J. DIFFERENTIAL GEOMETRY
12 (1977) 1-45

DEFORMATION OF COMPLEX STRUCTURES ON MANIFOLDS WITH BOUNDARY. I: THE STABLE CASE

RICHARD S. HAMILTON

This is the first of a series of papers generalizing the theory of deformation of complex structures (which can be found for example in Morrow and Kodaira [7]) to the case of manifolds with boundary. It is necessary to impose some mild restrictions on the number of negative eigenvalues of the Levi form on the boundary in order to guarantee the finite dimensionality of certain cohomology groups (as in Kohn and Folland [1]); aside from this the results will be completely general. In this paper we consider only the stable case $H^1(X; \mathscr{T}X) = 0$, where $\mathscr{T}X$ is the holomorphic tangent bundle, so that all deformations are trivial. In the second paper we discuss in very general terms families of linear non-coercive boundary value problems and develop the required estimates and operators to make the theorems in this paper work. In the third paper we will discuss the extension of complex structures across the boundary. The fourth paper will deal with the general case $H^1(X; \mathscr{T}X) \neq 0$ and the construction of a universal family.

In this paper we prove the following result. Let Y be a complex manifold and X a compact subset whose boundary ∂X is smooth. We suppose that the Levi form on ∂X never has exactly one negative eigenvalue; that is, either all are strictly positive or else at least two are strictly negative. This implies that

$$\dim H^1(X; \mathscr{T}X) < \infty .$$

Theorem. *Suppose $H^1(X; \mathscr{T}X) = 0$. Then for any complex structure μ on X sufficiently close to the given structure we can find a map $f: X \to Y$ close to the identity so that f is analytic from X with the new structure μ to Y with the given structure. Thus any small deformation of the complex structure on X can be induced by a small motion of X in Y.*

To be precise, an almost complex structure μ on X is represented by a vector valued one-form $\mu^\alpha_{\bar\beta}$ which is a section of the bundle $L(\bar{\mathscr{T}}X, \mathscr{T}X)$. In the above theorem μ and f are C^∞ (smooth) functions on X, up to and including ∂X. A complex structure on X means an integrable almost complex structure, one with $\bar\partial\mu - \frac{1}{2}[\mu, \mu] = 0$. The conclusion holds for all μ sufficiently close to 0;

Received February 18, 1975. Research partially supported by NSF grant #GP-28251A3.

that is, for all μ in a neighborhood $\|\mu\|_{C^r(X)} < \varepsilon$ in the topology of $C^\infty(X)$. Thus it is only necessary that some finite number of derivatives be small. That f is analytic on X with the structure μ means that f satisfies the Cauchy-Riemann equation $\partial f^{-1} \circ \bar{\partial} f = \mu$. Hence proving the theorem amounts to solving a non-linear over-determined subelliptic boundary value problem. This is done using a generalization of the Nash-Moser inverse function theorem [2]. The work involved is to prove estimates on how the solution of the $\bar{\partial}_\mu$-Neumann problem depends on the complex structure μ.

We mention two applications. The first is when Y is a Stein manifold and ∂X is strictly pseudo-convex. In this case $H^1(X; \mathscr{T}X) = 0$ automatically and the theorem applies. This case was considered previously in [3]. The second is when $Y = \boldsymbol{C}^n$ and X is the region $1 \leq |z| \leq 2$ between two balls. If $n \geq 3$ then the Levi form has $n - 1 \geq 2$ strictly negative eigenvalues on the inner boundary and all strictly positive on the outer boundary. Moreover $H^1(X; \mathscr{T}X) = 0$, so again the theorem applies. Hence any small deformation of the complex structure cannot grow an isolated singularity inside. The situation is quite different in $\boldsymbol{C}^2$; see Rossi [8].

We wish especially to thank Masatake Kuranishi for his invaluable assistance in the preparation of this series of papers. His article [6], dealing with the parallel case of deformation of complex structures on the boundary, has provided a model for our case. We have borrowed several important ideas from that paper; in particular, the treatment of nonzero cohomology groups using spectral theory, and the use of approximate splittings of cohomology sequences in connection with the Nash-Moser inverse function theorem. We also wish to thank J. J. Kohn, who suggested that the results in [3] for strictly pseudoconvex domains extend to the case of sufficiently many negative eigenvalues of the Levi form.

1. Deformation of complex structures

1.1. Complex structures on vector spaces. We begin with some linear algebra. Let E be a real vector space of finite dimension. Write $\boldsymbol{C}E = \boldsymbol{C} \otimes_{\boldsymbol{R}} E$ for the complexification of E. There is a natural real-linear inclusion $E \xrightarrow{j} \boldsymbol{C}E$ given by $v \to 1 \otimes v$. There is also a natural conjugation $\boldsymbol{C}E \to \boldsymbol{C}E$ given by $\overline{c \otimes v} = \bar{c} \otimes v$. The image $j(E) \subseteq \boldsymbol{C}E$ is the subspace $\boldsymbol{R}E$ of real vectors, those which are self-conjugate.

If E also has the structure of a complex vector space, there is a natural multiplication $m: \boldsymbol{C}E \to E$ given by $m(c \otimes v) = cv$. The kernel is a complex-linear subspace which we call $\bar{\mathscr{E}}$, for reasons which will become clear. There is a natural exact sequence

$$0 \longrightarrow \bar{\mathscr{E}} \longrightarrow \boldsymbol{C}E \xrightarrow{m} E \longrightarrow 0 .$$

Since the composition $E \xrightarrow{j} CE \xrightarrow{m} E$ is the identity, $\bar{\mathscr{E}}$ is complementary to $\boldsymbol{R}E$.

Conversely, suppose E is a real vector space, and we are given a complex linear subspace $\bar{\mathscr{E}}$ of $\boldsymbol{C}E$ complementary to $\boldsymbol{R}E$. Then there exists a unique complex structure on E such that $\bar{\mathscr{E}}$ is the kernel of the multiplication $m\colon \boldsymbol{C}E \to E$. To show the uniqueness, observe that if $\bar{\mathscr{E}} = \ker m$ then the map

$$E \xrightarrow{j} \boldsymbol{C}E \xrightarrow{\bar{\pi}} \boldsymbol{C}E/\bar{\mathscr{E}}$$

is a complex linear isomorphism, since $1 \otimes iv - i \otimes v \in \bar{\mathscr{E}} = \ker m$ for any $v \in E$. To show the existence for a given $\bar{\mathscr{E}}$, we give E the complex structure which makes the real-linear isomorphism $\bar{\pi} j$ a complex-linear isomorphism. This means that for each $v \in E$ we define iv to be the unique element with $1 \otimes iv - i \otimes v \in \bar{\mathscr{E}}$. Then these elements span both $\bar{\mathscr{E}}$ and $\ker m$, so $\bar{\mathscr{E}} = \ker m$ for this complex structure.

The space $S(E)$ of complex structures on E can therefore be identified with an open subset of a Grassmannian manifold. It is convenient to have natural local coordinates on $S(E)$ in a neighborhood if a reference point $\bar{\mathscr{E}} \in S(E)$. For this it is necessary to choose a complex-linear complementary subspace. Fortunately there is a natural way to do this. Namely, if $\bar{\mathscr{E}} \in S(E)$ then the conjugate of $\bar{\mathscr{E}}$ is a complex linear subspace $\mathscr{E}$ of $\boldsymbol{C}E$ complementary to $\bar{\mathscr{E}}$, and we have a direct sum decomposition

$$\boldsymbol{C}E = \mathscr{E} \oplus \bar{\mathscr{E}} \ .$$

Note that there is a natural complex-linear isomorphism of $\mathscr{E}$ to E given by

$$\mathscr{E} \longrightarrow \boldsymbol{C}E \xrightarrow{m} E \ .$$

Other authors sometimes write E' and E'' instead of $\mathscr{E}$ and $\bar{\mathscr{E}}$. Suppose then that $\bar{\mathscr{E}} \in S(E)$ is a reference structure. There is a local one-to-one correspondence between complex structures $\bar{\mathscr{E}}_\mu$ near $\bar{\mathscr{E}}$ and small complex-linear maps

$$\mu\colon \bar{\mathscr{E}} \to \mathscr{E} \quad \text{given by} \quad \bar{\mathscr{E}}_\mu = \{v - \mu v : v \in \bar{\mathscr{E}}\} \ .$$

If π and $\bar{\pi}$ are the projections of $\boldsymbol{C}E$ onto $\mathscr{E}$ and $\bar{\mathscr{E}}$, then μ is determined by the commutative diagram

$$\begin{tikzcd} & \bar{\mathscr{E}}_\mu \arrow[dl, "-\bar{\pi}"'] \arrow[dr, "\pi"] & \\ \bar{\mathscr{E}} \arrow[rr, "\mu"'] & & \mathscr{E} \end{tikzcd}$$

since $\bar{\pi}\colon \bar{\mathscr{E}}_\mu \to \bar{\mathscr{E}}$ is an isomorphism when $\bar{\mathscr{E}}_\mu$ is close to $\bar{\mathscr{E}}$. The diagram commutes since for $v - \mu v \in \bar{\mathscr{E}}_\mu$

$$\mu\bar{\pi}(v - \mu v) = \mu v = -\pi(v + \mu v) \; .$$

Thus we have constructed natural local coordinates on the Grassmannian manifold $S(E)$ with values in the vector space $L(\bar{\mathscr{E}}, \mathscr{E})$.

1.2. Almost complex structures on manifolds. All the preceeding generalizes immediately to manifolds. If X is a manifold with tangent bundle TX we can form the complexified tangent bundle $\boldsymbol{C}TX = \boldsymbol{C} \otimes_{\boldsymbol{R}} TX$, and a complex structure on each fibre defines a subbundle $\bar{\mathscr{T}}X = \operatorname{Ker} m \colon \boldsymbol{C}TX \to TX$ and a direct sum decomposition $\boldsymbol{C}TX = \mathscr{T}X \oplus \bar{\mathscr{T}}X$ where $\mathscr{T}X$ is the conjugate of $\bar{\mathscr{T}}X$. An almost complex structure on X is defined as a smooth complex-linear subbundle $\bar{\mathscr{T}}X$ of $\boldsymbol{C}TX$ complementary to the real subbundle $\boldsymbol{R}TX$. By the previous argument an almost complex structure on X can be identified with a smooth section of the fibre bundle $S(TX)$ obtained by applying the functor S to each fibre. If $\bar{\mathscr{T}}X \in S(TX)$ is a reference structure, we can choose local coordinates on $S(TX)$ with values in the vector bundle $L(\bar{\mathscr{T}}X, \mathscr{T}X)$. Hence an almost complex structure μ close to the reference structure corresponds to a small smooth section μ of the vector bundle $L(\bar{\mathscr{T}}X, \mathscr{T}X)$.

If $z^1, \cdots, z^n$ are complex coordinates on X, the bundles $\mathscr{T}X$ and $\bar{\mathscr{T}}X$ are spanned by the $\partial/\partial z^\alpha$ and $\partial/\partial \bar{z}^\alpha$ respectively. An almost complex structure close to it is represented by a tensor $\mu = \mu_\beta^\alpha \, d\bar{z}^\beta \otimes \partial/\partial z^\alpha$, and $\bar{\mathscr{T}}X_\mu$ is spanned by $\partial/\partial \bar{z}^\beta + \mu_\beta^\alpha(\partial/\partial z^\alpha)$.

1.3. The integrability condition. Let μ be an almost complex structure corresponding to the subbundle $\bar{\mathscr{T}}X_\mu$. We say that μ is integrable if $\bar{\mathscr{T}}X_\mu$ is integrable. This means that for any two vector fields v and w with values in $\bar{\mathscr{T}}X_\mu$ the Lie bracket $[v, w]$ again has values in $\bar{\mathscr{T}}X_\mu$. Note that the Lie bracket is defined for complex valued vector fields, which are just sections of $\boldsymbol{C}TX$. In general there will be an obstruction $J(\mu)$ which is a smooth section of the bundle $\Lambda^2(\bar{\mathscr{T}}X_\mu, \mathscr{T}X_\mu)$ of alternating 2-forms on $\bar{\mathscr{T}}X_\mu$ with values in $\mathscr{T}X_\mu$, such that if v and w are smooth sections of $\bar{\mathscr{T}}X_\mu$ then

$$[v, w] \equiv J(\mu)(v, w) \qquad \operatorname{mod} \bar{\mathscr{T}}X_\mu \; .$$

Note that although the Lie bracket is an operator of degree 1, the error $J(\mu)(v, w)$ is an operator of degree 0, i.e., a pointwise multiplication. We can regard J as a partial differential operator of degree 1 as follows. $S(X)$ is a fibre bundle over X, and $\Lambda^2(\bar{\mathscr{T}}X_\mu, \mathscr{T}X_\mu)$ is a vector bundle over $S(X)$ whose fibre at $\mu(x)$ is $\Lambda^2(\bar{\mathscr{T}}X_{\mu(x)}, \mathscr{T}X_{\mu(x)})$. If μ is a section of $S(X)$, then $J(\mu)$ is a section of $\Lambda^2(\bar{\mathscr{T}}X_\mu, \mathscr{T}X_\mu)$ lying over μ.

We can compute $J(\mu)$ explicitly in terms of a ccomplex reference structure $z^1, \cdots, z^n$. It is more convenient to compute an equivalent tensor $Q(\mu)$ which is a section of $\Lambda^2(\bar{\mathscr{T}}X, \mathscr{T}X)$ defined by

$$J(\mu)(v, w) = \pi_\mu Q(\mu)(\bar{\pi}v, \bar{\pi}w) \; ,$$

where $\bar{\pi}: \bar{\mathscr{T}}X_\mu \to \bar{\mathscr{T}}X$ and $\pi_\mu: \mathscr{T}X \to \mathscr{T}X_\mu$ are isomorphisms induced by the projections

$$\mathscr{T}X \oplus \bar{\mathscr{T}}X \to \bar{\mathscr{T}}X , \qquad \mathscr{T}X_\mu \oplus \bar{\mathscr{T}}X_\mu \to \mathscr{T}X_\mu .$$

Suppose $\bar{\pi}v = p = p^\alpha(\partial/\partial\bar{z}^\alpha)$ and $\bar{\pi}w = q = q^\alpha(\partial/\partial\bar{z}^\alpha)$. Then

$$v = p^\alpha\left(\frac{\partial}{\partial\bar{z}^\alpha} + \mu_\alpha^\beta\frac{\partial}{\partial z^\beta}\right), \qquad w = q^\gamma\left(\frac{\partial}{\partial\bar{z}^\gamma} + \mu_\gamma^\delta\frac{\partial}{\partial z^\delta}\right),$$

and $Q(\mu)(p, q)$ is determined by

$$Q(\mu)(p, q) \equiv [v, w] \qquad \operatorname{mod} \bar{\mathscr{T}}X_\mu .$$

Now clearly

$$[v, w] \equiv p^\alpha q^\gamma\left[\frac{\partial}{\partial\bar{z}^\alpha} + \mu_\alpha^\beta\frac{\partial}{\partial z^\beta}, \frac{\partial}{\partial\bar{z}^\gamma} + \mu_\gamma^\delta\frac{\partial}{\partial z^\delta}\right] = p^\alpha q^\gamma Q(\mu)_{\alpha\gamma}^\theta\frac{\partial}{\partial z^\theta},$$

where

$$Q(\mu)_{\alpha\gamma}^\theta = \frac{\partial\mu_\gamma^\theta}{\partial\bar{z}^\alpha} - \frac{\partial\mu_\alpha^\theta}{\partial\bar{z}^\gamma} + \mu_\alpha^\beta\frac{\partial\mu_\gamma^\theta}{\partial z^\beta} - \mu_\gamma^\delta\frac{\partial\mu_\alpha^\theta}{\partial z^\delta} .$$

Thus $Q(\mu)$ is a nonlinear partial differential operator of degree 1. If $\Lambda^p(\bar{\mathscr{T}}X, \mathscr{T}X)$ denotes p-linear alternating forms on $\bar{\mathscr{T}}X$ with values in $\mathscr{T}X$, then μ is a section of $\Lambda^1(\bar{\mathscr{T}}X, \mathscr{T}X) = L(\bar{\mathscr{T}}X, \mathscr{T}X)$ and $Q(\mu)$ is a section of $\Lambda^2(\bar{\mathscr{T}}X, \mathscr{T}X)$, so

$$Q: \Lambda^1(\bar{\mathscr{T}}X, \mathscr{T}X) \to \Lambda^2(\bar{\mathscr{T}}X, \mathscr{T}X) ,$$

and μ is integrable if and only if $Q(\mu) = 0$.

If the almost complex structure μ is induced by a complex coordinate system $z^1, \cdots, z^n$, then clearly μ is integrable since $[\partial/\partial\bar{z}^\alpha, \partial/\partial\bar{z}^\beta] = 0$. The classical theorem of Newlander and Nirenberg asserts the converse; if μ is an integrable almost complex structure then μ is induced locally by a complex coordinate system. Thus an integrable almost complex manifold is a complex manifold, at least in the interior; this argument breaks down at the boundary.

1.4. The $\bar{\partial}_\mu$ complex. Let X be a complex manifold. A vector valued p-form is a section $\varphi \in C^\infty(X; \Lambda^p(\bar{\mathscr{T}}X, \mathscr{T}X)) = \lambda^p(X)$ given locally by

$$\varphi = \varphi_A^\alpha\, d\bar{z}^A \otimes \frac{\partial}{\partial z^\alpha} ,$$

where $A = (\alpha_1, \cdots, \alpha_p)$ is a multi-index, and $d\bar{z}^A = d\bar{z}^{\alpha_1} \wedge \cdots \wedge d\bar{z}^{\alpha_p}$, and the summation ranges over all strictly increasing indices A. We define ε_A^B to be ± 1 if A is a permutation of B of sign ± 1, and define ε_B^A to be 0 otherwise. The complex

$$\cdots \longrightarrow \lambda^{p-1}(X) \xrightarrow{\bar\partial} \lambda^p(X) \xrightarrow{\bar\partial} \lambda^{p+1}(X) \longrightarrow \cdots$$

is defined locally by the formula

$$\bar\partial \varphi_A^\alpha = \varepsilon_A^{\beta B} \partial \varphi_B^\alpha / \partial \bar z^\beta$$

with summation over β and all strictly increasing B. It is easy to check that this definition is invariant under a complex-analytic change of coordinates, and that $\bar\partial\bar\partial = 0$.

There is also a Lie bracket operation on the $\lambda^p(X)$ which agrees with the ordinary Lie bracket on $\lambda^\circ(X) = C^\infty(X; \mathscr{T}X)$ and acts as a combination Lie bracket and wedge product on higher order forms. If

$$\varphi = \varphi_A^\alpha \, d\bar z^A \otimes \frac{\partial}{\partial z^\alpha}, \qquad \psi = \psi_B^\beta \, d\bar z^B \otimes \frac{\partial}{\partial z^\beta},$$

we define

$$[\varphi, \psi] = \varepsilon_C^{AB}\left(\varphi_A^\alpha \frac{\partial \psi_B^\beta}{\partial z^\alpha} - \psi_B^\alpha \frac{\partial \varphi_A^\beta}{\partial z^\alpha}\right) d\bar z^C \otimes \frac{\partial}{\partial z^\beta} .$$

We can then verify the following rules (see Morrow and Kodaira [7, p. 152]). Let $p = \deg \varphi$, $q = \deg \psi$, $r = \deg \tau$. Then

$$[\psi, \varphi] = -(-1)^{pq}[\varphi, \psi] ,$$
$$\bar\partial[\varphi, \psi] = [\bar\partial\varphi, \psi] + (-1)^p[\varphi, \bar\partial\psi] ,$$
$$(-1)^{pr}[\varphi, [\psi, \tau]] + (-1)^{pq}[\psi, [\tau, \varphi]] + (-1)^{rq}[\tau, [\varphi, \psi]] = 0 .$$

The first is the antisymmetry relation, the second is the formula for the derivative of a product, and the third is Jacobi's identity. Using these formulas we can write the integrability condition as

$$Q(\mu) = \bar\partial\mu - \tfrac{1}{2}[\mu, \mu] .$$

Suppose now that μ is a complex structure, so that $Q(\mu) = 0$. Then writing $\lambda_\mu^p(X) = C^\infty(X; \Lambda^p(\bar{\mathscr{T}}X_\mu, \mathscr{T}X_\mu))$ we will have a complex $\bar\partial_{[\mu]}$ which is just $\bar\partial$ in the structure μ

$$\cdots \longrightarrow \lambda_\mu^{p-1}(X) \xrightarrow{\bar\partial_{[\mu]}} \lambda_\mu^p(X) \xrightarrow{\bar\partial_{[\mu]}} \lambda^{p+1}(X) \longrightarrow \cdots$$

and again $\bar\partial_{[\mu]}\bar\partial_{[\mu]} = 0$.

There is another complex associated to μ for any almost complex structure close to the reference structure. Namely, we define

$$\cdots \longrightarrow \lambda^{p-1}(X) \xrightarrow{\bar\partial_\mu} \lambda^p(X) \xrightarrow{\bar\partial_\mu} \lambda^{p+1}(X) \longrightarrow \cdots$$

by the formula

$$\bar{\partial}_\mu \varphi = \bar{\partial}\varphi - [\mu, \varphi] .$$

It is an easy consequence of the relations for the Lie bracket that

$$\bar{\partial}_\mu \bar{\partial}_\mu \varphi + [Q(\mu), \varphi] = 0 ,$$

where $Q(\mu) = \bar{\partial}\mu - \frac{1}{2}[\mu, \mu]$ is the integrability condition. Thus $\bar{\partial}_\mu \bar{\partial}_\mu = 0$ if and only if $Q(\mu) = 0$. We can also write the integrability condition as $Q(\mu) = \bar{\partial}_{\mu/2}\mu$. There is a simple relation between $\bar{\partial}_\mu$ and $\bar{\partial}_{[\mu]}$. Recall that we have isomorphisms

$$\bar{\pi} : \bar{\mathscr{T}}X_\mu \to CTX = \mathscr{T}X \oplus \bar{\mathscr{T}}X \to \bar{\mathscr{T}}X ,$$
$$\pi_\mu : \mathscr{T}X \to CTX = \mathscr{T}X_\mu \oplus \bar{\mathscr{T}}X_\mu \to \mathscr{T}X_\mu .$$

These induce an isomorphism

$$\Lambda^p(\bar{\pi}, \pi_\mu) : \Lambda^p(\bar{\mathscr{T}}X, \mathscr{T}X) \to \Lambda^p(\bar{\mathscr{T}}X_\mu, \mathscr{T}X_\mu) .$$

Write $c_\mu = C^\infty(X ; \Lambda^p(\bar{\pi}, \pi_\mu))$ for the induced isomorphism

$$c_\mu : \lambda^p(X) \to \lambda^p_\mu(X) .$$

Then the relation between $Q(\mu)$ and $J(\mu)$ is expressed by $J(\mu) = c_\mu Q(\mu)$.

Theorem 1. *There is a commutative diagram* $c_\mu \bar{\partial}_\mu = \bar{\partial}_{[\mu]} c_\mu$

$$\begin{tikzcd}
\lambda^p(X) \arrow[r, "\bar{\partial}_\mu"] \arrow[d, "c_\mu"'] & \lambda^{p+1}(X) \arrow[d, "c_\mu"] \\
\lambda^p_\mu(X) \arrow[r, "\bar{\partial}_{[\mu]}"] & \lambda^{p+1}_\mu(X)
\end{tikzcd}$$

for every complex structure μ.

1.5. Local coordinates.

Proof. Let $z^1, \cdots, z^n$ be complex coordinates in the reference structure, and $w^1, \cdots, w^n$ complex coordinates in the new structure μ. Then

$$\frac{\partial w^\alpha}{\partial \bar{z}^\beta} = \frac{\partial w^\alpha}{\partial z^\gamma}\mu^\gamma_\beta .$$

It follows that

$$\frac{\partial w^\alpha}{\partial z^\gamma}\left\{\frac{\partial z^\gamma}{\partial \bar{w}^\delta} + \mu^\gamma_\beta \frac{\partial \bar{z}^\beta}{\partial \bar{w}^\delta}\right\} = \frac{\partial w^\alpha}{\partial \bar{z}^\delta} = 0 .$$

When μ is small, $\partial w^\alpha / \partial z^\gamma$ will be invertible so we must have

$$\frac{\partial z^\tau}{\partial \bar{w}^\delta} + \mu^\tau_\beta \frac{\partial \bar{z}^\beta}{\partial \bar{w}^\delta} = 0 .$$

Now

$$\frac{\partial}{\partial \bar{w}^\tau} = \frac{\partial \bar{z}^\beta}{\partial \bar{w}^\tau}\frac{\partial}{\partial \bar{z}^\beta} + \frac{\partial z^\eta}{\partial \bar{w}^\tau}\frac{\partial}{\partial z^\eta} ,$$

so by the previous relation

$$\frac{\partial}{\partial \bar{w}^\tau} = \frac{\partial \bar{z}^\beta}{\partial \bar{w}^\tau}\left\{\frac{\partial}{\partial \bar{z}^\beta} - \mu^\eta_\beta \frac{\partial}{\partial z^\eta}\right\} ,$$

which is the fundamental relation in $\mathscr{T}X_\mu$.

Next observe that by definition

$$\bar{\pi}\frac{\partial}{\partial \bar{w}^\alpha} = \frac{\partial \bar{z}^\beta}{\partial \bar{w}^\alpha}\frac{\partial}{\partial \bar{z}^\beta} , \qquad \pi_\mu \frac{\partial}{\partial z^\alpha} = \frac{\partial w^\beta}{\partial z^\alpha}\frac{\partial}{\partial w^\beta} .$$

Suppose $\psi = c_\mu \varphi$. Write

$$\varphi = \varphi^\alpha_A \, d\bar{z}^A \otimes \frac{\partial}{\partial z^\alpha} , \qquad \psi = \psi^\beta_B \, d\bar{w}^B \otimes \frac{\partial}{\partial w^\beta} .$$

If $A = (\alpha_1, \cdots, \alpha_p)$ and $B = (\beta_1, \cdots, \beta_p)$, we put

$$\frac{\partial z^A}{\partial w^B} = \sum_\pi (-1)^\pi \frac{\partial z^{\pi(\alpha_1)}}{\partial w^{\beta_1}} \cdots \frac{\partial z^{\pi(\alpha_p)}}{\partial w^{\beta_p}} .$$

Then locally c_μ is expressed by

$$\psi^\beta_B = \frac{\partial \bar{z}^A}{\partial \bar{w}^B}\frac{\partial w^\beta}{\partial z^\alpha}\varphi^\alpha_A .$$

By definition

$$\bar{\partial}_{[\mu]}\psi^\beta_C = \varepsilon^{\tau B}_C \, \partial \psi^\beta_B / \partial \bar{w}^\tau ,$$

which in turn is a sum of three terms. The first of these is

$$\varepsilon^{\tau B}_C \frac{\partial \bar{z}^A}{\partial \bar{w}^B}\frac{\partial w^\beta}{\partial z^\alpha}\frac{\partial \varphi^\alpha_A}{\partial \bar{w}^\tau} = \varepsilon^{\tau B}_C \frac{\partial \bar{z}^\theta}{\partial \bar{w}^\tau}\frac{\partial \bar{z}^A}{\partial \bar{w}^B}\frac{\partial w^\beta}{\partial z^\alpha}\left\{\frac{\partial \varphi^\alpha_A}{\partial \bar{z}^\theta} - \mu^\eta_\theta \frac{\partial \varphi^\alpha_A}{\partial z^\eta}\right\} ,$$

using the formula for $\partial/\partial \bar{w}^\tau$. Next note that

$$\begin{aligned} \frac{\partial}{\partial \bar{w}^\tau}\frac{\partial w^\beta}{\partial z^\alpha} &= \frac{\partial \bar{z}^\theta}{\partial \bar{w}^\tau}\left\{\frac{\partial}{\partial \bar{z}^\theta} - \mu^\eta_\theta \frac{\partial}{\partial z^\eta}\right\}\frac{\partial w^\beta}{\partial z^\alpha} \\ &= \frac{\partial \bar{z}^\theta}{\partial \bar{w}^\tau}\left[\frac{\partial}{\partial z^\alpha}\left\{\frac{\partial w^\beta}{\partial \bar{z}^\theta} - \mu^\eta_\theta \frac{\partial w^\beta}{\partial z^\eta}\right\} + \frac{\partial \mu^\eta_\theta}{\partial z^\alpha}\frac{\partial w^\beta}{\partial z^\eta}\right] = \frac{\partial \bar{z}^\theta}{\partial \bar{w}^\tau}\frac{\partial w^\beta}{\partial z^\eta}\frac{\partial \mu^\eta_\theta}{\partial z^\alpha} , \end{aligned}$$

since the term in braces is zero. Therefore the second term is

$$\varepsilon_C^{\gamma B} \frac{\partial \bar{z}^\theta}{\partial \bar{w}^\gamma} \frac{\partial \bar{z}^A}{\partial \bar{w}^B} \frac{\partial w^\beta}{\partial z^\eta} \frac{\partial \mu_\theta^\eta}{\partial z^\alpha} \varphi_A^\alpha .$$

For the third term observe that

$$\frac{\partial}{\partial \bar{w}^\gamma} \frac{\partial \bar{z}^A}{\partial \bar{w}^B} = \varepsilon^A_{\rho R} \varepsilon_B^{\sigma S} \frac{\partial \bar{z}^R}{\partial \bar{w}^S} \frac{\partial^2 \bar{z}^\rho}{\partial \bar{w}^\gamma \partial \bar{w}^\sigma} .$$

Hence the third term is

$$\varepsilon_C^{\gamma B} \varepsilon^A_{\rho R} \varepsilon_B^{\sigma S} \frac{\partial \bar{z}^R}{\partial \bar{w}^S} \frac{\partial^2 \bar{z}^\rho}{\partial \bar{w}^\gamma \partial \bar{w}^\sigma} \frac{\partial w^\beta}{\partial z^\alpha} \varphi_A^\alpha = \varepsilon_C^{\gamma\sigma S} \frac{\partial^2 \bar{z}^\rho}{\partial \bar{w}^\gamma \partial \bar{w}^\sigma} \varepsilon^A_{\rho R} \frac{\partial \bar{z}^R}{\partial \bar{w}^S} \frac{\partial w^\beta}{\partial z^\alpha} \varphi_A^\alpha = 0 ,$$

because $\varepsilon_C^{\gamma\sigma S}$ is antisymmetric in γ and σ, while $\partial^2 \bar{z}^\rho / \partial \bar{w}^\gamma \partial \bar{w}^\sigma$ is symmetric, and the summation convention applies. Also we observe that

$$\varepsilon_C^{\gamma B} \frac{\partial \bar{z}^\theta}{\partial \bar{w}^\gamma} \frac{\partial \bar{z}^A}{\partial \bar{w}^B} = \varepsilon_D^{\theta A} \frac{\partial \bar{z}^D}{\partial \bar{w}^C} .$$

Hence we have

$$\bar{\partial}_{[\mu]} \psi_B^\beta = \varepsilon_D^{\theta A} \frac{\partial \bar{z}^D}{\partial \bar{w}^C} \frac{\partial w^\beta}{\partial z^\alpha} \left\{ \frac{\partial \varphi_A^\alpha}{\partial \bar{z}^\theta} - \mu_\theta^\eta \frac{\partial \varphi_A^\alpha}{\partial z^\eta} + \frac{\partial \mu_\theta^\alpha}{\partial z^\eta} \varphi_A^\eta \right\} ,$$

$$\bar{\partial}_{[\mu]} \psi_B^\beta = \frac{\partial \bar{z}^D}{\partial \bar{w}^C} \frac{\partial w^\beta}{\partial z^\alpha} \{ \bar{\partial} \varphi - [\mu, \varphi] \}_D^\alpha .$$

Therefore $\bar{\partial}_{[\mu]} \psi = c_\mu \{ \bar{\partial} \varphi - [\mu, \varphi] \}$ if $\psi = c_\mu \varphi$. This proves the theorem.

1.6. Induced complex structures. Suppose now that Y is a complex manifold and X is a compact subset with smooth boundary ∂X. If $f \colon X \to Y$ is close to the identity we define the induced structure

$$\mu = P(f) = \partial f^{-1} \circ \bar{\partial} f ,$$

or equivalently $\bar{\partial} f = \partial f \circ \mu$. Here $df \colon TX \to TY$ has complexification $Cdf \colon CTX \to CTY$, and under the direct sum decompositions $CTX = \mathscr{T}X \oplus \bar{\mathscr{T}}X$ and $CTY = \mathscr{T}Y \oplus \bar{\mathscr{T}}Y$ the map is represented by a matrix

$$Cdf = \begin{pmatrix} \partial f & \bar{\partial} f \\ \partial \bar{f} & \overline{\partial f} \end{pmatrix} ,$$

where $\partial f \colon \mathscr{T}X \to \mathscr{T}Y$ and $\bar{\partial} f \colon \bar{\mathscr{T}}X \to \mathscr{T}Y$. Thus $\mu = \partial f^{-1} \circ \bar{\partial} f \colon \bar{\mathscr{T}}X \to \mathscr{T}X$ is a complex structure. In local coordinates

$$\frac{\partial f^\alpha}{\partial \bar{z}^\beta} = \frac{\partial f^\alpha}{\partial z^\gamma} \mu^\gamma_\beta \, .$$

If we write $\mathscr{F}(X, Y)$ for the manifold of maps of X into Y, then M is a non-linear partial differential operator of degree 1:

$$M: \mathscr{F}(X, Y) \to \lambda^1(X) \, .$$

We wish to compute the derivative $DM(f)g$. Here $g \in T_f\mathscr{F}(X, Y)$ is an infinitesimal variation in the map f, which can be regarded as a section of the pull-back bundle $f^*\mathscr{T}Y$. Suppose $DM(f)g = \nu$. Then a variation of g in f must accompany a variation of ν in μ. Applying this in local coordinates to the equation

$$\frac{\partial f^\alpha}{\partial \bar{z}^\beta} = \frac{\partial f^\alpha}{\partial z^\gamma} \mu^\gamma_\beta \, ,$$

we must have

$$\frac{\partial g^\alpha}{\partial \bar{z}^\beta} = \frac{\partial g^\alpha}{\partial z^\gamma} \mu^\gamma_\beta + \frac{\partial f^\alpha}{\partial z^\gamma} \nu^\gamma_\beta \, .$$

Define $\chi \in \lambda^\circ(X)$ by the equation

$$g = \partial f \circ \chi \, ,$$

which in local coordinates is

$$g^\alpha = \frac{\partial f^\alpha}{\partial z^\theta} \chi^\theta \, .$$

If f is near the identity, then $\partial f^\alpha / \partial z^\gamma$ is invertible so g determines χ. We have

$$\frac{\partial f^\alpha}{\partial z^\theta} \frac{\partial \chi^\theta}{\partial \bar{z}^\beta} + \frac{\partial^2 f^\alpha}{\partial \bar{z}^\beta \partial z^\theta} \chi^\theta = \frac{\partial f^\alpha}{\partial z^\theta} \frac{\partial \chi^\theta}{\partial z^\gamma} \mu^\gamma_\beta + \frac{\partial^2 f^\alpha}{\partial z^\gamma \partial z^\theta} \chi^\theta \mu^\gamma_\beta + \frac{\partial f^\alpha}{\partial z^\gamma} \nu^\gamma_\beta \, .$$

However

$$\frac{\partial f^\alpha}{\partial \bar{z}^\beta} = \frac{\partial f^\alpha}{\partial z^\gamma} \mu^\gamma_\beta \, ,$$

so differentiating with respect to z^θ

$$\frac{\partial^2 f^\alpha}{\partial \bar{z}^\beta \partial z^\theta} = \frac{\partial^2 f^\alpha}{\partial z^\gamma \partial z^\theta} \mu^\gamma_\beta + \frac{\partial f^\alpha}{\partial z^\gamma} \frac{\partial \mu^\gamma_\beta}{\partial z^\theta} \, .$$

If we interchange γ and θ in some terms, we have

$$\frac{\partial f^\alpha}{\partial z^\theta}\left\{\frac{\partial \chi^\theta}{\partial \bar{z}^\beta} - \frac{\partial \chi^\theta}{\partial z^\tau}\mu^\tau_\beta + \frac{\partial \mu^\theta_\beta}{\partial z^\tau}\chi^\tau\right\} = \frac{\partial f^\alpha}{\partial z^\theta}\nu^\theta_\beta .$$

Since $\partial f^\alpha / \partial z^\theta$ is invertible for f near the identity, we have

$$\frac{\partial \chi^\theta}{\partial \bar{z}^\beta} - \frac{\partial \chi^\theta}{\partial z^\tau}\mu^\tau_\beta + \frac{\partial \mu^\theta_\beta}{\partial z^\tau}\chi^\tau = \nu^\theta_\beta .$$

But this is just the local expression for

$$\bar{\partial}_\mu \chi = \bar{\partial}\chi - [\mu, \chi] = \nu .$$

Hence the derivative of M is given by

$$DM(f)g = \bar{\partial}_\mu(\partial f^{-1} \circ g) .$$

The manifold of maps $\mathscr{F}(X, Y)$ is modeled on the vector space $\lambda^\circ(X)$ near the identity. To accomplish this we choose a spray $\sigma: \mathscr{T}X \to Y$ and define the local coordinate chart

$$S: \lambda^\circ(X) \to \mathscr{F}(X, Y)$$

by composition

$$S(\varphi) = \sigma \circ \varphi = f .$$

In local coordinates

$$f^\alpha(z^\beta) = \sigma^\alpha(z^\beta, \varphi^\tau(z^\beta)) ,$$

where the $\sigma^\alpha(z^\beta, v^\tau)$ are functions of variables $z^1, \cdots, z^n$ on X and $v^1, \cdots, v^n$ defining the tangent directions, i.e., $v^\tau = dz^\tau$. We can make $\sigma(z, 0) = z$, $(\partial\sigma^\alpha / \partial v^\tau)(z, 0) = \delta^\alpha_\tau$ and $(\partial\sigma^\alpha / \partial \bar{v}^\tau)(z, 0) = 0$ by a suitable choice of σ. The map S has a derivative

$$DS(\varphi)\psi = g$$

given in local coordinates by

$$g^\alpha = \frac{\partial \sigma^\alpha}{\partial v^\tau}(z, \varphi)\psi^\tau + \frac{\partial \sigma^\alpha}{\partial \bar{v}^\tau}(z, \varphi)\bar{\psi}^\tau .$$

The composition $P = MS$ is a nonlinear partial differential operator of degree 1:

$$P: \lambda^\circ(X) \to \lambda^1(X) .$$

Its derivative is given by the Chain Rule

$$DP(\varphi)\psi = DM(f)DS(\varphi)\psi = \nu \ .$$

Recall that

$$g^\alpha = \frac{\partial f^\alpha}{\partial z^\theta}\chi^\theta \ , \qquad f^\alpha = \sigma^\alpha(z^\beta, \varphi^\gamma(z^\beta)) \ ,$$

$$\frac{\partial f^\alpha}{\partial z^\theta} = \frac{\partial \sigma^\alpha}{\partial z^\theta}(z, \varphi) + \frac{\partial \sigma^\alpha}{\partial v^\gamma}(z, \varphi)\frac{\partial \varphi^\gamma}{\partial z^\theta} + \frac{\partial \sigma^\alpha}{\partial \bar{v}^\gamma}(z, \varphi)\frac{\partial \bar{\varphi}^\gamma}{\partial z^\theta} \ .$$

Therefore ψ and χ are related by the equation

$$\frac{\partial \sigma^\alpha}{\partial v^\gamma}\psi^\gamma + \frac{\partial \sigma^\alpha}{\partial \bar{v}^\gamma}\bar{\psi}^\gamma = \left\{\frac{\partial \sigma^\alpha}{\partial z^\theta} + \frac{\partial \sigma^\alpha}{\partial v^\gamma}\frac{\partial \varphi^\gamma}{\partial z^\theta} + \frac{\partial \sigma^\alpha}{\partial \bar{v}^\gamma}\frac{\partial \bar{\varphi}^\gamma}{\partial z^\theta}\right\}\chi^\theta \ .$$

If φ is close to zero, these equations can be solved either way. Let us write

$$\chi = a_\varphi \psi \ .$$

Then "a" is an operator

$$a\colon (U \subseteq \lambda^\circ(X)) \times \lambda^\circ(X) \to \lambda^\circ(X) \ ,$$

which is nonlinear of degree 1 in φ and linear of degree 0 in ψ. Moreover for small φ the linear map a_φ is invertible, and the solution

$$\psi = a_\varphi^{-1}\chi$$

defines an operator

$$a^{-1}\colon (U \subseteq \lambda^\circ(X)) \times \lambda^\circ(X) \to \lambda^\circ(X) \ ,$$

which is also nonlinear of degree 1 in φ and linear of degree 0 in χ. We now have the formula

$$DP(\varphi)\psi = \bar{\partial}_\mu a_\varphi \psi \qquad \text{if } \mu = P(\varphi) \ .$$

1.7. The nonlinear complex. The operators P and Q define a nonlinear complex

$$U \subseteq \lambda^\circ(X) \xrightarrow{P} \lambda^1(X) \xrightarrow{Q} \lambda^2(X)$$

where $P(\varphi) = MS(\varphi)$ and $Q(\mu) = \bar{\partial}_{\mu/2}\mu$. Since the complex structure on Y is integrable and $P(\varphi)$ is its pull-back under the map $f = S(\varphi)$, it follows that $\mu = P(\varphi)$ is always integrable so $Q(\mu) = 0$. Thus $QP(\varphi) = 0$ for all φ. We wish to assert that this nonlinear complex is exact.

Theorem 2. *If $\mu \in \lambda^1(X)$ is sufficiently small and $Q(\mu) = 0$, then there exists a $\varphi \in \lambda^\circ(X)$ with $P(\varphi) = \mu$.*

Corollary. *There exists an* $f = S(\varphi) \in \mathcal{F}(X, Y)$ *with* $M(f) = \mu$. *Thus every integrable almost complex structure* μ *on* X *close enough to the given structure can be obtained by a small wiggle* f *of* X *in* Y.

This is the main result of this paper.

1.8. The Nash-Moser theorem. We shall prove Theorem 2 using a version of the Nash-Moser inverse function theorem which is proved in § 2. We state the theorem briefly here. A grading on a Fréchet space E is an increasing sequence of norms $\|\ \|_n$ $(n = 0, 1, 2, \cdots)$ which define the topology. Two gradings are said to be equivalent if for some r

$$\|\ \|_n^1 \leq C\|\ \|_{n+r}^2 , \qquad \|\ \|_n^2 \leq C\|\ \|_{n+r}^1 .$$

A graded Fréchet space is defined as a Fréchet space with an equivalence class of gradings. We also assume the existence of smoothing operators. If X is a compact manifold with boundary and B is a vector bundle over X, then $C^\infty(X; B)$ is a graded Fréchet space with smoothing operators. We say that a map

$$P: U \subseteq E \to V \subseteq F$$

is tame if every $x_0 \in U$ has a neighborhood on which for some number r we have estimates

$$\|P(x)\|_n \leq C(\|x\|_{n+r} + 1) .$$

We say P is smooth if all its derivatives exist, and we call P a smooth tame map if P and all its derivatives are tame. Every nonlinear differential operator

$$P: U \subseteq C^\infty(X; B) \to V \subseteq C^\infty(X; C)$$

is a smooth tame map. Also the composition of two smooth tame maps is a smooth tame map.

The Nash-Moser inverse function theorem says the following. Suppose $0 \in U$ and

$$P: U \subseteq E \to V \subseteq F$$

is a smooth tame map with $P(0) = 0$ whose derivative

$$DP: (U \subseteq E) \times E \to F$$

is invertible everywhere in U, and suppose also that the family of inverses VP defined by the relation

$$VP(f)h = g \Longleftrightarrow DP(f)g = h$$

is a smooth tame map

$$VP\colon (U \subseteq E) \times F \to E \,.$$

Then for possibly smaller neighborhoods U' and V' of the origin, $P\colon U' \to V'$ is invertible and the inverse $P^{-1}\colon V' \to U'$ is also a smooth tame map.

We shall use a generalization which is the Nash-Moser theorem for nonlinear exact sequences. Suppose E, F and G are graded Fréchet spaces (with smoothing operators as always) and U, V and W are neighborhoods of the origin, and we have a nonlinear complex

$$U \subseteq E \xrightarrow{P} V \subseteq F \xrightarrow{Q} W \subseteq G \,,$$

where P and Q are smooth tame maps with $QP(f) = 0$ for all $f \in U$. We wish to find a condition under which the complex is exact, i.e., $\operatorname{Im} P = \operatorname{Ker} Q$.

We assume that for each $f \in U$

$$\operatorname{Im} DP(f) = \operatorname{Ker} DQ(Pf) \,,$$

so that the linearized complex is exact everywhere in U. We assume moreover that we can find a smooth tame splitting.

Theorem 3. *Suppose there exist smooth tame maps*

$$VP\colon (U \subseteq E) \times F \to E \,, \qquad VQ\colon (U \subseteq E) \times G \to F \,,$$

such that $VP(f)h$ and $VQ(f)k$ are linear in h and k, and split the linearized complex in the sense that

$$DP(f)VP(f)h + VQ(f)DQ(Pf)h = h \,.$$

Then the nonlinear complex is exact at 0, *i.e., we can find a possibly smaller neighborhood V' of the origin such that if $y \in V'$ and $Q(y) = 0$ then $y = P(x)$ for some $x \in U$. Moreover we can find a smooth tame map*

$$S\colon V' \subseteq F \to U \subseteq E$$

such that if $y \in V'$ then

$$PSy = y \qquad \textit{whenever } Qy = 0 \,.$$

In order to apply the Nash-Moser theorem it is necessary to construct the smooth tame splitting maps. This is done in § 5, where we prove their existence under very general conditions. Suppose that P and Q are nonlinear partial differential operators of degree 1 on a compact manifold with boundary. If we choose families of hermitian metrics (which may depend on f) we can form the adjoint operator $D^*P(f)h$ dual to $DP(f)g$. There will also be a boundary condition $d^*p(f)h$ such that

$$\langle\!\langle DP(f)g, h\rangle\!\rangle + \langle\!\langle g, D^*P(f)h\rangle\!\rangle = 0$$

for all g if $d^*p(f)h = 0$. Suppose that for each $f \in U$ the derivatives $DP(f)g$ and $DQ(Pf)h$ form an elliptic complex. The important fact to verify is that we have a uniform persuasive (or subelliptic) estimate; for all $f \in U$ and all h with $d^*p(f)h = 0$ on ∂X we have

$$\int_{\partial X} |h|^2 dS \lesssim \int_X |D^*P(f)h|^2 dV + \int_X |DQ(Pf)h|^2 dV + \int_X |h|^2 dV ,$$

where "$\lesssim$" means "$\leq$ a constant times". This guarantees that

$$\operatorname{Ker} DQ(Pf) / \operatorname{Im} DP(f)$$

is finite dimensional. Suppose in addition that $\operatorname{Im} DP(0) = \operatorname{Ker} DQ(0)$. Then $\operatorname{Im} DP(f) = \operatorname{Ker} DQ(Pf)$ for all f in a neighborhood of 0, and there exist smooth tame splitting maps VP and VQ as required. Consequently the theorem applies.

The philosophy behind this method is clear. In dealing with coercive problems it is sufficient that the derivative at the origin be coercive, for any problem close enough to a coercive problem is again coercive. For noncoercive problems it is necessary to assume that all the derivatives remain uniformly within some tractable class of problems. From there on, invertibility or exactness at the origin will imply the same in a neighborhood of the origin, and we can crank out the tame estimates needed for the Nash-Moser theorem.

In applying this theorem to the present problem it is somewhat more aesthetic to work with the $\bar{\partial}_\mu$ complex than with the $DP - DQ$ complex. They are essentially the same since

$$DP(\varphi)\psi = \bar{\partial}_\mu a_\varphi \psi , \qquad DQ(P\varphi)\chi = \bar{\partial}_\mu \chi ,$$

the only difference being the operator a_φ. But a_φ is invertible with a smooth tame inverse a_φ^{-1}; in fact it acts pointwise on ψ. Therefore, if K_μ and L_μ are a smooth tame splitting for the $\bar{\partial}_\mu$ complex so that

$$\bar{\partial}_\mu K_\mu + L_\mu \bar{\partial}_\mu = I ,$$

then $a_\varphi^{-1} K_{P\varphi}$ and $L_{P\varphi}$ are a smooth tame splitting for the $DP - DQ$ complex, so that

$$DP(\varphi) a_\varphi^{-1} K_{P\varphi} + L_{P\varphi} DQ(P\varphi) = I .$$

We proceed to verify the uniform persuasive estimate for the $\bar{\partial}_\mu$ complex; this is known classically as Morrey's estimate.

1.9. The Levi form. Let Y be a complex manifold, and X a compact subset with smooth boundary ∂X. The complex structure on X induces a decomposition

$$CTX = \mathscr{T}X \oplus \bar{\mathscr{T}}X$$

as before. At the boundary there is a more refined structure. We define the distinguished tangent space

$$\mathscr{D}X = \mathscr{T}X \cap CT\partial X$$

and its conjugate

$$\bar{\mathscr{D}}X = \bar{\mathscr{T}}X \cap CT\partial X \ .$$

Then $\mathscr{D}X$ has complex codimension 1 in $\mathscr{T}X$, and $\mathscr{D}X \oplus \bar{\mathscr{D}}X$ has complex codimension 1 in $CT\partial X$. Now $\mathscr{D}X$ is an integrable subbundle of $CT\partial X$ in the sense that the Lie bracket of two vector fields in $\mathscr{D}X$ lies again in $\mathscr{D}X$, and so is $\bar{\mathscr{D}}X$, but the direct sum $\mathscr{D}X \oplus \bar{\mathscr{D}}X$ is not in general. The obstruction to integrability is the Levi Form

$$\Lambda\colon \mathscr{D}X \times \bar{\mathscr{D}}X \to CT\partial X/\mathscr{D}X \oplus \bar{\mathscr{D}}X$$

defined by the relation that if v is a vector field in $\mathscr{D}X$ and $\bar{w}$ a vector field in $\bar{\mathscr{D}}X$ then

$$\Lambda(v, \bar{w}) \equiv i[v, \bar{w}] \qquad \bmod \mathscr{D}X \oplus \bar{\mathscr{D}}X \ ,$$

where $[v, \bar{w}]$ is the Lie bracket. The space $CT\partial X/\mathscr{D}X \oplus \bar{\mathscr{D}}X$ is equipped with a natural conjugation operation, and the Levi form is hermitian-symmetric

$$\Lambda(w, \bar{v}) = \overline{\Lambda(v, \bar{w})} \ .$$

It therefore makes sense to speak of the "number of positive, zero and negative eigenvalues" of Λ as true invariants, even though the actual value of the eigenvalues would depend on the choice of a basis in $\mathscr{T}X$.

Suppose that $L_1, \cdots, L_{n-1}, L_n$ form a basis locally for the vector fields in $\mathscr{T}X$ with $L_1, \cdots, L_{n-1}$ forming a basis for $\mathscr{D}X$. Then we can write

$$[L_l, \bar{L}_j] = a_{lj}^k L_k - \bar{a}_{jl}^k \bar{L}_k \ .$$

Suppose that we also choose L_n so that $i(L_n - \bar{L}_n) \in T\partial X$. Since $[L_l, \bar{L}_j] \in CT\partial X$ for $l, j < n$, we must have $a_{lj}^n = \bar{a}_{jl}^n$ and $i[L_l, \bar{L}_j] \equiv a_{lj}^n i(L_n - \bar{L}_n)$, mod $\mathscr{D}X \oplus \bar{\mathscr{D}}X$ for $l, j < n$. Therefore the hermitian-symmetric matrix $\{a_{lj}^n : l, j < n\}$ represents the Levi form Λ in the basis $L_1, \cdots, L_{n-1}, L_n$.

Let $z^1, \cdots, z^n$ be local coordinates. We introduce the notation that $a \overset{\circ}{=} b$ means that $a = b$ at the origin. By a proper choice of coordinates we can make

$$T\partial X \stackrel{\circ}{=} \{x^n = 0\} ,$$

where $z^n = x^n + iy^n$. Then we can choose

$$L_l = v_l^k \frac{\partial}{\partial z^k} ,$$

so that $v_l^k \stackrel{\circ}{=} \delta_l^k$; thus $L_l \stackrel{\circ}{=} \partial/\partial z^l$. In this case $i(L_n - \bar{L}_n) = \partial/\partial y^n \in T\partial X$. Let ρ be a smooth real valued function with $\rho = 0$ on ∂X and $\partial\rho/\partial z^n \stackrel{\circ}{=} 1$. Since L_l is parallel to ∂X for $l < n$, we have

$$L_l \rho = v_l^k \frac{\partial \rho}{\partial z^k} = 0 \qquad \text{on } \partial X .$$

Then if also $j < n$ we have

$$v_l^k \frac{\partial^2 \rho}{\partial z^k \partial \bar{z}^j} + \frac{\partial v_l^n}{\partial \bar{z}^j} \stackrel{\circ}{=} 0 .$$

Thus

$$\begin{aligned} i[L_l, \bar{L}_j] &= i\left[v_l^k \frac{\partial}{\partial z^k}, \bar{v}_j^m \frac{\partial}{\partial \bar{z}^m}\right] \\ &= i\left(v_l^k \frac{\partial \bar{v}_j^n}{\partial z^k} \frac{\partial}{\partial \bar{z}^n} - \bar{v}_j^m \frac{\partial v_l^n}{\partial \bar{z}^m} \frac{\partial}{\partial z^n}\right) + \cdots \\ &= \frac{\partial^2 \rho}{\partial z^k \partial \bar{z}^j} v_l^k \bar{v}_j^n \cdot i\left(\frac{\partial}{\partial z^n} - \frac{\partial}{\partial \bar{z}^n}\right) + \cdots \\ &\stackrel{\circ}{=} \frac{\partial^2 \rho}{\partial z^l \partial \bar{z}^j} \frac{\partial}{\partial y^n} + \cdots . \end{aligned}$$

Therefore in local coordinates with $T\partial X \stackrel{\circ}{=} \{x^n = 0\}$ if ρ is a real function with $\rho = 0$ on ∂X and $\partial\rho/\partial z^n \stackrel{\circ}{=} 1$ then the Levi form at 0 is given by the matrix

$$\frac{\partial^2 \rho}{\partial z^l \partial \bar{z}^j} \qquad (l, j < n) .$$

1.10. Adjoint operators. We now choose a hermitian metric $h = \langle\ ,\ \rangle$ on X. Let $L_1, \cdots, L_{n-1}, L_n$ be a basis for $\mathscr{T}X$ as before, so that $L_1, \cdots, L_{n-1}$ are a basis for $\mathscr{D}X = \mathscr{T}X \cap CT\partial X$. Let $\omega^1, \cdots, \omega^n$ be the dual basis of forms. Then the local representatives of h are the matrices $h_{l_j} = \langle L_l, \bar{L}_j\rangle$ and $h^{lj} = \langle \omega^l, \bar{\omega}^j\rangle$ which are inverse to each other. If A is a multi-index $A = (a_1, \cdots, a_q)$, we write

$$\omega^A = \omega^{a_1} \wedge \omega^{a_2} \wedge \cdots \wedge \omega^{a_q} .$$

For two multi-indices A and B, we let $\varepsilon_B^A = \pm 1$ if B is a permutation of A with sign ± 1, and otherwise $\varepsilon_B^A = 0$. In particular $\varepsilon_b^a = 1$ if $a = b$, and $\varepsilon_b^a = 0$ if $a \neq b$. Then

$$\varepsilon_B^A = \sum_\pi (-1)^\pi \varepsilon_{\pi(b_1)}^{a_1} \varepsilon_{\pi(b_2)}^{a_2} \cdots \varepsilon_{\pi(b_q)}^{a_q} ,$$

where the sum ranges over all permutations π, and $(-1)^\pi$ is the sign of π. By analogy let

$$h^{AB} = \sum_\pi (-1)^\pi h^{a_1 \pi(b_1)} h^{a_2 \pi(b_2)} \cdots h^{a_q \pi(b_q)} .$$

If we choose coordinates with $h^{ab} \overset{\circ}{=} \varepsilon_b^a$ at the origin, then $h^{AB} \overset{\circ}{=} \varepsilon_B^A$ at the origin.

Any vector valued q-form $\varphi \in \lambda^q(X)$, which is a section of the bundle $\Lambda^q(\bar{\mathscr{T}}X, \mathscr{T}X)$, can be written locally as

$$\varphi = \varphi_A^l \bar{\omega}^A \otimes L_l .$$

Here and later we adopt the convention that summation is only over strictly increasing multi-indices. There is an induced Hermitian metric on the bundle $\Lambda^q(\bar{\mathscr{T}}X, \mathscr{T}X)$ given by

$$\langle \varphi, \psi \rangle = \varphi_A^l \bar{\psi}_B^j h_{lj} h^{BA} .$$

(Note that A is a conjugate index so it comes second in h^{BA}.) If dV is the volume element arising from the hermitian metric, then

$$dV = \det h = h_{MN} \omega^M \wedge \bar{\omega}^N ,$$

and there is an inner product on $\lambda^q(X)$ given by

$$\langle\!\langle \varphi, \psi \rangle\!\rangle = \iint_X \varphi_A^l \bar{\psi}_B^j h_{lj} h^{BA} \, dV .$$

The operator $\bar{\partial}_\mu$ is given in local coordinates as

$$\bar{\partial}_\mu \varphi_A^l = \varepsilon_A^{cC} (\bar{L}_c \varphi_C^l - \mu_c^a L_a \varphi_C^l) + \cdots ,$$

where the dots denote terms of degree zero. This leads us to define

$$\bar{L}_c^\mu = \bar{L}_c - \mu_c^a L_a ,$$

which we observe is a vector field in $\bar{\mathscr{T}}X_\mu$. Then

$$\bar{\partial}_\mu \varphi_A^l = \varepsilon_A^{cC} \bar{L}_c^\mu \varphi_C^l + \cdots .$$

Next we wish to calculate the adjoint $\bar{\partial}_\mu^*$. We let the hermitian metric h_μ depend smoothly on μ. Then we have

$$\begin{aligned}\langle\!\langle \bar\partial_\mu\varphi, \psi\rangle\!\rangle_\mu &= \iint_X \bar\partial_\mu \varphi_A^l \cdot \bar\psi_B^j h_{lj}^\mu h_\mu^{BA}\, dV_\mu \\ &= \iint_X \varepsilon_A^{cC} \bar L_c^\mu \varphi_C^l \cdot \bar\psi_B^j h_{lj}^\mu h_\mu^{BA}\, dV_\mu + \cdots \\ &= \iint_X \bar L_c^\mu \varphi_C^l \cdot \varepsilon_{dD}^B h_\mu^{dc} \bar\psi_B^j h_{lj}^\mu h_\mu^{DC}\, dV_\mu + \cdots ,\end{aligned}$$

since $\varepsilon_A^{cC} h_\mu^{BA} = \varepsilon_{dD}^B h_\mu^{dc} h_\mu^{DC}$. Now we can move $\bar L_c^\mu$ to the other side provided the boundary integral vanishes. Let ρ be any real function with $\rho = 0$ on ∂X but nonzero gradient; say $L_n\rho \neq 0$. Then the direction of the vector $\bar L_c^\mu \rho$ is independent of the choice of ρ, and the boundary integral vanishes if and only if

$$\bar L_c^\mu \rho \cdot \varphi_C^l \cdot \varepsilon_{dD}^B h_\mu^{dc} \bar\psi_B^j h_{lj}^\mu h_\mu^{DC} = 0$$

on ∂X. We introduce the dual operators

$$L_\mu^d = h_\mu^{cd} L_c^\mu ,$$

where L_c^μ is the conjugate of $\bar L_c^\mu$. We also let $\nu_\mu^d = L_\mu^d \rho$. Then the boundary integral vanishes for all φ if and only if

$$\varepsilon_{dD}^B \psi_B^j \nu_\mu^d = 0 \qquad \text{on } \partial X .$$

Write $\nu_\mu = \nu_\mu^d \bar L_d$. Then ν_μ is a vector field in $\mathscr{T}X$. Also if we let $n^*\psi_D^j = \varepsilon_{dD}^B \psi_B^j \nu_\mu^d$, then n^* is the contraction map on ν_μ:

$$\begin{aligned}n_\mu^* &: \Lambda^{q+1}(\mathscr{T}X, \mathscr{T}X) \to \Lambda^q(\mathscr{T}X, \mathscr{T}X) , \\ n_\mu^* \psi(v_1, \cdots, v_q) &= \psi(\nu_\mu, v_1, \cdots, v_q) .\end{aligned}$$

Suppose $n_\mu^*\psi = 0$ on ∂X. Then moving $\bar L_c^\mu$ to the other side in the previous integral we have

$$\begin{aligned}\langle\!\langle \bar\partial_\mu\varphi, \psi\rangle\!\rangle_\mu &= \iint_X \varphi_C^l \cdot \varepsilon_{dD}^B \bar L_\mu^d \bar\psi_B^j h_{lj}^\mu h_\mu^{DC}\, dV_\mu + \cdots \\ &= \langle\!\langle \varphi, \bar\partial_\mu^* \psi\rangle\!\rangle_\mu ,\end{aligned}$$

where

$$\bar\partial_\mu^* \psi_D^j = \varepsilon_{dD}^B L_\mu^d \psi_B^j + \cdots .$$

If we are careful we can arrange things so that the boundary condition $n_\mu^*\psi = 0$ is independent of μ. For this condition does not depend upon the actual choice of the vector ν_μ but only upon its direction. Therefore we must make the direction of ν_μ independent of μ. Recall that

$$\nu_\mu^d = L_\mu^d \rho = h_\mu^{cd} L_c^\mu \rho .$$

When $\mu = 0$ we have $L_c\rho = 0$ for $c < n$, since $L_1, \cdots, L_{n-1}$ are parallel to the boundary. Moreover we can choose L_n to be orthogonal to $L_1, \cdots, L_{n-1}$ in the metric $h = h_0$ for $\mu = 0$. Then $h^{nd} = 0$ for $d < n$, so $\nu^d = 0$ for $d < n$, and $\nu = \bar{L}_n$. Let $L^\mu\rho = L^\mu_c\rho\cdot\omega^c$; thus $L^\mu\rho$ is a covector field in $\mathscr{T}^*X$. We now choose the hermitian metric h_μ to vary smoothly with μ in such a way that the direction of the covector $L^\mu\rho$ is always dual to the direction of $\bar{L}_n$. Thus we want

$$\langle v, \bar{L}_n\rangle_\mu = 0 \Longleftrightarrow L^\mu\rho(v) = 0$$

for all $v \in \mathscr{T}X$. Since the direction of $\bar{L}_n$ is that orthogonal to $\mathscr{D}X$ in the metric h_0, this is the same as requiring

$$\{v : L^\mu_\rho(v) = 0\} \perp_\mu \{w : w \perp_0 \mathscr{D}X\} .$$

It is clear that this requirement can be fulfilled even globally, with h_μ depending smoothly upon μ. In this case we have

$$\nu^d_\mu = L^d_\mu\rho = 0 \qquad \text{for } d < n .$$

This implies that the operators L^d_μ are all parallel to the boundary for $d < n$.

The boundary condition $n^*\psi = 0$ is independent of μ. Since

$$n^*\psi(v_1, \cdots, v_q) = \psi(\nu, v_1, \cdots, v_q) ,$$

we see that n^* defines a complex

$$\cdots \longrightarrow \Lambda^{p+1}(\mathscr{T}X, \mathscr{T}X) \xrightarrow{n^*} \Lambda^p(\mathscr{T}X, \mathscr{T}X) \xrightarrow{n^*} \Lambda^{p-1}(\mathscr{T}X, \mathscr{T}X) \longrightarrow \cdots$$

which is exact, i.e., $\operatorname{Im} n^* = \operatorname{Ker} n^*$. Write

$$\mathscr{N}^q = \operatorname{Ker} n^* \subseteq \Lambda^q(\mathscr{T}X, \mathscr{T}X) .$$

Then $\mathscr{N}^q$ is a vector subbundle and $n^*\psi = 0 \Leftrightarrow \psi \in \mathscr{N}^q$. Also $n^* : \Lambda^q(\mathscr{T}X, \mathscr{T}X) \to \mathscr{N}^{q-1}$ is a surjective bundle morphism. We can consider $n^*\psi$ as having its values in $\mathscr{N}^{q-1} \subseteq \Lambda^{q-1}(\mathscr{T}X, \mathscr{T}X)$. In local coordinates $\nu^d = 0$ for $d < n$. Therefore $n^*\psi = 0 \Leftrightarrow \psi_A = 0$ whenever $n \in A$.

1.11. The uniform Morrey estimate. This estimate was first proved for the $\bar\partial$ complex by Morrey in the pseudoconvex case and by Hörmander in the general case. We show that the estimate holds for the complex $\bar\partial_\mu$ uniformly in μ. This involves no new techniques, only a casual glance at the effect of μ. We follow more or less the argument of Kohn and Folland [1]. In particular we adopt their convention that

$$\text{“}f(x) \lesssim g(x)\text{”} \quad \text{means} \quad \text{“}\exists C\ \forall x\ f(x) \leq Cg(x)\text{”} .$$

Let $\| \ \|$ and $| \ |$ denote the L_2 norm on X and ∂X respectively. We say ∂X satisfies condition $Z(q)$ if the Levi form Λ never has q negative eigenvalues; i.e., at every point Λ either has at least $n - q$ strictly positive or else at least $q + 1$ strictly negative eigenvalues.

Uniform Morrey estimate. *Suppose ∂X satisfies condition $Z(q)$. Then for all $\varphi \in \lambda^q(X)$ and all μ in a neighborhood of zero*

$$|\varphi| \lesssim \|\bar{\partial}_\mu \varphi\| + \|\bar{\partial}_\mu^* \varphi\| + \|\varphi\| \qquad \textit{when } n^*\varphi = 0$$

with a constant independent of φ and μ.

Proof. It is sufficient to prove the estimate for forms with support in a single coordinate chart, since we can patch together with a partition of unity. Indeed if $\sum \sigma_i^2 = 1$, then we will have

$$\begin{aligned}|\varphi|^2 = \sum |\sigma_i \varphi|^2 &\lesssim \sum \|\bar{\partial}_\mu \sigma_i \varphi\|^2 + \|\bar{\partial}_\mu^* \sigma_i \varphi\|^2 + \|\varphi\|^2 \\ &\lesssim \|\bar{\partial}_\mu \varphi\|^2 + \|\bar{\partial}_\mu^* \varphi\|^2 + \|\varphi\|^2 .\end{aligned}$$

We choose as before a basis $L_1, \cdots, L_{n-1}, L_n$ for $\mathscr{T}X$ with $L_1, \cdots, L_{n-1}$ a basis for $\mathscr{D}X$. Moreover we suppose for simplicity that the L_l are orthonormal in the metric h_0 for $\mu = 0$, and that the matrix for the Levi form with respect to the basis $L_1, \cdots, L_{n-1}, L_n$ is diagonal at the origin of the coordinate system. From the previous section we have the formulas

$$\bar{\partial}_\mu \varphi_A^l = \varepsilon_A^{cC} \bar{L}_c^\mu \varphi_C^l + \cdots , \qquad \bar{\partial}_\mu^* \varphi_D^j = \varepsilon_{dD}^B L_\mu^d \psi_B^j + \cdots ,$$

where the dots denote terms of degree zero. Therefore

$$\iint_X \varepsilon_A^{cC} \varepsilon_B^{dD} \cdot \bar{L}_c^\mu \varphi_C^l \cdot L_d^\mu \bar{\varphi}_D^j \cdot h_{lj}^\mu h_\mu^{BA} \, dV_\mu \lesssim \|\bar{\partial}_\mu \varphi\|^2 + \|\varphi\|^2 .$$

We have the identity

$$\varepsilon_A^{cC} \varepsilon_B^{dD} h_\mu^{BA} = h_\mu^{fc} h_\mu^{de} \{h_{fe}^\mu h^{DC} - \varepsilon_{eE}^C \varepsilon_{fF}^D h_\mu^{FE}\} .$$

(In order to verify it, imagine a new coordinate system with $h^{lj} = \varepsilon_j^l$.) Since $h_\mu^{fc} \bar{L}_c^\mu = \bar{L}_\mu^f$ and $h_\mu^{de} L_d^\mu = L_\mu^e$, we can rewrite the previous integral as a difference of two integrals

$$\iint_X \bar{L}_\mu^f \varphi_C^l \cdot L_\mu^e \bar{\varphi}_D^j \cdot h_{lj}^\mu h_{fe}^\mu h_\mu^{CD} \, dV_\mu - \iint_X \varepsilon_{eE}^C \varepsilon_{fF}^D \cdot \bar{L}_\mu^f \varphi_C^l \cdot L_\mu^e \bar{\varphi}_D^j \cdot h_{lj}^\mu h_\mu^{FE} \, dV_\mu .$$

The first of these may be appropriately called $\|\bar{L}_\mu \varphi\|^2$. For we have in that case the relation

$$\|\bar{L}_\mu \varphi\|^2 \lesssim \sum \|\bar{L}_\mu^f \varphi_C^l\|^2 \lesssim \|\bar{L}_\mu \varphi\|^2$$

with a sum over all f, l, C. We return to this integral later. Meanwhile we consider the other. We claim we can always move the operator L_μ^e from $\bar{\varphi}_D^j$ onto

$\bar{L}^f_\mu \varphi^l_C$ without introducing any boundary integrals when $n^*\varphi = 0$. If $e < n$ then L^e_μ is parallel to the boundary, and we can do it. When $e = n$ we have $n \in C$ so $\varphi^l_C = 0$ on ∂X by the boundary condition $n^*\varphi = 0$. If $f < n$ then $\bar{L}^f_\mu$ is parallel to the boundary so $\bar{L}^f_\mu \varphi^l_C = 0$ on ∂X, which would kill any boundary integral. If $f = n$ then $n \in D$ and $\bar{\varphi}^j_D = 0$ on ∂X, which would do the same. Thus we never get any boundary integrals. In integrating by parts we will produce some lower order terms from L^e_μ falling on the metric; however these are all clearly bounded $\lesssim \|\bar{L}_\mu \varphi\| \cdot \|\varphi\|$. The new integral is

$$\iint_X \varepsilon^C_{eE} \varepsilon^D_{fF} \cdot L^e_\mu \bar{L}^f_\mu \varphi^l_C \cdot \bar{\varphi}^j_D h^\mu_{lj} h^{FE}_\mu \, dV_\mu \, .$$

Using the commutator we write

$$L^e_\mu \bar{L}^f_\mu = \bar{L}^f_\mu L^e_\mu + [L^e_\mu, \bar{L}^f_\mu] \, .$$

This produces two integrals. The first is

$$\iint_X \varepsilon^C_{eE} \varepsilon^D_{fF} \cdot \bar{L}^f_\mu L^e_\mu \varphi^l_C \cdot \bar{\varphi}^j_D h^\mu_{lj} h^{FE}_\mu \, dV_\mu \, .$$

Now we claim we can transfer $\bar{L}^f_\mu$ back from $L^e_\mu \varphi^l_C$ onto $\bar{\varphi}^j_D$ without introducing any boundary integrals. For if $f < n$ then $\bar{L}^f_\mu$ is parallel to the boundary, and if $f = n$ then $f \in D$ and $\bar{\varphi}^j_D = 0$ on the boundary. Again we will have some integrals of lower order of the form

$$\iint_X \varepsilon^C_{eE} \varepsilon^D_{fF} L^e_\mu \varphi^l_C \cdot \bar{\varphi}^j_D \cdot \; \cdots \; dV_\mu \, .$$

We claim that in these integrals we can always move L^e_μ to the other side. For if $e < n$ then L^e_μ is parallel to the boundary, and if $e = n$ then $n \in C$ and $\varphi^l_C = 0$ on the boundary. Therefore the lower order integrals are $\lesssim \|\bar{L}_\mu \varphi\| \cdot \|\varphi\| + \|\varphi\|^2$. Then, since

$$\bar{\partial}^*_\mu \varphi^l_E = \varepsilon^C_{eE} L^e_\mu \varphi^l_C + \cdots ,$$

we can bound the main integral

$$\iint_X \varepsilon^C_{eE} \varepsilon^D_{fF} L^e_\mu \varphi^l_C \cdot \bar{L}^f_\mu \bar{\varphi}^j_D h^\mu_{lj} h^{FE}_\mu \, dV_\mu \lesssim \|\bar{\partial}^*_\mu \varphi\|^2 + \|\varphi\|^2 \, .$$

We then still have the integral from the commutator. We can write

$$[L^e_\mu, \bar{L}^f_\mu] = a^{ef}_{\mu g} L^g_\mu - \bar{a}^{fe}_{\mu g} \bar{L}^g_\mu \, .$$

The integrals involving $\bar{L}^g_\mu$ will all be bounded $\lesssim \|\bar{L}_\mu \varphi\| \cdot \|\varphi\|$. For the integrals with L^g_μ, when $g < n$ we can move this operator to the other side as before and bound the term $\lesssim \|\bar{L}_\mu \varphi\| \cdot \|\varphi\| + \|\varphi\|^2$. We are left with

$$\iint_X \varepsilon^C_{eE}\varepsilon^D_{fF}a^{ef}_{\mu n}L^n_\mu\varphi^l_C\cdot\bar{\varphi}^j_D h^\mu_{lj}h^{FE}_\mu\, dV_\mu\ .$$

Now if $e = n$ then $n \in C$ and $\varphi^l_C = 0$ on the boundary, and we can move L^n_μ to the other side and bound the term as above; similarly if $f = n$ then $n \in D$ and $\bar{\varphi}^j_D = 0$ on the boundary. Hence we only need to consider those terms $a^{ef}_{\mu n}$ with $e, f < n$. Now when $\mu = 0$, $h^{lj} = \varepsilon^l_j$ and $L^e = L_e$. Therefore for $\mu = 0$ the matrix a^{ef}_n is just the matrix of the Levi form Λ, as we showed in the section on the Levi form. Moveover at the origin this matrix is diagonal. When we move the operator L^n_μ to the other side we produce a boundary integral

$$\sum_{e,f<n}\int_{\partial X} \varepsilon^C_{eE}\varepsilon^D_{fF}a^{ef}_{\mu n}L^n_\mu\rho_\mu\cdot\varphi^l_C\cdot\bar{\varphi}^j_D h^\mu_{lj}h^{FE}_\mu\, dS_\mu\ ,$$

where ρ_μ is the distance to the boundary in the metric h_μ, and dS_μ is the induced volume on the boundary. In particular at $\mu = 0$ we have $L^n\rho = 1$. To summarize the situation so far, we have shown

$$\begin{aligned}\iint_X \bar{L}^f_\mu\varphi^l_C\cdot L^e_\mu\bar{\varphi}^j_D\cdot h^\mu_{lj}h^\mu_{fe}h^{CD}_\mu\, dV_\mu \\ + \sum_{e,f<n}\int_{\partial X} \varepsilon^C_{eE}\varepsilon^D_{fF}a^{ef}_{\mu n}L^n_\mu\rho_\mu\cdot\varphi^l_C\bar{\varphi}^j_D h^\mu_{lj}h^{FE}_\mu\, dS_\mu \\ \lesssim \|\bar{\partial}_\mu\varphi\|^2 + \|\bar{\partial}^*_\mu\varphi\|^2 + \|\bar{L}_\mu\varphi\|\cdot\|\varphi\| + \|\varphi\|^2\ .\end{aligned}$$

Now for any $\varepsilon > 0$ we can bound

$$\|\bar{L}_\mu\varphi\|\cdot\|\varphi\| \le \varepsilon\|\bar{L}_\mu\varphi\|^2 + \frac{1}{\varepsilon}\|\varphi\|^2\ .$$

Moreover if we choose the neighborhood of the origin on X and the neighborhood U of 0 for μ to be sufficiently small, then the errors introduced by replacing h^{lj}_μ and $a^{ef}_{\mu n}L^n_\mu\rho_\mu$ by their values at the origin for $\mu = 0$ will be bounded $\lesssim \varepsilon\|\bar{L}_\mu\varphi\|^2 + \varepsilon|\varphi|^2$. For $\mu = 0$ and at the origin we have $h^{lj}_\mu = \varepsilon^l_j$ while $a^{ef}_{\mu n}L^n_\mu\rho_\mu$ becomes a diagonal matrix of eigenvalues λ^e. Thus

$$\begin{aligned}\sum_{e,l,C}\iint_X \bar{L}^e_\mu\varphi^l_C\cdot L^e_\mu\bar{\varphi}^l_C\, dV + \sum_{l,C}\left\{\sum_{e\in C}\lambda^e\right\}\int_{\partial X}\varphi^l_C\bar{\varphi}^l_C\, dS \\ \lesssim \|\bar{\partial}_\mu\varphi\|^2 + \|\bar{\partial}^*_\mu\varphi\|^2 + \varepsilon\|\bar{L}_\mu\varphi\|^2 + \varepsilon|\varphi|^2 + \frac{1}{\varepsilon}\|\varphi\|^2\ .\end{aligned}$$

Now with any term $\bar{L}^e_\mu\varphi^l_C$ we can argue as follows (with no summation); if $e < n$

$$\begin{aligned}\iint_X \bar{L}^e_\mu \varphi^l_C \cdot L^e_\mu \bar{\varphi}^l_C \, dV &= -\iint_X L^e_\mu \bar{L}^e_\mu \varphi^l_C \cdot \bar{\varphi}^l_C \, dV \\ &= -\iint_X [L^e_\mu, \bar{L}^e_\mu] \varphi^l_C \cdot \bar{\varphi}^l_C \, dV + \iint_X L^e_\mu \varphi^l_C \cdot \bar{L}^e_\mu \bar{\varphi}^l_C \, dV \\ &\geq \int_{\partial X} \lambda^e \varphi^l_C \bar{\varphi}^l_C \, dV\end{aligned}$$

with errors $\lesssim \varepsilon \| \bar{L}_\mu \varphi \|^2 + \varepsilon |\varphi|^2 + \|\varphi\|^2/\varepsilon$. We apply this argument only for those eigenvalues $\lambda^e < 0$. For the others we use

$$\iint_X \bar{L}^e_\mu \varphi^l_C \cdot L^e_\mu \bar{\varphi}^l_C \, dV \geq 0 \, .$$

Also we must hold out $\varepsilon \| \bar{L}_\mu \varphi \|^2$ to cancel the term on the right. This however will produce an error bounded by $\varepsilon |\varphi|^2$ in the result. Therefore we have

$$\begin{aligned}&\sum_{l,C} \left\{ \sum_{e \in C} \lambda^e - \sum_{neg} \lambda^e \right\} \int_{\partial X} \varphi^l_C \bar{\varphi}^l_C \, dS \\ &\qquad \lesssim \|\bar{\partial}_\mu \varphi\|^2 + \|\bar{\partial}^*_\mu \varphi\|^2 + \varepsilon |\varphi|^2 + \frac{1}{\varepsilon} \|\varphi\|^2 \, .\end{aligned}$$

Here we let $\sum_{neg}$ (resp. $\sum_{pos}$) denote the sum only over negative (resp. positive) eigenvalues. Now

$$\sum_{e \in C} \lambda^e - \sum_{neg} \lambda^e = \sum_{\substack{pos \\ e \in C}} \lambda^e - \sum_{\substack{neg \\ e \notin C}} \lambda^e \, .$$

Thus

$$\begin{aligned}&\sum_{C} \left\{ \sum_{\substack{pos \\ e \in C}} \lambda^e - \sum_{\substack{neg \\ e \notin C}} \lambda^e \right\} \int_{\partial X} \varphi^l_C \bar{\varphi}^l_C \, dS \\ &\qquad \lesssim \|\bar{\partial}_\mu \varphi\|^2 + \|\bar{\partial}^*_\mu \varphi\|^2 + \varepsilon |\varphi|^2 + \frac{1}{\varepsilon} \|\varphi\|^2 \, .\end{aligned}$$

Now suppose ∂X satisfies condition $Z(q)$. If there are $n - q$ strictly positive eigenvalues, then every multi-index C must contain some e with $\lambda^e > 0$. On the other hand if there are $q + 1$ strictly negative eigenvalues, then every multi-index C of length q must omit some e with $\lambda^e < 0$. In either case

$$\sum_{\substack{pos \\ e \in C}} \lambda^e - \sum_{\substack{neg \\ e \notin C}} \lambda^e > 0 \, .$$

Therefore the term on the left is $\geq \eta |\varphi|^2$ for some $\eta > 0$. Then taking the neighborhood so small that we can cancel the $\varepsilon |\varphi|^2$ term we have for all φ with $n^* \varphi = 0$ on ∂X the estimate

$$|\varphi|^2 \lesssim \|\bar{\partial}_\mu \varphi\|^2 + \|\bar{\partial}_\mu^* \|^2 + \|\varphi\|^2$$

uniformly for all μ in a neighborhood of zero. This proves the uniform Morrey estimate.

2. The Nash-Moser theorem for nonlinear exact sequences

2.1. Near-projections. We assume the reader is familiar with [2]. A Nash-Moser interation algorithm is based on a near-projection. Let E be a graded Fréchet space which we always assume to admit smoothing operators, and let U be an open set in E. A projection is a smooth tame map

$$P: U \subseteq E \to U \subseteq E \quad \text{with} \quad P \circ P = P .$$

The fixed point set $\mathscr{F}(P)$ is defined as

$$\mathscr{F}(P) = \{x \in U : P(x) = x\} .$$

For simplicity suppose U is convex. Define a smooth tame map

$$\Lambda: (U \subseteq E) \times E \times E \to E ,$$

$$\Lambda(x)(v, w) = \int_{t=0}^{1} D^2P(x + t[P(x) - x])(v, w)\, dt .$$

Note that $\Lambda(x)(v, w)$ is bilinear in v and w. By Taylor's formula with integral remainder we have

$$P(P(x)) = P(x) + DP(x)(P(x) - x) + \Lambda(x)(P(x) - x, P(x) - x) .$$

If P is a projection, then $P(P(x)) = P(x)$ and

$$DP(x)(P(x) - x) + \Lambda(x)(P(x) - x, P(x) - x) = 0 .$$

This motivates the following definition: Let $G: U \subseteq E \to E$. We say that G is a near-projection if there exists a smooth tame map $\Lambda: (U \subseteq E) \times E \times E \to E$ with $\Lambda(x)(v, w)$ bilinear in v and w such that

$$DG(x)(G(x) - x) + \Lambda(x)(G(x) - x, G(x) - x) = 0 .$$

Thus every projection is a near-projection. Note that Λ represents a quadratic error.

Given a near projection G with fixed point set

$$\mathscr{F}(G) = \{x \in U : G(x) = x\}$$

we would like to find a true projection P with the same fixed point set. The classical way to do this would be by iteration

$$P = G \circ G \circ G \circ \cdots .$$

However for Féchet spaces this may not converge due to the "loss of derivatives". Therefore we modify the iteration by inserting smoothing operators. We set up the following algorithm. Choose starting values $x_0 \in U \subseteq E$ and real $t_0 \geq 3$.

$$\text{Algorithm:} \quad t_{n+1} = t_n^{3/2} ,$$
$$x_{n+1} = [I - S(t_n)]x_n + S(t_n)G(x_n) .$$

Thus x_{n+1} is a weighted average between x_n and $G(x_n)$ which tends rapidly towards the $G(x_n)$ side as $n \to \infty$. Note that if $x_0 \in \mathscr{F}(G)$, then $x_n = x_0$ for all n so $x_\infty = \lim_{n \to \infty} x_n = x_0$.

Of course in general the x_n may not all be defined, and even if they are they may not converge. However if they are all defined and do converge we say

$$P(x_0) = x_\infty = \lim_{n \to \infty} x_n .$$

This defines a map P on some set including $\mathscr{F}(G)$.

Theorem. *We can find an open set V containing $\mathscr{F}(G)$ such that P is defined on all of V, P maps V into itself, and P is a projection with the same fixed point set as G, i.e., $\mathscr{F}(P) = \mathscr{F}(G)$. Moreover P is a smooth tame map $P: V \subseteq E \to V \subseteq E$.*

It suffices to prove the theorem in a neighborhood of each point in $\mathscr{F}(G)$; therefore we may assume U is convex. We can rewrite the algorithm as

$$x_{n+1} = x_n + \Delta x_n , \quad \Delta x_n = S(t_n)z_n , \quad z_n = G(x_n) - x_n .$$

Here z_n represents the error, and Δx_n the correction. We can derive a recursion relation for z_n. Define a map

$$\Phi: (U \subseteq E) \times (U \subseteq E) \times E \times E \to E ,$$
$$\Phi(x, y)(v, w) = \int_{t=0}^{1} D^2G((1 - t)x + ty)(v, w)\, dt .$$

Then by Taylor's formula we have

$$G(y) = G(x) + DG(x)(y - x) + \Phi(x, y)(y - x, y - x) .$$

Moreover $\Phi(x, y)(v, w)$ is a smooth tame map and is bilinear in v and w. Also

$$G(x_{n+1}) = G(x_n) + DG(x_n)\Delta x_n + \Phi(x_n, x_{n+1})(\Delta x_n, \Delta x_n) .$$

Then

$$
\begin{aligned}
z_{n+1} &= G(x_{n+1}) - x_{n+1} \\
&= G(x_n) - x_n - \Delta x_n + DG(x_n)z_n - DG(x_n)[I - S(t_n)]z_n \\
&\quad + \Phi(x_n, x_{n+1})(\Delta x_n, \Delta x_n) .
\end{aligned}
$$

Then, since

$$DG(x_n)(G(x_n) - x_n) + \Lambda(x_n)(G(x_n) - x_n, G(x_n) - x_n) = 0 ,$$

we have

$$
\begin{aligned}
z_{n+1} = {} & [I - DG(x_n)][I - S(t_n)]z_n \\
& - \Lambda(x_n)(z_n, z_n) + \Phi(x_n, x_{n+1})(S(t_n)z_n, S(t_n)z_n) .
\end{aligned}
$$

Thus z_{n+1} is a sum of three terms; the first should go to zero rapidly since $I - S(t_n)$ does; the second and third are quadratic in z_n and should also go to zero rapidly.

2.2. Low norm estimates. The following estimates will hold uniformly in t_0 for all $t_0 \geq 3$. We shall use the following simple fact.

Lemma. *If $t_0 \geq 3$, then $\sum_0^\infty t_n^{-1} \leq 1$.*

Proof. We have $t_n \geq 3^{-(3/2)^n}$. Now $(3/2)^n = (1 + 1/2)^n \geq 1 + n/2$, so $t_n^{-1} \leq 3^{-(1+n/2)} \leq (\sqrt{3})^{-n}/3$. Then

$$\sum_0^\infty t_n^{-1} \leq \frac{1}{3} \sum_0^\infty \left(\frac{1}{\sqrt{3}}\right)^n \leq \frac{1}{3} \cdot \frac{1}{1 - (1/\sqrt{3})} = \frac{1}{3 - \sqrt{3}} \leq 1 .$$

Pick a base point x_b in the fixed point set $\mathscr{F}(G)$. Then $G(x_b) = x_b$. Since G, DG, Λ, Φ are smooth tame maps, we can find $\theta > 0$ and numbers k, s such that for all x in the set

$$N = \{x : \|x - x_b\|_k \leq 2\theta\}$$

we have the following estimates for all $l \geq k$:

$$
\begin{aligned}
\|G(x)\|_{l-s} &\leq C(\|x\|_l + 1) , \\
\|DG(x)v\|_{l-s} &\leq C(\|v\|_l + \|x\|_l \|v\|_{k-s}) , \\
\|\Lambda(x)(v, w)\|_{l-s} &\leq C(\|v\|_l \|w\|_{k-s} + \|v\|_{k-s} \|w\|_l \\
&\qquad + \|x\|_l \|v\|_{k-s} \|w\|_{k-s}) ,
\end{aligned}
$$

and if also $\|y - x_b\|_k \leq 2\theta$, then

$$
\begin{aligned}
\|\Phi(x, y)(v, w)\|_{l-s} \leq C(&\|v\|_l \|w\|_{k-s} + \|v\|_{k-s} \|w\|_l \\
&+ \|x\|_l \|v\|_{k-s} \|w\|_{k-s} + \|y\|_l \|v\|_{k-s} \|w\|_{k-s}) .
\end{aligned}
$$

For simplicity we always take $s \geq 2$. We can deduce the following estimate.

Lemma. $\|G(x) - G(y)\|_{l-s} \leq C\|x - y\|_l$.

Proof. We have

$$G(x) - G(y) = \int_{t=0}^{1} DG(tx + (1 - t)y)(x - y)\, dt\ .$$

We apply the estimate for DG and integrate. Observe that if $x, y \in N$, then $tx + (1 - t)y \in N$ for $0 \leq t \leq 1$.

We will also have estimates on the smoothing operators; if $l \leq m$, then

$$\|S(t)x\| \leq Ct^{m-l+s}\|x\|_l\ ,$$
$$\|[I - S(t)]x\|_l \leq Ct^{l-m+s}\|x\|_m\ .$$

From now on we assume that $x_0 \in N$ and $t_0 \geq 3$. C will denote various constants independent of x_0 and t_0. We suppose that some members $x_0, x_1, \cdots, x_n$ of the sequence can be defined and lie in N. As soon as some x_n falls outside of N we terminate the algorithm.

Lemma. *For all* $l \geq k$

$$\|x_n - x_b\|_l \leq Ct_n^{5s}\|x_0 - x_b\|_l\ .$$

Proof. When $n = 0$ this is trivial. We proceed by induction. Suppose we can find a constant A_n so that whenever x_n is defined we have

$$\|x_n - x_b\|_l \leq A_n t_n^{5s}\|x_0 - x_b\|_l\ .$$

Suppose $x_n \in N$ so that x_{n+1} is defined. Then

$$\begin{aligned}
&\|x_{n+1} - x_b\|_l \leq \|x_n - x_b\|_l + \|\Delta x_n\|_l\ ,\\
&\|\Delta x_n\|_l = \|S(t_n)z_n\|_l \leq Ct_n^{2s}\|z_n\|_{l-s}\ ,\\
&\|z_n\|_{l-s} = \|G(x_n) - x_n\|_{l-s}\\
&\qquad\quad \leq \|G(x_n) - G(x_b)\|_{l-s} + \|x_n - x_b\|_{l-s}\ ,\\
&\|G(x_n) - G(x_b)\|_{l-s} \leq C\|x_n - x_b\|_l\ .
\end{aligned}$$

Thus

$$\|x_{n+1} - x_b\|_l \leq CA_n t_n^{7s}\|x_0 - x_b\|_l\ .$$

Then

$$\|x_{n+1} - x_b\|_l \leq A_{n+1}t_{n+1}^{5s}\|x_0 - x_b\|_l\ ,$$

provided $CA_n t_n^{7s} \leq A_{n+1}t_{n+1}^{5s}$.

Now $t_{n+1} = t_n^{3/2}$ and we took $s \geq 2$, so $t_n^{7s}t_{n+1}^{-5s} = t_n^{-s/2} \leq t_n^{-1}$. Thus we need

$$A_{n+1} \geq CA_n t_n^{-1}\ .$$

But $t_n \to 0$, so we can satisfy this with a sequence $A_n \to 0$, uniformly for all $t_0 \geq 3$. Since $A_n \to 0$ we have $A_n \leq C$. This proves the Lemma.

Corollary. $\|z_n\|_{l-s} \leq Ct_n^{5s} \|x_0 - x_b\|_l$.

Proof. We saw $\|z_n\|_{l-s} \leq C\|x_n - x_b\|_l$.

Lemma. *We can choose $\varepsilon > 0$, $\eta > 0$ sufficiently small so that if $t_0 \geq 3$ and*

$$\|x_0 - x_b\|_{k+25s} \leq \eta , \qquad \|G(x_0) - x_0\|_{k-s} \leq \varepsilon t_0^{-12s} ,$$

then x_n is defined and belongs to N for all n, and we have estimates

$$\|G(x_n) - x_n\|_{k-s} \leq \varepsilon t_n^{-12s} , \qquad \|\Delta x_n\|_k \leq \theta t_n^{-10s} .$$

Proof. We proceed by induction on n. Suppose that $x_0, x_1, \cdots, x_n$ are all defined and belong to N, and that

$$\|G(x_n) - x_n\|_{k-s} \leq \varepsilon t_n^{-12s} .$$

This says $\|z_n\|_{k-s} \leq \varepsilon t_n^{-12s}$. Then

$$\|\Delta x_n\|_k = \|S(t_n)z_n\|_k \leq Ct_n^{2s}\|z_n\|_k \leq C\varepsilon t_n^{-10s} .$$

Thus $\|\Delta x_n\|_k \leq \theta t_n^{-10s}$ provided $\varepsilon > 0$ is so small that $C\varepsilon \leq \theta$. Now if η is sufficiently small, we will have

$$\|x_0 - x_b\|_k \leq \theta ,$$

and then

$$\|x_{n+1} - x_b\|_k \leq \|x_0 - x_b\|_k + \sum_{k=0}^{m} \|\Delta x_j\|_k \leq 2\theta ,$$

(using $\sum t_j^{-10s} \leq \sum t_j^{-1} \leq 1$). This shows that $x_{n+1} \in N$ also and the algorithm will continue.

Now recall the error recursion formula

$$\begin{aligned} z_{n+1} = {} & [I - DG(x_n)][I - S(t_n)]z_n \\ & - A(x_n)(z_n, z_n) + \Phi(x_n, x_{n+1})(S(t_n)z_n, S(t_n)z_n) . \end{aligned}$$

We can make the following estimates:

$$\begin{aligned} &\|[I - DG(x_n)][I - S(t_n)]z_n\|_{k-s} \leq C\|[I - S(t_n)]z_n\|_k , \\ &\|A(x_n)(z_n, z_n)\|_{k-s} \leq C\|z_n\|_k\|z_n\|_{k-s} , \\ &\Phi(x_n, x_{n+1})(S(t_n)z_n, S(t_n)z_n)\|_{k-s} \leq C\|S(t_n)z_n\|_k\|S(t_n)z_n\|_{k-s} , \end{aligned}$$

since $\|x_n\|_k \leq C$ and $\|x_n\|_{k+1} \leq C$.

The A term is handled in this way. Write

$$z_n = [I - S(t_n)]z_n - S(t_n)z_n \ ,$$
$$\|z_n\|_k \leq \|[I - S(t_n)]z_n\|_k + \|S(t_n)z_n\|_k \ .$$

Then we use $\|S(t_n)z_n\|_k \leq Ct_n^{2s}\|z_n\|_{k-s}$ and $\|z_n\|_{k-s} \leq C$. Thus

$$\|\Lambda(x_n)(z_n, z_n)\|_{k-s} \leq C\|[I - S(t_n)]z_n\|_k + Ct_n^{2s}\|z_n\|_{k-s}^2 \ .$$

For the Φ term we have

$$\|\Phi(x_n, x_{n+1})(S(t_n)z_n, S(t_n)z_n)\|_{k-s} \leq Ct_n^{3s}\|z_n\|_{k-s}^2 \ .$$

Thus all in all

$$\|z_{n+1}\|_{k-s} \leq C\|[I - S(t_n)]z_n\|_k + Ct_n^{3s}\|z_n\|_{k-s}^2 \ .$$

Now

$$\|[I - S(t_n)]z_n\|_k \leq Ct_n^{-23s}\|z_n\|_{k+24s} \ .$$

By the previous corollary we have

$$\|z_n\|_{k+24s} \leq Ct_n^{5s}\|x_0 - x_b\|_{k+25s} \ ,$$
$$\|[I - S(t_n)]z_n\|_k \leq Ct_n^{-18s}\|x_0 - x_b\|_{k+25s} \ .$$

Using the induction hypothesis we obtain

$$\|z_{n+1}\|_{k-s} \leq C\eta t_n^{-18s} + C\varepsilon^2 t_n^{-21s} \leq \varepsilon t_n^{-18s} \ ,$$

provided ε is so small that $C\varepsilon \leq 1/2$ and η is so small that $C\eta \leq \varepsilon/2$. Then $t_{n+1} = t_n^{3/2}$, so

$$\|z_{n+1}\|_{k-s} \leq \varepsilon t_{n+1}^{-12s} \ .$$

This verifies the induction step and proves the Lemma.

2.3. High norm estimates.

Lemma. *Suppose as before that* $t_0 \geq 3$ *and*

$$\|x_0 - x_b\|_{k+25s} \leq \eta \ , \qquad \|G(x_0) - x_0\|_{k-s} \leq \varepsilon t_0^{-12s} \ .$$

Then for every $l \geq k$ *we have estimates*:

$$\|x_n\|_l \leq C(\|x_0\|_{l+18s} + 1) \ ,$$
$$\|\Delta x_n\|_l \leq Ct_n^{-5s}(\|x_0\|_{l+18s} + 1) \ ,$$
$$\|G(x_n) - x_n\|_{l-s} \leq Ct_n^{-7s}(\|x_0\|_{l+18s} + 1) \ .$$

Proof. We proceed by induction on n. Suppose that for $0 \leq j \leq n$ we have estimates

$$\|z_j\|_{l-s} \leq A_j t_j^{-7s}(\|x_0\|_{l+18s} + 1)$$

with an increasing sequence of constants $1 \leq A_0 \leq A_1 \leq \cdots \leq A_n$. Then we have

$$\|\Delta x_j\|_l = \|S(t_j)z_j\|_l \leq Ct_j^{2s}\|z_j\|_{l-s} ,$$
$$\|\Delta x_j\|_l \leq CA_j t_j^{-5s}(\|x_0\|_{l+18s} + 1) .$$

Since $\sum t_j^{-5s} \leq \sum t_j^{-1} \leq 1$, we have

$$\sum_{j=0}^{m} \|\Delta x_j\|_l \leq CA_n(\|x_0\|_{l+18s} + 1)$$

using the fact that the A_j are increasing. Then

$$\|x_n\|_l + \|x_{n+1}\|_l \leq CA_n(\|x_0\|_{l+18s} + 1) .$$

Again we use the error recursion formula

$$\begin{aligned} z_{n+1} = {} & [I - DG(x_n)][I - S(t_n)]z_n \\ & - \Lambda(x_n)(z_n, z_n) + \Phi(x_n, x_{n+1})(S(t_n)z_n, S(t_n)z_n) . \end{aligned}$$

Now we claim that

$$\|z_{n+1}\|_{l-s} \leq CA_n t_n^{-11s}(\|x_0\|_{l+18s} + 1) .$$

To show this we must show that every term in the estimate for z_{n+1} has this as bound.

First

$$\begin{aligned} & \|[I - DG(x_n)][I - S(t_n)]z_n\|_{l-s} \\ & \quad \leq \|[I - S(t_n)]z_n\|_l + \|x_n\|_l \|I - S(t_n)z_n\|_{k-s} . \end{aligned}$$

We have

$$\begin{aligned} \|[I - S(t_n)]z_n\|_l & \leq Ct_n^{-16s}\|z_n\|_{l+17s} \\ & \leq Ct_n^{-16s}t_n^{5s}\|x - x_b\|_{l+18s} \\ & \leq Ct_n^{-11s}(\|x_0\|_{l+18s} + 1) \end{aligned}$$

since $\|x_b\|_{l+18s}$ is a constant C. Also

$$\begin{aligned} \|x_n\|_l \|I - S(t_n)]z_n\|_{k-s} & \leq Ct_n^s\|x_n\|_l \|z_n\|_{k-s} \\ & \leq CA_n t_n^{-11s}(\|x_0\|_{l+18s} + 1) , \end{aligned}$$

since $\|z_n\|_{k-s} \leq Ct_n^{-12s}$ by the previous lemma.

Next

$$\|A(x_n)(z_n, z_n)\|_{l-s} \le C\|x_n\|_l \|z_n\|_{k-s}^2 + C\|z_n\|_l \|z_n\|_{k-s} .$$

Now

$$\|x_n\|_l \|z_n\|_{k-s}^2 \le CA_n t_n^{-24s}(\|x_0\|_{l+18s} + 1) .$$

For the second we use

$$\|z_n\|_l \le \|[I - S(t_n)]z_n\|_l + \|S(t_n)z_n\|_l .$$

The term $\|[I - S(t_n)]z_n\|_l \|z_n\|_{k-s}$ is easier to bound than $\|[I - S(t_n)]z_n\|_l$ which we handled before. The term

$$\begin{aligned}\|S(t_n)z_n\|_l \|z_n\|_{k-s} &\le Ct_n^{2s}\|z_n\|_{l-s}\|z_n\|_{k-s} \\ &\le CA_n t_n^{-17s}(\|x_0\|_{l+18s} + 1)\end{aligned}$$

using $\|z_n\|_{l-s} \le A_n t_n^{-7s}(\|x_0\|_{l+18s} + 1)$ which is the induction hypothesis, and $\|z_n\|_{k-s} \le Ct_n^{-12s}$.

Lastly we have

$$\begin{aligned}&\|\Phi(x_n, x_{n+1})(S(t_n)z_n, S(t_n)z_n)\|_{l-s} \\ &\quad \le C\|S(t_n)z_n\|_l \|S(t_n)z_n\|_{k-s} + C(\|x_n\|_l + \|x_{n+1}\|_l)\|S(t_n)z_n\|_{k-s}^2 .\end{aligned}$$

Then

$$\begin{aligned}\|S(t_n)z_n\|_l \|S(t_n)z_n\|_{k-s} &\le Ct_n^{3s}\|z_n\|_{l-s}\|z_n\|_{k-s} \\ &\le CA_n t_n^{-16s}(\|x_0\|_{l+18s} + 1) ,\end{aligned}$$

and $\|S(t_n)z_n\|_{k-s} \le Ct_n^{s}\|z_n\|_{k-s}$ so

$$(\|x_n\|_l + \|x_{n+1}\|_l)\|S(t_n)z_n\|_{k-s}^2 \le CA_n t_n^{-20s}(\|x_0\|_{l+18s} + 1) .$$

Thus all the terms have been bounded as we claim and

$$\|z_{n+1}\| \le CA_n t_n^{-11s}(\|x_0\|_{l+18s} + 1) .$$

Now $t_n^{-11s} = t_n^{-s/2} t_{n+1}^{-7s}$. Thus

$$\|z_{n+1}\|_{l-s} \le A_{n+1} t_{n+1}^{-7s}(\|x_0\|_{l+18s} + 1) ,$$

provided $A_{n+1} \ge CA_n t_n^{-s/2}$. But as soon as $Ct_n^{-s/2} \le 1$ we can take $A_{n+1} = A_n$ (recall the A_n were to be increasing). Hence the sequence A_n is bounded, and

$$\|z_n\|_{l-s} \le Ct_n^{-7s}(\|x_0\|_{l+18s} + 1) .$$

The other estimates follow from the first part of the argument. Let V_b^0 be the set of all x_0 with

$$\|x_0 - x_b\|_{k+25s} < \eta\ , \qquad \|G(x_0) - x_0\|_{k-s} < \varepsilon t_0^{-12s}\ .$$

Define $P_n^0(x_0) = x_n$. Then P_n^0 is a smooth tame map $P_n^0\colon V_b^0 \subseteq E \to E$. For we have seen that if $x_0 \in V_b^0$ then x_n is defined for all n, and from the algorithm it is clear that x_{n+1} is a smooth tame function of x_n. Moreover

$$\sum_{n=0}^{\infty} \|\Delta x_n\|_l \leq C(\|x_0\|_{l+18s} + 1)\ .$$

Therefore, if $x_0 \in V_b^0$, the sequence x_n must converge to an element $x_\infty \in E$, and

$$\|x_n\|_l \leq C(\|x_0\|_{l+18s} + 1)\ ,$$

so

$$\|x_\infty\|_l \leq C(\|x_0\|_{l+18s} + 1)\ .$$

We write $P_\infty^0(x_0) = x_\infty$. Then P_∞^0 is a map $P_\infty^0\colon V_b^0 \subseteq E \to E$. By the above estimate we see that P_∞^0 is tame. Moreover

$$\|G(x_n) - x_n\|_{l-s} \leq C t_n^{-7s}(\|x_0\|_{l+18s} + 1)\ .$$

Therefore in the limit $G(x_\infty) = x_\infty$. Hence $\operatorname{Im} P_\infty^0 \subseteq \mathscr{F}(G)$. Also if $x_0 \in V_b^0 \cap \mathscr{F}(G)$, then $x_n = x_0$ for all n, so $x_\infty = x_0$ or $P_\infty^0(x_0) = x_0$. Hence $\mathscr{F}(P_\infty^0) = \mathscr{F}(G) \cap V_b^0$

Lemma. *$P_\infty^0\colon V^0 \subseteq E \to E$ is continuous.*

Proof. We have $P_\infty^0 = \lim_{n\to\infty} P_n^0$, and the P_n^0 are surely continuous. Moreover since

$$\sum_{n=0}^{\infty} \|\Delta x_n\|_l \leq C(\|x_0\|_{l+18s} + 1)\ ,$$

we see that the convergence is uniform on every bounded set of x_0 and hence on every compact set of x_0. Therefore the $P_n^0 \to P_\infty^0$ uniformly on compact sets. But $V_b^0 \subseteq E$, and E is a Fréchet space and therefore metrizable. Thus V_b^0 is a k-space (see Kelley [4, p. 231]), so P_∞^0 is continuous.

We can also let V_b^m be the set of all x_m such that

$$\|x_m - x_b\|_{k+25s} < \eta\ , \qquad \|G(x_m) - x_m\|_{k-s} < \varepsilon t_m^{-12s}\ .$$

If $x_m \in V_b^m$, then by the same argument $x_{m+1}, \cdots, x_n, \cdots$ are all defined and we have the same sort of estimates, so again $x_n \to x_\infty$. Let us write (for $m \leq n$)

$$x_n = P_n^m(x_m)\ .$$

Then $P_n^m\colon V_b^m \subseteq E \to E$ is a smooth tame map, and $P_\infty^m = \lim_{n\to\infty} P_n^m$ exists in the sense of uniform convergence on compact sets, and $P^m_\infty\colon V^m_b \subseteq E \to E$ is a continuous tame map. Moreover, if $x_0 \in V^0_b$ and $x_m = P^0_m(x_0) \in V^m_b$, then for all $n \geq m$ we have

$$P^m_n(x_m) = x_n = P^0_n(x_0) \ ,$$

and hence in the limit

$$P^m_\infty(x_m) = x_\infty = P^0_\infty(x_0) \ .$$

Therefore $P^0_\infty = P^m_\infty P^0_m$ (at least where the composition is defined).

2.4. More rapid convergence. If we are willing to involve arbitrarily high norms of x_0, we can make $x_n \to x_\infty$ as fast as any power of the t_n. Let $C(x_0)$ denote a constant which may depend on x_0.

Lemma. *For any c we have*

$$\|z_n\|_{k-s} \leq C(x_0)t_n^{-c}$$

Proof. We proceed by induction on c. We already know the lemma holds for $c = 12s$. Suppose that for some c the estimate

$$\|z_n\|_{k-s} \leq C(x_0)t_n^{-c} \ ,$$

holds for all n. In the earlier argument we saw that

$$\|z_{n+1}\|_{k-s} \leq C\,\|[I - S(t_n)]z_n\|_k + Ct_n^{3s}\,\|z_n\|^2_{k-s} \ .$$

Now

$$\begin{aligned}\|[I - S(t_n)]z_n\|_k &\leq Ct_n^{-2c-2s}\,\|z_n\|_{k+2c+3s} \\ &\leq Ct_n^{-2c+3s}\,\|x_0 - x_b\|_{k+2c+4s} \leq C(x_0)t_n^{-2c+3s} \ ,\end{aligned}$$

using $\|z_n\|_{k+2c+3s} \leq Ct_n^{-5s}\,\|x_0 - x_b\|_{k+2c+4s}$ as was seen in an early lemma. Also using the induction hypothesis

$$t_n^{3s}\,\|z_n\|^2_{k-s} \leq C(x_0)t_n^{-2c+3s} \ .$$

Therefore

$$\|z_{n+1}\|_{k-s} \leq C(x_0)t_n^{-2c+3s} \leq C(x_0)t_{n+1}^{-c-s} \ ,$$

provided that $2(-2c + 3s) \leq 3(-c - s)$ which holds if $c \geq 9s$. But we start from $c = 12s$. Therefore we conclude that

$$\|z_n\|_{k-s} \leq C(x_0)t_n^{-c-s}$$

for all $n \geq 1$. However it clearly holds also for $n = 0$, i.e., $\|z_0\|_{k-s} \leq C(x_0)$.

Therefore, if the lemma holds for c, it also holds for $c + s$. By induction (in steps of s) it holds for all c.

Lemma. *For every l and every c*

$$\|z_n\|_{l-s} \leq C(x_0)t_n^{-c} .$$

Proof. We estimate $\|z_n\|_{l-s}$ as before from the error recursion formula. We consider one at a time the terms which arose. The first was

$$\|[I - S(t_n)]z_n\|_l \leq Ct_n^{-2c-5s}\|z_n\|_{l+2c+6s} \leq C(x_0)t_n^{-2c} ,$$

since $\|z_n\|_{l+2c+6s} \leq Ct_n^{5s}\|x_0 - x_b\|_{l+2c+7s}$ by a previous lemma. The next term was

$$t_n^s\|x_n\|_l\|z_n\|_{k-s} \leq C(x_0)t_n^{-2c} ,$$

since $\|z_n\|_{k-s} \leq C(x_0)t_n^{-2c-7s}$ by the previous lemma and

$$\|x_n - x_b\|_l \leq Ct_n^{5s}\|x_0 - x_b\|_l \quad \text{so} \quad \|x_n\|_l = C(x_0) .$$

We can deal with all the remaining terms in a similar fashion. They are:

$$\begin{aligned} &\|x_n\|_l\|z_n\|_{k-s}^2 , \\ &t_n^{2s}\|z_n\|_{l-s}\|z_n\|_{k-s} , \\ &t_n^{3s}\|z_n\|_{l-s}\|z_n\|_{k-s} , \\ &t_n^{2s}(\|x_n\|_l + \|x_{n+1}\|_l)\|z_n\|_{k-s}^2 . \end{aligned}$$

In each case $\|x_n\|_l \leq C(x_0)t_n^{5s}$, $\|z_n\|_{l-s} \leq C(x_0)t_n^{5s}$, $\|x_{n+1}\|_l \leq C(x_0)t_{n+1}^{5s} \leq C(x_0)t_n^{8s}$ and $\|z_n\|_{l-c} \leq C(x_0)t_n^{-2c-10s}$ (or any power of t_n), so surely each term above is bounded by $C(x_0)t_n^{-2c}$. Thus

$$\|z_{n+1}\|_{l-s} \leq C(x_0)t_n^{-2c} \leq C(x_0)t_{n+1}^{-c} .$$

This proves $\|z_n\|_{l-s} \leq C(x_0)t_n^{-c}$ for all $n \geq 1$, but again it is trivial for $n = 0$. Thus the lemma holds.

Corollary. *For every l and every c*

$$\|\Delta x_n\|_l \leq C(x_0)t_n^{-c} , \qquad \|x_n - x_\infty\|_l \leq C(x_0)t_n^{-c} .$$

Proof. $\|\Delta x_n\|_l \leq Ct_n^{2s}\|z_n\|_{l-s}$ and $\|x_n - x_\infty\|_l \leq \sum_{j=n}^{\infty}\|\Delta x_j\|_l$.

2.5. Derivatives of the projection. We now wish to show that on the entire set V_b^0 the map P_b^0 is smooth and all its derivatives are continuous and tame. We begin by showing that this is true in a neighborhood of any fixed point.

Instead of working with derivatives it is convenient to work with tangent functors. Recall that if $F: U \subseteq E \to G$, then the tangent of F is the map

$$TF: (U \subseteq E) \times E \to G \times G ,$$
$$TF(x, v) = (F(x), DF(x)v) .$$

The second tangent T^2F is the tangent of the tangent of F; i.e.,

$$T^2F = T(TF): (U \subseteq E) \times E \times E \times E \to G \times G \times G \times G ,$$
$$T^2F(x, v, w, z) = (F(x), DF(x)v, DF(x)w, DF(x)z + D^2F(x)(v, w)) .$$

Similarly the kth tangent T^kF is defined as $T^kF = T(T^{k-1}F)$. It is a map

$$T^kF: (U \subseteq E) \times E \times E \times \cdots \times E \to G \times G \times G \times \cdots \times G .$$

If the function F and its derivatives up to D^kF are continuous and tame, then T^kF will be continuous and tame. Conversely, if T^kF is continuous and tame, then so will be F and its derivatives up to D^kF, since we can solve for the D^jF $(0 \leq j \leq k)$ in terms of the components of T^kF.

More precisely, let $x_0 \in U$, and suppose $T^kF(x, v, w, \cdots, z)$ is continuous and tame in a neighborhood of

$$\tilde{x}_0 = (x_0, 0, 0, \cdots, 0) ,$$

say of the form $(\varepsilon > 0)$

$$\|x - x_0\|_l < \varepsilon , \quad \|v\|_l < \varepsilon , \quad \|w\|_l < \varepsilon , \quad \cdots , \quad \|z\| < \varepsilon .$$

Then each $D^jF(x)(v, w, \cdots, z)$ will be continuous and tame in a similar but possibly smaller neighborhood (say $\varepsilon' > 0$). However $D^jF(x)(v, w, \cdots, z)$ is multilinear in $v, w, \cdots, z$. From this it follows that $D^jF(x)(v, w, \cdots, z)$ is continuous and tame on the set $\|x - x_0\|_l < \varepsilon'$ for all $v, w, \cdots, z$ without restriction. Then the same will be true for $T^kF(x, v, w, \cdots, z)$.

The advantage of tangent functors is that they simplify the statement of the chain rule. Namely we have

$$T(F_1 \circ F_2) = TF_1 \circ TF_2 ,$$

and more generally

$$T^k(F_1 \circ F_2) = T^kF_1 \circ T^kF_2 .$$

Now consider a near-projection $G: U \subseteq E \to E$. Recall that this means that G is a smooth tame map, and we can find a smooth tame map

$$\Lambda: (U \subseteq E) \times E \times E \to E$$

such that $\Lambda(x)(v, w)$ is bilinear in v and w, with

$$DG(x)(G(x) - x) + \Lambda(x)(G(x) - x, G(x) - x) = 0 .$$

Lemma. *If $G\colon U \subseteq E \to E$ is a near-projection, then so is*

$$TG\colon (U \subseteq E) \times E \to E \times E\ .$$

Proof. Temporarily we write a tangent vector as $\begin{pmatrix} x \\ v \end{pmatrix}$ instead of (x, v). We define

$$\psi\begin{pmatrix} x \\ v \end{pmatrix}\left(\begin{pmatrix} w \\ z \end{pmatrix}, \begin{pmatrix} u \\ y \end{pmatrix}\right) = \begin{pmatrix} \Lambda(x)(w, u) \\ D\Lambda(x)(v)(w, u) + \Lambda(x)(w, y) + \Lambda(x)(z, u) \end{pmatrix}.$$

Clearly ψ is a smooth tame map and is bilinear in $\begin{pmatrix} w \\ z \end{pmatrix}$ and $\begin{pmatrix} u \\ y \end{pmatrix}$. Now

$$TG\begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} G(x) \\ DG(x)v \end{pmatrix},$$

$$DTG\begin{pmatrix} x \\ v \end{pmatrix}\begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} DG(x)w \\ D^2G(x)(w, v) + DG(x)w \end{pmatrix},$$

$$TG\begin{pmatrix} x \\ v \end{pmatrix} - \begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} G(x) - x \\ DG(x)v - v \end{pmatrix},$$

$$\begin{aligned} DTG\begin{pmatrix} x \\ v \end{pmatrix}\left(TG\begin{pmatrix} x \\ v \end{pmatrix} - \begin{pmatrix} x \\ v \end{pmatrix}\right) &= DTG\begin{pmatrix} x \\ v \end{pmatrix}\begin{pmatrix} G(x) - x \\ DG(x)v - v \end{pmatrix} \\ &= \begin{pmatrix} DG(x)(G(x) - x) \\ D^2G(x)(G(x) - x, v) + DG(x)(DG(x)v - v) \end{pmatrix}. \end{aligned}$$

If we differentiate the identity

$$DG(x)(G(x) - x) + \Lambda(x)(G(x) - x, G(x) - x) = 0\ ,$$

we have

$$\begin{aligned} &D^2G(x)(G(x) - x, v) + DG(x)(DG(x)v - v) + D\Lambda(x)(v)(G(x) - x, G(x) - x) \\ &\quad + \Lambda(x)(DG(x)v - v, G(x) - x) + \Lambda(x)(G(x) - x, DG(x)v - v) = 0\ . \end{aligned}$$

Therefore

$$\begin{aligned} &DTG\begin{pmatrix} x \\ v \end{pmatrix}\left(TG\begin{pmatrix} x \\ v \end{pmatrix} - \begin{pmatrix} x \\ v \end{pmatrix}\right) \\ &\quad + \psi\begin{pmatrix} x \\ v \end{pmatrix}\left(TG\begin{pmatrix} x \\ v \end{pmatrix} - \begin{pmatrix} x \\ v \end{pmatrix}, TG\begin{pmatrix} x \\ v \end{pmatrix} - \begin{pmatrix} x \\ v \end{pmatrix}\right) = 0\ . \end{aligned}$$

This proves that TG is also a near-projection.

Corollary. *For all k, T^kG is a near-projection.*

Next consider the algorithm for G:

$$x_{n+1} = [I - S(t_n)]x_n + S(t_n)G(x_n) .$$

We may define smoothing operators on $E \times E$ by the obvious formula

$$S(t)\begin{pmatrix} x \\ v \end{pmatrix} = \begin{pmatrix} S(t)x \\ S(t)v \end{pmatrix} .$$

Then the algorithm for TG is

$$\begin{pmatrix} x_{n+1} \\ v_{n+1} \end{pmatrix} = [I - S(t_n)]\begin{pmatrix} x_n \\ v_n \end{pmatrix} + S(t_n)TG\begin{pmatrix} x_n \\ v_n \end{pmatrix} .$$

Then, if $x_n = P_n^0(x_0)$, we will have

$$\begin{pmatrix} x_n \\ v_n \end{pmatrix} = TP_n^0\begin{pmatrix} x_0 \\ v_0 \end{pmatrix} .$$

Hence the approximate projections for the algorithm of TG are just the tangents TP_n^0 of the approximate projections P_n^0 for the algorithm of G. Now if x_b is a fixed point for G, then $\begin{pmatrix} x_b \\ 0 \end{pmatrix}$ is surely a fixed point for TG. Therefore TP_n^0 will converge (uniformly on compact sets) in a neighborhood of $\begin{pmatrix} x_b \\ 0 \end{pmatrix}$ to a continuous tame map which is clearly TP_∞^0 by the next lemma.

Lemma. *If F_n is a sequence of continuously differentiable maps, and if $F_n \to F$ and $DF_n \to G$ uniformly on compact sets, then F is continuously diferetiable and $DF = G$.*

Proof. By the fundamental theorem of calculus

$$F_n(x + \Delta x) - F_n(x) = \int_{t=0}^{1} DF_n(x + t\Delta x)\Delta x \, dt .$$

Since $F_n \to F$ and $DF_n \to G$ uniformly on the compact set $\{x + t\Delta x : 0 \leq t \leq 1\}$ we surely have

$$F(x + \Delta x) - F(x) = \int_{t=0}^{1} G(x + t\Delta x)\Delta x dt .$$

Now $DF_n(x)\Delta x$ is linear in Δx, so $G(x)\Delta x$ must be also. Then

$$\frac{1}{h}\{F(x + h\Delta x) - F(x)\} = \int_{t=0}^{1} G(x + th\Delta x)\Delta x \, dt .$$

Since G is continuous (being a uniform limit on compact sets of a sequence of continuous functions on a metrizable space) we have

$$\lim_{h\to 0}\frac{1}{h}\{F(x+h\Delta x)-F(x)\}=G(x)\Delta x\ .$$

Thus F is continuously differentiable, and $DF(x)\Delta x = G(x)\Delta x$ as claimed.

Returning to the argument, we see that TP^0_∞ exists and is continuous and tame, at least in a neighborhood of $\begin{pmatrix} x_b \\ 0 \end{pmatrix}$. Then by our previous reasoning the same is true for all $\begin{pmatrix} x \\ v \end{pmatrix}$ with x in a neighborhood $\|x - x_b\|_l < \varepsilon$ and v unrestricted. Exactly the same argument applies to each $T^kP^0_\infty$. We would be done, except that the neighborhood may shrink to a point as $k \to \infty$, and we need to show that all the $T^kP^0_\infty$ are continuous and tame on some fixed neighborhood of x_b. This follows by a slightly more complicated reasoning.

Lemma. *$T^kP^0_\infty(x, v, w, \cdots, z)$ is continuous and tame for all $x \in V^0_b$ and all $v, w, \cdots, z$ without restriction.*

Proof. By our previous argument, it is enough to show that $T^kP^0_\infty$ is continuous and tame in a neighborhood of $(x_0, 0, 0, \cdots, 0)$ for each $x_0 \in V^0_b$. Fix such an x_0, and let $x_\infty = P^0_\infty(x_0)$. Then x_∞ is a fixed point of G. Now we apply our previous reasoning, not to G near x_b, but to T^kG near x_∞. It follows that we can find numbers $\tilde{k}$ and $\tilde{s}$ and $\tilde{\varepsilon} > 0$, $\tilde{\eta} > 0$ such that if $\tilde{V}^m_\infty$ is the set

$$\|x_m - x_\infty\|_{\tilde{k}+25\tilde{s}} < \tilde{\eta}\ ,$$
$$\|v_m\|_{\tilde{k}+25\tilde{s}} < \tilde{\eta},\ \cdots,\ \|z_m\|_{\tilde{k}+25\tilde{s}} < \tilde{\eta}\ ,$$
$$\|T^kG(x_m, v_m, \cdots, z_m) - (x_m, v_m, \cdots, z_m)\|_{\tilde{k}-\tilde{s}} < \tilde{\varepsilon}t_m^{-12\tilde{s}}\ ,$$

then the maps $T^kP^m_n$ are all defined for $(x_m, v_m, \cdots, z_m) \in \tilde{V}^m_\infty$ and converge (uniformly on compact sets) as $n \to \infty$ to a continuous tame map which must be $T^kP^m_\infty$. Thus P^m_∞ and its derivatives of order up to k exist and are continuous and tame on the set $\tilde{V}^m_\infty$. Now

$$x_m = P^0_m(x_0)\ , \qquad (x_m, 0, 0, \cdots, 0) = T^kP^0_m(x_0, 0, \cdots, 0)\ ,$$

and we have seen that for all l and c

$$\|G(x_m) - x_m\|_l = \|z_m\|_l \le Ct_m^{-c}\ .$$

Since $T^kG(x_m, 0, \cdots, 0) = (G(x_m), 0, \cdots, 0)$, for all l and c we have

$$\|T_kG(x_m, 0, \cdots, 0) - (x_m, 0, \cdots, 0)\|_l \le Ct_m^{-c}\ ,$$

and therefore $(x_m, 0, \cdots, 0) \in \tilde{V}^m_\infty$ when m is sufficiently large. Then $T^kP^m_\infty$ exists and is a continuous tame map in a neighborhood of $(x_m, 0, \cdots, 0)$. Also we surely know that $T^kP^0_m$ exists and is a continuous tame map in a neighborhood of $(x_0, 0, \cdots, 0)$. Since

$$P^0_\infty = P^m_\infty \cdot P^0_m ,$$

it follows that $T^k P^0_\infty$ exists in a neighborhood of $(x_0, 0, \cdots, 0)$ and

$$T^k P^0_\infty = T^k P^m_\infty \cdot T^k P^0_m .$$

But a composition of continuous tame maps is a continuous tame map. Thus $T^k P^0_\infty$ is a continuous tame map in a neighborhood of $(x_0, 0, \cdots, 0)$ for any $x_0 \in V^0_b$. As we observed before, this implies that P^0_∞ is a smooth tame map on V^0_b. This proves the theorem in § 1.

2.6. Nonlinear exact sequences. We are now in a position to prove the Nash-Moser theorem for nonlinear exact sequences. Let E, F, G be graded Fréchet spaces which admit smoothing operators. Let $U \subseteq E$, $V \subseteq F$, $W \subseteq G$ be three open sets, and let P and Q be two smooth tame maps

$$U \subseteq E \xrightarrow{P} V \subseteq F \xrightarrow{Q} W \subseteq G$$

such that the composition $QP = 0$.

Theorem. *Suppose we can find two smooth tame maps*

$$VP: (U \subseteq E) \times F \to E ,$$
$$VQ: (U \subseteq E) \times G \to F ,$$

such that $VP(x)v$ and $VQ(x)w$ are linear in v and w for each x, with

$$DP(x)VP(x)v + VQ(x)DQ(Px)v = v$$

for all $x \in U$ and all $v \in F$. Then for any $x_0 \in U$ we can find a smooth tame map

$$S: V' \subseteq F \to U \subseteq E$$

on some (possibly smaller) neighborhood V' of Px_0 in F such that

$$PSy = y \qquad \textit{whenever } Qy = 0 .$$

It follows that $\operatorname{Im} P = \operatorname{Ker} Q$, at least in a neighborhood of Px_0, i.e., $\operatorname{Im} P \cap V' = \operatorname{Ker} Q \cap V'$. Also let

$$V'' = \{y \in V' ; PSy \in V'\} .$$

Then V'' is also an open neighborhood of Px_0, and $PS: V'' \to V''$ is a smooth tame projection onto $\operatorname{Im} P \cap V'' = \operatorname{Ker} Q \cap V''$. We make the following definition.

Definition. *A* set $X \subseteq E$ is a local smooth tame retract if for every $x \in X$ we can find an open neighborhood V of x and a smooth tame projection $\pi: V \to V$ with $\pi \circ \pi = \pi$ and $\operatorname{Im} \pi = X \cap V$.

Corollary. *Under the above hypotheses* $\operatorname{Im} P$ *is a local smooth tame retract.*

Observe that in the category of Banach spaces every local smooth retract is a submanifold. We have been unable to show that $\operatorname{Im} P$ is a submanifold; in fact there is reason to doubt it in general. We hope that the notion of a local smooth tame retract will be an adequate substitute. For example, a local smooth tame retract has a well-defined "tangent bundle"; namely if locally $X = \operatorname{Im} \pi$, where π is a projection, then $T\pi$ is also a projection and we put $TX = \operatorname{Im} T\pi$. It is not hard to see that $TX \subseteq E \times E$ is independent of the choice of π. Also TX is again a local smooth tame retract in $E \times E$.

Before we prove the theorem we make the following observation.

Lemma. *We may assume*

$$VP(x)VQ(x)w = 0$$

for all $x \in U$ *and* $w \in G$.

Proof. We know that

$$DP(x)VP(x) + VQ(x)DQ(Px) = I ,$$

and that $DQ(Px)DP(x) = 0$ since $QP = 0$. Then $DQ(Px)VQ(x)DQ(Px) = DQ(Px)$ and $DP(x)VP(x)DP(x) = DP(x)$. Therefore

$$I = [DP(x)VP(x) + VQ(x)DQ(Px)]^2 = I + DP(x)VP(x)VQ(x)DQ(Px) ,$$

or $DP(x)VP(x)VQ(x)DQ(Px) = 0$. Now we may replace VP and VQ by two other smooth tame maps

$$\widetilde{VP}(x)v = VP(x)DP(x)VP(x)v , \qquad \widetilde{VQ}(x)w = VQ(x)DQ(Px)VQ(x)w .$$

We then have again

$$DP(x)\widetilde{VP}(x)v + \widetilde{VQ}(x)DQ(Px)v = v$$

and now also

$$\widetilde{VP}(x)\widetilde{VQ}(x) = 0 .$$

Corollary. *We have*

$$\begin{aligned} DP(x)VP(x) &= I \quad && \textit{on } \operatorname{Im} DP(x) , \\ VP(x)DP(x) &= I \quad && \textit{on } \operatorname{Im} VP(x) , \\ DQ(Px)VQ(x) &= I \quad && \textit{on } \operatorname{Im} DQ(Px) , \\ VQ(x)DQ(Px) &= I \quad && \textit{on } \operatorname{Im} VQ(x) . \end{aligned}$$

Proof of the Theorem. We set up the following algorithm. Let

$$\Gamma : U \times V \times W \subseteq E \times F \times G \to E \times F \times G$$

be defined by

$$\Gamma \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - VP(x)(Px - y) \\ y - VQ(x)Qy \\ z - DQ(Px)(Px - y) \end{pmatrix} .$$

Lemma. *Γ is a near-projection.*

Proof. Let

$$\begin{aligned} \Delta x &= VP(x)(Px - y) , \\ \Delta y &= VQ(x)Qy , \\ \Delta z &= DQ(Px)(Px - y) . \end{aligned}$$

We must show that there exists a smooth tame map Φ, bilinear in the last two arguments, such that

$$D\Gamma \begin{pmatrix} x \\ y \\ z \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = \Phi \begin{pmatrix} x \\ y \\ z \end{pmatrix} \left\{ \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}, \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} \right\} .$$

First observe that

$$Px - y = DP(x)\Delta x + VQ(x)\Delta z .$$

Hence any expression which is bilinear in

$$\Delta x, \Delta y, \Delta z, Px - y$$

has the required form. Now we have

$$Q(y) = Q(w) + DQ(w)(y - w) + \Phi(w, y)(y - w, y - w)$$

from Taylor's formula with integral remainder, where

$$\Phi(w, y)(u, v) = \int_{t=0}^{1} D^2Q((1 - t)w + ty)(u, v)dt$$

is a smooth tame map bilinear in u and v. Apply this with $w = Px$ and we have

$$Q(y) = DQ(Px)(y - Px) + \Phi(Px, y)(y - Px, y - Px) .$$

Hence $Q(y) + \Delta z = \Phi(Px, y)(y - Px, y - Px)$ has the form of an admissible quadratic error.

$$D\Gamma\begin{pmatrix} x \\ y \\ z \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix}$$
$$= \begin{pmatrix} \Delta x - VP(x)(DP(x)\Delta x - \Delta y) - DVP(x)(\Delta x, Px - y) \\ \Delta y - VQ(x)DQ(y)\Delta y - DVQ(x)(\Delta x, Qy) \\ \Delta z - DQ(Px)(DP(x)\Delta x - \Delta y) - D^2Q(Px)(DP(x)\Delta x, Px - y) \end{pmatrix} .$$

Now $DVP(x)(\Delta x, Px - y)$ and $D^2Q(Px)(DP(x)\Delta x, Px - y)$ are admissible quadratic errors. Since Qy differs from Δz by an admissible quadratic error, the term $DVQ(x)(\Delta x, Qy)$ is also an admissible quadratic error. For the other terms, we see that

$$\Delta x = VP(x)DP(x)\Delta x$$

since $\Delta x \in \operatorname{Im} VP(x)$ and $VP(x)\Delta y = VP(x)VQ(x)Qy = 0$ because we may assume $VP(x)VQ(x) = 0$. Also

$$\Delta y - VQ(x)DQ(Px)\Delta y = 0 ,$$

since $\Delta y \in \operatorname{Im} VQ(x)$. This leaves a term

$$VQ(x)[DQ(y) - DQ(Px)]\Delta y .$$

However

$$\begin{aligned}[DQ(y) - DQ(Px)]v &= \int_{t=0}^{1} D^2Q((1 - t)y + tPx)(y - Px, v)dt \\ &= \Phi(Px, y)(y - Px, v) ,\end{aligned}$$

where Φ is a smooth tame map. Therefore

$$[DQ(y) - DQ(Px)]\Delta y = \Phi(Px, y)(y - Px, \Delta y)$$

is an admissible quadratic error.

Also $DQ(Px)DP(x)\Delta x = 0$. The last remaining term is

$$\Delta z + DQ(Px)\Delta y .$$

Now we have already seen that

$$Q(y) + \Delta z$$

it an admissible quadratic error. Since $\Delta y = VQ(x)Qy$ it follows that

$$\Delta y + VQ(x)\Delta z$$

is an admissible quadratic error. Therefore we are left with

$$\Delta z - DQ(Px)VQ(x)\Delta z = 0 ,$$

since $\Delta z \in \operatorname{Im} DQ(Px)$. This proves that Γ is a near-projection.

It follows from the theorem in § 1 that on a neighborhood $U' \times V' \times W'$ of $(0, 0, 0)$ the algorithm

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \\ z_{n+1} \end{pmatrix} = [I - S(t_n)] \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} + S(t_n)\Gamma \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix}$$

converges to a smooth tame projection π. Write

$$\pi \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} S(x, y, z) \\ T(x, y, z) \\ U(x, y, z) \end{pmatrix} ,$$

and let $S(y) = S(0, y, 0)$. Then $S\colon V' \subseteq F \to U \subseteq E$ is a smooth tame map (on a sufficiently small neighborhood V' of 0).

Lemma. *Let $y \in V'$. Then $PSy = y$ if $Qy = 0$.*

Proof. We know that

$$\pi \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} x_\infty \\ y_\infty \\ z_\infty \end{pmatrix} = \lim_{n \to \infty} \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} .$$

Now suppose $Qy_0 = 0$. Then

$$\begin{aligned} \Delta y_0 &= VQ(x_0)Qy_0 = 0 , \\ y_1 &= [I - S(t_0)]y_0 + S(t_0)(y_0 - \Delta y_0) = y_0 . \end{aligned}$$

By induction we see that $y_n = y_0$ for all n, so $y_\infty = y_0$. We also know that

$$\Gamma \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} - \begin{pmatrix} x_n \\ y_n \\ z_n \end{pmatrix} \to 0 ,$$

so $\Delta x_n \to 0$, $\Delta y_n \to 0$, $\Delta z_n \to 0$. Then

$$Px_n - y_n = DP(x_n)\Delta x_n + VQ(x_n)\Delta z_n \longrightarrow 0 ,$$

so $Px_\infty = y_\infty$. Now put $x_0 = 0$ and $z_0 = 0$, with $Qy_0 = 0$ as above. Then

$$x_\infty = S(0, y_0, 0) = S(y_0) , \qquad PS(y_0) = Px_\infty = y_\infty = y_0 .$$

Thus $PS(y_0) = y_0$ if $y_0 \in V'$ and $Qy_0 = 0$. This proves the lemma and hence also the theorem.

Bibliography

[1] G. B. Folland & J. J. Kohn, *The Neumann problem for the Cauchy-Riemann complex,* Ann. Math. Studies, No. 75, Princeton University Press, Princeton, 1972, 146.

[2] R. S. Hamilton, *The inverse function theorem of Nash and Moser,* to appear.

[3] ——, *Deformation of complex structures on pseudoconvex domains,* preprint, Cornell University.

[4] J. L. Kelley, *General topology,* Van Nostrand, Princeton, 1955, 298.

[5] J. J. Kohn & L. Nirenberg, *Non-coercive boundary value problems,* Comm. Pure Appl. Math. **18** (1965) 443–492.

[6] M. Kuranishi, *Deformations of isolated singularities and* $\bar{\partial}_b$, to appear.

[7] J. Morrow & K. Kodaira, *Complex manifolds,* Holt Rinehart and Winston, New York, 1971, 192.

[8] H. Rossi, *Attaching analytic spaces to an analytic space along a pseudo-concave boundary,* Proc. Conf. Complex Analysis Minneapolis, Springer, Berlin, 1965, 242.

CORNELL UNIVERSITY