n-1} |A^i|_{L^2_{1,0}(T_t)} + |B|_{L^2(T_t)}$.

Applying Gronwall's inequality (5.c.5) to this estimate gives an *a priori* estimate for $|u|_{L^2(T_t)}$ in terms of $|f|_{L^2(T_t)}$ and the coefficients of P. An immediate consequence is the uniqueness of u. However, since what we really want are Sobolev estimates, we will defer the use of lemma (5.c.5).

Proof of (5.c.3):

$$\frac{1}{2} \frac{\partial}{\partial t} (u(x,t), u(x,t)) = \left(u, \frac{\partial u}{\partial t} \right) = \left(u, -A^i \frac{\partial u}{\partial x^i} - Bu + f \right)$$

$$= -\frac{1}{2} \frac{\partial}{\partial x^i} (u, A^i u) + \left(u, \left[\frac{1}{2} \sum_{i=1}^{n-1} \frac{\partial A^i}{\partial x^i} - B \right] u \right) + (u, f).$$

Integrate each side over T_t to get:

$$|u|_{L^2(T_t)} \frac{d}{dt} |u|_{L^2(T_t)} = \int_{T_t} \left(u, \left[\frac{1}{2} \sum_{i=1}^{n-1} \frac{\partial A^i}{\partial x^i} - B \right] u \right) dx + \int_{T_t} (u, f) \, dx$$

$$\leqslant c(t)|u|^2_{L^2(T_t)} + |u|_{L^2(T_t)} |f|_{L^2(T_t)} \quad \text{Q.E.D.}$$

Next, we extend this estimate to Sobolev norms involving only derivatives in the x^i's.

(5.c.4) PROPOSITION. *Let u solve (5.c.2) with any initial data. Then*

$$\frac{d}{dt} |u|_{L^2_{k,0}(T_t)} \leqslant c_k(t)|u|_{L^2_{k,0}(T_t)} + \tilde{c}_k(t)|u|_{L^2_{1,0}(T_t)} + |f|_{L^2_{k,0}(T_t)}, \qquad k \geqslant 0, \quad 0 \leqslant t \leqslant 1;$$

where

$$c_k(t) = a_k \left(\sum_{i=1}^{n-1} |A^i|_{L^2_{1,0}(T_t)} + |B|_{L^2(T_t)} \right),$$

$$\tilde{c}_k(t) = \tilde{a}_k \left(\sum_{i=1}^{n-1} |A^i|_{L^2_{k,0}(T_t)} + |B|_{L^2_{k,0}(T_t)} \right),$$

and $a_k, \tilde{a}_k$ are constants independent of A^i, B, f, and u.

Proof. For $k = 0$, we have (5.c.3).

For $k = 1$, differentiate (5.c.2) with respect to x^j, $j = 1, \ldots, n - 1$. Use the resulting set of $(n - 1)m$ equations and the original m equations to construct an nm by nm symmetric hyperbolic system for which the vector $(u, (\partial u/\partial x^1), \ldots, (\partial u/\partial x^{n-1}))$ is a solution. Applying Proposition (5.c.3) to this new system, we get the estimate we want.

For $k \geqslant 2$, we will use the calculus lemmas proved in §V(b). Fix a multi-index $\alpha \in \mathbf{Z}^{n-1}$. Let $u^{(\alpha)} = D_x^\alpha u$. Apply D_x^α to (5.c.2) to get:

$$\frac{\partial u^{(\alpha)}}{\partial t} + A^i \frac{\partial u^{(\alpha)}}{\partial x^i} = D^\alpha f - D_x^\alpha \left(A^i \frac{\partial u}{\partial x^i} \right) + A^i D_x^\alpha \frac{\partial u}{\partial x^i} - D_x^\alpha (Bu)$$

This is a symmetric hyperbolic system; we can apply Proposition (5.c.3) and the triangle inequality, obtaining

$$\frac{d}{dt} |u^{(\alpha)}|_{L^2(T_t)} < c(t) |u^{(\alpha)}|_{L^2(T_t)} + |D^\alpha f|_{L^2(T_t)}$$

$$+ \sum_{i=1}^{n-1} \left| D_x^\alpha \left(A^i \frac{\partial u}{\partial x^i} \right) - D_x^\alpha A^i \frac{\partial u}{\partial x^i} \right|_{L^2(T_t)} + |D_x^\alpha (Bu)|_{L^2(T_t)} .$$

We now use lemma (5.b.1) to estimate the last two terms on the right hand side, and sum both sides over all $|\alpha| \leqslant k$, proving the proposition.   Q.E.D.

(5.c.5) LEMMA (Gronwall's inequality).   *Let $\varphi(t), K(t), h(t)$ be smooth positive functions satisfying*

$$\frac{d\varphi}{dt} \leqslant K\varphi + h.$$

Then $\varphi(t) \leqslant e^{C(t)} [\varphi(0) + \int_0^t e^{-C(\tau)} h(\tau) d\tau]$, where $C(t) = \int_0^t K(\tau) d\tau$.

(5.c.5) is proved by solving the corresponding differential equation explicitly. Using this lemma, we can prove the following:

(5.c.6) PROPOSITION.   *Fix $R > 0$. Let $P = \partial/\partial t + A^i(\partial/\partial x^i) + B$ be a symmetric hyperbolic operator with smooth coefficients on X, satisfying*

$$\sum_{i=1}^{n-1} |A^i|_{L^2_{1,0}(X)} + |B|_{L^2_{1,0}(X)} < R$$

Let u solve

$$Pu = f, \qquad f \in C^\infty(X, \mathbf{R}^m).$$

Then for any $k \geqslant 1$, the following estimate holds:

$$|u|_{L^2_{k,0}(T_t)} \leqslant C_{k,0}\left[|u|_{L^2_{k,0}(T_0)} + |f|_{L^2_{k,0}(X)} \right.$$

$$\left. + \left(\sum_{i=1}^n |A^i|_{L^2_{k,0}(X)} + |B|_{L^2_{k,0}(X)} \right)\left(|u|_{L^2_{1,0}(T_0)} + |f|_{L^2_{1,0}(X)} \right) \right]$$

where $C_{k,0}$ depends only on R.

Proof. We first prove for $k = 1$,

$$|u|_{L^2_{1,0}(T_t)} \leqslant C_{1,0}\left[|u|_{L^2_{1,0}(T_0)} + |f|_{L^2_{1,0}(X)} \right].$$

Apply (5.c.5) to (5.c.4) when $k = 1$, obtaining

$$|u|_{L^2_{1,0}(T_t)} \leqslant e^{C_1(t)}\left[|u|_{L^2_{1,0}(T_0)} + \int_0^t e^{-C_1(\tau)} |f|_{L^2_{1,0}(T_\tau)}\, d\tau \right]$$

where

$$C_1(t) = \int_0^t c_1(\tau) + \tilde{c}_1(\tau)\, d\tau \leqslant \int_0^1 c_1(\tau) + \tilde{c}_1(\tau)\, d\tau.$$

One can easily show that:

$$\int_0^1 |g|_{L^2_{k,0}(T_\tau)}\, d\tau \leqslant |g|_{L^2_{k,0}(X)}.$$

Therefore

$$C_1(t) \leqslant (a_1 + \tilde{a}_1) \int_0^1 \sum |A^i|_{L^2_{1,0}(T_\tau)} + |B|_{L^2(T_\tau)}\, d\tau < (a_1 + \tilde{a}_1)R.$$

Similarly

$$\int_0^t e^{-C_1(\tau)} |f|_{L^2_{1,0}(T_\tau)}\, d\tau \leqslant \int_0^1 |f|_{L^2_{1,0}(T_\tau)}\, d\tau \leqslant |f|_{L^2_{1,0}(X)}.$$

We may set $C_{1,0} = e^{(a_1 + \tilde{a}_1)R}$, proving the lemma when $k = 1$. For $k \geqslant 2$, the proof proceeds in a similar fashion. Apply Gronwall's inequality to (5.c.4), obtaining

$$|u|_{L^2_{k,0}(T_t)} \leqslant e^{C_k(t)}\left[|u|_{L^2_{k,0}(T_0)} + \int_0^t e^{-c_k(\tau)}\left(|f|_{L^2_{k,0}(T_\tau)} + \tilde{c}_k(\tau)|u|_{L^2_{1,0}(T_\tau)} \right) d\tau \right]$$

where

$$C_k(t) = \int_0^t c_k(\tau)\, d\tau \leqslant a_k R,$$

$$\leqslant C_k' \left[|u|_{L_{k,0}^2(T_0)} + \int_0^1 |f|_{L_{k,0}^2(T_\tau)}\, d\tau + \int_0^1 \tilde{c}_k(\tau) |u|_{L_{1,0}^2(T_\tau)}\, d\tau \right]$$

$$\leqslant C_k' \left[|u|_{L_{k,0}^2(T_0)} + |f|_{L_{k,0}^2(X)} + \left(\int_0^1 |\tilde{c}_k(\tau)|^2\, d\tau \right)^{1/2} \left(\int |u|_{L_{1,0}^2(T_\tau)}^2\, d\tau \right)^{1/2} \right]$$

$$= C_k' \left[|u|_{L_{k,0}^2(T_0)}^2 + |f|_{L_{k,0}^2(X)} + \tilde{a}_k \left(\sum_{i=1}^{n-1} |A^i|_{L_{k,0}^2(X)} + |B|_{L_{k,0}^2(X)} \right) |u|_{L_{1,0}^2(X)} \right].$$

The lemma is then proved by putting in the estimate for $|u|_{L_{1,0}^2(X)}$ and absorbing $\tilde{a}_k C_{1,0}$ into C_k' to define a new constant $C_{k,0}$.   Q.E.D.

We now want to derive Sobolev estimates which include $\partial/\partial t$'s. However, we also want to suppress the contribution to the estimates from the initial data. Therefore, we will assume zero initial data, and take f to be identically zero in a neighborhood of the initial surface T_0. We get the following result:

(5.c.7) PROPOSITION. *Fix* $R > 0$. *Let* $P = \partial/\partial t + A^i(\partial/\partial x^i) + B$ *be a symmetric hyperbolic operator which satisfies*

$$\sum_{i=1}^{n-1} |A^i|_{L_1^2(X)} + |B|_{L_1^2(X)} < R.$$

Let $f \in C^\infty(X, \mathbf{R}^m)$ *satisfy the following: There exists* $\epsilon > 0$ *such that*

$$f|_{T^{n-1} \times [0,\epsilon]} = 0.$$

Let u *solve the initial value problem*:

$$Pu = f, \qquad u|_{T_0} = 0.$$

Then the following holds:

$$|u|_{L_k^2(X)} \leqslant \hat{C}_k \left[|f|_{L_k^2(X)} + \left(\sum_{i=1}^{n-1} |A^i|_{L_k^2(X)} + |B|_{L_k^2(X)} \right) |f|_{L_1^2(X)} \right].$$

Proof. We prove the following estimates by induction: For any $k \geqslant l \geqslant 0$,

$$|u|_{L_{k,l}^2(X)} \leqslant \hat{C}_k \left[|f|_{L_k^2(X)} + \left(\sum_{i=1}^{n-1} |A^i|_{L_k^2(X)} + |B|_{L_k^2(X)} \right) |f|_{L_1^2(X)} \right]. \qquad (5.\text{c}.8)$$

This will prove the proposition.

To prove $(5.c.8)_{k,0}$, square both sides of the estimate in (5.c.6), integrate over t, and take the square root. The right hand side doesn't even depend on t and doesn't change.

Now fix integers $0 < l \leqslant k$. Assume that $(5.c.8)_{r,s}$ has been proved for the following pairs of integers (r,s):

(1) $s \leqslant r < k$

(2) $s < l, r = k$.

Let $\alpha \in \mathbf{Z}^{n-1}$ be any multi-index, $|\alpha| = k - l$.

$$\left| \frac{\partial^l}{\partial t^l} D_x^\alpha u \right|_{L^2(X)} = \left| D_u^\alpha \frac{\partial^{l-1}}{\partial t^{l-1}} \left(f - A^i \frac{\partial u}{\partial x^i} - Bu \right) \right|_{L^2(X)}$$

$$\leqslant |f|_{L^2_{k-1,l-1}(X)} + \left| D_x^\alpha \frac{\partial^{l-1}}{\partial t^{l-1}} \left(A^i \frac{\partial u}{\partial x^i} \right) \right|_{L^2(X)} + \left| D_x^\alpha \frac{\partial^{l-1}}{\partial t^{l-1}} (Bu) \right|_{L^2(X)}.$$

We estimate each of the last two terms:

$$\left| D_x^\alpha \frac{\partial^{l-1}}{\partial t^{l-1}} (Bu) \right|_{L^2(X)} \leqslant G_k \big[|B|_{L^2_{k-1}(X)} |u|_{L^2(X)} + |B|_{L^2(X)} |u|_{L^2_{k-1}(X)} \big],$$

by Lemma (5.b.1a).

$$\left| D_x^\alpha \frac{\partial^{l-1}}{\partial t^{l-1}} \left(A^i \frac{\partial u}{\partial x^i} \right) \right|_{L^2(X)}$$

$$\leqslant \left| D_x^\alpha \frac{\partial^{l-1}}{\partial t^{l-1}} \left(A^i \frac{\partial u}{\partial x^i} \right) - A^i D_x^\alpha \frac{\partial^l u}{\partial x^i \partial t^{l-1}} \right|_{L^2(X)} + \left| A^i D_x^\alpha \frac{\partial^l u}{\partial x^i \partial t^{l-1}} \right|_{L^2(X)}$$

$$\leqslant G_{k-1} \left[\left(\sum_{i=1}^{n-1} |A^i|_{L^2_{k-1}(X)} \right) |u|_{L^2_1(X)} + \left(\sum_{i=1}^{n-1} |A^i|_{L^2_1(X)} \right) |u|_{L^2_{k-1}(X)} \right]$$

$$+ \left(\sum |A^i|_{L^2(X)} \right) |u|_{L^2_{k,l-1}(X)}.$$

Here, we have used (5.b.1b) to estimate the first term and the Cauchy–Schwarz inequality for the second.

We can now substitute in the assumed estimates for $|u|_{L^2_{k-1}(X)}$ and $|u|_{L^2_{k,l-1}(X)}$. We also need the following:

$$|u|^2_{L^2_1(X)} = |u|^2_{L^2_{1,0}(X)} + \left| \frac{\partial u}{\partial t} \right|^2_{L^2(X)} = |u|^2_{L^2_{1,0}(X)} + \left| f - A^i \frac{\partial u}{\partial x^i} - Bu \right|^2_{L^2(X)}$$

$$\leqslant C \left\{ \left[\left(1 + \sum_{i=1}^{n-1} |A^i|_{L^2(X)} + |B|_{L^2(X)} \right) |u|_{L^2_{1,0}(X)} \right]^2 + |f|^2_{L^2(X)} \right\}.$$

Now substitute in the estimate for $|u|_{L^2_{1,0}(X)}$. The resulting estimate can in turn be

put into the estimate for $|D^{\alpha}(\partial^{l}/\partial t^{l})u|_{L^{2}(X)}$. Summing these estimates over $|\alpha| = k - l$ proves $(5.c.8)_{k,l}$.   Q.E.D.

(d)   We give a proof of Theorem G. Given a symmetric hyperbolic operator

$$P = A^{n}\frac{\partial}{\partial t} + A^{i}\frac{\partial}{\partial x^{i}} + B$$

on M and $f \in C^{\infty}(M, \mathbf{R}^{m})$, we can solve the equation

$$Pu = f$$

by first transforming P into an operator of the form

$$\tilde{P} = \frac{\partial}{\partial t} + \tilde{A}^{i}\frac{\partial}{\partial x^{i}} + \tilde{B}, \qquad \tilde{A}^{i} \text{ symmetric,}$$

extending $\tilde{P}$ to a symmetric hyperbolic operator $\hat{P}$ on X, solving an initial value problem on X, and restricting the solution back to M. This procedure defines a right inverse Q to P. The estimate (5.a.7) is then proved using Proposition (5.c.7), Lemma (5.b.1), and Lemma (5.b.2).

Let $\mathscr{E}' : L^{2}(M) \to L^{2}(\mathbf{R}^{n})$ be the extension operator given by the Stein extension theorem (see §IV(c)). We can then define an extension operator $\mathscr{E} : L^{2}(M) \to L^{2}(X)$ as

$$\mathscr{E}u = (\varphi \mathscr{E}'u)|_{[0,1]^{n}}, \qquad u \in L^{2}(M),$$

where $\varphi \in C^{\infty}(\mathbf{R}^{n})$ satisfies the following:
   (a) support of $\varphi \subset (0, 1)^{n}$
   (b) $\varphi \equiv 1$ on M.
The function $\varphi \mathscr{E}'u$ may be viewed as a function on X since it is identically zero on a neighborhood of the boundary of $[0, 1]^{n}$.

Given a smooth matrix-valued function A^{n} which is always positive definite symmetric, the Gram–Schmidt procedure can be used to define a smooth matrix-valued functio $S : M \to \mathbf{R}^{m^{2}}$ satisfying:

$${}^{t}S(x, t)A^{n}(x, t)S(x, t) = I.$$

Let

$$\tilde{P} = {}^{t}SPS = \frac{\partial}{\partial t} + \tilde{A}^{i}\frac{\partial}{\partial x^{i}} + \tilde{B},$$

where $\tilde{A}^{i} = {}^{t}SA^{i}S$, $\tilde{B} = {}^{t}SA^{n}(\partial S/\partial t) + {}^{t}SA^{i}(\partial S/\partial x^{i}) + {}^{t}SBS$. We then extend $\tilde{P}$ to an operator on X,

$$\hat{P} = \frac{\partial}{\partial t} + (\mathscr{E}\hat{A}^{i})\frac{\partial}{\partial x^{i}} + (\mathscr{E}B).$$

Since P is symmetric hyperbolic on X, a right inverse Q may be defined as follows: Given $f \in C^{\infty}(X, \mathbf{R}^m)$, set Qf equal to the solution u of the initial value problem:

$$\hat{P}\hat{u} = \hat{f}$$

$$\hat{u}|_{T_0} = 0.$$

The right inverse Q to P is then defined to be

$$Q = S\mathscr{R}\hat{Q}\mathscr{E}^t S,$$

where $\mathscr{R}: L^2(X) \to L^2(M)$ is the restriction operator, since

$$PQ = PS\mathscr{R}\hat{Q}\mathscr{E}^t S = {}^t S^{-1}\tilde{P}\mathscr{R}\hat{Q}\mathscr{E}^t S = {}^t S^{-1}R\hat{P}\hat{Q}\mathscr{E}^t S = I.$$

We begin the proof of the estimate (5.a.7) with a lemma.

(5.d.1) LEMMA. *Fix $b > a > 0$. Let $A \in C^{\infty}(M, S^2(\mathbf{R}^m))$ satisfy the following conditions (corresponding to (5.a.5) and (5.a.6))*:
  (1) $|A(x,t)| \leqslant b$, $(x,t) \in M$.
  (2) $v \cdot A(x,t)v \geqslant a|v|^2 v \in \mathbf{R}^m$, $(x,t) \in M$.
Then there exists a smooth function $S(x,t)$ such that

$$^t SAS = I,$$

and such that for any $f \in C^{\infty}(M, \mathbf{R}^m)$ and $k \geqslant 0$, the following estimate holds:

$$|Sf|_{L_k^2(M)} \leqslant N_k(|A|_{L_k^2(M)}|f|_{L^2(M)} + |f|_{L_k^2(M)}).$$

N_k *depends only on a and b.*

Proof. Let $\mathscr{P}_{a,b} = \{A \in S^2(\mathbf{R}^m) \mid a \leqslant (v, Av) \leqslant b\}$. $\mathscr{P}_{a,b}$ is clearly a compact domain in $S^2(\mathbf{R}^m) = \{$symmetric $m \times m$ matrices$\}$. The Gram–Schmidt process defines a smooth function $\psi: \mathscr{P}_{a,b} \to GL(m, \mathbf{R})$ such that

$$^t\psi(A)A\psi(A) = I, \qquad A \in \mathscr{P}_{a,b}.$$

We then define

$$S(x,t) = (\psi \circ A)(x,t).$$

Linear algebra shows that

$$|S(x,t)| \leqslant \frac{1}{\sqrt{a}}, \quad \text{implying that} \quad |S|_{L^2(M)} \leqslant \frac{1}{\sqrt{a}}.$$

On the other hand, Lemma (5.b.2) gives

$$|S|_{L_k^2(M)} \leqslant G_k|A|_{L_k^2(M)}.$$

Therefore, using Lemma (5.b.1a) and these bounds for the norms of S, we get

$$|Sf|_{L_k^2(M)} \leqslant G_k\left[|S|_{L_k^2(M)}|f|_{L^2(M)} + |S|_{L^2(M)}|f|_{L_k^2(M)}\right]$$

$$\leqslant G_k\left[M_k|A|_{L_k^2(M)}|f|_{L^2(M)} + \frac{1}{\sqrt{a}}|f|_{L_k^2(M)}\right]$$

The proof is completed by setting $N_k = \max(G_k M_k, G_k(1/\sqrt{a}))$.   Q.E.D.

Proof of (5.a.7).   Observe that $|\mathscr{E}u|_{L_k^2(X)} \leqslant E_k|u|_{L_k^2(X)}$ and $|\mathscr{R}\hat{u}|_{L_k^2(M)} \leqslant |\hat{u}|_{L_k^2(X)}$. Recall that

$$P = S^n \frac{\partial}{\partial t} + A^i \frac{\partial}{\partial x^i} + B$$

is assumed to satisfy (5.a.5) and (5.a.6). In particular, (5.a.5) implies that

$$|A^n|_{L_\alpha^2(M)} \leqslant b.$$

Since $\alpha > [n/2]$, the Sobolev lemma implies that there is a constant b' such that

$$|A^n(x,t)| \leqslant b', \qquad (x,t) \in M.$$

Therefore, we may use Lemma (5.d.1) to obtain the following estimate.

$$|\tilde{A}^i|_{L_k^2(M)} = |{}^tSA^iS|_{L_k^2(M)} \leqslant N_k\left[|A^n|_{L_k^2(M)}|A^iS|_{L^2(M)} + |A^iS|_{L_k^2(M)}\right]$$

$$\leqslant N_k\left[\frac{b}{\sqrt{a}}|A^n|_{L_k^2(M)} + N_k\left(|A^n|_{L_k^2(M)}|A^i|_{L^2(M)} + |A^i|_{L_k^2(M)}\right)\right]$$

$$\leqslant N_k'|P|_{L_k^2(M)} \qquad \text{(recall definition (5.a.3)).}$$

A similar argument shows that

$$|\tilde{B}|_{L_k^2(M)} \leqslant N_k''|P|_{L_{k+1}^2(M)} .$$

Then given $f \in C^\infty(M, \mathbf{R}^m)$,

$$|Qf|_{L_k^2(M)} = |S\mathscr{R}\hat{Q}\mathscr{E}^tSf|_{L_k^2(M)}$$

$$\leqslant N_k\left[|A^n|_{L_k^2(M)}|\hat{Q}\mathscr{E}^tSf|_{L^2(M)} + |\hat{Q}\mathscr{E}^tSf|_{L_k^2(M)}\right] \qquad \text{by (5.d.1).}$$

Using (5.a.5) and the bounds for $\tilde{A}^i, \tilde{B}$ we just derived, we see that $R > 0$ can be chosen, depending only on a and b, so that $\hat{P}$ satisfies the assumptions of Proposition (5.c.7). Observe as well that by the definition of $\mathscr{E}$, $\mathscr{E}^tSf$ meets the

condition given in (5.c.7). Therefore, we have the following:

$$|\hat{Q}\mathscr{E}'Sf|_{L^2(M)} \leqslant C_0 E_1 |'Sf|_{L_1^2(M)} \leqslant C_0' |f|_{L_1^2(M)} .$$

$$|\hat{Q}\mathscr{E}'Sf|_{L^2(M)} \leqslant \hat{C}_k \Bigg[E_k \big(|A^n|_{L_k^2(M)} |f|_{L^2(M)} + |f|_{L_k^2(M)} \big)$$

$$+ \bigg(\sum_{i=1}^{n-1} |\mathscr{E}\tilde{A}^i|_{L_k^2(M)} + |\mathscr{E}\tilde{B}|_{L_k^2(M)} \bigg) |\mathscr{E}'Sf|_{L_k^2(M)} \Bigg].$$

We get the estimate by substituting in the estimates for $|\tilde{A}^i|_{L_k^2(M)}, |\tilde{B}|_{L_k^2(M)},$ $|'Sf|_{L_1^2(M)},$ and observing that $|A|_{L_1^2(M)} \leqslant |P|_{L_1^2(M)} \leqslant |P|_{L_\alpha^2(M)} < b.$ Q.E.D.

A.V. Appendix: Estimates for linear strictly hyperbolic systems.

(a) This appendix contains the propositions necessary to prove theorem G' and their proofs.

An L^2-inequality similar to the one in Proposition (5.c.3) is proved in §AV(c) using a pseudo-differential operator to "symmetrize" the operator P. In using such machinery, more derivatives of the coefficients A^i, B will be required than in the symmetric hyperbolic case. The argument used to extend the basic inequality to a Sobolev inequality is essentially the same as before.

The new wrinkle here is the use of pseudo-differential operators on T^{n-1}. A precise representation of the composition of two pseudo-differential operators and of the adjoint of a pseudo-differential operator is required. Also, we need to estimate the norm of an operator in terms of its symbol. §AV(b) contains a brief exposition of pseudo-differential operators on T^n (we use T^n instead of T^{n-1} for convenience). The operators are defined globally using Fourier series rather than Fourier transform. This avoids the use of local co-ordinate charts, and a partition of unity which are necessary in the standard construction of pseudo-differential operators on a compact manifold. Moreover, the statements and formulas obtained are simpler than even those for operators on $\mathbb{R}^n$ since there one has to worry about the operators being properly supported. However, the results in §AV(b) do hold in some sense for any compact manifold. Although the corresponding statements are not invariant, they can be made with respect to some fixed choice of co-ordinate charts and partition of unity. Using this, Theorem G can be proved for a strictly hyperbolic operator on $I \times S$, S a compact manifold without boundary.

The last piece needed to complete the proof of Theorem G' is showing how to extend a strictly hyperbolic operator on M to one on X. This we do in §AV(d); the rest of the discussion in §V(d) carries over without change to complete the proof.

(b) We will only use matrix-valued pseuo-differential operators; there is essentially no difference in the following propositions and the corresponding ones for scalar-valued pseudodifferential operators. Let V denote the space of

complex $m \times m$ matrices. We will use the standard matrix norm on V, i.e., as a linear operator on $\mathbb{C}^m$. We denote $D_\xi = (\partial/\partial\xi_1, \ldots, \partial/\partial\xi_n)$ and $D_x = (\partial/\partial x^1, \ldots, \partial/\partial x^n)$; α, β are multi-indices.

(A.5.b.1) *Definition.* Denote by $\mathscr{S}^k(T^n, V)$, $k \in \mathbb{R}$, the space of smooth functions $a: T^n \times \mathbb{R}^n \to V$ satisfying the following property:

For any multi-indices α, β, there is a constant $C_{\alpha,\beta} > 0$ such that

$$|D_x^\alpha D_\xi^\beta a(x,\xi)| \leqslant C_{\alpha,\beta}(1 + |\xi|)^{k-|\beta|} \qquad \text{(A.5.b.2)}$$

An element $a(x,\xi) \in \mathscr{S}^k(T^n, V)$ is called a *V-valued symbol of order k.*

We can take the Fourier transform of $a(x,\xi) \in \mathscr{S}^k(T^n, V)$ as a function of x. Since it is a function on the torus, we obtain a Fourier series as follows:

$$\hat{a}(\eta,\xi) = \int_{T^n} e^{-2\pi i x \cdot \eta} a(x,\xi)\, dx, \qquad \eta \in \mathbb{Z}^n, \quad \xi \in \mathbb{R}^n;$$

$$a(x,\xi) = \sum_{\eta \in \mathbb{Z}^n} e^{2\pi i x \cdot \eta} a(\eta,\xi), \qquad x \in T^n, \quad \xi \in \mathbb{R}^n.$$

Here, $T^n = \mathbb{R}^n/\mathbb{Z}^n$.

Since a is smooth, $\hat{a}$ is rapidly decreasing in the first variable, i.e., it satisfies the following estimates:

$$|D_\xi^\alpha \hat{a}(\eta,\xi)| \leqslant \hat{C}_{\alpha,l}(1 + |\eta|)^{-l}(1 + |\xi|)^{k-|\alpha|}, \qquad \text{(A.5.b.3)}$$

where l is any nonnegative integer, α is a multi-index.

A collection of seminorms is then defined on $\mathscr{S}^k(T^n, V)$ by setting $|a|_{m,l}$ equal to the infinum of all $\hat{C}_{\alpha,l}$ satisfying (A.5.b.3) for $|\alpha| \leqslant m$.

Given $a(x,\xi) \in \mathscr{S}^k(T^n, V)$, a pseudo-differential operator, denoted $\mathrm{Op}\, a$, is defined as follows:

$$(\mathrm{Op}\, a)u(x) = \sum_{\eta \in \mathbb{Z}^n} e^{2\pi i x \cdot \eta} a(x,\eta)\hat{u}(\eta), \qquad u \in C^\infty(T^n, \mathbb{C}^m), \quad \text{(A.5.b.4)}$$

where $\hat{u}$ is the Fourier series of u, i.e.

$$\hat{u}(\eta) = \int_{T^n} e^{-2\pi i x \cdot \eta} u(x)\, dx, \qquad \eta \in \mathbb{Z}^n.$$

Since u is smooth, $|\hat{u}(\eta)|$ is a rapidly decreasing function of η. It is then easily checked that the sum (A.5.b.4) and all its derivatives are absolutely convergent, showing that $(\mathrm{Op}\, a)u \in C^\infty(T^n, \mathbb{R}^m)$.

Furthermore, $\mathrm{Op}\, a$ can be uniquely extended to a continuous linear map $\mathrm{Op}\, a: L^2_{s+k}(T^n) \to L^2_s(T^n)$. More specifically, we have the following:

(A.5.b.5) PROPOSITION. *For any* $a \in \mathscr{S}^k(T^n, V)$, $u \in C^\infty(T^n, \mathbb{C}^m)$,

$$|(\mathrm{Op}\, a)u|_{L^2_s(T^n)} \leqslant C_{s,k}|a|_{0,s+n+1}|u|_{L^2_{s+k}(T^n)}, \qquad \text{k, s nonnegative integers.}$$

The proof presented is essentially the standard one for continuity of pseudo-differential operators on $\mathbf{R}^n$. We first recall some basic facts on the Sobolev norm of a function.

Given a function $\hat{f}: \mathbf{Z}^n \to \mathbf{C}^m$, we will denote

$$|\hat{f}|_{L^2} = \left(\sum_{\eta \in \mathbf{Z}^n} |\hat{f}(\eta)|^2 \right)^{1/2}.$$

The Plancherel theorem states that for any $u \in L^2(T^n, \mathbf{C}^m)$ and $\hat{u}$ its Fourier transform,

$$|u|_{L^2(T^n)} = |\hat{u}|_{L^2}.$$

Basic fact. For any nonnegative integer k, there are constants C_k', C_k'' such that

$$|u|_{L_k^2(T^n)} \leqslant C_k' |(1 + |\eta|)^k u(\eta)|_{L^2} \leqslant C_k'' |u|_{L_k^2(T^n)} \qquad \text{(A.5.b.6)}$$

Proof of proposition.

$$\widehat{(\mathrm{Op}\,a)u}(\eta) = \int_{T^n} e^{-2\pi i x \cdot \eta} (\mathrm{Op}\,a)u(x)\,dx$$

$$= \int_{T^n} \sum_{\xi \in \mathbf{Z}^n} e^{-2\pi i x \cdot (\eta - \xi)} a(x, \xi) \hat{u}(\xi)\,dx$$

$$= \sum_{\xi \in \mathbf{Z}^n} \left[\int_{T^n} e^{-2\pi i x \cdot (\eta - \xi)} a(x, \xi)\,dx \right] \hat{u}(\xi)$$

$$= \sum_{\xi \in \mathbf{Z}^n} \hat{a}(\eta - \xi, \xi) \hat{u}(\xi).$$

By (A.5.b.6) we have

$$|(\mathrm{Op}\,a)u|_{L_s^2(T^n)} \leqslant C_s' |(1 + |\eta|)^s \widehat{(\mathrm{Op}\,a)u}(\eta)|_{L^2}$$

$$= C_s' \left| (1 + |\eta|)^s \sum_{\xi \in \mathbf{Z}^n} \hat{a}(\eta - \xi, \xi) \hat{u}(\xi) \right|_{L^2}$$

$$\leqslant C_s' \left| (1 + |\eta|)^s \sum_{\xi \in \mathbf{Z}^n} |\hat{a}(\eta - \xi, \xi)| \, |\hat{u}(\xi)| \right|_{L^2}$$

$$\leqslant C_{s,l}' |a|_{0,l} \left| (1 + |\eta|)^s \sum_{\xi \in \mathbf{Z}^n} (1 + |\eta - \xi|)^{-l} (1 + |\xi|)^k |\hat{u}(\xi)| \right|_{L^2}.$$

Now using the fact that $1 + |\eta| \leqslant 1 + |\xi| + |\eta - \xi| \leqslant (1 + |\xi|)(1 + |\eta - \xi|)$, we

obtain

$$\leqslant C'_{s,l}|a|_{0,l}\left\|\sum_{\xi\in\mathbf{Z}^n}(1+|\eta-\xi|)^{s-l}(1+|\xi|)^{s+k}|\hat{u}(\xi)|\right\|_{L^2}$$

$$\leqslant C'_{s,l}|a|_{0,l}\left(\sum_{\eta\in\mathbf{Z}^n}(1+|\eta|)^{s-l}\right)|(1+|\eta|)^{s+k}\hat{u}(\eta)|_{L^2},$$

by Hölder's inequality. By setting $l = s + n + 1$, the sum in the second factor converges. Using (A.5.b.6) we obtain the desired estimate.   Q.E.D.

We next develop the symbol calculus of pseudo-differential operators up to "zeroth order." However, an error term, analogous to that of a Taylor series, is obtained.

(A.5.b.7) PROPOSITION.   Let $a\in\mathscr{S}^k(T^n,V)$, $b\in\mathscr{S}^l(T^n,V)$. Then $ab\in\mathscr{S}^{k+l}(T^n,V)$ and

$$(\operatorname{Op}a)\circ(\operatorname{Op}b)=\operatorname{Op}ab+\operatorname{Op}c,$$

where the Fourier transform of $c\in\mathscr{S}^{k+l-1}(T^n,V)$ is given as follows:

$$\hat{c}(\eta,\xi)=\frac{1}{2\pi i}\sum_{\delta\in\mathbf{Z}^n}\int_0^1\widehat{D_\xi a}(\eta-\delta,\xi+t\delta)\cdot\widehat{D_x b}(\delta,\xi)\,dt.$$

Proof.

$$(\operatorname{Op}a)(\operatorname{Op}b)u(x)=\sum_{\eta\in\mathbf{Z}^n}e^{2\pi i x\cdot\eta}a(x,\eta)\widehat{(\operatorname{Op}b)u}(\eta)$$

$$=\sum_{\eta\in\mathbf{Z}^n}\sum_{\xi\in\mathbf{Z}^n}e^{2\pi i x\cdot\eta}a(x,\eta)\hat{b}(\eta-\xi,\xi)\hat{u}(\xi).$$

By the fundamental theorem of calculus,

$$a(x,\eta)=a(x,\xi)+(\eta-\xi)\cdot r(x,\xi,\eta-\xi),$$

where

$$r(x,\xi,\delta)=\int_0^1 D_\xi a(x,\xi+t\delta)\,dt.$$

Then $(\operatorname{Op}a)(\operatorname{Op}b)u(x)=S_1(x)+S_2(x)$, where

$$S_1(x)=\sum_{\eta,\xi\in\mathbf{Z}^n}e^{2\pi i x\cdot\eta}a(x,\xi)\hat{b}(\eta-\xi,\xi)\hat{u}(\xi)$$

and

$$S_2(x) = \sum_{\eta,\xi \in \mathbf{Z}^n} e^{2\pi i x \cdot \eta} (\eta - \xi) \cdot r(x,\xi,\eta - \xi) \hat{b}(\eta - \xi,\xi) \hat{u}(\xi).$$

$$S_1(x) = \sum_{\xi \in \mathbf{Z}^n} e^{2\pi i x \cdot \xi} a(x,\xi) \left[\sum_{\eta \in \mathbf{Z}^n} e^{2\pi i x \cdot (\eta - \xi)} \hat{b}(\eta - \xi,\xi) \right] \hat{u}(\xi)$$

$$= \sum_{\xi \in \mathbf{Z}^n} e^{2\pi i x \cdot \xi} a(x,\xi) b(x,\xi) \hat{u}(\xi) = (\mathrm{Op}\, ab) u(x).$$

$$\hat{S}_2(\delta) = \int_{T^n} e^{-2\pi i x \cdot \delta} S_2(x)\, dx$$

$$= \sum_{\eta,\xi \in \mathbf{Z}^n} \left[\int_{T^n} e^{-2\pi i x \cdot (\delta - \eta)} r(x,\xi,\eta - \xi)\, dx \right] \cdot (\eta - \xi) \hat{b}(\eta - \xi,\xi) \hat{u}(\xi)$$

$$= \sum_{\eta,\xi \in \mathbf{Z}^n} \hat{r}(\delta - \eta,\xi,\eta - \xi) \cdot (\eta - \xi) \hat{b}(\eta - \xi,\xi) \hat{u}(\xi)$$

where

$$\hat{r}(\eta,\xi,\delta) = \int_0^1 D_\xi \hat{a}(\eta,\xi + t\delta)\, dt.$$

Since $\widehat{D_x b}(\eta - \xi,\xi) = 2\pi i (\eta - \xi) \hat{b}(\eta - \xi,\xi)$, we get

$$\hat{S}_2(\delta) = \sum_{\eta,\xi \in \mathbf{Z}^n} \left[\int_0^1 D_\xi \hat{a}(\delta - \eta,\xi + t(\eta - \xi)) \cdot \widehat{D_x b}(\eta - \xi,\xi)\, dt \right] \hat{u}(\xi)$$

$$= \sum_{\epsilon,\xi \in \mathbf{Z}^n} \left[\int_0^1 D_\xi \hat{a}(\delta - \xi - \epsilon,\xi + t\epsilon) \cdot \widehat{D_x b}(\epsilon,\xi)\, dt \right] \hat{u}(\xi)$$

$$= \widehat{(\mathrm{Op}\, c)} u(\delta).$$

We still need to check that c is in fact a symbol. It suffices to show that $\hat{c}$ satisfies (A.5.b.3) for all α and l. Only the case $\alpha = 0$ will be proved here; the general estimate for $\alpha \neq 0$ is proved in the same way.

$$|\hat{c}(\eta,\xi)| \leq \frac{1}{2\pi} \sum_{\delta \in \mathbf{Z}^n} \int_0^1 |\widehat{D_\xi a}(\eta - \delta,\xi + t\delta)| \, |\widehat{D_x b}(\eta,\xi)|\, dt$$

$$\leq |a|_{1,m} |b|_{0,j} \sum_{\delta \in \mathbf{Z}^n} \int_0^1 (1 + |\eta - \delta|)^{-m} (1 + |\xi + t\delta|)^{k-1} (1 + |\eta|)^{-j+1} (1 + |\xi|)^l\, dt$$

$$\leq |a|_{1,m} |b|_{0,j} (1 + |\eta|)^{-j+1} (1 + |\xi|)^l \sum_{\delta \in \mathbf{Z}^n} \int_0^1 (1 + |\eta - \delta|)^{-m} (1 + |\xi + t\delta|)^{k-1}\, dt$$

$$\leq |a|_{1,m} |b|_{0,j} (1 + |\eta|)^{-j+1} (1 + |\xi|)^{k+l-1} \sum_{\delta \in \mathbf{Z}^n} (1 + |\eta - \delta|)^{-m} \int_0^1 (1 + t|\delta|)^{k-1}\, dt.$$

For $k \geqslant 1$, we obtain

$$|c(\eta, \xi)| \leqslant |a|_{1,m} |b|_{0,j} (1 + |\eta|)^{-j+1} (1 + |\xi|)^{k+l-1} \sum_{\delta \in Z^n} (1 + |\eta - \delta|)^{-m} (1 + |\delta|)^{k-1}$$

$$\leqslant |a|_{1,m} |b|_{0,j} (1 + |\eta|)^{-j+k} (1 + |\xi|)^{k+l-1} \sum_{\delta \in Z^n} (1 + |\eta - \delta|)^{-m+k-1}.$$

Set $m = n + k$ to make the sum converge, $s = j - k$ obtaining

$$|c|_{0,s} \leqslant C |a|_{1,n+k} |b|_{0,s+k} \mathscr{D}; \qquad C = \sum_{\delta \in Z^n} (1 + |\delta|)^{-n-1}. \qquad \text{(A.5.b.8)}$$

For $k \leqslant 0$, the integral drops right out, then setting $s = j - 1$ and $m = n + 1$, the following is obtained:

$$|c|_{0,s} \leqslant C |a|_{1,n+1} |b|_{0,s+1}. \qquad \text{(A.5.b.8')}$$

In particular, $c(x, \xi)$ satisfies the symbol estimates of order $k + l - 1$.   Q.E.D.

We will denote the adjoint operator to $\operatorname{Op} a$ with respect to the L^2 inner product as $(\operatorname{Op} a)^*$.

(A.5.b.9) PROPOSITION.   *For any symbol $a \in \mathscr{S}^k (T^n, V)$, $(\operatorname{Op} a)^*$ is also a pseudo-differential operator $\operatorname{Op} a^*$ of order k. Moreover, its symbol is given as follows*:

$$\hat{a}^*(\eta, \xi) = {}^t\bar{a}(-\eta, \xi) + \int_0^1 {}^t\eta \cdot D_\xi {}^t\overline{a(-\eta, \xi + t\eta)} \, dt$$

Proof.   By Plancherel's theorem,

$$\langle u, (\operatorname{Op} a)v \rangle = \sum_{\eta \in Z^n} \hat{u}(\eta) \overline{(\widehat{\operatorname{Op} a)v}(\eta)} = \sum_{\eta \in Z^n} \sum_{\xi \in Z^n} \hat{u}(\eta) \overline{\hat{a}(\eta - \xi, \xi)} \, \overline{\hat{v}(\xi)}$$

Thus,

$$\widehat{(\operatorname{Op} a)^*}u(\eta) = \sum_{\xi \in Z^n} {}^t\overline{\hat{a}(\xi - \eta, \eta)} \, \hat{u}(\xi) \qquad \text{(A.5.b.10)}$$

As before, use the fundamental theorem of calculus to say

$$\hat{a}(\xi - \eta, \eta) = \hat{a}(\xi - \eta, \xi) + \int_0^1 (\eta - \xi) \circ D_\xi \hat{a}(\xi - \eta, \xi + t(\eta - \xi)) \, dt$$

Substituting this into (A.5.b.10), we obtain the desired expression for $(\operatorname{Op} a)^*u$. It is easily checked that $a^*(x, \xi) \in \mathscr{S}^k (T^n, V)$.   Q.E.D.

(c)   Fix $\mathscr{A}$ and $\mathscr{H}$ as defined in §V(a). Throughout this section, we will also fix a set $\mathscr{A}' \subset \mathscr{H}$ such that
(1) $\mathscr{A} \subset$ interior $(\mathscr{A}')$
(2) $\mathscr{A}' \cong$ closed ball in $\mathbb{R}^{(n-1)m^2}$.

Let $P = \partial/\partial t + A^i(\partial/\partial x^i) + B$ be a strictly hyperbolic operator on X. We will always take $(A^1, \ldots, A^{n-1}) \in C^\infty(X, \mathscr{A}')$ and $B \in C^\infty(X, \mathbb{R}^m \otimes \mathbb{R}^m)$. Denote

$$A_t = A^i(x,t)\frac{\partial}{\partial x^i}\bigg|_{T_t^{n-1}} : C^\infty(T, \mathbb{R}^m) \to C^\infty(T, \mathbb{R}^m)$$

Setting $a_t(x,\xi) = 2\pi i \sum_{i=1}^{n-1} A^i(x,t)\xi_i$, $A_t = \mathrm{Op}\, a_t$.

P strictly hyperbolic implies that $a_t(x,\xi)$, $\xi \neq 0$, has m distinct imaginary eigenvalues, denoted $2\pi i\lambda_1(x,t,\xi), \ldots, 2\pi i\lambda_m(x,t,\xi)$, where $\lambda_1 < \lambda_2 < \ldots < \lambda_m$. Using this property, a symmetrizer for A_t can be constructed as follows:

(A.5.c.1) PROPOSITION. *Let* $(A^i) \in C^\infty(X, \mathscr{H})$,

$$a_t(x,\xi) = 2\pi i \sum_{j=1}^{n-1} A^j(x,t)\xi_j,$$

and

$$A_t = \mathrm{Op}\, a_t = \sum_{j=1}^{n-1} A^j \frac{\partial}{\partial x^j}.$$

Then there exists a pseudo-differential operator $R_t = \mathrm{Op}\, r_t$ *of order zero satisfying the following*:

$$R_t^* = R_t \tag{A.5.c.2}$$

$$\langle u, R_t u \rangle \geqslant \frac{1}{2m^2}|u|^2_{L^2(T^{n-1})}, \qquad u \in C^\infty(T^{n-1}, \mathbb{R}^m), \tag{A.5.c.3}$$

where $\langle u, v \rangle = \int_{T^n} (u(x), v(x))\,dx$.

$$A_t^* R_t + R_t A_t \text{ is an operator of order zero.} \tag{A.5.c.4}$$

Futhermore, if we assume that $(A^i) \in C^\infty(X, \mathscr{A}')$, *then the following estimates hold*:

$$|R_t u|_{L^2(T^n)} \leqslant M\left(\sum_{j=1}^{n-1} |A^i|_{L^2_{n+1}(X)} + 1\right)^2 |u|_{L^2(T)} \tag{A.5.c.5}$$

$$\left|\frac{dR_t}{dt}u\right|_{L^2(T)} \leqslant M\left(\sum_{j=1}^{n-1} |A^i|_{L^2_{n+2}(X)} + 1\right)^2 |u|_{L^2(T)} \tag{A.5.c.6}$$

Proof. For each eigenvalue $2\pi i\lambda_\nu(x,t,\xi)$, $\xi \neq 0$, define the matrix

$$p_\nu(x,t,\xi) = \frac{1}{2\pi i}\oint_{\gamma_\nu(x,t,\xi)} (zI - a_t(x,\xi))^{-1}\,dz,$$

where $\gamma_\nu(x,t,\xi)$ is a contour, smoothly varying with (x,t,ξ), enclosing

$2\pi i\lambda_\nu(x,t,\xi)$ and no other eigenvalue of $a_t(x,\xi)$. p_ν is a well-defined, smooth, real matrix-valued function on $X \times \mathbf{R}^{n-1}\backslash\{0\}$. As a linear transformation on V, p_ν projects a vector onto its component in the $\lambda_\nu(x,t,\xi)$-eigenspace. It is then seen that

(1) p_ν is homogeneous of degree zero.

(2) p_ν commutes with a_t.

Now set

$$r'_t(x,\xi) = \sum_{\nu=1}^{m} {}^t p_\nu(x,t,\xi)p_\nu(x,t,\xi);$$

$r'_t(x,\xi)$ is a positive definite symmetric matrix. Furthermore, since a_t is pure imaginary, $a_t^* = -{}^t a_t$. Using this, it is easy to check that $r'_t a_t + a_t^* r'_t = 0$.

However, we need a positive definite, self-adjoint operator R_t and having a positive definite symmetric symbol does not necessarily give one. One can be constructed, in the following manner:

A bound on the lowest eigenvalue of r'_t is given by:

$$|v|^2 = \left| \sum_{\nu=1}^{m} p_\nu(x,t,\xi)v \right|^2 \leqslant \left(\sum_{\nu=1}^{m} |p_\nu(x,t,\xi)v| \right)^2$$

$$\leqslant |m^2 \sum_{\nu=1}^{m} |p_\nu(x,t,\xi)v|^2 = m^2(v, r'_t(x,\xi)v),$$

where $v \in \mathbf{R}^m$, $(x,t,\xi) \in X \times \mathbf{R}^{n-1}\backslash\{0\}$.

Therefore, $r'_t(x,\xi) - (1/2m^2)I$, $I =$ identity matrix, is also positive definite symmetric. It then has a smoothly varying "square root", $b'_t(x,\xi)$, satisfying

$$r'_t(x,\xi) = {}^t b'_t(x,\xi)b'_t(x,\xi) + \frac{1}{2m^2}I.$$

b'_t, like all the other functions defined thus far, is not a symbol since it is not smooth when $\xi = 0$. This is remedied by fixing a function $\varphi \in C^\infty(\mathbf{R})$, $\varphi|_{[0,1/2]} \equiv 0$, $\varphi|_{[1,\infty)} \equiv 1$ and setting $b_t(x,\xi) = \varphi(|\xi|)b'_t(x,\xi)$. Since b_t is smooth on $T^{n-1} \times \mathbf{R}^{n-1}$, for fixed t, and homogeneous of degree zero in ξ for $|\xi| \geqslant 1$, $b_t \in \mathscr{S}^0(T,V)$.

The desired symmetrizing operator is then given as:

$$R_t = (\operatorname{Op} b_t)^* \circ (\operatorname{Op} b_t) + \frac{1}{2m^2}I,$$

where $I : C^\infty(T,\mathbf{R}^m) \to C^\infty(T,\mathbf{R}^m)$ is the identity map. R_t clearly satisfies (A.5.c.2) and (A.5.c.3). Using the standard symbol calculus of pseudo-differential operators, as seen in propositions (A.5.b.7) and (A.5.b.9), (A.5.c.4) follows easily. Observe that the cut off function adds only an error of infinitely negative order.

It remains to prove (A.5.c.5), (A.5.c.6). For this, the following lemma is needed:

(A.5.c.7) LEMMA. *Let $p \in C^\infty(X, V)$ be such that for fixed t, $p(\cdot, t, \cdot)$ $\in \mathscr{S}^k(T^{n-1}, V)$. Then p satisfies the following inequality*:

$$|p(\cdot, t, \cdot)|_{0,s} \leqslant M_s \sup_{\xi \in \mathbb{R}} (1 + |\xi|)^{-k} |p(x, t, \xi)|_{L^2_{s+1}(X)}.$$

Proof.

$$|p|_{0,s} = \sup_{\substack{\eta \in \mathbb{Z}^{n-1} \\ \xi \in \mathbb{R}^{n-1}}} (1 + |\eta|)^s (1 + |\xi|)^{-k} \int_{T^n} e^{-2\pi i x \cdot \eta} p(x, t, \xi) \, dx$$

$$\leqslant M_s' \sup_{\xi \in \mathbb{R}^{n-1}} (1 + |\xi|)^{-k} \int_{T^n} \sum_{|\alpha| \leqslant s} |D_x^\alpha p(x, t, \xi)| \, dx$$

Fact. For any $f \in C^\infty([0, 1])$, $|f(x)| \leqslant \int_0^1 |f(t)| + |f'(t)| \, dt$, $x \in [0, 1]$.

We leave the proof as an elementary exercise. Using this fact and the Cauchy–Schwartz inequality, we get

$$\leqslant M_s' \sup_{\xi \in \mathbb{R}^{n-1}} (1 + |\xi|)^{-k} \int_{T^n} \int_0^1 \sum_{|\alpha| \leqslant s} |D_x^\alpha p(x, t, \xi)| + \left| \frac{\partial}{\partial t} D_x^\alpha p(x, t, \xi) \right| dx \, dt$$

$$\leqslant M_s' \sup_{\xi \in \mathbb{R}^{n-1}} (1 + |\xi|)^{-k} |p(x, t, \xi)|_{L^2_{s+1}(X)}. \quad \text{Q.E.D.}$$

For fixed ξ, the function $b_t(x, \xi)$ is obtained by composing a fixed smooth map $\mathscr{H} \times \mathbb{R}^{n-1} \to \mathbb{R}^m \otimes \mathbb{R}^m$ with $(A^1, \ldots, A^{n-1}, \xi)$. In particular, if we restrict to $(A^1, \ldots, A^{n-1}) \in \mathscr{A}'$, a compact set in $\mathscr{H}$, lemma (5.b.2) gives

$$|b_t(x, \xi)|_{L^2_{s+1}(X)} \leqslant C_{s+1,\xi}' |a_t(x, \xi)|_{L^2_{s+1}(X)}, \quad \text{for each } \xi.$$

Since b_t is homogeneous of degree zero in ξ,

$$\sup_{\xi \in \mathbb{R}^{n-1}} |b_t(x, \xi)|_{L^2_{s+1}(X)} \leqslant \sup_{|\xi|=1} C_{s+1,\xi}' |a_t(x, \xi)|_{L^2_{s+1}(X)} \leqslant C_{s+1} \sum_{j=1}^{n-1} |A^j|_{L^2_{s+1}(X)}$$

Combining this estimate with lemma (A.5.c.7), we obtain

$$|b_t|_{0,n} \leqslant M' \sum_{j=1}^{n-1} |A^j|_{L^2_{n+1}(X)}. \tag{A.5.c.8}$$

Finally we estimate the norm of R_t:

$$\langle v, R_t u \rangle = \left\langle v, \left[(\operatorname{Op} b_t)^* \circ (\operatorname{Op} b_t) + \frac{1}{2m^2} \right] u \right\rangle$$

$$= \langle (\operatorname{Op} b_t) v, (\operatorname{Op} b_t) u \rangle + \frac{1}{2m^2} \langle v, u \rangle$$

$$\leqslant |b_t|_{0,n}^2 |v|_{L^2(T)} |u|_{L^2(T)} + \frac{1}{2m^2} |v|_{L^2(T)} |u|_{L^2(T)}$$

Setting $v = R_t u$ and using (A.5.c.8), we get

$$|R_t u|_{L^2(T)} \leqslant \left(\sum_{j=1}^{n-1} |A^j|_{L^2_{n+1}(X)} + 1 \right)^2 |u|_{L^2(T)}.$$

Repeating the argument above, it can be proved that

$$\left| \frac{\partial b_t}{\partial t} \right|_{0,n} \leqslant M_1' \sum_{j=1}^{n-1} |A^j|_{L^2_{n+2}(X)}.$$

Then by differentiating $\langle v, R_t u \rangle$ with respect to t, we obtain the estimate

$$\left| \frac{\partial R_t}{\partial t} u \right|_{L^2(T)} \leqslant M_1 \left(\sum_{j=1}^{n-1} |A^j|_{L^2_{n+2}(X)} + 1 \right)^2 |u|_{L^2(T)}. \quad \text{Q.E.D.}$$

Using R_t, we can prove an energy integral inequality. It will be convenient to define the following inner product and norms on $L^2(T, \mathbf{R}^m)$:

$$\langle u, v \rangle_{R_t} = \langle u, R_t v \rangle$$

$$|u|_{R_t} = \langle u, u \rangle_{R_t}^{1/2}.$$

The norm $|\ |_{R_t}$, by (A.5.c.3) and (A.5.c.5), is equivalent to the L^2-norm.

(A.5.c.9) PROPOSITION.   *Let* $P = \partial / \partial t + A^i(\partial / \partial x^i) + B$ *be a strictly hyperbolic operator on* X *with* $(A^i) \in C^\infty(X, \mathscr{A}')$ *and* $B \in C^\infty(X, \mathbf{R}^m \otimes \mathbf{R}^m)$. *Given any* $f, u \in C^\infty(X, \mathbf{R}^m)$ *satisfying* $Pu = f$, *the following inequality holds*:

$$\frac{d}{dt} |u|_{R_t} \leqslant c|u|_{R_t} + |f|_{R_t}, \qquad t \in [0, 1],$$

where c *is a constant depending on the coefficients* A^i. *However, if we fix a constant* $b > 0$ *and assume that*

$$\sum_{j=1}^{n-1} |A^j|_{L^2_{n+3}} + |B|_{L^2_n} < b,$$

then c *may be chosen so as to depend only on* $\mathscr{A}'$ *and* b *and to be otherwise independent of* P.

   Proof.

$$\frac{d}{dt} |u|^2_{R_t} = \left\langle \frac{\partial u}{\partial t}, R_t u \right\rangle + \left\langle u, R_t \frac{\partial u}{\partial t} \right\rangle + \left\langle u, \frac{dR_t}{dt} u \right\rangle$$

$$= \langle f - A_t u - Bu, R_t u \rangle + \langle u, R_t(f - A_t u - Bu) \rangle + \left\langle u, \frac{dR_t}{dt} u \right\rangle$$

$$= 2 \langle f, R_t u \rangle + \left\langle u, \frac{dR_t}{dt} u \right\rangle - 2 \langle u, R_t Bu \rangle - \langle u, (R_t A_t + A_t^* R_t)u \rangle$$

We estimate each term separately:

$$\langle f, E_t u \rangle_t \leqslant |f|_{R_t} |u|_{R_t} \qquad \text{by Cauchy–Schwartz.}$$

$$\left\langle u, \frac{dR_t}{dt} u \right\rangle_t \leqslant M_1 \left(\sum_{j=1}^{n-1} |A^j|_{L^2_{n+2}(X)} + 1 \right)^2 |u|^2_{L^2(T_t)}, \qquad \text{by (A.5.c.6)},$$

$$\leqslant 2m^2 M_1 \left(\sum_{j=1} |A^j|_{L^2_{n+2}(X)} + 1 \right)^2 |u|^2_{R_t}, \qquad \text{by (A.5.c.3).}$$

$$\langle u, R_t Bu \rangle = \langle R_t u, Bu \rangle \leqslant |R_t u|_{L^2(T_t)} |Bu|_{L^2(T_t)}$$

$$\leqslant M \left(\sum_{j=1}^{n-1} |A^j|_{L^2_{n+1}(X)} + 1 \right)^2 |B|_{L^2(T_t)} |u|^2_{L^2(T_t)}$$

$$\leqslant 2m^2 M \left(\sum_{j=1}^{n-1} |A^j|_{L^2_{n+1}(X)} + 1 \right)^2 |B|_{L^2(T_t)} |u|^2_{R_t}$$

A crude bound on $|B|_{L^2(T_t)}$ is given by the Sobolev lemma:

$$|B|_{L^2(T_t)} \leqslant \left(\sup_{x \in T} |B(x,t)| \right) \text{vol}(T_t) \leqslant \sup_{(x,t) \in X} |B(x,t)| \leqslant |B|_{L^2_n(X)}.$$

To estimate the last term, the symbol of $R_t A_t + A_t^* R_t$ must be estimated. First, the symbol of A_t^* is given as follows:

$$A_t^* = \left(A^j \frac{\partial}{\partial x^j} \right)^* = -\frac{\partial}{\partial x^j} \circ {}^t A^j = -\sum_{j=1}^{n-1} \frac{\partial {}^t A^j}{\partial x^j} - {}^t A^j \frac{\partial}{\partial x^j}.$$

Therefore,

$$a_t^*(x, \xi) = -2\pi i \sum_{j=1}^{n-1} A^j \xi_j - \sum_{j=1}^{n-1} \frac{\partial {}^t A^j}{\partial x^j}.$$

Using (A.5.b.8), (A.5.b.8′), and (A.5.b.9), we see that

$$\langle u, (R_t A_t + A_t^* R_t) u \rangle \leqslant C |r_t|_{1,n+1} (|a_t|_{1,n+1} + |a_t^*|_{1,n+1}) |u|^2_{L^2(T_t)}$$

$$\leqslant C_1 \left(\sum_{j=1}^{n-1} |A^j|_{L^2_{n+3}(X)} + 1 \right)^3 |u|^2_{L^2(T_t)}$$

$$\leqslant 2m^2 C_1 \left(\sum_{j-1}^{n-1} |A^j|_{L^2_{n+3}(X)} + 1 \right)^3 |u|^2_{R_t}$$

The proposition follows immediately.   Q.E.D.

The extension of the estimate to Sobolev estimates proceeds exactly as in §V(c). The only difference is that a different norm is used, namely

$$|u|_{R_t,x} = \left[\int_0^1 |u|^2_{R_t} \, dt \right]^{1/2}$$

The final estimate, of the form given in Proposition (5.c.7), can be converted back to the desired estimate using the L_k^2-norms, using (A.5.c.3) and (A.5.c.5). This gives the following result:

(A.5.c.10) PROPOSITION. *Fix* $b' > 0$ *and* $\mathscr{A}' \subset \mathscr{H}$ *as described earlier. Let* $P = \partial/\partial t + A^i(\partial/\partial x^i) + B$ *be a strictly hyperbolic operator on* X *with* $(A^i) \in C^\infty(X, \mathscr{A}')$, $B \in C^\infty(X, \mathbb{R}^m \otimes \mathbb{R}^m)$,

$$\sum_{j=1}^{n-1} |A^j|_{L^2_{n+3}(X)} + |B|_{L^2_n(X)} < b'$$

Let $f \in C^\infty(X, \mathbb{R}^m)$ satisfy the following: There exists $\epsilon > 0$ such that $f|_{T^{n-1} \times [0,\epsilon]} = 0$. Let u solve the initial value problem:

$$Pu = f, \qquad u|_{T_0} = 0.$$

The following estimate holds:

$$|u|_{L^2_k(X)} \leqslant \hat{C}_k \left[|f|_{L^2_k(X)} + \left(\sum_{j=1}^{n-1} |A^j|_{L^2_k(X)} + |B|_{L^2_k(X)} \right) |f|_{L^2_1(X)} \right]$$

when $\hat{C}_k$ depends on b' and $\mathscr{A}'$ but is otherwise independent of u, f, and P.

We now give a similar argument as in §V(d) to indicate how a strictly hyperbolic system of the form

$$A^n \frac{\partial u}{\partial t} + A^i \frac{\partial u}{\partial x^i} + Bu = f$$

on M is solved using the results of §AV(c) and an extension operator. The procedure is slightly more involved than for symmetric hyperbolic systems.

Let $\mathscr{A}$, b, and α be fixed as in the statement of Theorem b', given in §V(a). Also, fix $\mathscr{A}' \supset \mathscr{A}$ as described in §AV(c).

Consider a strictly operator $P = A^n(\partial/\partial t) + A^i(\partial/\partial x^i) + B$ on M satisfying the assumptions of Theorem b. Denote

$$\tilde{P} = (A^n)^{-1}P = \frac{\partial}{\partial t} + \tilde{A}^i \frac{\partial}{\partial x^i} + \tilde{B}, \qquad \tilde{A}^i = (A^n)^{-1}A, \quad \tilde{B} = (A^n)^{-1}B.$$

We want to extend $\tilde{P}$ to a strictly hyperbolic operator on X. Fix an extension operator, as described in §IV(c), $\mathscr{E} : L^2(M, \mathbb{R}^{(n-1)m^2}) \to L^2(X, \mathbb{R}^{(n-1)m^2})$. Recall that $\mathscr{E}$ is bounded in the L^∞ norm. Therefore, by embedding $\mathscr{A}'$ appropriately

as a ball in $\mathsf{R}^{(n-1)m^2}$ and $\mathscr{A}$ as a much smaller ball, we can define $\tilde{\mathscr{E}} : L^2(M, \mathscr{A}) \to L^2(X, \mathscr{A}')$. Moreover, it is clear that this definition of $\tilde{\mathscr{E}}$ still gives a bounded operator on L_k^2, $k \geqslant 0$. Now set $\hat{A}^i = \tilde{\mathscr{E}} \tilde{A}^1$. Extend $\tilde{B}$ to a map $\hat{B} : X \to \mathsf{R}^m \otimes \mathsf{R}^m$ as in §V(d). This defines a strictly hyperbolic operator $\hat{P}$ satisfying the assumptions of Proposition (A.5.c.10). Observe that b' depends only on b and $\mathscr{E}$. Let $\hat{Q}$ be the right inverse of $\hat{P}$ as given in (A.5.c.10). Then a right inverse of P is given by

$$Q = \mathscr{R}\hat{Q}\mathscr{E}(A^n)^{-1}f,$$

when $\mathscr{R} : L^2(X, \mathsf{R}^m) \to L^2(M, \mathsf{R}^m)$ is the restriction operator. That Q is a right inverse of P and satisfies (5.a.8) is proved in the same fashion as Theorem G was proved in §V(d). We do not repeat the argument here. This completes the proof of Theorem G'.

VI. Local existence of smooth isometric embeddings.

(a)   In this section, all the various pieces presented in §§III–V are assembled together to prove a local existence theorem for smooth isometric embeddings of a three-manifold in R^6.

Throughout this section; we will use the manifold $M^3 \cong \Omega^2 \times [\frac{1}{4}, \frac{3}{4}]$, Ω a compact convex domain in R^2.

(6.a.1) *Definition.*   A smooth embedding $u : M^3 \to \mathsf{R}^6$ is *hyperbolic* if the following equivalent conditions hold:

(1)   The differential operator given by the left-hand side of (4.d.5) is hyperbolic as defined by (5.a.1) and (5.a.2).

(2)   The characteristic variety of the linearized isometric embedding system is a cubic with two components; *and* the line $\mathsf{R} dx^3 \in T_x^* M$ lies inside the oval for each $x \in M$.

(6.a.2) *Remark.*   The first half of (2) corresponds to the definition of hyperbolicity given in §III(c). The second part corresponds to the x^3-direction in M being a "timelike direction." Using (1), Theorems G and G', and the discussion in §IV(d), we see that the linearized isometric embedding system (4.d.3) can be solved whenever u is hyperbolic. This then leads to the following result:

THEOREM H.   *Let* $u_0 : M \to \mathsf{R}^6$ *be a smooth hyperbolic embedding. There exists* $s \in \mathsf{Z}^+$, $\delta \in \mathsf{R}^+$ *such that for any metric* $g \in C^\infty(S^2 T^* M)$ *satisfying*

$$|g - \Phi(u_0)|_{L_s^2(M)} \leqslant \delta,$$

there is a smooth embedding $u : M \to \mathsf{R}^6$ *inducing the metric* g; *i.e.,*

$$\Phi(u) = g.$$

Remark.   Theorem H also holds for any manifold $M' \cong S \times I$ where S is a compact surface and I a closed interval. The proof is exactly as for M, but

messier since there are no global co-ordinates. To obtain the estimates analogous to Theorems G and G', M' is embedded in $X' = S \times I'$, $I \subset\subset I'$; and the initial value problem is solved in X'.

The proof of this theorem consists of checking that the assumptions of the Nash–Moser–Schwartz–Sergeraert theorem hold for a suitable choice of the Banach spaces and the functional. The conclusion then follows directly.

The scales of Banach spaces we will use are given as follows:

$$E_k = L^2_{k+1}(M, \mathbb{R}^6), \qquad F_k = L^2_k(S^2 T^* M); \qquad k \geqslant 0.$$

The norms for $\{F_k\}$ will be taken to be the usual L^2_k-norms. However, we will fix $\epsilon > 0$ and scale the norms for $\{E_k\}$ as follows:

$$|u|_k = \epsilon^{-1} |u|_{L^2_{k+1}(M)}, \qquad u \in E_k.$$

The constant ϵ will be specified more precisely later in the proof.

Denote by $D_k(u_0)$ the unit ball in E_k centered at u_0.

The functional to which the Nash–Moser–Schwartz–Sergeraert theorem is to be applied is

$$\Phi_0(v) = \Phi(u_0 + v) - \Phi(u_0), \qquad v \in L^2_1(M, \mathbb{R}^6)$$

when Φ is the functional defined by (4.d.1).

(1) Clearly we have $\Phi_0(0) = 0$.

(2) To see that Φ_0 is twice Frechet differentiable, it suffices to show that Φ is. The derivatives of Φ are easily seen to be as follows:

$$\Phi'(u)y = 2(du, dy) = 2\left(\frac{\partial u}{\partial x^i}, \frac{\partial u}{\partial x^j} \right) dx^i dx^j$$

$$(\Phi''(u)y, w) = \Phi'(y)w = 2(dy, dw)$$

(3) The estimate (4.b.3) is proved using lemma (5.b.1) as follows: let $u = u_0 + v \in D_0(u_0)$.

$$|\Phi_0'(v)y|_k = 2\left| \left(\frac{\partial v}{\partial x^i}, \frac{\partial y}{\partial x^j} \right) dx^i dx^j \right|_{L^2_k(M)}$$

$$\leqslant 2G_k \left[\sum_{i,j=1}^{3} \left| \frac{\partial v}{\partial x^i} \right|_{L^2_k(M)} \left| \frac{\partial y}{\partial x^j} \right|_{L^2(M)} + \left| \frac{\partial v}{\partial x^i} \right|_{L^2(M)} \left| \frac{\partial y}{\partial x^j} \right|_{L^2_k(M)} \right]$$

$$\leqslant 18 G_k \left[|v|_{L^2_{k+1}(M)} |y|_{L^2_1(M)} + |v|_{L^2_1(M)} |y|_{L^2_{k+1}(M)} \right]$$

$$\leqslant C_k' (|v|_k |y|_0 + |v|_0 |y|_k) \leqslant C_k' (|v|_k |y|_0 + |y|_k),$$

since $|v|_0 = |u - u_0|_0 \leqslant 1$.

(4.b.4) is proved the same way but is even easier since Lemma (5.b.1) is not used.

(4) The rest of this section will be devoted to constructing a right inverse $Q(u)$ to $\Phi'(u)$ and showing that the estimate (4.b.5) holds. Equivalently we want to show that given an embedding $u \in D_\alpha(u_0) \cap E_\infty$, α some fixed integer, $h \in F_\infty$, we can always solve for $y : M \to \mathbf{R}^6$ such that

$$\Phi'(u)y = h \tag{6.a.3}$$

and

$$|y|_{k-\alpha} \leqslant K_k(|u|_k|h|_\alpha + |h|_k), \qquad k \geqslant \alpha. \tag{6.a.4}$$

We showed in §IV(d) that to solve for y, it suffices to find a 1-form $\varphi = \varphi_i dx^i$ satisfying the system of partial differential equations (4.d.5). The object of the rest of this section is to show first that if u is sufficiently close to u_0 in the C^2-norm, the system (4.d.5) is a hyperbolic system of the form studied in §V. Theorems G and G' then give a solution for φ and hence for y. Then, in order to prove the required estimate, we must bound the norms of g_{ij}, Γ_{ij}^k, b_r^{ij}, and H_{ij} in terms of norms of u. This is done by observing that all these geometric quantities are obtained by composing fixed smooth functions with the 2-jet of u. Lemma (5.b.2) then bounds the L_k^2-norms of these functions by the L_k^2-norm of the 2-jet of u; but this is simply the L_{k+2}^2 norm of u. Finally applying Lemmas (5.b.1) and (5.b.2) a few more times, the desired estimate for $|y|_{L_k^2(M)}$ is obtained.

We now provide the details of this argument.

Given a smooth map $u : M \to \mathbf{R}^6$, we denote by $j^2(u)$ the 2-jet of u, the associated section of the bundle $\mathbf{R}^6 \otimes J^2(M)$. Let V be the fiber of this bundle. The co-ordinates on M determine a trivialization of $\mathbf{R}^6 \otimes J^2(M)$; we can then identify

$$V = \mathbf{R}^6 \otimes (\mathbf{R} \oplus \mathbf{R}^3 \oplus S^2\mathbf{R}^3);$$

and the 2-jet of u is given as follows:

$$j^2(u)(x) = \left(u(x), \frac{\partial u}{\partial x^1}(x), \ldots, \frac{\partial u}{\partial x^3}(x), \frac{\partial^2 u}{(\partial x^1)^2}(x), \ldots, \frac{\partial^2 u}{(\partial x^3)^2}(x) \right).$$

The property that $u : M^3 \to \mathbf{R}^6$ be a general immersion is a linear algebraic condition on $j^2(u)(x)$, $x \in M$. In particular, u is an immersion at x if the vectors $(\partial u/\partial x^1)(x)$, $(\partial u/\partial x^2)(x)$, $(\partial u/\partial x^3)(x)$ are linearly independent. It is not hard to show that an embedding u is non-degenerate, as defined in §I(a), if at each $x \in M$, the vectors $(\partial u/\partial x^1)(x)$, $(\partial u/\partial x^2)(x)$, $(\partial u/\partial x^3)(x)$, $(\partial^2 u/(\partial x^1)^2)(x)$, $\ldots$, $(\partial u^2/(\partial x^3)^2)(x)$ span all of $\mathbf{R}^6$. We can rephrase this as follows: An embedding u is non-degenerate if and only if the 2-jet of u maps into $V_1 \subset V$, where we define

$$V_1 = \left\{ (u, u_i, u_{ij}) \mid \{u_i\} \text{ are linearly independent, } \{u_i, y_{ij}\} \text{ span } \mathbf{R}^6 \right\}.$$

On V_1, we define the following functions:

$$\hat{g}_{ij}(u, u_r, u_{st}) = (u_i, u_j)$$

Let $\hat{g}(u, u_r, u_{st}) = (\hat{g}_{ij}(u, u_r, u_{st}))$ denote the full positive definite symmetric matrix;

$$\hat{g}_{ij}^{-1}(u, u_r, u_{st}) = (i, j)\text{th component of } \hat{g}(u, u_r, u_{st})^{-1};$$

$$\hat{\Gamma}_{ij}^k(u, u_r, u_{st}) = \sum_{l=1}^{3} \hat{g}_{kl}(u, u_r, u_{st})(u_l, u_{ij});$$

$$\hat{H}_{ij}(u, u_r, u_{st}) = u_{ij} - \hat{\Gamma}_{ij}^k(u, u_r, u_{st})u_k.$$

Notice that H_{ij} is $\mathbf{R}^6$-valued.

It is clear from the discussion in §IV(d) that the functions $g_{ij} = \hat{g}_{ij} \circ j^2(u)$, $\Gamma_{ij}^k = \hat{\Gamma}_{ij}^k \circ j^2(u)$, and $H_{ij} = \hat{H}_{ij} \circ j^2(u)$ give the metric, Christoffel symbols and the second fundamental form associated to the embedding u. The span of the vectors $\{H_{ij}(x)\}$ give the normal bundle of M as a subbundle of $T_x\mathbf{R}^6$.

We can associate a cubic form to any point in V_1 as follows: The vectors $H_{ij}(u, u_r, u_{st})$ span a 3-dimensional subspace of $\mathbf{R}^6$. Choose any basis of this subspace. The equation (4.a.3) then defines a continuous map.

$$[F]: V_1 \to \mathbf{P}(S^3\mathbf{R}^3)$$

$[F]$ is well-defined since a change of basis of the subspace only changes F by a scalar factor.

Let $\mathscr{C} \subset \mathbf{P}(S^3\mathbf{R}^3)$ denote the space of cubics whose associated complex cubics are smooth. $\mathscr{C}$ is clearly an open set in $\mathbf{P}(S^3\mathbf{R}^3)$. Lemma (4.a.9) defines a continuous function $\sigma: \mathscr{C} \to \mathbf{R}$ which is a projective invariant of each cubic.

Let $V_2 \subset V_1$ denote the set of (u, u_r, u_{st}) which satisfy the following:

(1) $[F](u, u_r, u_{st}) \in \mathscr{C}$.

(2) $\sigma([F](u, u_r, u_{st})) \in (-\infty, -\frac{1}{2}) \cup (-\frac{1}{2}, 0) \cup (1, \infty)$.

(3) The point $[0, 0, 1] \in \mathbf{R}P^3$ lies inside the oval of the cubic $[F](u, u_r, u_{st})$.

V_2 is an open set of V_1 and hence of V. Moreover, V_2 can be characterized as follows:

(6.a.5) An embedding u is a hyperbolic embedding of M^3 in $\mathbf{R}^6$ if and only if $j^2(u): M \to V_2$.

(6.a.6) LEMMA. *Let* $\beta \in \mathbf{Z}^+$, $\beta \geqslant 3$. *There exists* $\epsilon > 0$ *such that for any* u *satisfying*

$$|u - u_0|_{L_{\beta+1}^2(M)} \leqslant \epsilon,$$

u is a hyperbolic embedding.

Proof. By the Sobolev lemma, there is a fixed constant C_β such that

$$|j^2(u)(x) - j^2(u_0)(x)| \leqslant C_\beta|u - u_0|_{L_{\beta+1}^2(M)} \leqslant C_\beta\epsilon, \qquad x \in M.$$

Here, we are using the norm on V induced by the standard one on $\mathbf{R}^3$.

By choosing ϵ small enough, u will clearly be an embedding. Moreover, since V_2 is open, we have $j^2(u): M \to V_2$. By (6.a.5), the conclusion of the lemma follows.   Q.E.D.

Fix $\beta \geqslant 3$ and $\epsilon > 0$ as determined by the lemma. Use ϵ to define the norms $| \ |_k$ for the scale $\{E_k\}$ as in (6.a.2). Then by construction, $D_\beta(u_0)$ contains only hyperbolic embeddings satisfying (6.a.11).

To use Lemma (5.b.2), $j^2(u)$ must map to a fixed compact set. The set we use is the following:

$$B(u) = \left\{ (u, u_r, u_{st}) \in V \,|\, (u, u_r, u_{st}) = j^2(u)(x) \text{ for some } u \in D_\beta(u_0), \, x \in M \right\}.$$

That $B(u_0)$ is in fact compact follows from the Sobolev lemma and the fact that M is compact.

We still need one more set of functions on V_2, namely the ones that give a basis for $\mathrm{II}^\perp$. Following §IV(b), we make the basis a symmetric third order tensor, and for convenience, we lower all the indices. Thus, given (u, u_r, u_{st}), set $(b_{ijk}(u, u_r, u_{st}))$ equal to the solution of the following system:

$$b_{ijk} = b_{jik} = b_{ikj}, \qquad 1 \leqslant i, j, k \leqslant 3$$

$$\sum_{i,j=1}^{3} b_{kij} H_{ij} = 0, \qquad 1 \leqslant k \leqslant 3.$$

At each point $(u, u_r, u_{st}) \in V_2$, this gives 26 independent equations for 27 unknowns $b_{ijk}(u, u_r, u_{st})$, determining the solution only up to an arbitrary scale factor. However, V_2—in particular, condition (3)—implies that the matrix $(b_{3ij}(u, u_r, u_{st}))$ has real, nonzero eigenvalues. We can then obtain well-defined smooth functions $b_{ijk}(u, u_r, u_{st})$ by demanding that the eigenvalue of (b_{3ij}) with the least magnitude be equal to 1. If (u, u_r, u_{st}) is a "symmetric hyperbolic" point, then $b_{3ih}(u, u_r, u_{st})$ will in fact then be positive definite.

We now begin with estimates on the solution of the "intrinsic", piece of the linearized system. Fix for the rest of the section an integer $\gamma \geqslant 9$.

(6.a.7) PROPOSITION. *Let* $u \in D_\gamma(y_0) \cap C^\infty(M, \mathbf{R}^6)$. *Given any* $h = h_{ij} dx^i dx^j \in C^\infty(S^2 T^* M)$, *there exists a smooth 1-form* $\varphi = \varphi_i dx^i$ *solving* (4.d.5) *and satisfying the following estimate*:

$$|\varphi|_{L_k^2(M)} \leqslant M_k \big(|u|_{L_{k+\gamma}^2(M)} |h|_{L_1^2(M)} + |h|_{L_k^2(M)} \big),$$

where M_k *is a constant independent of* u, h, *and* φ.

Proof. This uses Theorems G and G' with the operator

$$P = \sum_{k=1}^{3} A^k \frac{\partial}{\partial x^k} + B, \qquad A^k = (b_{kij}), \quad \text{and} \quad B = (b_{irs} \Gamma_{rs}^j).$$

We wish to solve the system $P\varphi = b_{krs} h_{rs}$. First, $|P|_{L_k^2(M)}$ is bounded in terms

of u:

$$|P|_{L^2_k(M)} = \sum_{j=1}^{3} |A^j|_{L^2_k(M)} + |B|_{L^2_k(M)} \leqslant \sum_{i,j,k=1}^{3} |b_{ijk}|_{L^2_k(M)} + \sum_{i,j=1}^{3} |b_{irs}\Gamma^j_{rs}|_{L^2_k(M)}$$

Since $b_{ijk} = \hat{b}_{ijk} \circ j^2(u)$ and $\Gamma^j_{rs} = \hat{\Gamma}^j_{rs} \circ j^2(u)$, Lemma (5.b.2) may be applied, obtaining:

$$|P|_{L^2_k(M)} \leqslant M'_k |u|_{L^2_{k+2}(M)}. \tag{6.a.8}$$

In particular, let $\alpha = \gamma - 2 > n + 3 = 6$. Then

$$|P|_{L^2_\alpha(M)} \leqslant M'_\alpha |u|_{L^2_\gamma(M)} \leqslant M',$$

where M' is a constant independent of $u \in D_\gamma(u_0)$.

Suppose P is symmetric hyperbolic. Then by construction, $\lambda(P) = 1$. Therefore, both (5.a.5) and (5.a.6) hold with $b = M'$, $a = 1$. Therefore, a solution φ satisfying (5.a.7) exists. Substituting (6.a.8) into the estimate proves the proposition when P is symmetric hyperbolic.

On the other hand if u_0 defines a strictly hyperbolic operator P_0, then any $u \in D_\gamma(u_0)$ will also define a strictly hyperbolic operator. Moreover, there exists a closed ball $\mathcal{A} \subset \mathcal{H}$ such that the operator P defined by an embedding $u \in D_\gamma(u_0)$ satisfies (5.a.8). Theorem G' and (6.a.8) now combine to prove the proposition.   Q.E.D.

In §IV(d), a solution y to (6.a.3) is obtained from φ by setting

$$y = y' + y'',$$

where

$$y' = g_{jk}^{-1} \varphi_k \frac{\partial u}{\partial x^j}$$

and at $x \in M$, $y''(x)$ is the unique solution to the following system of linear equations:

$$\left(\frac{\partial u}{\partial x^i}(x), y''(x) \right) = 0,$$

$$\left(\frac{\partial^2 u}{\partial x^j \partial x^k}(x), y''(x) \right) = \frac{1}{2}\left(\frac{\partial \varphi_j}{\partial x^k}(x) + \frac{\partial \varphi_k}{\partial x^j}(x) \right) - \Gamma^m_{jk}(x)\varphi_m(x) - h_{jk}(x),$$

$$1 \leqslant i, j, k \leqslant 3.$$

y'' can be written in the form:

$$y''(x) = \sum_{j,k=1}^{3} J_{jk}(x)\left[\frac{1}{2}\left(\frac{\partial \varphi_j}{\partial x^k} + \frac{\partial \varphi_k}{\partial x^j} \right) - \Gamma^m_{jk}\varphi_m - h_{jk} \right],$$

where the functions $J_{jk} \in C^\infty(M, \mathbf{R}^6)$ are defined in terms of $\partial u/\partial x^i$,

$\partial^2 u / \partial x^j \partial x^k$. In other words, there are fixed functions $\hat{J}_{jk} \in C^\infty(V_2, \mathbf{R}^6)$ such that $J_{jk} = \hat{J}_{jk} \circ j^2(u)$. The following proposition gives the last estimate needed for the Nash–Moser–Schwartz–Sergeraert theorem:

(6.a.9) PROPOSITION. *Given* $u \in D_\gamma(u_0) \cap C^\infty(M, \mathbf{R}^6)$, $\Phi'(u)$ *has a right inverse* $Q(u) \colon C^\infty(S^2 T^* M) \to C^\infty(M, \mathbf{R}^6)$ *such that given*

$$Q'(u)h = y,$$

the estimate (6.a.4) *holds for any fixed* $\alpha \geqslant \gamma + 1$.

Proof. Given $h \in C^\infty(S^2 T^* M)$, solve for y as described above, and set

$$Q(u)h = y.$$

We will prove that the L_k^2-norms of y' and y'' are bounded by the right-hand side of (6.a.10). This will prove the proposition.

$$|y'|_{L_k^2(M)} = \left| \left(g_{ij}^{-1} \frac{\partial u}{\partial x^j} \right) \varphi_i \right|_{L_k^2(M)}$$

$$\leqslant G_k \left(\sum_{i=1}^{3} \left| g_{ij}^{-1} \frac{\partial u}{\partial x^j} \right|_{L_k^2(M)} |\varphi|_{L^2(M)} + \left| g_{ij}^{-1} \frac{\partial u}{\partial x^j} \right|_{L^2(M)} |\varphi|_{L_k^2(M)} \right)$$

$$\leqslant C_k''' \left(|u|_{L_{k+2}^2(M)} |\varphi|_{L^2(M)} + |u|_{L_2^2(M)} |\varphi|_{L_k^2(M)} \right)$$

$$\leqslant C_k'' \left(|u|_{L_{k+2}^2(M)} |h|_{L_1^2(M)} + |u|_{L_{k+\gamma}^2(M)} |h|_{L_1^2(M)} + |h|_{L_k^2(M)} \right)$$

$$\leqslant C_k' \left(|u|_{L_{k+\gamma}^2(M)} |h|_{L_1^2(M)} + |h|_{L_k^2(M)} \right)$$

Replacing k by $k - \alpha + 1$, this becomes

$$|y'|_{k-\alpha} \leqslant C_k (|u|_k |h|_\alpha + |h|_k).$$

$$|y''|_{L_k^2(M)} \leqslant \sum_{i,j=1}^{3} \left| J_{ij} \left[\frac{1}{2} \left(\frac{\partial \varphi_i}{\partial x^j} + \frac{\partial \varphi_j}{\partial x^i} \right) - \Gamma_{ij}^m \varphi_m - h_{ij} \right] \right|_{L_k^2(M)}$$

$$\leqslant G_k \left[\sum_{i,j=1}^{3} |J_{ij}|_{L_k^2(M)} \left| \frac{1}{2} \left(\frac{\partial \varphi_i}{\partial x^j} + \frac{\partial \varphi_j}{\partial x^i} \right) - \Gamma_{ij}^m \varphi_m - h_{ij} \right|_{L^2(M)} \right.$$

$$\left. + \sum_{i,j=1}^{3} |J_{ij}|_{L^2(M)} \left| \frac{1}{2} \left(\frac{\partial \varphi_i}{\partial x^j} + \frac{\partial \varphi_j}{\partial x^i} \right) - \Gamma_{ij}^m \varphi_m - h_{ij} \right|_{L_k^2(M)} \right]$$

$$\leqslant C_k'''' \left[|u|_{L_{k+2}^2(M)} \left(|\varphi|_{L_1^2(M)} + |h|_{L^2(M)} \right) \right.$$

$$\left. + |u|_{L^2(M)} \left(|u|_{L_{k+2}^2(M)} |\varphi|_{L^2(M)} + |\varphi|_{L_{k+1}^2(M)} + |h|_{L_k^2(M)} \right) \right]$$

$$\leqslant C_k''' \left[|u|_{L_{k+2}^2(M)} |h|_{L_1^2(M)} + |u|_{L_{k+2}^2(M)} |h|_{L_1^2(M)} \right.$$

$$\left. + |u|_{L_{k+\gamma+1}^2(M)} |h|_{L_1^2(M)} + |h|_{L_{k+1}^2(M)} \right]$$

Collecting terms and replacing k by $k - a + 1$, the estimate becomes

$$|y''|_{k-\alpha} \leqslant C_k''(|u|_\alpha + |h|_k). \quad \text{Q.E.D.}$$

All of the assumptions of the Nash–Moser–Schwartz–Sergeraert theorem have been confirmed; Theorem H then follows immediately.

Theorems F and H now combine to give us the Main Theorem which we restate here using some different letters to minimize confusion (remember that $M = \Omega \times [\frac{1}{4}, \frac{3}{4}]$).

MAIN THEOREM. *Let S be a smooth Riemannian three-manifold. Let $p \in S$ be such that the Einstein tensor at $p, \mathscr{G}(p)$, is neither zero nor a perfect square $L^2 = l_i l_j dx^i dx^j$, $L = l_i dx^i \in T_p^* S$. Then there exists a neighborhood N of p with a smooth isometric embedding*

$$u : N \to \mathbb{R}^6.$$

Proof. Let $g = g_{ij} dx^i dx^j$ be the metric on S. The basic idea is to find a smooth hyperbolic embedding u_0 of some neighborhood of p such that $\Phi(u_0)$ is very close to g. Theorem H would then give the desired conclusion. However, this argument does not quite work because we are varying the embedding u_0 which is fixed in Theorem H. Therefore, we don't know what happens to the δ given by Theorem H as u_0 or the neighborhood is varied in order to make g very close to $\Phi(u_0)$. The trick is to fix u_0 appropriately and then to define a new metric g' which agrees with g in a small neighborhood of p *and* is sufficiently close to $\Phi(u_0)$.

We begin by solving formally for the Taylor series of a map $u_0 : S \to \mathbb{R}^6$ such that $\Phi(u_0)$ agrees with g at p up to infinite order. We do this so that $j^2(u)(p) \in V_2$; Theorem F shows that this can be done. That the series exists is a consequence of the proof of the local analytic isometric embedding theorem which uses the Cauchy–Kowalewski theorem; here, we only need the formal part of the Cauchy–Kowalewski theorem, without considering convergence. Given such a Taylor series, we can then use the Borel theorem to find a neighborhood N of p and a smooth embedding $u_0 : N \to \mathbb{R}^6$ such that the following hold:

(1) $\Phi(u_0)$ agrees with g up to arbitrarily high order at p.

(2) $N \cong M$, as defined in §IV(a).

(3) u_0 is hyperbolic.

Now let $\varphi : \mathbb{R} \to \mathbb{R}$ be a smooth function such that

$$\varphi|_{[0,1/3]} \equiv 1 \quad \text{and} \quad \varphi|_{[2/3,\infty]} \equiv 0.$$

Given $0 < \rho < 1$, let $\varphi_\rho(t) = \varphi(t/\rho)$.

Now fix co-ordinates $x : N \to B \times [-1, 1]$, where B is the closed unit ball in $\mathbb{R}^2$ and $x(p) = 0$. For any $\rho > 0$, define a metric

$$g_\rho(x) = \varphi_\rho(|x|) g(x) + (1 - \varphi_\rho(|x|)) g_0(x)$$

Observe that on a sufficiently small neighborhood of p, $g_\rho = g$. Therefore, if we can isometrically embed N with the metric g_ρ, for some $\rho > 0$, we will have proved the Main Theorem. The existence of such an embedding is given by the following proposition and Theorem H.

(6.b.1) PROPOSITION. *Given a positive integer k_0 and a positive real number δ, there exists $\rho > 0$ such that*

$$|g_\rho - g_0|_{L^2_{k_0}(N)} < \delta.$$

Proof.

$$|g_\rho - g_0|_{L^2_{k_0}(N)} = |\varphi_\rho(|x|)(g - g_0)|_{L^2_{k_0}(N)}$$

$$= |\varphi_\rho(g - g_0)|_{L^2_{k_0}(B_\rho)}, \qquad \text{where } B_\rho = \text{ball of radius } \rho,$$

$$\leqslant \left[|\varphi_\rho|_{L^2_{k_0}(B_\rho)}|g - g_0|_{L^2(B_\rho)} + |\varphi_\rho|_{L^2(B_\rho)}|g - g_0|_{L^2_{k_0}(B_\rho)} \right]$$

We need to show that each term on the right-hand side goes to zero as $\rho \to 0$. For the second term, it suffices to observe that

$$|\varphi_\rho|_{L^2(B_\rho)} \leqslant \text{vol}(B_\rho) = \pi\rho^3$$

and $|g - g_0|_{L^2_k(B_\rho)} \leqslant |g - g_0|_{L^2_{k_0}(N)}$, a fixed quantity. On the other hand,

$$|\varphi_\rho|_{L^2_{k_0}(B_\rho)} \leqslant C_{k_0} \sum_{|\alpha| \leqslant k_0} |D^\alpha \varphi_\rho|_{L^2(B_\rho)} \leqslant C_{k_0} \sum_{|\alpha| \leqslant k_0} \rho^{-|\alpha|}\left(\sup_{x \in B} |D^\alpha \varphi(|x|)| \right)\text{vol}(B_\rho)$$

$$\leqslant C''\rho^{-k_0+3}, \qquad \text{for} \quad \rho < 1.$$

Since $g - g_0$ vanishes to infinite order at $x = 0$, we also have the following:

$$|g(x) - g_0(x)| \leqslant C_k |x|^k, \qquad k \geqslant 0.$$

Therefore, the first term can be estimated as follows:

$$|\varphi_\rho|_{L^2_{k_0}(B_\rho)}|g - g_0|_{L^2(B_\rho)} \leqslant C''\rho^{-k_0+3}\left(\sup_{x \in B_\rho} |g(x) - g_0(x)| \right)\text{vol}(B_\rho)$$

$$\leqslant C'_k\rho^{k-k_0+6}, \qquad \text{for any} \quad k \geqslant 0.$$

In particular, if we choose $k > k_0 - 6$, we see that the first term also becomes arbitrarily small as $\rho \to 0$.   Q.E.D.

Corrections added in proof.

A. These statements and proofs of Theorems G and G′ are not exactly correct, because Lemmas (5.b.1) and (5.b.2) are incorrect. When the correct versions of the lemmas are used, there is a larger loss of derivatives than indicated here. This, however, does not affect the proofs of Theorem H and the Main Theorem.

B. The estimates given here are incorrect. n derivatives should be added to all the norms of f. The

proof given is clearly wrong, but, using the Sobolev lemma, is easily modified to give the corrected estimate.

C. The statement and proof of this lemma are wrong. If the proof is corrected using the Hölder inequality and the Sobolev lemma, the estimate given is correct when the L_k^2 norm of u is replaced by the L_{k+n}^2 norm.

References

1. J. F. ADAMS, P. D. LAX, AND R. S. PHILLIPS, *On matrices whose real linear combinations are nonsingular*, Proceedings of Amer. Math. Soc. **16** (1965), 318–322.
2. E. ARBARELLO, M. CORNALBA, P. GRIFFITHS, AND J. HARRIS, *Geometry of algebraic curves*, Springer-Verlag Ergebnisse Series, to appear in 1984.
3. E. BERGER, R. BRYANT, AND P. GRIFFITHS, *Some isometric embedding and rigidity results for Riemannian manifolds*, Proc. Nat. Acad. Sci. **78** (1981), 4657–4660.
4. E. BERGER, R. BRYANT, AND P. GRIFFITHS, *Characteristics and rigidity of isometric embeddings*, Duke Math. J. **50** (1983), 803–892.
5. R. BRYANT, P. GRIFFITHS, AND D. YANG, "Characteristic varieties of exterior differential systems", to appear in *Essays on Exterior Differential Systems* with S. S. Chern, R. Gardner, and H. Goldschmidt.
6. K. O. FRIEDERICHS, *Symmetric hyperbolic linear differential equations*, Comm. Pure Appl. Math. **7** (1954), 345–392.
7. ROBERT E. GREENE, *Isometric embeddings of Riemannian and pseudo-Riemannian manifolds*, Memoirs of the Amer. Math. Soc. **97** (1970).
8. P. GRIFFITHS AND J. HARRIS, *Principles of Algebraic Geometry*, J. Wiley and Sons, New York, 1978.
9. HAROLD HILTON, *Plane Algebraic Curves*, Oxford University Press, London, 1932.
10. LARS HÖRMANDER, *Linear Partial Differential Operators*, Springer-Verlag, New York, 1963.
11. HOWARD JACOBOWITZ, *Local isometric embeddings*, Ann. of Math. Studies **102** (1982), 381–393.
12. E. KANEDA AND N. TANAKA, *Rigidity for isometric embeddings*, J. Math. Kyoto U. **18** (1978), 1–70.
13. SERGIU KLAINERMAN, *Global existence for nonlinear wave equations*, Comm. Pure Appl. Math. **33** (1980), 43–101.
14. S. KLEIMAN AND J. LANDOLFI, *Geometry and deformation of special Schubert varieties*, Compositio Math. **23** (1971), 407–434.
15. JURGEN MOSER, *A new technique for the construction of solutions of nonlinear differential equations*, Proc. Nat. Acad. Sci. **47** (1961), 1824–1831.
16. J. NASH, *The imbedding problem for Riemannian manifolds*, Ann. of Math. **63** (1956), 20–63.
17. J. T. SCHWARTZ, *Nonlinear Functional Analysis*, New York University, New York, 1964.
18. J. T. SCHWARTZ, *On Nash's implicit functional theorem*, Comm. Pure Appl. Math. **13** (1960), 509–530.
19. M. F. SERGERAERT, *Une généralisation du théorème des fonctions implicites de Nash*, Comptes Rendus Acad. Sc. Paris **270A** (1970), 861–863.
20. J.-P. SERRE, *Faisceaux algébriques cohérents*, Ann. of Math. **61** (1955), 197–278.
21. ELIAS STEIN, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.
22. SHLOMO STERNBERG, Lecture notes on PDE's from course given at University of Pennsylvania, 1967.
23. MICHAEL TAYLOR, *Pseudodifferential Operators*, Princeton University Press, Princeton, 1981.
24. KETI TENENBLAT, *A rigidity theorem for three dimensional submanifolds in Euclidean six-space*, J. Diff. Geom. **14** (1979), 187–203.

GRIFFITHS: DEPARTMENT OF MATHEMATICS, HARVARD UNIVERSITY, CAMBRIDGE, MASS. 02138
BRYANT: DEPARTMENT OF MATHEMATICS, RICE UNIVERSITY, HOUSTON, TEXAS 77001
YANG: COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, NEW YORK 10012