Unfoldings of discrete dynamical systems

JOEL W. ROBBIN

Mathematics Department, University of Wisconsin, Madison, WI 53706, USA

Abstract. A universal unfolding of a discrete dynamical system f_0 is a manifold F of dynamical systems such that each system g sufficiently near f_0 is topologically conjugate to an element f of F with the conjugacy φ and the element f depending continuously on f_0. An infinitesimally universal unfolding of f_0 is (roughly speaking) a manifold F transversal to the topological conjugacy class of f_0. Using Nash–Moser iteration we show infinitesimally universal unfoldings are universal and (in part II) give a class of examples relating to moduli of stability introduced by Palis and De Melo.

PART I. GENERAL THEORY

0. *The philosophy of unfoldings*

(0.1) Let us imagine we are studying a space $\mathcal{D}$ of objects which we call 'systems'. We suppose this space is endowed with an equivalence relation $g \sim f$ and a topology. For each $f \in \mathcal{D}$ we denote its equivalence class by:

$$C(f) = \{g \in \mathcal{D} : g \sim f\}.$$

Our object is to classify elements of $\mathcal{D}$ up to equivalence. As a first step we should classify elements in some neighbourhood $\mathcal{G} \subset \mathcal{D}$ of a particular element $f_0 \in \mathcal{D}$ which we understand. The simplest case will be when there is only one element in the classification:

$$\mathcal{G} \subset C(f_0). \tag{1}$$

In this case one says that f_0 is *rigid* or *structurally stable*. When $\mathcal{D}$ is a smooth manifold and the equivalence classes $C(f) \subset \mathcal{D}$ are smooth manifolds one often solves (1) by solving the 'linearized equation'

$$T_f \mathcal{D} = T_f C(f) \tag{2}$$

and applying (a suitable version of) the implicit function theorem.

When f_0 is not rigid the local classification problem is attacked as follows. One considers (germs of) submanifolds $F \subset \mathcal{D}$ at f_0 (i.e. $f_0 \in F$). Such an F is called a *deformation* or *unfolding* of f_0. When we can solve

$$\mathcal{G} \subset \bigcup_{f \in F} C(f) \tag{3}$$

we say that F is *universal*. (Some authors say 'versal' as the element $f \in F$ such that $g \in C(f)$ need not be uniquely determined by $g \in \mathcal{G}$. In other contexts, F is called

a 'slice'.) Equation (3) says that the elements of F are 'normal forms' for perturbations g of f_0.

The appropriate linearized version of (3) is that $C(f)$ and F intersect transversally:

$$C(f) \pitchfork_f F$$

which in terms of tangent spaces is expressed by the equation

$$T_f \mathscr{D} = T_f C(f) + T_f F. \tag{4}$$

Again one often proves (4)$\Rightarrow$(3) using the implicit function theorem.

(0.2) Two particularly successful examples of this philosophy are Thom's Elementary Catastrophe theory (henceforth called E.C. theory) and Mather's theory of singularities of maps. (For a good exposition of E.C. theory see [25]; for an exposition of Mather's theory see [4] and the original references cited there.)

Let us review the high points of E.C. theory. Take for the space $\mathscr{D}$ of systems the ring $\mathscr{D} = \mathscr{M}_n^2 \subset \mathscr{E}_n$ of germs at the origin of smooth real valued functions which vanish, together with the first derivatives at the origin. We take the equivalence relation to be right-equivalence:

$$C(f) = \{f \circ \varphi : \varphi \in \mathrm{Diff}_n\}$$

where Diff_n denotes the group of germs of diffeomorphisms which fix the origin. Thus g and f are equivalent if they differ by a smooth change of coordinates. We may take as topology the topology induced by the map $\mathscr{E}_n \to J^\infty(n)$ which sends a germ to its formal power series at the origin ($J^\infty(n)$ is a product of infinitely many copies of $\mathbb{R}$). The tangent space to $C(f)$ at f is given by the product of the maximal ideal with the so-called 'Jacobian ideal' of f, i.e. the ideal generated by the partials of f:

$$T_f C(f) = \mathscr{M}_n \cdot (\partial_1 f, \dots, \partial_n f).$$

To see this note that a typical element $\hat{f} \in T_t C(f)$ has form

$$\hat{f} = \frac{d}{dt} f \circ \varphi_t \bigg|_{t=0}$$

where $\varphi_t \in \mathrm{Diff}_n$ is a curve through the identity: $\varphi_0(x) = x$. Thus

$$\hat{f} = \sum (\partial_i f) \cdot \xi_i$$

where $\xi_i(0) = 0$ (i.e. $\xi_i \in \mathscr{M}_n$) as $\varphi_t(0) = 0$. We see that f is infinitesimally rigid – i.e. satisfies (2) – exactly when the matrix of second partials is non-degenerate. We expect that in that case f should be rigid as well. Indeed, this is so: that is precisely the content of the Morse lemma (f is right equivalent to any quadratic form having the same signature as $\sum \partial_i \partial_j f(0) x_i x_j$).

Now let us consider an unfolding of form $F = \{f_\gamma : \gamma \in \Gamma\}$ where Γ is an open neighbourhood of the origin in $\mathbb{R}^p$. We then have for $f = f_\gamma$ the tangent space $T_f F = A(\gamma)\mathbb{R}^p$ where $A(\gamma) : \mathbb{R}^p \to \mathscr{D}$ is the linear map given by

$$A(\gamma)\hat{\gamma} = \frac{d}{dt} f_{\gamma+t\hat{\gamma}} \bigg|_{t=0}.$$

The transversality condition (4) can be expressed by saying that the p-dimensional real vector space spanned by the partials $\partial f_\gamma / \partial \gamma_k$ is a vector space complement to the ideal generated by $x_j \partial_i f_\gamma$. According to the C^∞ preparation theorem, if this is true for $\gamma = 0$ it is true for γ in a neighbourhood of 0. As an application we conclude that (with $n = 1$) the unfolding

$$f_\gamma(x) = \gamma_1 x^2 + \gamma_2 x^3 + \cdots + \gamma_p x^{p+1} + x^{p+2}$$

is universal since $x^2, x^3, \ldots, x^{p+1}$ span a complement to the ideal generated by x^{p+2} (by Taylor's formula).

(0.3) Another instructive example is given by the theory of Lie groups. Here we take $\mathcal{D} = G$ a finite dimensional Lie group (e.g. a closed group of matrices) and the equivalence relation to be conjugacy in the group:

$$C(f) = \{\varphi \circ f \circ \varphi^{-1} : \varphi \in G\}.$$

There are no rigid elements since a conjugacy class can never be open (eigenvalues are invariant under conjugacy) but universal unfoldings play a prominent role in the theory. In fact, if f_0 is a regular element in a cartan subgroup $H \subset G$, and $\Gamma \subset H$ is a neighbourhood of the identity then the unfolding $F = \{\gamma \cdot f_0 : \gamma \in \Gamma\}$ is universal. For example if we take $G = \text{GL}(n)$, f_0 to be diagonal with distinct eigenvalues, and H to be the subgroup of all diagonal matrices, this says that any matrix near f_0 can be diagonalized. See e.g. [23] where this approach is exploited.

(0.4) In this paper we study the case where the systems f are discrete dynamical systems on a compact smooth manifold M, $\mathcal{D} = \text{Diff}(M)$, where $\text{Diff}(M)$ denotes the group of smooth diffeomorphisms $f : M \to M$. There are several equivalence relations which have been studied in this context; it is instructive to see what happens in simple examples if we take the obvious equivalence relation, *viz.* conjugacy in the group. (We call this *smooth conjugacy.*) We endow $\text{Diff}(M)$ with the C^∞ topology.

Suppose $f, g \in \text{Diff}(M)$ are smoothly conjugate. Then there exists $\varphi \in \text{Diff}(M)$ with

$$g^n = \varphi \circ f^n \circ \varphi^{-1}$$

for $n = 1$ and hence all integers n. Suppose x is a periodic point of period n of $f : f^n(x) = x$. Then $y = \varphi(x)$ is a periodic point of g and

$$Dg(y)^n = D\varphi(x) Df(x)^n D\varphi(x)^{-1}.$$

This shows that the eigenvalues of $Df(x)^n$ are invariant under smooth conjugacies. It follows that no rigid system can have any periodic points.

It is unknown whether any rigid systems exist (the author believes they do) but the most well studied systems have periodic points and so cannot be rigid.

Hence let us seek 'universal' unfoldings, say for the simplest systems: the gradient-like systems. If f is gradient-like its non-wandering set consists of finitely many hyperbolic fixed points, and so by Sternberg's linearization theorem M can be covered by finitely many coordinate charts on which f is linear. It is tempting to

imagine that a universal unfolding can be constructed by simply varying the eigenvalues at the fixed points, but this is not so.

To see why let us suppose that $M = [0, 1]$ and $f: M \to M$ satisfies:

$$f(0) = 0; \qquad f(1) = 1;$$
$$0 < a = f'(0) < 1 < b = f'(1);$$
$$f(t) < t, \qquad 0 < t < 1. \tag{5}$$

Choose diffeomorphisms $\alpha: [0, 1) \to [0, \infty)$ and $\beta: (0, 1] \to [0, \infty)$ which linearize f:

$$(\alpha \circ f \circ \alpha^{-1})(x) = ax$$
$$(\beta \circ f \circ \beta^{-1})(y) = by$$

and note that α and β are unique if we impose the conditions $\alpha'(0) = 1 = \beta'(1)$. Indeed if $\psi: [0, \infty) \to [0, \infty)$ satisfies $\psi(ax) = a\psi(x)$ then $\psi(0) = 0$ and

$$\psi(x) = \lim_{n \to \infty} a^{-n} \psi(a^n x) = \psi'(0)x.$$

It follows that we can assign an invariant $\gamma_f \in \text{Diff}((0, \infty))$ to f via the formula $\gamma_f = \beta \circ \alpha^{-1}$. We note that γ_f satisfies the condition $\gamma_f(ax) = b\gamma_f(x)$ but is subject to no further restriction. Thus the smooth conjugacy class of f has infinite codimension in $\text{Diff}(M)$. Note further that f interpolates a flow (i.e. there is a group homomorphism $\mathbb{R} \to \text{Diff}(M): t \to f^t$ with $f = f^1$) if and only if the conjugacy invariant has form $\gamma_f(x) = x^\tau$, where $a^\tau = b$ (i.e. $\tau = \log a / \log b$). This is essentially lemma 3 (page 172) of [8].

This phenomenon occurs quite generally, *viz.* any time the basins of a sink and source overlap. Thus (at the present state of knowledge) one must use a different equivalence relation to obtain finite dimensional universal unfoldings.

(0.5) Hence we impose a coarser equivalence relation and seek rigid systems and universal unfoldings for it. This relation is *topological conjugacy* i.e. conjugacy in the group $\text{Homeo}(M)$ of all homeomorphisms of M. Hence the topological conjugacy class of $f \in \text{Diff}(M)$ is defined by:

$$C(f) = \{\varphi \circ f \circ \varphi^{-1}: \varphi \in \text{Homeo}(M)\} \cap \text{Diff}(M).$$

There is no reason to suppose that $C(f)$ is in any sense a submanifold of $\text{Diff}(M)$; note that is not even the orbit of a group action since $\varphi \circ f \circ \varphi^{-1}$ need not be smooth. Nonetheless, we shall see that the infinitesimal methods described above apply.

For the equivalence relation of topological conjugacy, rigid discrete dynamical systems are called *structurally stable*. Note that the map $f: [0, 1] \to [0, 1]$ of the interval defined above is structurally stable; indeed if g also satisfies (5) we may define a conjugacy φ by choosing any homeomorphism $\varphi: [f(\frac{1}{2}), \frac{1}{2}] \to [g(\frac{1}{2}), \frac{1}{2}]$ and extending it to a homeomorphism $\varphi: [0, 1] \to [0, 1]$ by the equations:

$$\varphi(0) = 0, \qquad \varphi(1) = 1,$$
$$\varphi(t) = g^n \circ \varphi \circ f^{-n}(t),$$

for $0 < t < 1$ when n is chosen so that $f^{-n}(t) \in [f(\frac{1}{2}), \frac{1}{2}]$. Reflect on the fact that φ can be a diffeomorphism on the open interval and that one can estimate quite precisely

how rapidly the derivatives of φ blow up as we approach the endpoints. This is the key to the constructions described below.

(0.6) We attempt to prove $f \in \text{Diff}(M)$ is structurally stable using the implicit function theorem. Thus for each $g \in \text{Diff}(M)$ we must solve

$$g = \varphi \circ f \circ \varphi^{-1} \tag{6}$$

for $\varphi \in \text{Homeo}(M)$. Rewrite this in the form

$$f \circ g^{-1} \circ \varphi = f \circ \varphi \circ f^{-1}. \tag{7}$$

If we take $g = g_t$ where $t \to g_t$ is a curve through f we should be able to solve for $\varphi = \varphi_t$ where $t \to \varphi_t$ is a curve through the identity. Hence differentiating at $t = 0$ gives

$$(1 - f_\#)\zeta = w \tag{8}$$

where

$$\zeta = \frac{d}{dt}\varphi_t \bigg|_{t=0}$$

$$w = -\frac{d}{dt} f \circ g_t^{-1} \bigg|_{t=0}$$

$$f_\# \zeta = Tf \circ \zeta \circ f^{-1}.$$

Here w, ζ are elements of $\mathscr{X}(M)$, a space of continuous vector fields on M, and $f_\#: \mathscr{X}(M) \to \mathscr{X}(M)$ is a linear operator on this space. Hence the solvability of (8) for ζ in terms of w is the appropriate linearized version of structural stability; if $f \in \text{Diff}(M)$ satisfies this then f is called *infinitesimally stable*. Since the maps $C^0(M, M) \to C^0(M, M)$, $\varphi \to f \circ g^{-1} \circ \varphi, f \circ \varphi \circ f^{-1}$ are differentiable (as maps of Banach manifolds) the hypothesis that f is infinitesimally stable (with $\mathscr{X}(M) = \mathscr{X}^0(M)$, the space of continuous vector fields) implies that we can solve (8) for φ in terms of g. This is carried out in [16]. Actually Robbin uses a slightly different space $\mathscr{X}(M) = \mathscr{X}_f(M)$ of vector fields which enables him to prove that the solution of (7) is a homeomorphism so that (6) holds. This is all quite easy; the hard part is to prove certain geometric conditions (*viz.* Axiom A and strong transversality) imply infinitesimal stability. The conjecture that Axiom A and strong transversality is equivalent to structural stability is due to Smale [21]; the other direction is still open. Mañé [10] however has proved that infinitesimal stability (with $\mathscr{X}(M) = \mathscr{X}^0(M)$) is equivalent to Axiom A and strong transversality.

(0.7) Now we apply the same reasoning to an unfolding F. To prove it universal we must still solve (6) but now $\varphi \in \text{Homeo}(M)$ and $f \in F$ are the unknowns. Reasoning as before with an additional curve $t \to f_t \in F$ we obtain a linearized equation

$$(1 - f_\#)\zeta + \beta = w \tag{9}$$

where

$$\beta = \frac{d}{dt} f_t \circ f^{-1} \bigg|_{t=0}$$

so that $\beta \in \mathscr{X}^{\infty}(M)$ and $\beta \circ f \in T_f F$. Note that (when F is finite dimensional) the solvability of (9) for ζ and β in terms of w implies that $1 - f_{\#}$ has finite dimensional co-kernel; i.e. that its image has finite codimension.

We must confront two difficulties that weren't present before. In the first place the map $F \times C^0(M, M) \to C^0(M, M)$, $(f, \varphi) \to f \circ \varphi \circ f^{-1}$ is not differentiable as a map of Banach manifolds since φ is only continuous. (However the map $(f, \varphi) \to f \circ g^{-1} \circ \varphi$ is differentiable.)

However the map

$$F \times C^{r+s}(M, M) \to C^r(M, M),$$

$$(f, \varphi) \to f \circ \varphi \circ f^{-1},$$

is of class C^s: since there is a loss of differentiability one is tempted to try to solve using Nash–Moser iteration; but this would give a smooth conjugacy φ and hence, as we have seen, cannot cover any known examples.

The second difficulty seems to be related: Mañé [10] shows that when $\mathscr{X}(M) = \mathscr{X}^0(M)$ the condition that $1 - f_{\#}$ should have finite dimensional co-kernel implies Axiom A and strong transversality and hence that f is structurally stable. This seems to suggest that the method cannot lead to non-trivial universal unfoldings.

The solution is to use a space $\mathscr{X}(M)$ of vector fields with

$$\mathscr{X}^{\infty}(M) \subset \mathscr{X}(M) \subset \mathscr{X}^0(M).$$

This space must be 'adapted' to the unfolding F which means that equation (7) satisfies the hypotheses of the Nash–Moser implicit function theorem. In particular, the map $F \times \mathscr{X}(M) \to \mathscr{X}(M)$, $(f, \zeta) \to f_{\#}\zeta$ is C^{∞} although there is a 'loss of differentiability'. (The topology of $\mathscr{X}(M)$ is given by a countable family of norms like that of $\mathscr{X}^{\infty}(M)$.)

(0.8) In part I of the present paper we implement this program. For the convenience of the reader we have included a proof of the Nash–Moser implicit function theorem in the form in which we use it. In § 13 we state the conjecture on universal unfoldings analogous to Smale's structural stability conjecture in that it asserts that certain geometric conditions on an unfolding imply that it is universal. In part II we work out in detail an example and present some additional theoretical material (e.g. the 'tame' preparation theorem) which we feel is not yet in its final form.

It is probably not necessary to use the implicit function theorem to attack problems in this area. Thus Robinson [17] proved the structural stability theorem of Robbin [16] mentioned above without linearizing and achieved a sharper result to boot. We prefer the linearization technique on the grounds that it is always easier to solve a linear equation than a non-linear one. This is because one can combine 'local solutions' by adding; to do the analogous thing in the non-linear case one must change the equation. Against this point one could object that the present paper proves more than we are interested in; we only wanted a homeomorphism φ but the one we construct in part II is in fact a diffeomorphism on a large open set. Perhaps the direct approach yields a homeomorphism with less work (without proving the additional smoothness).

The examples in part II are essentially versions of the examples of Palis [14] and De Melo [3]. Note however that these authors do not construct unfoldings in our sense although their arguments can undoubtedly be improved to do so. The reason is that the definition of unfolding requires the map $g \to (\varphi, f)$ to be continuous and this implies 'continuous unfolding' of the non-transversal intersection of stable and unstable manifolds in these examples. The existence of (even merely continuous) unfolding of singularities is relatively deep and requires some form of the C^∞-preparation theorem.

1. *Notation*

(1.1) Throughout $\mathbb{N}$ denotes the natural numbers (= non-negative integers), $\mathbb{Z}$ denotes the integers, $\mathbb{R}$ denotes the real numbers, and $\mathbb{E}, \mathbb{F}, \mathbb{H}$ (possibly with subscripts) denote vector spaces over $\mathbb{R}$ (often infinite dimensional). We denote by $L^n(\mathbb{E}_1, \ldots, \mathbb{E}_n; \mathbb{F})$ the space of multi-linear maps from $\mathbb{E}_1 \times \cdots \times \mathbb{E}_n$ to $\mathbb{F}$; thus using the canonical identification (these days called 'Currying' by the computer scientists) we have

$$L^0(\mathbb{F}) = \mathbb{F}$$
$$L^n(\mathbb{E}_1, \mathbb{E}_2, \ldots, \mathbb{E}_n; \mathbb{F}) = L(\mathbb{E}_1, L^{n-1}(\mathbb{E}_2, \ldots, \mathbb{E}_n; \mathbb{F}))$$

where $L(\mathbb{E}, \mathbb{F})$ denotes the space of linear maps from $\mathbb{E}$ to $\mathbb{F}$. Note that even when $\mathbb{E}$ and $\mathbb{F}$ are topological vector spaces we do *not* impose a continuity condition for membership of $L(\mathbb{E}, \mathbb{F})$ and we do *not* give $L(\mathbb{E}, \mathbb{F})$ a topology. This is in contrast with Banach space theory where $L(\mathbb{E}, \mathbb{F})$ denotes the space of continuous linear maps and is itself a Banch space. The reason is that in the tame category (defined in § 2) the tame linear operators do not form a tame linear space in any reasonable sense. We use the abbreviations

$$L(\mathbb{E}) = L(\mathbb{E}, \mathbb{E})$$
$$L^n(\mathbb{E}; \mathbb{F}) = L^n(\underbrace{\mathbb{E}, \ldots, \mathbb{E}}_{n}; \mathbb{F})$$

and we denote by $L_s^n(\mathbb{E};\mathbb{F})$ the space of symmetric multilinear maps; i.e. for $A \in L_s^n(\mathbb{E};\mathbb{F})$ and σ a permutation of $1, 2, \ldots, n$ we have

$$Av_1 \cdots v_n = Av_{\sigma(1)} \cdots v_{\sigma(n)},$$

for $v_1, \ldots, v_n \in \mathbb{E}$.

(1.2) If M is a topological space and $X \subset M$ we denote by $\mathrm{cl}\,(X)$ and $\mathrm{int}\,(X)$ the closure and interior of X in M respectively. If $f: M \to N$ and $u: N \to \mathbb{E}$ we denote by f^*u the map $f^*u: M \to \mathbb{E}$ defined by

$$(f^*u)(x) = u(f(x))$$

for $x \in M$. When f is bijective we denote by f_* the inverse of f^*: $(f_*v)(y) = v(f^{-1}(y))$ for $v: M \to \mathbb{E}$ and $y \in N$.

When $f: M \to N$ is a smooth map of smooth manifolds we denote derivative of f at $x \in M$ by:

$$Df(x): T_xM \to T_{f(x)}N$$

so that

$$Df(x)v = \frac{d}{dt} f(c(t))\Big|_{t=0}$$

where $c: \mathbb{R} \to M$ is any curve *through* x (i.e. $c(0) = x$) and *tangent* to v (i.e. $\dot{c}(0) = v$). We denote by Tf the *tangent map* of f; i.e. $Tf: TM \to TN$ covers f and $Tf|\,T_xM = Df(x)$.

Let $\mathscr{X}^0(M)$ denote the space of continuous vector fields and assume $f: M \to N$ is a diffeomorphism. We then have isomorphisms $f_\#: \mathscr{X}(M) \to \mathscr{X}(N), f^\#: \mathscr{X}(N) \to \mathscr{X}(M)$ defined by

$$f_\# v = Tf \circ v \circ f^{-1},$$
$$f^\# w = (Tf)^{-1} \circ w \circ f,$$

for $v \in \mathscr{X}^0(M)$ and $w \in \mathscr{X}^0(N)$. These operators (like f_* and f^*) are inverse to one another.

If $Z \subset \mathbb{R}^m$ is open we may identify TZ with $Z \times \mathbb{R}^m$. Strictly speaking a vector field on Z is a map $w: Z \to TZ = Z \times \mathbb{R}^m$ of form $w(x) = (x, w_0(x))$ where $w_0: Z \to \mathbb{R}^m$. However we shall not distinguish w and w_0 and hence we identify $\mathscr{X}^0(Z) = C^0(Z) \otimes \mathbb{R}^m$, where $C^0(Z)$ is the ring of continuous functions on Z. Thus if $z: M_0 \to Z$ is a coordinate patch on M (i.e. $M_0 \subset M$ is open and z is a diffeomorphism) each $w \in \mathscr{X}^0(M)$ (in fact each $w \in \mathscr{X}^0(M_0)$) determines a map $z_\# w: Z \to \mathbb{R}^m$.

(1.3) A vector bundle is a generalization of a vector space i.e. a vector space is a vector bundle over a point. With that in mind we generalize notations used with vector spaces to vector bundles. For example if $E_1 \to M, \ldots, E_n \to M, F \to M$ are vector bundles over M, $L^n(E_1, \ldots, E_n; F) \to M$ denotes the vector bundle whose fibre at $x \in M$ is given by:

$$L^n(E_1, \ldots, E_n; F)_x = L^n(E_{1x}, \ldots, E_{nx}; F_x).$$

For example a section Φ of the vector bundle $L(TM) \to M$ is the same as a vector bundle endomorphism $\Phi: TM \to TM$ covering the identity.

Now let $\mathscr{F}(M)$ be a ring of functions on a manifold M. Generally we assume $C^\infty(M) \subset \mathscr{F}(M) \subset C^0(M)$, where $C^r(M)$ $(r = 0, 1, \ldots, \infty)$ denotes the ring of r-times continuously differentiable real valued functions on M. Now if $\mathbb{E}$ is a finite dimensional vector space then $\mathscr{F}(M) \otimes \mathbb{E}$ would denote the $\mathscr{F}(M)$-module of maps $M \to \mathbb{E}$ such that each component (with respect to any basis) lies in $\mathscr{F}(M)$. For example $\mathscr{F}(M) \otimes \mathbb{R}^n$ denotes the space of all maps $u = (u_1, \ldots, u_n): M \to \mathbb{R}^n$ such that $u_1, \ldots, u_n \in \mathscr{F}(M)$. Thus if $E \to M$ is a vector bundle, it is natural to denote by $\mathscr{F}(M) \otimes E$ the space of sections of E represented by elements of $\mathscr{F}(M)$ in local coordinates. (This is meaningful as $C^\infty(M) \subset \mathscr{F}(M)$ and $\mathscr{F}(M)$ is a ring and hence closed under multiplication $C^\infty(M) \times \mathscr{F}(M) \to \mathscr{F}(M)$.) Alternatively we can define $\mathscr{F}(M) \otimes E$ by embedding E in a trivial bundle: $E \subset M \times \mathbb{R}^n$ (so that $E_x \subset \mathbb{R}^n$ for $x \in M$) and defining:
$$\mathscr{F}(M) \otimes E = \{w \in \mathscr{F}(M) \otimes \mathbb{R}^n : w(x) \in E_x \quad \text{for all } x \in M\}.$$
The definition is independent of the embedding as $\mathscr{F}(M)$ is closed under multiplication by an element of $C^\infty(M)$. We generally use the abbreviations
$$\mathscr{X}(M) = \mathscr{F}(M) \otimes TM$$
$$\mathscr{L}(M) = \mathscr{F}(M) \otimes L(TM)$$
$$\mathscr{X}^r(M) = C^r(M) \otimes TM$$
so that $\mathscr{X}^r(M)$ is the space of C^r-vector fields on M.

(1.4) We adhere to the convention that $|u|_r$ denotes (some variant of) the C^r-sup norm of u while $\|u\|_r$ denotes a more exotic norm determined by the derivatives of order $\leq r$. Thus the norms $|u|_r, r \in \mathbb{N}$, determine the topology of $C^\infty(M)$ and the norms $\|u\|_r, r \in \mathbb{N}$, determine the topology of $\mathscr{F}(M)$.

(1.5) For $f \in \text{Diff}(M)$ (the space of C^∞ diffeomorphisms from M onto M) and $X \subset M$ we denote by X^f, X^{f+}, X^{f-} the *orbit*, the *positive orbit*, and the *negative orbit* of X by f:
$$X^f = \bigcup_{n \in \mathbb{Z}} f^n(X)$$
$$X^{f+} = \bigcup_{n \in \mathbb{N}} f^n(X)$$
$$X^{f-} = \bigcup_{n \in \mathbb{N}} f^{-n}(X).$$

2. *The tame category*
(2.1) Let $\mathbb{E}$ be a vector space. An $\mathbb{N}$-*norm* on $\mathbb{E}$ is a sequence of norms
$$\mathbb{E} \to [0, \infty): u \to \|u\|_r, \quad r \in \mathbb{N}$$
on $\mathbb{E}$ which are non-decreasing:
$$\|u\|_0 \leq \|u\|_1 \leq \|u\|_2 \leq \cdots.$$
Two $\mathbb{N}$-norms are *tamely equivalent* iff there is an integer p such that for each $k \in \mathbb{N}$ there exists $c = c(k) \in \mathbb{R}$ with
$$\|u\|'_k \leq c\|u\|_{k+p},$$
$$\|u\|_k \leq c\|u\|'_{k+p},$$

for all $k \in \mathbb{N}$; the $\mathbb{N}$-norms are *equivalent* iff this holds with $p = 0$. An $\mathbb{N}$-*normed space* is a vector space $\mathbb{E}$ together with an $\mathbb{N}$-norm on $\mathbb{E}$. (All our results are invariant under the replacing of an $\mathbb{N}$-norm by a tamely equivalent $\mathbb{N}$-norm; to emphasize this fact we may define an $\mathbb{N}$-*normable space* to be a vector space $\mathbb{E}$ together with a tame equivalence class of $\mathbb{N}$-norms on $\mathbb{E}$ but generally we shall eschew this terminology.)

A subset U of an $\mathbb{N}$-normed space $\mathbb{E}$ is called *open of degree p* iff for every $u_0 \in U$ there exists $\varepsilon > 0$ such that

$$\|u - u_0\|_p < \varepsilon \Rightarrow u \in U.$$

The set U is *open* iff it is the union (as p varies) of sets which are open of degree p. This makes the $\mathbb{N}$-normed space $\mathbb{E}$ into a topological vector space. We shall always assume that $\mathbb{E}$ is complete in this topology.

Let $\mathbb{E}$ and $\mathbb{F}$ be $\mathbb{N}$-normed spaces and $U \subset \mathbb{E}$ be open. A map $P: U \to \mathbb{F}$ *satisfies a tame estimate of degree p* iff it is continuous and for every $k \in \mathbb{N}$ there exists $c = c(k)$ such that we have the 'tame estimate':

$$\|P(u)\|_k \le c(1 + \|u\|_{p+k})$$

for $u \in U$; the map P is *tame of degree p* iff every point $u_o \in U$ has a neighbourhood $U_0 \subset U$ such that $P | U_0$ *satisfies a tame estimate of degree p*; it is *tame* iff it is tame of degree p for some p.

PROPOSITION 1. *A composition of tame maps is tame.*

In other words, we have a category whose objects are open sets in $\mathbb{N}$-normed spaces and whose morphisms are tame maps. This category is a generalization of the category of open sets of Banach spaces and continuous maps between them.

(2.2) If $\mathbb{E}$ and $\mathbb{F}$ are $\mathbb{N}$-normed spaces so is the product $\mathbb{E} \times \mathbb{F}$; the $\mathbb{N}$-norm is defined to be:

$$\|(u, v)\|_k = \|u\|_k + \|v\|_k$$

for $(u, v) \in \mathbb{E} \times \mathbb{F}$, $k \in \mathbb{N}$.

Given $\mathbb{N}$-normed spaces $\mathbb{E}_1, \ldots, \mathbb{E}_m$ define a 'multi-norm' by:

$$\|u_1 : \cdots : u_m\|_{p,k} = \sum_{i=1}^m \|u_i\|_{p+k} \|u_1\|_p \cdots \|\hat{u}_i\|_p \cdots \|u_m\|_p$$

for $u_1 \in \mathbb{E}_1, \ldots, u_m \in \mathbb{E}_m$. (The hat indicates that the factor $\|u_i\|_p$ is omitted.) For example:

$$\|u_1\|_{p,k} = \|u_1\|_{p+k}$$

$$\|u_1 : u_2\|_{p,k} = \|u_1\|_{p+k} \|u_2\|_p + \|u_1\|_p \|u_2\|_{p+k}.$$

If p is not present, 0 is assumed:

$$\|u_1 : \cdots : u_m\|_k = \|u_1 : \cdots : u_m\|_{0,k}.$$

PROPOSITION 2. *A multi-linear map $A \in L^m(\mathbb{E}_1, \ldots, \mathbb{E}_m; \mathbb{F})$ is tame of degree p if and only if for each k there exists $c = c(k)$ such that*

$$\|Au_1 \cdots u_m\|_k \le c \|u_1 : \cdots : u_m\|_{p,k}$$

for $u_1 \in \mathbb{E}_1, \ldots, u_m \in \mathbb{E}_m$.

Proof. See [**5**, p. 141].

(2.3) There seems to be no way to make the space of tame multi-linear maps into an $\mathbb{N}$-normed space, but we achieve the same effect by focusing on the evaluation map. Define

$$\llbracket u: u_1: \cdots : u_m \rrbracket_{p,k} = (1 + \|u\|_{p+k}) \|u_1\|_p \cdots \|u_m\|_p + \|u_1: \cdots : u_m\|_{p,k},$$

for $u \in \mathbb{E}$ and $u_1 \in \mathbb{E}_1, \ldots, u_m \in \mathbb{E}_m$. Call a map $A: U \subset \mathbb{E} \to L^m(\mathbb{E}_1, \ldots, \mathbb{E}_m; \mathbb{F})$ a *continuous* [resp. *tame*; resp. *tame of degree p*] *family* iff the evaluation map

$$U \times \mathbb{E}_1 \times \cdots \times \mathbb{E}_m \to \mathbb{F}$$

$$(u, u_1, \ldots, u_m) \to A(u)u_1 \cdots u_m$$

is continuous [resp. tame; resp. tame of degree p].

PROPOSITION 3. *The family A is a tame family of degree p if and only if it is a continuous family and for each $u_0 \in U$ there exists a neighbourhood U_0 of u_0 in U such that for each $k \in \mathbb{N}$ there is $c = c(k)$ such that*

$$\|A(u)u_1 \cdots u_m\|_k \le c(k) \llbracket u: u_1: \cdots : u_m \rrbracket_{p,k}$$

for $u \in U_0$, $u_1 \in \mathbb{E}_1, \ldots, u_m \in \mathbb{E}_m$.

Proof. See [**5**, lemma 2.1.7].

(2.4) A map $P: U \subset \mathbb{E} \to \mathbb{F}$ is *tame-C^r* [*of degree p*] iff there exist (necessarily unique) tame [of degree p] families $D^m P: U \to L_s^m(\mathbb{E}, \mathbb{F})$, $m = 0, 1, \ldots, r$, with $D^0 P = P$ such that

$$D^m P(u)u_1 \cdots u_m = \frac{d}{dt} D^{m-1} P(u + tu_1)u_2 \cdots u_m \Big|_{t=0}$$

for $u \in U$; $u_1, \ldots, u_m \in \mathbb{E}$; $m = 1, \ldots, r$.

PROPOSITION 4. *Assume P is tame-C^{r+1} of degree p and for $u, u + h \in U$ define $Q(u, h) \in \mathbb{F}$ by*

$$P(u + h) = \sum_{m=1}^{r} \frac{D^m P(u)}{m!} h^m + Q(u, h).$$

Then for each $k \in \mathbb{N}$ there exists $c = c(k)$ such that

$$\|Q(u, h)\|_k \le c \llbracket u: \underbrace{h: \cdots : h}_{r+1} \rrbracket_{p,k}.$$

Proof. Use Taylor's formula.

(2.5) A splitting $\mathbb{E} = \mathbb{E}_1 \oplus \mathbb{E}_2 \oplus \cdots \oplus \mathbb{E}_m$ of an $\mathbb{N}$-normed space is called *tame* iff the associated projections $e_i \in L(\mathbb{E})$ determined by

$$1 = e_1 + e_2 + \cdots + e_m$$

$$\mathbb{E}_i = e_i \mathbb{E}$$

are tame. More generally let $\mathcal{N}$ be an open subset of an $\mathbb{N}$-normed space. A *tame family of splittings* of $\mathbb{E}$ indexed by $\mathcal{N}$ is a function which assigns to each $z \in \mathcal{N}$ a splitting

$$\mathbb{E} = \mathbb{E}_1(z) \oplus \mathbb{E}_2(z) \oplus \cdots \oplus \mathbb{E}_m(z)$$

such that for each $i = 1, \ldots, m$ the map $e_i: \mathcal{N} \to L(\mathbb{E})$ is a tame linear family.

(2.6) A *family of smoothing operators* on the $\mathbb{N}$-normed space $\mathbb{E}$ is a map $S:[1,\infty) \to L(\mathbb{E})$ such that there exists $d \in \mathbb{N}$ (called the *offset* of the family) and constants $c = c(q)$, $q \in \mathbb{N}$, satisfying

$$\|S(t)u\|_q \le ct^{q-p+d}\|u\|_p,$$
$$\|u - S(t)u\|_p \le ct^{p-q+d}\|u\|_q,$$

for $u \in \mathbb{E}$ and $p \le q$.

The $\mathbb{N}$-normed space $\mathbb{E}$ is said to satisfy the *interpolation inequalities* iff for each triple $p, q, r \in \mathbb{N}$ with $p \le q \le r$ there is a constant $c = c(p, q, r)$ such that

$$\|u\|_q^{r-p} \le c\|u\|_p^{r-q}\|u\|_r^{q-p},$$

for $u \in \mathbb{E}$. Some authors call these 'convexity inequalities' since they have the form $\|u\|_q \le c\|u\|_p^{\lambda}\|u\|_r^{1-\lambda}$ where $0 \le \lambda = (r-q)/(r-p) \le 1$.

PROPOSITION 5. *If $\mathbb{E}$ admits a family of smoothing operators with offset $d = 0$, then it satisfies the interpolation inequalities.*

Proof. Write $u = S(t)u + (1 - S(t))u$ to obtain:

$$\|u\|_q \le c(t^{q-p}\|u\|_p + t^{q-r}\|u\|_r).$$

Then take $t = (\|u\|_r/\|u\|_p)^{1/(r-p)}$ so that

$$t^{q-p}\|u\|_p = t^{q-r}\|u\|_r = \|u\|_p^{\lambda}\|u\|_r^{1-\lambda}.$$

Example. Let $X \subset \mathbb{R}^m$ be open and denote by $\mathscr{B}(X) \subset C^{\infty}(X)$ the $\mathbb{N}$-normed space given by

$$u \in \mathscr{B}(X) \Leftrightarrow |u|_r < \infty \quad \text{for all } r \in \mathbb{N},$$

where $|u|_r = \sup\{|\partial^{\alpha}u(x)|: x \in X, |\alpha| \le r\}$. For $X = \mathbb{R}^m$ one can construct (using convolutions) a family of smoothing operators with offset $d = 0$. If X is the interior of a smooth compact manifold, we construct smoothing operators $u \to S(t)u$ by smoothing the function $x \to u(x + t^{-1}\xi(x))$ where ξ is a vector field which points in on the boundary. This family has offset $d = 1$. It seems to be unknown which spaces $\mathscr{B}(X)$ admit smoothing operators and what is the smallest possible offset. On the other hand, regardless of the nature of the boundary of X the space $\mathscr{B}(X)$ satisfies the interpolation inequalities if X is convex. (See [**5**, p. 143].)

(2.7) *Notes.* Our definitions follow [**5**] and [**19**]. Hamilton calls $\mathbb{N}$-normed spaces 'graded Frechet spaces' but we prefer the term '$\mathbb{N}$-normed space' as it suggests the possibility of generalizing to other index sets. Sergeraert (essentially) calls open subsets of $\mathbb{N}$-normed spaces 'L-objects' and tame maps 'L-morphisms'. We do not adopt Hamilton's notion of 'tame space' as his tame spaces all admit smoothing operators with offset $d = 0$ and it is not clear that this will be true in the applications we have in mind. Sergeraert also assumes offset $d = 0$ but the original treatment of Moser [**11**] allowed arbitrary offset. The proof of proposition 5 comes from [**19**, p. 604] or [**26**, p. 118]. Note the similarity to the abstract interpolation theory exposed e.g. in [**15**, p. 32]; see also [**12**, p. 272].

3. *The tame implicit function theorem*

Suppose $\mathbb{H}$, $\mathbb{E}$, $\mathbb{F}$ are $\mathbb{N}$-normed spaces, $Z \subset \mathbb{H}$, $U \subset \mathbb{E}$ are open, $z_0 \in Z$, $u_0 \in U$, and $P: Z \times U \to \mathbb{F}$.

THEOREM. *Assume P is tame-C^2, that*

$$P(z_0, u_0) = 0, \tag{1}$$

and that there is a tame family $L: Z \times U \to L(\mathbb{F}, \mathbb{E})$ such that

$$D_2 P(z, u) L(z, u) v = v$$

for $z \in Z$, $u \in U$, $v \in \mathbb{F}$. Assume that $\mathbb{E}$ admits a family $\{S(t): t \geq 1\}$ of smoothing operators and satisfies the interpolation inequalities. Then there is a neighbourhood Z_0 of z_0 in Z and a continuous map $G: Z_0 \to U$ such that $G(z_0) = u_0$, $P(z, G(z)) = 0$ for $z \in Z_0$.

Proof. We may assume that U has form

$$U = \{ u \in \mathbb{E}: \|u - u_0\|_p < M^{-1} \} \tag{2}$$

where $p \in \mathbb{N}$ and M is a large positive number. Note that we may increase p and M (thus shrinking U) in the course of the proof without loss of generality; we may also shrink Z. The hypothesis that P is C^2 gives inequalities

$$\|P(z, u)\|_q \leq M \tag{3}$$

$$\|D_2 P(z, u)\hat{u}\|_q \leq M \|\hat{u}\|_p \tag{4}$$

$$\|P(z, u + \hat{u}) - P(z, u) - D_2 P(z, u)\hat{u}\|_q \leq M \|\hat{u}\|_p^2 \tag{5}$$

for $u, u + \hat{u} \in U$ and $z \in Z$; here $p - q$ is the degree of P.

We shall specify an integer $m \in \mathbb{N}$ below. The hypothesis that P is tame allows us to conclude that

$$\|P(z, u)\|_{q+m} \leq M(1 + \|u\|_{p+m}) \tag{6}$$

for $z \in Z$ and $u \in U$ provided that we assume that $\|z\|_{p+m}$ is bounded on Z, i.e. that Z satisfies

$$Z \subset \{ z \in \mathbb{H}: \|z\|_{p+m} < M \}.$$

We now define an 'approximate right inverse' to $D_2 P$. Let $L: [1, \infty) \times Z \times U \to L(\mathbb{F}, \mathbb{E})$ be given by $L(t, z, u) = S(t) L(z, u)$. We then have inequalities

$$\|L(t, z, u)v\|_p \leq M t^b \|v\|_q \tag{7}$$

$$\|L(t, z, u)v\|_{p+m} \leq M t^b [\![u: v]\!] \tag{8}$$

$$\|v - D_2 P(z, u) L(t, z, u)v\|_q \leq M t^{-a} [\![u: v]\!] \tag{9}$$

for $z \in Z$, $u \in U$, $v \in \mathbb{F}$ where $[\![u: v]\!] = (1 + \|u\|_{p+m})\|v\|_q + \|v\|_{q+m}$, $b = p - r + d$, $a = m - b$; the degree of the tame map $Z \times U \times \mathbb{F} \to \mathbb{E}$, $(z, u, v) \to L(z, u)v$ is $q - r$; and d is the offset of the family S of smoothing operators. We need (7), (8), (9) only when $v = P(z, u)$ in which case (enlarging M if necessary) they take the form:

$$\|L(t, z, u) P(z, u)\|_p \leq M t^b \|P(z, u)\|_q \tag{10}$$

$$\|L(t, z, u) P(z, u)\|_{p+m} \leq M t^b (1 + \|u\|_{p+m}) \tag{11}$$

$$\|P(z, u) - D_2 P(z, u) L(t, z, u) P(z, u)\|_q \leq M t^{-a} (1 + \|u\|_{p+m}) \tag{12}$$

for $t \in [1, \infty)$, $z \in Z$, $u \in U$. Note also that for each fixed t the map

$$Z \times U \to \mathbb{E} : (z, u) \to L(t, z, u) P(z, u) \tag{13}$$

is continuous. We take

$$m = 8(p - r + d) + 1 \tag{14}$$

so that $7b < a$. We define G by $G(z) = \lim_{n \to \infty} u_n$ where the sequence $u_n \in U$ is defined inductively by

$$u_{n+1} = u_n - L_n P_n \tag{15}$$

with the abbreviations $L_n = L(t_n, z, u_n)$, $P_n = P(z, u_n)$, and where $t_n \to \infty$ is a sequence of real numbers specified below. We shall define sequences $\delta_n \to 0$ and $\Delta_n \to \infty$ and establish inequalities

$$\|u_n - u_{n-1}\|_p \le \delta_n \qquad n \ge 1, \tag{$\delta : n$}$$

$$1 + \|u_n\|_{p+m} \le \Delta_n \qquad n \ge 0. \tag{$\Delta : n$}$$

In order to assure that the sequence u_n is well defined (i.e. that $u_n \in U$) we require the inequality

$$\sum_{n=1}^{\infty} \delta_n < M^{-1} \tag{16}$$

which together with $(\delta : k)$, $k = 1, \ldots, n-1$, (2), and the triangle inequality assures that $u_{n-1} \in U$ (so that u_n is well-defined).

Step 1. For a sufficiently small neighbourhood Z_0 of z_0 in Z we have $(\delta : 1)$ for $z \in Z_0$.

Proof. One takes $Z_0 = \{z \in Z : \|u_1 - u_0\|_p < \delta_1\}$ which is open by (13) and (15) (with $n = 0$).

Step 2. We have $(\Delta : 0)$ if

$$1 + \|u_0\|_{p+m} \le \Delta_0. \tag{17}$$

Step 3. We have $(\delta : n+1) \wedge (\Delta : n) \Rightarrow (\delta : n+2)$ if

$$2M^2 t_{n+1}^b t_n^{-a} \Delta_n \le \delta_{n+2}, \tag{18}$$

$$2M^2 t_{n+1}^b \delta_{n+1}^2 \le \delta_{n+2}. \tag{19}$$

Proof. By (15) and (10) $\|u_{n+2} - u_{n+1}\|_p \le M t_{n+1}^b \|P_{n+1}\|_q$ and by (15), (12), (5)

$$\|P_{n+1}\|_q \le M t_n^{-a} \Delta_n + M \delta_{n+2}^2. \tag{20}$$

Step 4: We have $(\Delta : n) \Rightarrow (\Delta : n+1)$ if

$$2\Delta_n \le \Delta_{n+1} \tag{21}$$

$$2M t_n^b \Delta_n \le \Delta_{n+1}. \tag{22}$$

Proof. By (15) and (11) $\|u_{n+1} - u_n\|_{p+m} \le M t_n^b (1 + \|u_n\|_{p+m})$. Now use

$$1 + \|u_{n+1}\|_{p+m} \le 1 + \|u_n\|_{p+m} + \|u_{n+1} - u_n\|_{p+m}.$$

Now note that (8) generalizes to

$$\|L(t, z, u)v\|_{p+k} \le M_k t^b [\![z, u : v]\!]_k \tag{8:k}$$

where

$$[\![z, u : v]\!]_k = (1 + \|z\|_{p+k} + \|u\|_{p+k}) \|v\|_q + \|v\|_{q+k},$$

which specializes (making M_k larger) to

$$\|L(t, z, u)P(z, u)\|_{p+k} \le M_k t^b (1 + \|z\|_{p+k} + \|u\|_{p+k}) \tag{11:k}$$

as before. This estimate holds for $z \in Z_0$, $u \in U_0$, and $t \in [1, \infty)$ without shrinking Z_0 or U_0 further and of course M_k is independent of t, z, u.

Step 5. For each k there is a constant C_k such that

$$1 + \|u_n\|_{p+k} \le C_k (1 + \|z\|_{p+k}) \Delta_n \tag{23}$$

provided that

$$\lim_{n \to \infty} \frac{t_n^b \Delta_n}{\Delta_{n+1}} = 0. \tag{24}$$

Proof.

$$
\begin{aligned}
1 + \|u_{n+1}\|_{p+k} &\le 1 + \|u_n\|_{p+k} + \|u_{n+1} - u_n\|_{p+k} \\
&= 1 + \|u_n\|_{p+k} + \|L_n P_n\|_{p+k} \\
&\le 1 + \|u_n\|_{p+k} + M_k t_n^b (1 + \|z\|_{p+k} + \|u_n\|_{p+k}) \\
&\le (1 + M_k t_n^b)(1 + \|u_n\|_{p+k}) + (1 + M_k t_n^b)(1 + \|z\|_{p+k}).
\end{aligned}
$$

Hence we can prove (23) inductively if $2(1 + M_k t_n^b)\Delta_n \le \Delta_{n+1}$ so by (24) the inequality (23) holds for all $n \ge n(k)$ if it holds for $n = n(k)$. But for any fixed value of n, (23) merely expresses the tame estimate on the map $z \to u_n$ and can be proved (with C_k depending on n) by induction on n.

Step 6. The sequence u_n converges uniformly in each norm $\|\cdot\|_{p+j}$ provided that there exists $l \in \mathbb{N}$ such that

$$\sum_{n=1}^{\infty} \delta_n^{(l-1)/l} \Delta_n^{1/l} < \infty. \tag{25}$$

Proof. Given j let $k = lj$. By interpolation

$$\|u_n - u_{n-1}\|_{p+j} \le C \|u_n - u_{n-1}\|_p^{1-(j/k)} \|u_n - u_{n-1}\|_{p+k}^{j/k}.$$

But

$$
\begin{aligned}
\|u_n - u_{n-1}\|_{p+k}^{j/k} &\le (\|u_n\|_{p+k} + \|u_{n-1}\|_{p+k})^{j/k} \\
&\le 2 C_k (1 + \|z\|_{p+k}) \Delta_n^{j/k}
\end{aligned}
$$

by step 5. The result is immediate by the Weierstrass M-test and the inequality $(\delta : n)$.

Step 7. The limit $u = G(z)$ satisfies $P(z, u) = 0$ as required.

Proof. $\lim \|P_n\|_q = 0$ by (16), (18), (20).

It remains to find the sequences L_n, δ_n, Δ_n satisfying (16)–(19), (21), (22), (24), (25). For this take $t_n = \exp(\gamma K^n)$, $\delta_n = \exp(-\alpha\gamma K^n)$, $\Delta_n = \exp(\beta\gamma K^n)$, with α, β, γ, K positive. (16) and (17) can be achieved by choosing γ large while (18), (19), (21), (22) take the form:

$$bK - a + \beta < -\alpha K^2 \tag{18'}$$

$$b - 2\alpha < -\alpha K \tag{19'}$$

$$1 < K \tag{21'}$$

$$b + \beta < \beta K. \tag{22'}$$

These in turn yield:

$$b(k-1)^{-1} < \beta, \quad b(2-K)^{-1} < \alpha, \quad \alpha K^2 + bK + \beta < a,$$

so:

$$(2-K)^{-1}K^2 + K + (K-1)^{-1} < a/b.$$

The expression on the left is minimized for $K = 4/3$ and takes the value 7. This leads to the solution:

$$K = 4/3, \quad \alpha = (3b + 2\varepsilon)/2, \quad \beta = 3b + \varepsilon,$$

where $3\varepsilon < a - 7b$.

Now (24) follows from (22′); for (25) we require

$$-\alpha(l-1) + \beta < 0 \tag{25′}$$

which is satisfied with $l = 3$.

Remark. The proof has been arranged so that all the steps follows from the hypotheses (1)–(9). Thus, if one can construct $L(t, z, u)v$ satisfying (7), (8), (9) one can conclude the existence of $G(z)$, even if $\mathbb{E}$ does not admit smoothing operators. (The interpolation inequalities and the hypotheses (8:k) are required only to obtain convergence and continuity in the higher norms.)

Notes. We have followed very closely the proof of Sergeraert [19] who in turn followed Schwartz [18], [18a]. This is also the argument of Moser [11]. Our formulation is slightly different from that of Sergeraert in that he proves the inverse function theorem whereas we prove the implicit function theorem. Also he proves more; namely that if P is tame-C^{r+2} (and L is tame-C^r), then G is C^r.

Note (as Sergeraert did) that we have not proved that G is tame; indeed, from step 6 we can only conclude $\|G(z)\|_{p+j} \leq C(1 + \|z\|_{p+3j})$. Hamilton [5] gives a proof which yields the tameness of G but his hypotheses would require that both P and L be tame-C^∞; this is tedious to verify (I think) in our applications. Hamilton uses the original method of Nash [13] defining the solution as the limit of the solution to a differential – rather than a difference – equation.

An earlier version of the proof presented here had the hypothesis $11b < a$ instead of $7b < a$. Thanks to Raz Stowe and Mike Field for pointing out the improvement and indicating that this is in some sense the best possible. As $7b < a$ leads to $K = 4/3$ which agrees with Sergeraert, I presume that he followed this reasoning as well.

4. *The tame contraction principle*

In this section we present another (easier) version of the tame implicit function theorem. It has stronger hypotheses and a stronger conclusion: *viz.* that the solution map G is tame.

Let $\mathbb{H}$ and $\mathbb{E}$ be $\mathbb{N}$-normed spaces, $Z \subset \mathbb{H}$ and $U \subset \mathbb{E}$ be open, $z_0 \in Z$ and $u_0 \in U$ and $S: Z \times U \to \mathbb{E}$ be C^2-tame of degree 0. Assume

$$S(z_0, u_0) = u_0 \tag{1}$$

and

$$\|D_2 S(z, u)^n \hat{u}\|_r \leq C\sigma^n \llbracket z, u: \hat{u} \rrbracket_r \tag{2}$$

for $z \in Z$; $u, u + \hat{u} \in U$. Here $[\![z, u : \hat{u}]\!]_r = (1 + \|z\|_r + \|u\|_r)\|\hat{u}\|_0 + \|\hat{u}\|_r$, $0 < \sigma < 1$, and the constants C and σ are independent of $z, u, \hat{u}$ and $n \in \mathbb{N}$. (They are as usual allowed to depend on $r \in \mathbb{N}$.) Note that were it not for the requirement that C and σ be independent of n, the inequality (2) would follow from the tame estimate on $D_2 S$.

THEOREM. *Under these hypotheses there exist neighbourhoods Z_0 of z_0 in Z and U_0 of u_0 in U and a tame map $G : Z_0 \to U_0$ such that*

$$S(z, u) = u \Leftrightarrow u = G(z)$$

for $z \in Z_0$ and $u \in U_0$.

Proof. We solve the equation $S(z, u) = u$ by Newton's method:

$$u_{n+1} = u_n - (1 - D_2 S(z, u_n))^{-1}(u_n - S(z, u_n)),$$
$$G(z) = \lim_{n \to \infty} u_n. \tag{3}$$

The inverse operator can be calculated via the Neumann series:

$$(1 - D_2 S(z, u))^{-1}\hat{u} = \sum_{n=0}^{\infty} D_2 S(z, u)^n \hat{u},$$

and (2) gives the tame estimate (enlarging C):

$$\|(1 - D_2 S(z, u))^{-1}\hat{u}\|_r \leq C[\![z, u : \hat{u}]\!]_r. \tag{4}$$

Now

$$u_n - S(z, u_n) = Q(z, u_{n-1}, u_n - u_{n-1}), \tag{5}$$

where Q is defined by

$$S(z, u + \hat{u}) = S(z, u) + D_2 S(z, u)\hat{u} + Q(z, u, \hat{u}),$$

so that (3), (4), and the appropriate tame estimate on Q yield:

$$\|u_{n+1} - u_n\|_r \leq C[\![z, u_{n-1} : u_n - u_{n-1}]\!]_r \|u_n - u_{n-1}\|_0. \tag{6}$$

In the case where $r = 0$ this becomes $\|u_{n+1} - u_n\|_0 \leq C\|u_n - u_{n-1}\|_0^2$ so that by induction:

$$\|u_{n+1} - u_n\|_0 \leq \delta_n, \tag{7}$$

provided that this holds for $n = 0$ and $\delta_n = (C\delta_0)^{2^n} C^{-1}$. Let $\mathbb{E}_0$ denote the completion of $\mathbb{E}$ in the norm $\|\cdot\|_0$ and for $\varepsilon > 0$ let

$$U_0 = \{u \in \mathbb{E} : \|u - u_0\|_0 < \varepsilon\}$$
$$U_0^* = \{u \in \mathbb{E}_0 : \|u - u_0\|_0 < \varepsilon\}.$$

Choose Z_0 and $\varepsilon > 0$ so small that $S(z, u) = u$ has a unique solution $u \in U_0^*$ for $z \in Z_0$ (e.g. by Banach's contraction principle) and shrink Z_0 further so that (7) holds for $n = 0$ when $z \in Z_0$. Assume δ_0 is so small that $\sum \delta_n < \varepsilon$.

It remains to prove that the sequence u_n converges in each norm $\|\cdot\|_r$ (so that $G(z) \in U_0$), that $u = G(z)$ solves $S(z, u) = u$, and that G is tame. First (6) and (7) imply

$$\|u_{n+1} - u_n\|_r \leq C(1 + \|z\|_r + \|u_n\|_r + \|u_{n-1}\|_r)\delta_{n-1} \tag{8}$$

(for sufficiently large n) which in turn implies

$$\|u_{n+1}\|_r \le (1 + \|z\|_r) + 2\|u_n\|_r + 2\|u_{n-1}\|_r$$

for sufficiently large n, say $n \ge m$. By induction we obtain the crude estimate:

$$\|u_n\|_r \le 2^{n-m}(1 + \|z\|_r + \|u_m\|_r + \|u_{m-1}\|_r). \tag{9}$$

But u_m, u_{m-1} are tame functions of z so that (8) and (9) give

$$\|u_{n+1} - u_n\|_r \le C 2^n (1 + \|z\|_r)\delta_n. \tag{10}$$

But $\sum 2^n \delta_n < \infty$ so that (10) gives both convergence and a tame estimate. The equation $G(z) - S(z, G(z)) = 0$ follows by replacing u_n and u_{n-1} by the limit $G(z)$ in (5).

5. $\mathbb{N}$-jet semi-Finslers

(5.1) Let M be a smooth manifold, not necessarily compact and let $E \to M$ be a vector bundle. A *semi-Finsler* [resp. a *Finsler*] on $E \to M$ is a continuous map $\Theta: E \to [0, \infty)$ such that for each $x \in M$ the restriction to the fibre $\Theta|E_x$ is a semi-norm [resp. a norm]. Two semi-Finslers Θ and Θ' on $E \to M$ are called *locally equivalent* iff there is a continuous function $c: M \to (0, \infty)$ such that

$$c(x)^{-1}\Theta(v) \le \Theta'(v) \le c(x)\Theta(v)$$

for $v \in E_x$, $x \in M$.

PROPOSITION. *Any two Finslers on $E \to M$ are locally equivalent. If M is compact, any two locally equivalent semi-Finslers on $E \to M$ are equivalent.*

(5.2) The following trivial observation is useful for constructing semi-Finslers with specific local properties.

Patching construction. Suppose $\{M_i: i \in I\}$ is a collection of open sets which cover M and Θ^i is a semi-Finsler on $E_i = E|M_i$. Then there is a semi-Finsler Θ on $E \to M$ such that $\Theta|E_i$ is locally equivalent to Θ^i for each $i \in I$ if and only if for each $i, j \in I$, $\Theta^i|E_i \cap E_j$ and $\Theta^j|E_i \cap E_j$ are locally equivalent semi-Finslers.

Proof. Use partitions of unity.

(5.3) An *r-jet semi-Finsler* [resp. *r-jet Finsler*] on M is a semi-Finsler [resp. Finsler] on the bundle $J^r(M) \to M$ of r-jets of real-valued functions on M. Given $u \in C^\infty(M)$ and $x \in M$ let

$$\|u\|_{r,x} = \Theta(j^r u(x))$$

and

$$\|u\|_r = \sup\{\|u\|_{r,x}: x \in M\}.$$

If we require more precision we may use Θ as a superscript, $\|u\|_{r,x}^\Theta = \|u\|_{r,x}$, $\|u\|_r^\Theta = \|u\|_r$, but generally it will be clear from the context which semi-Finsler we are using. When Θ is a Finsler we shall use single bars rather than double: $|u|_{r,x} = \|u\|_{r,x}$, $|u|_r = \|u\|_r$.

An $\mathbb{N}$-jet *semi-Finsler* [resp. $\mathbb{N}$-jet *Finsler*] is a non-decreasing sequence $\Theta = \{\Theta_r: r \in \mathbb{N}\}$ where Θ_r is an r-jet semi-Finsler [resp. r-jet Finsler] such that the

first one is a Finsler. Thus

$$\|u\|_{0,x} \leq \|u\|_{1,x} \leq \cdots \leq \|u\|_{r,x} \leq \cdots$$

and

$$\|u\|_{r,x} = 0 \Rightarrow u(x) = 0.$$

We call two $\mathbb{N}$-jet semi-Finslers Θ and Θ' *equivalent* [resp. *locally equivalent*] iff for each $r \in \mathbb{N}$ the semi-Finslers Θ_r, and Θ'_r are equivalent [resp. locally equivalent].

An $\mathbb{N}$-jet semi-Finsler determines an $\mathbb{N}$-norm on $C_0^\infty(M)$; we denote by $\mathcal{B}(M, \Theta)$ the completion. Thus

$$u \in \mathcal{B}(M, \Theta) \Leftrightarrow \|u\|_r < \infty \qquad \text{for all } r \in \mathbb{N}.$$

This is an $\mathbb{N}$-normed space. We denote by $\mathcal{F}(M, \Theta) = \mathcal{B}(M, \Theta)_{\text{loc}}$ the 'localization'; i.e.

$$u \in \mathcal{F}(M, \Theta) \Leftrightarrow \psi u \in \mathcal{B}(M, \Theta) \qquad \text{for all } \psi \in C_0^\infty(M).$$

When M is compact we have $\mathcal{F}(M, \Theta) = \mathcal{B}(M, \Theta)$ while if M is not compact, $\mathcal{F}(M, \Theta)$ is not an $\mathbb{N}$-normed space in any natural way. It is of course an '$\mathbb{N} \times C_0^\infty(M)$-normed space' where $\|u\|_{r,\psi} = \|\psi u\|_r$ for $r \in \mathbb{N}$, $\psi \in C_0^\infty(M)$.

PROPOSITION. *If Θ and Θ' are equivalent then the corresponding $\mathbb{N}$-norms are equivalent so that $\mathcal{B}(M, \Theta) = \mathcal{B}(M, \Theta')$. If Θ and Θ' are locally equivalent, then $\mathcal{F}(M, \Theta) = \mathcal{F}(M, \Theta')$. In particular, when Θ is an $\mathbb{N}$-jet Finsler we have $\mathcal{F}(M, \Theta) = C^\infty(M)$.*

Of course, we always have the inclusions

$$C_0^\infty(M) \subset \mathcal{B}(M, \Theta) \subset \mathcal{F}(M, \Theta) \subset C^0(M).$$

(5.4) When $U \subset \mathbb{R}^n$ is open the bundle $J^r(U) \to U$ is trivial

$$J^r(U) = U \times J^r(n)$$
$$J^r(n) = \mathbb{R} \times L(\mathbb{R}^n, \mathbb{R}) \times L_s^2(\mathbb{R}^n, \mathbb{R}) \times \cdots \times L_s^r(\mathbb{R}^n, \mathbb{R})$$

with

$$j^r u(x) = (u(x), Du(x), D^2 u(x), \ldots, D^r u(x))$$

for $u \in C^\infty(U)$, $x \in U$. There is correspondingly a 'trivial $\mathbb{N}$-jet Finsler' Θ_{triv} defined by

$$|u|_{r,x} = \sum_{k=0}^r |D^k u(x)|.$$

Thus $|u|_r = \sup\{|u|_{r,x} : x \in U\}$ is the usual C^r-sup norm. We define $\mathcal{B}(U) = \mathcal{B}(U, \Theta_{\text{triv}})$ so that

$$u \in \mathcal{B}(U) \Leftrightarrow |u|_r < \infty \qquad \text{for all } r \in \mathbb{R}.$$

(5.5) We can generalize this to define $\mathcal{B}(U)$ for certain open subsets $U \subset M$:

LEMMA. *Let $U \subset M$ be open with compact closure and let Θ and Θ' be $\mathbb{N}$-jet Finslers on M. Then the $\mathbb{N}$-jet Finslers $\Theta|U$ and $\Theta'|U$ are equivalent so that*

$$\mathcal{B}(U, \Theta|U) = \mathcal{B}(U, \Theta'|U).$$

Hence (when U has compact closure) we define $\mathscr{B}(U) = \mathscr{B}(U, \Theta|U)$. (The lemma is of course an easy variant on propositions 1 and 2.)

(5.6) Finally we introduce a bit more notation. Assume that M is compact and that $E \to M$ and $F \to M$ are vector bundles. Given open $V \subset E$ with compact closure we have $\mathscr{B}(V) \subset C^\infty(V)$ as above. Suppose F is embedded as a sub-bundle of a trivial bundle $F \subset M \times \mathbb{R}^n$ and define an $\mathbb{N}$-normed space $\mathscr{B}(V, F)$ by

$$q \in \mathscr{B}(V, F) \Leftrightarrow \begin{cases} q \in \mathscr{B}(V) \otimes \mathbb{R}^n \\ q(V \cap E_x) \subset F_x \quad \text{for all } x \in M. \end{cases}$$

(In other words, $q: V \subset E \to F$ a fibre preserving C^∞ map with bounded derivatives of all orders.) Note that each element $q \in \mathscr{B}(V, F) \subset \mathscr{B}(V) \otimes \mathbb{R}^n$ has form $q = (q_1, \ldots, q_n)$ where $q_i \in \mathscr{B}(V)$ so the $\mathbb{N}$-norm has form

$$|q|_r = \sup\{|q|_{r,y} : y \in V\}$$
$$|q|_{r,y} = |q_1|_{r,y} + \cdots + |q_n|_{r,y}.$$

It is easy to prove that the definition of $\mathscr{B}(V, F)$ is independent of the embedding (up to an equivalent renorming).

6. The composition inequality

(6.1) Let M be a compact manifold and $\mathscr{F}(M)$ be an $\mathbb{N}$-normed space of functions on M. We assume

$$C^\infty(M) \subset \mathscr{F}(M) \subset C^0(M)$$

with continuous inclusion on the right.

$$|u|_0 = \|u\|_0 \tag{1}$$

for $u \in \mathscr{F}(M)$. (As usual we denote the $\mathbb{N}$-norm of $\mathscr{F}(M)$ by $\|u\|_r$ and the usual C^r-sup norm by $|u|_r$.)

Given an open set $U \subset M \times \mathbb{R}^n$ define an open set $\mathscr{F}(M, U) \subset \mathscr{F}(M) \otimes \mathbb{R}^n$ by the requirement that

$$u \in \mathscr{F}(M, U) \Leftrightarrow (x, u(x)) \in U \quad \text{for all } x \in M.$$

Thus each $u \in \mathscr{F}(M, U)$ and $q \in \mathscr{B}(U)$ determines $v: M \to \mathbb{R}$ via the formula

$$v(x) = q(x, u(x)) \tag{2}$$

for $x \in M$.

We say that $\mathscr{F}(M)$ satisfies the *composition inequality* iff (for every U) we have $v \in \mathscr{F}(M)$ and for each $r \geq 1$ there is an inequality

$$\|v\|_r \leq c \sum |q|_k \|u\|_{j_1} \cdots \|u\|_{j_k}, \tag{3}$$

where $c = c(r, U)$ is independent of $q \in \mathscr{B}(U)$ and $u \in \mathscr{F}(M, U)$ and the sum is over all terms with $1 \leq k \leq r$ and $1 \leq j_1, \ldots, j_k \leq r$, and $j_1 + \cdots + j_k = r$.

THEOREM. *Assume $\mathscr{F}(M)$ satisfies the composition inequalities and the interpolation inequalities. Then the map*

$$\mathscr{B}(U) \times \mathscr{F}(M, U) \to \mathscr{F}(M)$$

$$(q, u) \to v$$

is tame-C^k of degree k for each $k \in \mathbb{N}$.

Proof. Interpolation gives estimates

$$|q|_k \leq c|q|_1^{(r-k)/(r-1)}|q|_r^{(k-1)/(r-1)}$$
$$\|u\|_j \leq c\|u\|_1^{(r-j)/(r-1)}\|u\|_r^{(j-1)/(r-1)}.$$

Hence

$$\|u\|_{j_1}\ldots\|u\|_{j_k} \leq c\|u\|_1^{(kr-r)/(r-1)}\|u\|_r^{(r-k)/(r-1)}$$

so that if $\mathcal{N} \subset \mathcal{B}(U) \times \mathcal{F}(M, U)$ satisfies

$$(q, u) \in \mathcal{N} \Leftrightarrow |q|_1 + \|u\|_1 < c$$

we have the estimate

$$\|v\|_r \leq c \sum_{k=1}^{r} |q|_r^{(k-1)/(r-1)} \|u\|_r^{(r-k)/(r-1)}.$$

Since the right hand side is homogeneous of degree one in $(|q|_r, \|u\|_r)$ (or by using the convexity of the log) this gives the tame estimate

$$\|v\|_r \leq c(|q|_r + \|u\|_r)$$

for $(q, u) \in \mathcal{N}$. The derivative $(q, u, \hat{q}, \hat{u}) \to \hat{v}$ of the map $(q, u) \to v$ is given by

$$\hat{v}(x) = \hat{q}(x, u(x)) + D_2 q(x, u(x))\hat{u}(x)$$

which is a map of the same form. The tame estimate on the derivative implies the continuity of the map by the mean value theorem; this completes the proof.

Remark. For *fixed q* one can use the interpolation inequality $\|u\|_j \leq c\|u\|_0^{(r-j)/r}\|u\|_r^{j/r}$ and obtain the global estimate $\|v\|_r \leq c(1 + \|u\|_r)$ for $u \in \mathcal{F}(M, U)$.

COROLLARY 1. *Under the hypotheses of the theorem, $\mathcal{F}(M)$ is a tame algebra; i.e.*

$$\|uv\|_r \leq c(\|u\|_0 \|v\|_r + \|u\|_r \|v\|_0),$$

for $u, v \in \mathcal{F}(M)$.

Proof. Take $n = 2$, $q(x, u, v) = uv$, and $U = \{(x, u, v) : |u| + |v| < 1\}$ in the remark and use proposition 2 of § 2.

COROLLARY 2. *Under the hypothesis of the theorem, each $\psi \in C^\infty(M)$ gives a tame linear map $\mathcal{F}(M) \to \mathcal{F}(M) : u \to \psi u$ of degree* 0.

Proof. $C^\infty(M) \subset \mathcal{F}(M)$.

(6.2) Now let $E \to M$ be a vector bundle and suppose there is a vector bundle embedding $E \subset M \times \mathbb{R}^n$ in a trivial bundle. We define

$$\mathcal{F}(M) \otimes E = \{u \in \mathcal{F}(M) \otimes \mathbb{R}^n : u(x) \in E_x \quad \text{for all } x \in M\}$$

so we obtain an $\mathbb{N}$-normed space of sections of E 'of class $\mathcal{F}(M)$'. By corollary 2 this space is well-defined; i.e. using a different embedding gives an equivalent $\mathbb{N}$-norm.

Suppose $F \to M$ is another vector bundle and $U \subset E$ is an open set with compact closure. Let

$$\mathcal{F}(M, U) = \{u \in \mathcal{F}(M) \otimes E : u(M) \subset U\}$$

and let $\mathcal{B}(U, F)$ be the $\mathbb{N}$-normed space of (5.6). One easily generalizes the composition theorem to:

COROLLARY 3. *The map*

$$\mathcal{B}(U, F) \times \mathcal{F}(M, U) \to \mathcal{F}(M) \otimes F$$

$$(q, u) \to q \circ u$$

is tame-C^r of degree r for every $r \in \mathbb{N}$.

(6.3) Now let Θ be an $\mathbb{N}$-jet semi-Finsler on M. We say that Θ satisfies the *composition inequalities* iff for every $\mathbb{N}$-jet Finsler on $M \times \mathbb{R}^n$ there exist continuous functions $c = c_r \colon M \times \mathbb{R}^n \to (0, \infty)$ such that for $q \in C^\infty(M \times \mathbb{R}^n)$, $u \in C^\infty(M) \otimes \mathbb{R}^n$, we have the pointwise estimate:

$$\|v\|_{r,x} \le c(y) \sum |q|_{k,y} \|u\|_{j_1,x} \cdots \|u\|_{j_k,x},$$

where v is given by (2) and the indices in the sum range as in (3) and $y = (x, u(x)) \in M \times \mathbb{R}^n$. (Note that the definition is meaningful even when M is non-compact.)

PROPOSITION. *If the $\mathbb{N}$-jet semi-Finsler Θ satisfies the composition inequalities and M is compact, then the $\mathbb{N}$-normed space $\mathcal{F}(M, \Theta)$ satisfies the composition inequalities.*

Proof. This is immediate from Faa di-Bruno's formula for the higher derivatives of a composition of maps. Note that in particular an $\mathbb{N}$-jet Finsler satisfies the composition inequality so we recover the result that the map

$$\mathcal{B}(U) \times C^\infty(M, U) \to C^\infty(M)$$

$$(q, u) \to v$$

is tame-C^r of degree r.

(6.4) *Notes.* This section follows [5, p. 147]. See also [6]. The formula of Faa di-Bruno is (in the one dimensional case)

$$D^r(q \circ u)(x) = \sum c D^k q(u(x)) D^{j_1} u(x) \cdots D^{j_k} u(x),$$

where the sum is over indices as in (3) and $c = c(k, j_1, \ldots, j_k, r)$ are certain integers. (See e.g. [7, p. 50].) In the general case the formula is true if the multiplication on the right is appropriately interpreted; see e.g. [1, p. 3]. Thanks to Bob Welland for pointing out that the formula there only holds on the diagonal; the authors neglected to symmetrize the right hand side (mea culpa).

7. Unfoldings

Let M be a compact manifold. An *unfolding* is a pair (F, f_0) where F is a finite dimensional submanifold of Diff (M) and $f_0 \in F$; we call f_0 the *centre* of the unfolding and say F is an *unfolding* of f_0.

Two unfoldings (F, f_0) and (F', f_0') are *germ-equivalent* iff $f_0 = f_0'$ and $\mathcal{G} \cap F = \mathcal{G} \cap F'$ for some neighbourhood $\mathcal{G}$ of f_0 in Diff (M). We are interested only in the *germ* (i.e. germ-equivalence class) of the unfolding. Thus we are always free to 'shrink F', i.e. to replace F by an open neighbourhood $\mathcal{G} \cap F$ of f_0 in F. We shall often abuse language and denote an unfolding by F rather than by (F, f_0).

We state the definition more precisely: a subset $F \subset \text{Diff}(M)$ is an *unfolding* of f_0 iff there is a finite dimensional vector space $\mathbb{F}$, an open neighbourhood Γ of the origin in $\mathbb{F}$, and a map $\Gamma \to F: \gamma \to f_\gamma$ satisfying the following conditions

  (i) the evaluation map $\Gamma \times M \to M: (\gamma, x) \to f_\gamma(x)$ is smooth;

  (ii) $f_\gamma|_{\gamma=0} = f_0$;

  (iii) the map $\gamma \to f_\gamma$ maps Γ bijectivity onto F;

  (iv) for each $\gamma \in \Gamma$ the linear map $A(\gamma): \mathbb{F} \to \mathscr{X}^\infty(M)$ defined by

$$A(\gamma)\hat{\gamma} = \frac{d}{dt} f_{\gamma+t\hat{\gamma}} \circ f_\gamma^{-1} \Big|_{t=0}$$

is injective.

Remark. We could if we like use a more general definition; namely we could define an unfolding with centre f_0 to be a map $\Gamma \to \text{Diff}(M): \gamma \to f_\gamma$ satisfying (i) and (ii). The reason is that if $\tilde{\Gamma}$ is a submanifold of Γ containing 0 such that for $\tilde{\mathbb{F}} = T_0\tilde{\Gamma} \subset \mathbb{F}$ the restriction $A(0)|\tilde{\mathbb{F}}$ is injective, then we may shrink $\tilde{\Gamma}$ and obtain an unfolding

$$\tilde{\Gamma} \to \tilde{F} = \{f_\gamma : \gamma \in \tilde{\Gamma}\}: \gamma \to f_\gamma$$

which satisfies (i)–(iv). (This is by the implicit function theorem.) The theorems we prove about unfoldings satisfying (i)–(iv) can be reformulated as (more general) theorems about unfoldings satisfying (i), (ii); the general case can then be reduced to the special case by judiciously choosing the submanifold $\tilde{\Gamma}$. (We work with the more restrictive definition since it is more suggestive geometrically and makes the notation somewhat less cumbersome.)

Now for $f \in \text{Diff}(M)$ let $C(f)$ denote the topological conjugacy class of f:

$$C(f) = \{\varphi \circ f \circ \varphi^{-1} | \varphi \in \text{Homeo}(M)\} \cap \text{Diff}(M).$$

Roughly speaking, the unfolding F is universal iff there is a neighbourhood $\mathscr{G}$ of f_0 in $\text{Diff}(M)$ such that

$$\mathscr{G} \subset \bigcup_{f \in F} C(f). \tag{1}$$

Make the (unwarranted) assumption that the sets $C(f)$ are submanifolds of finite codimension in $\text{Diff}(M)$ and that they fit together nicely (like the orbits of a group action). We can then solve (1) by the implicit function theorem, if we can solve it infinitesimally; i.e. if

$$T_f\mathscr{G} = T_fC(f) + T_fF. \tag{2}$$

(Note that (2) expresses the idea that the 'submanifolds' $C(f)$ and F are transverse at f.)

Let us examine (2) more closely. Suppose N is another smooth manifold (eventually we take $M = N$) and that $f \in C^\infty(M, N)$ i.e. that $f: M \to N$ is smooth. If $\mathbb{R} \to C^\infty(M, N): t \to f_t$ is a smooth curve through f (i.e. $f_t|_{t=0} = f$) we expect that $\hat{f}$ defined by

$$\hat{f} = \frac{d}{dt} f_t \Big|_{t=0} \tag{3}$$

lies in the tangent space $T_fC^\infty(M, N)$ to the manifold $C^\infty(M, N)$ at the point f.

This leads to the definition

$$T_f C^\infty(M, N) = \mathscr{X}(f),$$

where $\mathscr{X}(f)$ denotes the vector space of vector fields along f:

$$\hat{f} \in \mathscr{X}(f) \Leftrightarrow \begin{cases} \hat{f} \in C^\infty(M, TN) \\ \hat{f}(x) \in T_{f(x)}N & \text{for all } x \in M. \end{cases}$$

Now if $W \subset C^\infty(M, N)$ is a submanifold and $f \in W$ we define $T_f W \subset \mathscr{X}(f)$ by the condition that $\hat{f} \in T_f W$ iff there is a curve $t \to f_t$ as above with $f_t \in W$ for all $t \in \mathbb{R}$.

Now if f is a diffeomorphism, i.e. $f \in \text{Diff}(M, N)$ there is a linear isomorphism $f_* : \mathscr{X}(f) \to \mathscr{X}(N)$ given by

$$(f_* \hat{f})(x) = \hat{f}(f^{-1}(x))$$

for $x \in N$. Thus take $M = N$ and apply f_* to both·sides of (2):

$$f_*(T_f \mathscr{G}) = f_* T_f(C)(f) + f_* T_f F. \tag{2'}$$

Since $\mathscr{G} \subset \text{Diff}(M)$ is open we have

$$f_*(T_f \mathscr{G}) = \mathscr{X}(M). \tag{4}$$

For $w \in \mathscr{X}(M)$ we expect

$$w \in f_* T_f C(f) \Leftrightarrow w = \frac{d}{dt} \varphi_t \circ f \circ \varphi_t^{-1} \circ f^{-1}\big|_{t=0}$$

for some curve $\mathbb{R} \to \text{Homeo}(M)$ through the identity (i.e. $\varphi_t|_{t=0} = e = $ identity map of M). This gives

$$f_* F_f C(f) = (1 - f_\#)\mathscr{X}(M), \tag{5}$$

where the linear operator $f_\# : \mathscr{X}(M) \to \mathscr{X}(M)$ is defined by $f_\# \hat{\varphi} = Tf \circ \hat{\varphi} \circ f^{-1}$ for $\hat{\varphi} \in \mathscr{X}(M)$ (i.e. $f_\# \hat{\varphi}(x) = Df(y)\hat{\varphi}(y)$ where $y = f^{-1}(x)$, $x \in M$). Finally in the notation of (iv) note the equation

$$f_* T_f F = A(\gamma)\mathbb{F} \subset \mathscr{X}(M) \tag{6}$$

for $f = f_\gamma$, $\gamma \in \Gamma$. We thus reformulate (2') in the form

$$\mathscr{X}(M) = (1 - f_\#)\mathscr{X}(M) + f_* T_f F, \tag{2''}$$

which asserts that for any $w \in \mathscr{X}(M)$ there exists $\zeta = \hat{\varphi} \in \mathscr{X}(M)$ and $\beta = A(\gamma)\hat{\gamma} \in \mathscr{X}(M)$ satisfying

$$w = (1 - f_\#)\zeta + \beta, \qquad \beta \in f_* T_f F. \tag{7}$$

The solvability of (7) implies that the image of $(1 - f_\#)$ has finite codimension, i.e. that $\dim \text{coker}(1 - f_\#) \leq \dim(\mathbb{F}) < \infty$. (This as we have just seen expresses the idea that the conjugacy class $C(f)$ has finite codimension.) We may express this by requiring a linear map $\omega(f) : \mathscr{X}(M) \to \mathbb{F}$, where $\mathbb{F}$ is a finite dimensional vector space such that

$$\ker(\omega(f)) \subset \text{im}(1 - f_\#). \tag{8}$$

Given (8) we can solve (7) if the additional hypothesis

$$\omega(f)|f_* T_f F \to \mathbb{F} \text{ is surjective} \tag{*}$$

is satisfied.

Now as we indicated in (0.7) there is no hope of solving (7) if $\mathscr{X}(M)$ denotes some familiar space of vector fields say $\mathscr{X}(M) = C^0(M) \otimes TM$ or $\mathscr{X}(M) = C^\infty(M) \otimes TM$. We shall however show that sometimes one can solve (7) get $\mathscr{X}(M) = \mathscr{F}(M) \otimes TM$ with $\mathscr{F}(M)$ a suitable ℕ-normed space of functions between $C^\infty(M)$ and $C^0(M)$. Roughly, we call the unfolding F *preversal* if it satisfies (8), i.e. if $C(f)$ has finite codimension and *infinitesimally universal* if (in addition) it satisfies the *transversality condition* (∗). In §§ 8, 9 we show how to make these heuristic considerations precise and apply the tame implicit function theorem to prove:

THEOREM A. *An infinitesimally universal unfolding is universal.*

8. *Spaces adapted to an unfolding*

(8.1) Let M be a compact manifold, $F \subset \mathrm{Diff}\,(M)$ an unfolding, and $\mathscr{F}(M)$ an ℕ-normed space of functions on M:

$$C^\infty(M) \subset \mathscr{F}(M) \subset C^0(M).$$

Definition. The space $\mathscr{F}(M)$ is *adapted to the unfolding F* iff it satisfies the following conditions:
  (i) $\mathscr{F}(M)$ satisfies the composition inequalities of § 6;
  (ii) $\mathscr{F}(M)$ satisfies the interpolation inequalities of (2.6);
  (iii) $\mathscr{F}(M)$ admits a family of smoothing operators as in (2.6);
  (iv) the map $F \times \mathscr{F}(M) \to \mathscr{F}(M)$, $(f, u) \to f_* u$ is tame-C^2;
  (v)   there exists $r \in \mathbb{N}$ and $C > 0$ such that for $u \in \mathscr{F}(M)$, $x, y \in M$ there exists $n \in \mathbb{Z}$ with

$$\left| u(f^n(x)) - u(f^n(y)) \right| \le C \|u\|_r d(f^n(x), f^n(y)),$$

(where d is any metric on M compatible with the usual differential structure on M; i.e. such that the co-ordinate charts are Lipeomorphisms on compact sets).

As usual put $\mathscr{X}(M) = \mathscr{F}(M) \otimes TM$, $\mathscr{L}(M) = \mathscr{F}(M) \otimes L(TM)$, and note that (i), (ii) and (2.2) give a tame bilinear map

$$\mathscr{L}(M) \times \mathscr{X}(M) \to \mathscr{X}(M)$$
$$(\Phi, \zeta) \to \Phi \zeta.$$

(8.2) The purpose of this crucial definition is to enable us to rewrite the equation

$$g = \varphi \circ f \circ \varphi^{-1}, \tag{1}$$

which involves $g \in \mathrm{Diff}\,(M)$, $\varphi \in \mathrm{Homeo}\,(M)$, and $f \in F$ in the form

$$P(w, \zeta, \gamma) = 0, \tag{2}$$

which involves $w \in \mathscr{X}^\infty(M)$, $\zeta \in \mathscr{X}(M)$, and $\gamma \in \Gamma$. In other words we are going to rewrite an equation on a non-linear space as a (non-linear) equation on a linear space.

To do this let $\exp: TM \to M$ be a smooth exponential map: i.e. for each $x \in M$ $\exp_x = \exp | T_x M$ is a diffeomorphism onto an open subset of M and

$$\exp_x(0) = x$$
$$D \exp_x(0) \hat{x} = \hat{x}$$

for $x \in M, \hat{x} \in T_x M$. (We require no relation to sprays or geodesics.) Given $\zeta \in \mathscr{X}^0(M) = C^0(M) \otimes TM$ define $\exp(\zeta): M \to M$ by $\exp(\zeta)(x) = \exp(\zeta(x))$. Thus the equation

$$\varphi = \exp(\zeta) \tag{3}$$

establishes a bijective correspondence between $\mathscr{X}^0(M)$ and a neighbourhood of the identity map in $C^0(M, M)$ and the equation

$$g = \exp(w) \circ f_0 \tag{4}$$

(where f_0 is the centre of the unfolding F) establishes a bijective correspondence between a neighbourhood of 0 in $\mathscr{X}^\infty(M) = C^\infty(M) \otimes TM$ and a neighbourhood $\mathscr{G}$ of f_0 in $\mathrm{Diff}(M)$.

Fix a finite dimensional vector space $\mathbb{F}$, a neighbourhood Γ of 0 in F, and a bijection

$$\Gamma \to F: \gamma \to f = f_\gamma \tag{5}$$

as in § 7.

Fix a neighbourhood $\mathscr{H} \subset \mathscr{X}^\infty(M) \times \mathscr{X}(M) \times \Gamma$ of the origin. If $\mathscr{H}$ is sufficiently small there are maps $Q_i: \mathscr{H} \to \mathscr{X}(M)$, $i = 1, 2$, defined by

$$f \circ \varphi \circ f^{-1} = \exp(f_\# \zeta + Q_1(w, \zeta, \gamma)),$$
$$f \circ g^{-1} \circ \varphi = \exp(\zeta + Q_2(w, \zeta, \gamma)),$$

where φ, g, f are determined from ζ, w, γ by (3), (4), (5). (Q_1 does not depend on w.)

Define $P: \mathscr{H} \to \mathscr{X}(M)$ by:

$$P(w, \zeta, \gamma) = (1 - f_\#)\zeta + Q_2(w, \zeta, \gamma) - Q_1(w, \zeta, \gamma).$$

(8.3) We assume that $\mathscr{F}(M)$ is adapted to F and derive some basic properties of these constructions.

I. The maps

$$\mathscr{G} \subset \mathrm{Diff}(M) \to \mathscr{X}^\infty(M): g \to w$$
$$\mathscr{X}(M) \to C^0(M, M): \zeta \to \varphi$$
$$\Gamma \to F: \gamma \to f$$

determined by (3), (4), (5) are continuous.

(From now on assume (g, φ, f) and (w, ζ, γ) are related by (3), (4), (5).)

II. Equation (2) holds if and only if

$$f \circ \varphi \circ f^{-1} = f \circ g^{-1} \circ \varphi. \tag{1'}$$

Hence $P(0, 0, 0) = 0$.

III. The map P is tame-C^2.

Proof of III. Shrink $\mathscr{H}$ and construct a neighbourhood U of the zero section in TM so that the maps Q_i have form

$$Q_1(w, \zeta, \gamma) = q_1 \circ f_\# \zeta$$
$$Q_2(w, \zeta, \gamma) = q_2 \circ \zeta$$

where the maps $\gamma \to q_1$ and $(w, \gamma) \to q_2$ are tame-C^∞ into $\mathcal{B}(U, TM)$. (Thus q_1, q_2 are fibre preserving: $q_i(U \cap T_xM) \subset T_xM$ for $x \in M$.) Now use (i), (ii), (iii) and corollary 3 of (6.2).

IV. The partial derivative $D_2P: \mathcal{H} \to L(\mathcal{X}(M))$ has form

$$D_2P \cdot \hat{\zeta} = (1 + \Phi_2)\hat{\zeta} - (1 + \Phi_1)f_\#\hat{\zeta}$$

where $D_2P = D_2P(w, \zeta, \gamma)$, $\Phi_i = \Phi_i(w, \zeta, \gamma)$ and the Φ_i are given by maps $\Phi_i: \mathcal{H} \to \mathcal{L}(M)$ such that the evaluation maps

$$\mathcal{H} \times \mathcal{X}(M) \to \mathcal{X}(M);$$
$$(w, \zeta, \gamma, \hat{\zeta}) \to \Phi_i(w, \zeta, \gamma)\hat{\zeta}$$

are tame and $\Phi_i(0, 0, 0) = 0$.

Proof of IV. Φ_i is the derivative along the fibre of q_i. The vanishing follows from the equations $D_2Q_1(0, 0, 0) = 0$ and $Q_2(0, \zeta, 0) = 0$.

V. The partial derivative $D_3P: \mathcal{H} \to L(\mathbb{F}, \mathcal{X}(M))$, (which is tame by III) satisfies $D_3P(0, 0, 0)\hat{\gamma} = A(0)\hat{\gamma}$, where $A(\gamma): \mathbb{F} \to f_* T_fF$ is the isomorphism of § 7(iv).

Proof of V. $P(0, 0, \gamma) = Q_2(0, 0, \gamma)$, $\exp(Q_2(0, 0, \gamma)) = f \circ f_0^{-1}$.

VI. If (1′) holds and ζ is sufficiently small, then φ is a homeomorphism (so that (1) holds).

Proof of VI. This is the purpose of (v) of definition (8.1). (See [**16**, p. 462].) We shall sketch the proof under the simplifying hypothesis that M is an affine manifold (e.g. or torus) so that the exponential map has form $\exp_x(\hat{x}) = x + \hat{x}$. Assume (1′). Then

$$g^n \circ \varphi = \varphi \circ f^n \qquad (1'')$$

for $n \in \mathbb{Z}$. Since φ is homotopic to the identity and M is compact, φ is surjective so it is enough to prove φ is injective. Thus we assume $\varphi(x) = \varphi(y)$ and prove $x = y$. By (1″) we may replace x, y by $f^n(x), f^n(y)$ for some n to obtain:

$$|\zeta(x) - \zeta(y)| \le C\|\zeta\|_r|x - y|.$$

Thus if $C\|\zeta\|_r < 1$, $x = y$ follows from

$$x + \zeta(x) = \varphi(x) = \varphi(y) = y + \zeta(y).$$

9. *Universal, preversal, and infinitesimally universal unfoldings*

We continue the notation of §§ 7, 8.

Definition. The unfolding F is *universal* iff there is a neighbourhood $\mathcal{G}$ of f_0 in Diff (M) and continuous maps $\mathcal{G} \to$ Homeo $(M): g \to \varphi$, $\mathcal{G} \to F: g \to f$, such that

$$g = \varphi \circ f \circ \varphi^{-1}$$

for $g \in \mathcal{G}$. (The topology on Homeo (M) is the topology it inherits as a subset of $C^0(M, M)$.)

Now define $T: \mathcal{L}(M) \times F \to L(\mathcal{X}(M))$ by

$$T(\Phi, f)\zeta = (1 + \Phi)f_\#\zeta,$$

for $\Phi \in \mathcal{L}(M), f \in F, \zeta \in \mathcal{X}(M)$.

Definition. The unfolding F is *preversal* (with respect to the $\mathbb{N}$-normed space $\mathscr{F}(M)$ which is by assumption adapted to F) iff there is a neighbourhood $\mathscr{N} \subset \mathscr{L}(M) \times F$ of $(0, f_0)$, a finite dimensional vector space $\mathbb{F}$, a tame family $J: \mathscr{N} \to L(\mathscr{X}(M))$, and a continuous family $\omega: \mathscr{N} \to L(\mathscr{X}(M), \mathbb{F})$ such that we have

$$(1 - T(\Phi, f))\zeta = w \tag{1}$$

whenever $(\Phi, f) \in \mathscr{N}$ and $\zeta, w \in \mathscr{X}(M)$ satisfy

$$\zeta = J(\Phi, f)w, \tag{2}$$

$$\omega(\Phi, f)w = 0. \tag{3}$$

The unfolding is *infinitesimally universal* (with respect to $\mathscr{F}(M)$) iff in addition it satisfies the 'transversality condition'

$$\omega(\Phi, f)(f_* T_j F) = \mathbb{F}. \tag{*}$$

(By continuity (*) will hold for $(\Phi, f) \in \mathscr{N}$ with $\mathscr{N}$ sufficiently small if it holds at $(\Phi, f) = (0, f_0)$.)

THEOREM A. *An infinitesimally universal unfolding is universal.*

Proof. According to I, II, VI from § 8 it is enough to produce a neighbourhood $\mathscr{B}$ of 0 in $\mathscr{X}^\infty(M)$ and a continuous map $\mathscr{B} \subset \mathscr{X}^\infty(M) \to \mathscr{X}(M) \times \Gamma$, $w \to (\zeta, \gamma)$, such that $P(w, \zeta, \gamma) = 0$ for $w \in \mathscr{B}$. According to the tame implicit function theorem and III of § 8 it is enough to produce a tame map

$$\mathscr{H} \times \mathscr{X}(M) \to \mathscr{X}(M) \times \mathbb{F}$$

$$(w, \zeta, \gamma, \hat{w}) \to (\hat{\zeta}, \hat{\gamma})$$

which is linear in $\hat{w}$ and satisfies

$$D_2 P(w, \zeta, \gamma)\hat{\zeta} + D_3 P(w, \zeta, \gamma)\hat{\gamma} = \hat{w}.$$

By IV this may be rewritten in the form

$$(1 - T(\Phi, f))\hat{\zeta} + A(w, \zeta, \gamma)\hat{\gamma} = \hat{w}_1 \tag{4}$$

where

$$1 + \Phi = (1 + \Phi_2)^{-1}(1 + \Phi_1),$$

$$A(w, \zeta, \gamma) = (1 + \Phi_2(w, \zeta, \gamma))^{-1} D_3 P(w, \zeta, \gamma),$$

$$\hat{w}_1 = (1 + \Phi_2)^{-1}\hat{w}.$$

By continuity, V, and (*) we may choose $\hat{\gamma}$ so that

$$\omega(\Phi, f)A(w, \zeta, \gamma)\hat{\gamma} = \omega(\Phi, f)\hat{w}_1$$

and then solve (4) by taking $\hat{\zeta} = J(\Phi, f)(\hat{w}_1 - A(w, \zeta, \gamma)\hat{\gamma})$. The tame estimates follow from the tame estimates on J and IV of § 8.

10. *Level decompositions*

A *level decomposition respecting* $f \in \text{Diff}(M)$ is a collection

$$R_0, S_1, R_1, S_2, \ldots, S_l, R_l$$

of open subsets of M satisfying the following conditions:

   (i) $M \subset R \cup S$ where $R = \bigcup_{i=0}^{l} R_i$, $S = \bigcup_{j=1}^{l} S_j$;

(ii) $\operatorname{cl}(R_i) \cap \operatorname{cl}(R_j) = \varnothing$ $i \neq j$,

$\operatorname{cl}(S_i) \cap \operatorname{cl}(S_j) = \varnothing$ $i \neq j$,

$\operatorname{cl}(S_j) \cap \operatorname{cl}(R_i) = \varnothing$ $i \neq j, j-1$,

$\operatorname{cl}(S_j) \cap \operatorname{cl}(f^n(S_j)) = \varnothing$ $|n| > 1$;

(iii) $f(\operatorname{cl}(S_{i+1} \cup M_i^+)) \subset M_i^+$,

$f^{-1}(\operatorname{cl}(S_i \cup M_i^-)) \subset M_i^-$,

where $M_i^+ = \bigcup_{j=0}^{i} R_j \cup S_j$, $M_i^- = \bigcup_{j=1}^{l} R_{j+1} \cup S_j$ (in these unions take $S_0 = R_{l+1} = \varnothing$). A *level decomposition respecting* an unfolding F is one respecting each $f \in F$. Note that the conditions are open so that a decomposition respecting f_0 respects a neighbourhood F_0 of f_0 in F.

Now define

$$\Omega_i(f) = \bigcap_{n \in \mathbb{Z}} f^n(R_i),$$

$$W^s(\Omega_i, f) = \bigcup_{m \geq 0} \bigcap_{n \geq m} f^{-n}(R_i),$$

$$W^u(\Omega_i, f) = \bigcup_{m \geq 0} \bigcap_{n \geq m} f^n(R_i),$$

so that $\Omega_i(f) = W^s(\Omega_i, f) \cap W^u(\Omega_i, f)$ and we have two decompositions

$$M = \bigcup_{i=0}^{l} W^s(\Omega_i, f)$$

$$M = \bigcup_{i=0}^{l} W^u(\Omega_i, f)$$

of M into disjoint f-invariant sets.

These definitions are inspired by the 'spectral decompositions' and 'filtrations' of Smale [20] but we assume no hyperbolicity of Ω_i. For us the decomposition serves a 'bookkeeping' purpose. An orbit in $W^u(\Omega_j) \cap W^s(\Omega_i)$ (this implies $i \leq j$) can be thought to 'originate' in Ω_j and 'terminate' in Ω_i. On its journey it passes successively through the sets $R_j, S_j, R_{j-1}, S_{j-1}, \ldots, S_{i+1}, R_i$. It has at least one point and at most two points in each S_k where $j \geq k \geq i$ and so these strips serve as 'fundamental domains'. The recurrent behaviour occurs in the sets R_i. The strip S_i serves as an 'interface' between R_{i-1} and R_i.

By (iii) an orbit cannot leave and re-enter some R_i. Moreover (ii) and (iii) imply

$$\operatorname{cl}(f^{-1}(R_i) \cap R_i \cap f(R_i)) \subset R_i.$$

Thus $\Omega_i(f)$ is a (compact) isolated invariant set and R_i is an invariant set and an isolating neighbourhood for it. (The terminology is from [2].) Thus

$$W^u(\Omega_i, f) = \bigcap_{n=0}^{\infty} f^n(N_i)$$

$$W^s(\Omega_i, f) = \bigcap_{n=0}^{\infty} f^{-n}(N_i)$$

for any neighbourhood N_i of $\Omega_i(f)$ in R_i.

PROPOSITION 1. *Let N_i be any neighbourhood of $\Omega_i(f_0)$ in R_i. Then there is a neighbourhood F_0 of f_0 in F and open sets P_i^u, $P_i^s \subset R_i$ such that*

$$R_i = N_i \cup P_i^s \cup P_i^u,$$
$$W^u(\Omega_i, f) \cap \operatorname{cl}(P_i^s) = \varnothing,$$
$$W^s(\Omega_i, f) \cap \operatorname{cl}(P_i^u) = \varnothing,$$

for $f \in F_0$. Thus for a suitably large m

$$f^{-n}(\operatorname{cl}(P_i^s)) \cap \operatorname{cl}(R_i) = \varnothing,$$
$$f^n(\operatorname{cl}(P_i^u)) \cap \operatorname{cl}(R_i) = \varnothing,$$

for $f \in F_0$ and $n \geq m$.

PROPOSITION 2. *Given N_i as in proposition 1 (for each $i = 0, 1, \ldots, l$) there is an integer m and a neighbourhood F_0 of f_0 in F such that for each $f \in F_0$ and each $x \in M$ we have $f^n(x) \notin \bigcup_{i=0}^l N_i$ for at most m values of $n \in \mathbb{Z}$.*

Proof. From the open cover $M = \bigcup_{i=0}^l N_i^f$ extract a finite subcover $M = \bigcup_{i=0}^l \bigcup_{n=-m_i}^{m_i} f^n(N_i)$ and take $m = 2 \sum m_i$.

11. *Local preversality*

Continue the notation of § 9. In particular, assume $\mathscr{F}(M)$ is adapted to the unfolding F.

Definition. Let $N \subset M$ be open. Say that F is *preversal at N* (with respect to $\mathscr{F}(M)$) iff there is a neighbourhood $\mathscr{N} \subset \mathscr{L}(M) \times F$ of $(0, f_0)$, a finite dimensional vector space $\mathbb{E}$, a tame family $J: \mathscr{N} \to L(\mathscr{X}(M))$ and a continuous family $\delta: \mathscr{N} \to L(\mathscr{X}(M), \mathbb{F})$ such that we have

$$((1 - T(\Phi, f))\zeta)|N = w|N \tag{0}$$

whenever $(\Phi, f) \in \mathscr{N}$ and $\zeta, w \in \mathscr{X}(M)$ satisfy $\zeta = J(\Phi, f)w$, $\delta(\Phi, f)w = 0$.

Remark. It is enough to define $J(\Phi, f)w$ and $\delta(\Phi, f)w$ under the additional assumptions that $\operatorname{supp}(\Phi), \operatorname{supp}(w) \subset Z$, where $Z \subset M$ is an open set satisfying $\operatorname{cl}(N \cup f^{-1}(N)) \subset Z$ for $(\Phi, f) \in \mathscr{N}$. For then we may choose $\psi \in C_0^\infty(Z)$ with $\psi|N \cup f^{-1}(N) \equiv 1$ and take

$$J'(\Phi, f)w = J(\psi\Phi, f)(\psi w)$$
$$\delta'(\Phi, f)w = \delta(\psi\Phi, f)(\psi w).$$

Since $T(\Phi, f)\zeta|N = T(\psi\Phi, f)\zeta|N$, $w|N = \psi w|N$, for $\zeta, w \in \mathscr{X}(M)$ it follows that J' and δ' satisfy (0) if J and δ do. The linear maps $\Phi \to \psi\Phi$, $w \to \psi w$ are tame so that J', δ' are tame if J and δ are.

Now let $R_0, S_1, \ldots, R_l$ be a level decomposition respecting F.

Definition. Say that the level decomposition is *locally preversal* iff for each $i = 0, 1, \ldots, l$ there is a neighbourhood N_i of $\Omega_i(f_0)$ such that F is preversal at N_i.

Recall that $S = S_1 \cup \cdots \cup S_l$ and define $\mathscr{X}(S)_0 = \{w \in \mathscr{X}(M): \operatorname{supp}(w) \subset S\}$.

THEOREM. *Assume the level decomposition is locally preversal. Then their is a neighbourhood $\mathscr{N} \subset \mathscr{L}(M) \times F$ of $(0, f_0)$, a finite dimensional vector space $\mathbb{E}$, a tame family*

J: $\mathcal{N} \to L(\mathcal{X}(M))$ *and a continuous family* $\delta: \mathcal{N} \to L(\mathcal{X}(M), \mathbb{E})$ *such that we have*

$$(1 - T(\Phi, f))\zeta - w \in \mathcal{X}(S)_0$$

whenever $(\Phi, f) \in \mathcal{N}$, ζ, $w \in \mathcal{X}(M)$ *satisfy* $\zeta = J(\Phi, f)w$, $\delta(\Phi, f)w = 0$.

Proof. In other words we can solve the equation $(1 - T)\zeta = w$ on $M \backslash S$. Take $N = N_i$ in definition of 'preversal at N' to obtain $\mathcal{N}_i$, $\mathbb{E}_i$, J_i, δ_i and define

$$\mathcal{N} = \mathcal{N}_0 \cap \mathcal{N}_1 \cap \cdots \cap \mathcal{N}_l$$

$$\mathbb{E} = \mathbb{E}_0 \times \mathbb{E}_1 \times \cdots \times \mathbb{E}_l$$

$$\delta = (\delta_0, \delta_1, \ldots, \delta_l).$$

Shrink $\mathcal{N}$ further to obtain the decomposition $R_i = N_i \cup P_i^s \cup P_i^u$ as in § 10. This gives an open cover

$$M \backslash \bigcup_{i=0}^{l} N_i \subset \bigcup_{i=0}^{l} \{P_i^s \cup P_i^S\} \cup S$$

so a suitable partition of unity gives a decomposition

$$w - \sum_{i=0}^{l} w_i^* = \sum_{i=0}^{l} w_i + w', \tag{1}$$

where $w_i^* = (1 - T(\Phi, f)) J_i(\Phi, f)w$, $w_i = w_i^s + w_i^u$, supp $(w_i^s) \subset P_i^s$, supp $(w_i^u) \subset P_i^u$, and supp $(w') \subset S$. (We can assume that w_i^* vanishes on N_j for $i \neq j$ by multiplying by a suitable cut-off function; hence the left hand side of (1) vanishes on $N_0 \cup \cdots \cup N_l$.) Let

$$\zeta_i = \zeta_i^s + \zeta_i^s$$

where

$$\zeta_i^s = \sum_{n=0}^{\infty} T(\Phi, f)^n w_i^s,$$

$$\zeta_i^u = -\sum_{n=1}^{\infty} T(\Phi, f)^{-n} w_i^u.$$

The series converge pointwise since by proposition 1 of § 10 each contains only finitely many terms. (They do *not* converge uniformly however.) Clearly

$$\zeta_i - T(\Phi, f)\zeta_i = w_i, \tag{2}$$

$$\text{supp}\,(w_i) \subset R_i, \tag{3}$$

so we replace ζ_i by $\psi_i \zeta_i$ for a suitable cut-off function ψ_i and investigate the extent to which this invalidates (2).

We take $\psi_i \in C^{\infty}(M)$ satisfying

$$\psi | R_i \cup f^{-1}(R_i) \equiv 1, \tag{4}$$

$$f(\text{supp}\,(\psi)) \cup \text{supp}\,(\psi) \subset S_{i-1} \cup R_i \cup S_i; \tag{5}$$

(we take $S_{i-1} = \varnothing$ if $i = 0$). First note that (shrinking $\mathcal{N}$ if necessary) proposition 1 of § 10 gives an m independent of $(\Phi, f) \in \mathcal{N}$ with

$$\text{supp}\,(\psi_i) \cap f^{-n}(P_i^s) = \varnothing,$$

$$\text{supp}\,(\psi_i) \cap f^{n}(P_i^u) = \varnothing,$$

for $n \geq m$. Hence $\psi_i \zeta_i$ may be represented by a finite series so that the map $w \to \psi_i \zeta_i$ is tame. Let

$$A_i = (f(\operatorname{supp}(\psi_i)) \cup \operatorname{supp}(\psi_i)) \setminus R_i$$

so that $A_i \subset S_{i-1} \cup S_i$ and let w_i' denote the error in (2) when ζ_i is replaced by $\psi_i \zeta_i$:

$$(1 - T(\Phi, f))\psi_i \zeta_i = w_i + w_i'. \tag{6}$$

We claim $\operatorname{supp}(w_i') \subset A_i$. Clearly $w_i'(x) = 0$ whenever $\psi(x) = \psi(f^{-1}(x)) = 1$ since $T(\Phi, f)\psi_i \zeta_i = T(\Phi, f)\zeta_i$ at such a point x. Hence $w_i'|R_i \equiv 0$. In particular, $w_i'(x) = 0$ at any value of x where $w_i(x) = 0$. On the other hand $w_i'(x) = 0$ at any x where $\psi(x) = \psi(f^{-1}(x)) = 0$ and $w_i(x) = 0$ for then the three other terms of (6) vanish. Hence $w_i'(x) \neq 0$ only when $x \notin R_i$ and either $\psi(x) \neq 0$ or $\psi(f^{-1}(x)) \neq 0$ as required.

We now put

$$\zeta = J(\Phi, f)w = \sum_{i=0}^{l} (J_i(\Phi, f)w + \psi_i \zeta_i),$$

so that (1) and (6) give

$$(1 - T(\Phi, f))\zeta = \sum_{i=0}^{l} (w_i^* + w_i) + w''$$

where $w'' = w' + \sum_{i=0}^{l} w_i'$. Hence $\operatorname{supp}(w'') \subset S$ as required.

12. *The interface condition*

Continue the notation of §§ 9, 10, 11. Define $\mathbb{N}$-normed subspaces by:

$$\mathscr{X}(S_j)_0 = \{w \in \mathscr{X}(M): \operatorname{supp}(w) \subset S_j\}.$$

We say that the unfolding F satisfies the *interface condition* (with respect to the space $\mathscr{F}(N)$ and the level decomposition) iff there exist $p \in \mathbb{N}$, $\sigma \in (0, 1) \subset \mathbb{R}$, a neighbourhood $\mathscr{N} \subset \mathscr{L}(M) \times F$ of $(0, f_0)$, finite dimensional vector spaces $\mathbb{F}_1, \ldots, \mathbb{F}_l$, tame families $\pi_j^s, \pi_j^u : \mathscr{N} \to L(\mathscr{X}(S_j)_0, \mathscr{X}(M))$ and continuous families $\varepsilon_j : \mathscr{N} \to L(\mathscr{X}(S_j)_0, \mathbb{F}_j)$, $j = 1, \ldots, l$, satisfying:

   (i) $w = \pi_j^s(\Phi, f)w + \pi_j^u(\Phi, f)w$;

   (ii, s) $\|T(\Phi, f)^n \pi_j^s(\Phi, f)w\|_k \leq c\sigma^n [\![\Phi : w]\!]_{p,k}$;

   (ii, u) $\|T(\Phi, f)^{-n} \pi_j^u(\Phi, f)w\|_k \leq c\sigma^n [\![\Phi : w]\!]_{p,k}$;

whenever $(\Phi, f) \in \mathscr{N}$ and $w \in \mathscr{X}(S_j)_0$, satisfy $\varepsilon_j(\Phi, f)w = 0$. Here we use the abbreviation:

$$[\![\Phi : w]\!]_{p,k} = (1 + \|\Phi\|_{p+k})\|w\|_k + \|w\|_{p+k},$$

and $c = c(k)$ is independent of Φ, f, w and n.

THEOREM B. *Assume that the level decomposition is both locally preversal and satisfies the interface condition. Then the unfolding F is preversal (with respect to $\mathscr{F}(M)$). Hence if it also satisfies the additional transversality condition*

$$\omega(\Phi, f)(f_* T_f F) = \mathbb{F} \tag{*}$$

it is infinitesimally universal and hence (by theorem A) universal.

Proof. Let $\tilde{J}: \mathcal{N} \to L(\mathcal{X}(M))$, $\delta: \mathcal{N} \to L(\mathcal{X}(M), \mathbb{E})$, denote the maps constructed in § 11, so that $\delta(\Phi, f)w = 0$, $\tilde{\zeta} = \tilde{J}(\Phi, f)w$, imply that

$$\tilde{\zeta} - T(\Phi, f)\tilde{\zeta} = w + \sum_{j=1}^{l} w_j$$

where $w_j \in \mathcal{X}(S_j)_0$. Define

$$\mathbb{F} = \mathbb{E} \times \mathbb{F}_1 \times \cdots \times \mathbb{F}_l$$

$$\omega(\Phi, f)w = (\delta(\Phi, f)w, \varepsilon_1(\Phi, f)w_1, \ldots, \varepsilon_l(\Phi, f)w_l)$$

$$J(\Phi, f)w = \zeta = \tilde{\zeta} - \sum_{j=1}^{l} \zeta_j$$

where

$$\zeta_j = \zeta_j^s + \zeta_j^u$$

$$\zeta_j^s = \sum_{n=0}^{\infty} T(\Phi, f)^n \pi_j^s(\Phi, f)w_j$$

$$\zeta_j^u = -\sum_{n=1}^{\infty} T(\Phi, f)^{-n} \pi_j^u(\Phi, f)w_j.$$

The tame estimates

$$\|\zeta_j^s\|_k \leq c(1-\sigma)^{-1} [\![\Phi : w]\!]_{p,k},$$

$$\|\zeta_j^u\|_k \leq c\sigma(1-\sigma)^{-1} [\![\Phi : w]\!]_{p,k},$$

follow from (ii) and the formula for summing a geometric series. This gives the tameness of the map $\mathcal{N} \times \mathcal{X}(M) \to \mathcal{X}(M)$, $(\Phi, f, w) \to \zeta$. The continuity of the map $\mathcal{N} \times \mathcal{X}(M) \to \mathbb{E}$, $(\Phi, f, w) \to \omega(\Phi, f)w$ and the equation

$$(1 - T(\Phi, f))\zeta = w$$

are immediate.

We now give a criterion for verifying (∗) in terms of δ and $\varepsilon_1, \ldots, \varepsilon_l$.

Remark. Condition (∗) holds if $\delta(\Phi, f)(f_* T_f F) = \mathbb{E}$ and

$$\varepsilon_j(\Phi, f)\{(f_* T_f F) \cap \ker(\delta(\Phi, f)) \cap \mathcal{X}(S_j)_0\} = \mathbb{F}_j$$

for $j = 1, \ldots, l$. (As usual it is enough to check these conditions at $(\Phi, f) = (0, f_0)$.)

13. *Summary and prospects*

(13.1) Robbin [16] proves:

THEOREM A. *An infinitesimally stable diffeomorphism is structurally stable.*

THEOREM B. *A locally Anosov diffeomorphism is infinitesimally stable.*

THEOREM C. *A diffeomorphism satisfying Axiom A and the strong transversality condition is locally Anosov.*

We have extended two thirds of this trilogy to unfoldings.

THEOREM A. *An infinitesimally universal unfolding is universal.*

THEOREM B. *An unfolding which admits a locally preversal level decomposition satisfying the interface condition is preversal (and hence infinitesimally universal if it also satisfies the transversality condition).*

Conjecture C. An unfolding whose centre satisfies Axiom A, the no-cycle condition, and the strong finite codimension condition is isomorphic to an unfolding which has a locally preversal level decomposition satisfying the interface condition.

We define the terminology used in conjecture C.

Definition. Call two unfoldings (F, f_0) and (F', f_0') *isomorphic* iff $f_0 = f_0'$ and there exists a homeomorphism $(F, f_0) \to (F', f_0'): f \to f'$ and a continuous map $F \to \text{Homeo}(M): f \to \varphi$ such that

$$f' = \varphi \circ f \circ \varphi^{-1}$$

for $f \in F$.

PROPOSITION. *Suppose (F, f_0) and (F', f_0) are isomorphic. Then F is universal if and only if F' is.*

For references on 'Axiom A' and the 'no cycle condition' see [22] (the updated and annotated version of [20]). From our point of view the most important consequence is that the stable and unstable manifolds

$$W^s(x, f) = \{y \in M : \lim_{n \to \infty} d(f^n(x), f^n(y)) = 0\}$$

$$W^u(x, f) = W^s(x, f^{-1})$$

are actually (immersed) manifolds for each $x \in M$. This justifies the following.

Definition. A diffeomorphism f satisfies the *strong finite codimension* condition iff for each $x \in M$ (the sets $W^s(x, f)$ and $W^u(x, f)$ are immersed manifolds and) the codimension of the singularity of the intersection of $W^s(x, f)$ and $W^u(x, f)$ at x is finite. In particular, f satisfies the *strong transversality condition* iff for each x this codimension is zero, i.e. $W^s(x, f)$ and $W^u(x, f)$ intersect transversally at x.

(13.2) We now sketch the 'proof' of conjecture C.

Step 1. Using Axiom A and the no-cycle condition construct the 'spectral decomposition'

$$\Omega(f) = \Omega_0(f) \cup \Omega_1(f) \cup \cdots \cup \Omega_l(f)$$

of the non-wandering set of f; then construct a corresponding level decomposition. Here one follows [20]; a level decomposition is more or less the same thing as a 'filtration'. Fix neighbourhoods N_i of $\Omega_i(f)$.

Step 2. Construct the tame splittings

$$\mathcal{X}(N_i^f) = \mathcal{E}_i^s(f) \oplus \mathcal{E}_i^0(f) \oplus \mathcal{E}_i^u(f), \tag{1}$$

which are invariant under $f_\#$. Roughly speaking, the spectrum of $f_\# | \mathcal{E}_i^s(f), f_\# | \mathcal{E}_i^0(f), f_\# | \mathcal{E}_i^u(f)$ should lie inside, near, outside the unit circle but since we are using $\mathbb{N}$-normed spaces rather than Banach spaces this must be expressed in terms of tame asymptotic estimates. The splitting will in general be different from

the one determined by the hyperbolic structure of Ω_i. In particular, usually it will *not* come from a splitting of the tangent bundle. This means that $\mathscr{E}_i^s$, $\mathscr{E}_i^0$, $\mathscr{E}_i^u$ will *not* be invariant under multiplication by a real valued function and the projections onto the summands will *not* be local (i.e. pointwise defined).

Step 3. It is at this point we must define the space $\mathscr{F}(M)$. Introduce $\mathbb{N}$-jet semi-Finslers Θ_i on N_i^f and prove the asymptotic tame estimates. The Θ_i will be chosen so that certain derivatives can blow up at $W^u(\Omega_i, f)$ or $W^s(\Omega_i, f)$. Replace F by an isomorphic unfolding to achieve the conditions

$$W^e(\Omega_i, f) \cap N_i = W^e(\Omega_i, f_0) \cap N_i \tag{2}$$

for $e = s, u$, $i = 0, 1, \ldots, l$. Thus $\Theta_i | N_i$ will be independent of f.

Step 4. Use the tame contraction principle to define perturbed splittings

$$\mathscr{X}(N_i^f) = \mathscr{E}_i^s(\Phi, f) \oplus \mathscr{E}_i^0(\Phi, f) \oplus \mathscr{E}_i^u(\Phi, f) \tag{3}$$

which are invariant by $T(\Phi, f)$. Take $\mathbb{E}_i = \mathscr{E}_i^0(0, f_0)$ and define $\delta_i(\Phi, f): \mathscr{X}(N_i^f) \to \mathbb{E}_i$ to be the projection on $\mathscr{E}_i^0(\Phi, f)$ by (3) followed by the projection on $\mathbb{E}_i$ by (1). Thus

$$\ker (\delta_i(\Phi, f)) = \mathscr{E}_i^s(\Phi, f) \oplus \mathscr{E}_i^u(\Phi, f)$$

and the operator $J_i(\Phi, f): \mathscr{X}(N_i^f) \to \mathscr{X}(N_i^f)$ is defined using Neumann series. This verifies local preversality.

Step 5. At each point $x \in S_k$ of the interface set determine i and j with $i < k \le j$ so that

$$x \in W^s(\Omega_i, f_0) \cap W^u(\Omega_j, f_0). \tag{4}$$

Then use a tame version of the C^∞ preparation theorem of Malgrange-Mather (*et al*) to obtain a neighbourhood U of x and a decomposition

$$w | U = \pi_{i,j}^s(\Phi, f, U)w + \pi_{i,j}^u(\Phi, f, U)w \tag{5}$$

with

$$\pi_{ij}^s(\Phi, f, U)w \in \mathscr{E}_i^s(\Phi, f) | U \tag{6.s}$$

$$\pi_{ij}^u(\Phi, f, U)w \in \mathscr{E}_j^u(\Phi, f) | U. \tag{6.u}$$

Then piece these together to achieve a decomposition

$$w = \pi_k^s(\Phi, f)w + \pi_k^u(\Phi, f)w$$

for $w \in \mathscr{X}(S_k)_0$. The decomposition (5) will require a hypothesis

$$\varepsilon_x(\Phi, f)w = 0$$

where $\varepsilon_x(\Phi, f) \in L(\mathscr{X}(S_k)_0, \mathbb{F}_x)$ and the dimension of $\mathbb{F}_x$ is related to the codimension of the singularity of the intersection (5). (In particular, we should have $\dim (\mathbb{F}_x) = 0$ if the intersection is transversal.)

Step 6. Prove the following compatibility condition. Each point $x \in S_k$ has a neighbourhood U such that for $w \in \mathscr{X}(S_k)_0$ we have

$$\pi^s(\Phi, f)w | U \cap N_i^f \in \mathscr{E}_i^s(\Phi, f) \tag{7.s}$$

$$\pi^u(\Phi, f)w | U \cap N_j^f \in \mathscr{E}_j^u(\Phi, f) \tag{7.u}$$

for $i < k \leq j$. By proposition 2 of § 10 there is a bound independent of x on the number of points on the orbit of x which are not in $N_0 \cup N_1 \cup \cdots \cup N_k$. Whenever we are in some N_j a tame asymptotic estimate holds. This verifies the interface condition.

(13.3) In part II we shall carry out this program for a specific class of examples. The examples are restricted to make the technicalities as manageable as possible; our only aims are to show that the theory of part I is not vacuous and to provide some guidance in the general case. We see the following obstacles to proving conjecture C in full generality.

(13.3.1) The splittings are not local; hence equations (5), (6), (7) are meaningless. We need a weaker condition, say 'locality at the stable and unstable manifolds' to circumvent this difficulty.

(13.3.2) We do not know good general techniques for constructing $\mathbb{N}$-jet semi-Finslers. Finding smoothing operators seems to be a problem. In part II we have to work fairly hard to construct the $\mathbb{N}$-jet semi-Finsler near the source and the sink. This seems absurd; nothing can go wrong there. (Famous last words!) Mike Field has suggested a way around this. He proves a 'hybrid tame implicit function theorem'. (Of course, I prefer to do it my way.)

(13.3.3) I do not know the appropriate generalization of the tame preparation theorem (see § 16 in part II) from 1 to n dimensions. This is because the corresponding argument for germs loses too much differentiability in the induction step. Probably the way around this difficulty is to incorporate the loss in the $\mathbb{N}$-jet semi-Finsler. Thus suppose that for each $r \in \mathbb{N}$ there exists $k = k(r) > r$ so that for $x \in S$

$$\|u\|_{k,x} = 0 \Rightarrow j^r u(x) = 0.$$

Then $\mathscr{F}(S, \Theta) = C^\infty(S)$ as sets but the $\mathbb{N}$-normed spaces $\mathscr{F}(S, \Theta)$ and $C^\infty(S)$ will not be tamely equivalent.

PART II. ANALYSIS OF AN EXAMPLE

14. *Overview*

In part II we carry out the program sketched in the 'proof' of conjecture C given in § 13 for a simple example. Some general theory is also given (e.g. the 'tame preparation theorem' of § 16) but this theory is probably not yet sufficiently general to handle conjecture C in general (although it does suffice for the example). In this introductory section we show how to solve

$$(1 - f_\#)\zeta = w \tag{1}$$

under the hypothesis

$$\omega \cdot w = 0. \tag{2}$$

We shall worry neither about perturbing $f_\#$ to $T(\Phi, f)$ nor convergence. (The construction of the $\mathbb{N}$-normed space $\mathscr{F}(M)$ in which the various series converge is motivated by the heuristic argument and is given in detail in the sequel.)

(14.1) The diffeomorphism $f: M \to M$ has a non-wandering set consisting of a sink Ω_0, two saddles Ω_1 and Ω_2, and a source Ω_3. Thus M is a two-dimensional torus. We assume $W^u(\Omega_1)\setminus\Omega_1 \subset W^s(\Omega_0)$, $W^s(\Omega_2)\setminus\Omega_2 \subset W^u(\Omega_3)$, and define

$$W^{us} = W^s(\Omega_1) \cap W^u(\Omega_2).$$

Thus $\mathrm{cl}\,(W^{us}) = W^{us} \cup \Omega_1 \cup \Omega_2$ and

$$M = W^s(\Omega_0) \cup \mathrm{cl}\,(W^{us}) \cup W^u(\Omega_3).$$

(14.2) Now to solve (1) we may assume without loss of generality that $\mathrm{supp}\,(w) \subset Z$ where Z is a neighbourhood of $\mathrm{cl}\,(W^{us})$. Indeed if $w|Z \equiv 0$ then $w = w^u + w^s$ where $\mathrm{supp}\,(w^u) \subset W^u(\Omega_3)$, $\mathrm{supp}\,(w^s) \subset W^s(\Omega_0)$, so (1) is solved by

$$\zeta = \zeta^u + \zeta^s$$

where

$$\zeta^u = -\sum_{n=1}^{\infty} f_{\#}^{-n} w^u$$

$$\zeta^s = \sum_{n=0}^{\infty} f_{\#}^{n} w^s.$$

(14.3) Next we assume that there are coordinate systems $z_i = (x_i, y_i): M_i \to \mathbb{R}^2$, $i = 1, 2$, which linearize f at Ω_i. Thus $\Omega_i \in M_i$, $\mathrm{cl}\,(W^{us}) \subset Z \subset M_1 \cup M_2$, and:

$$f^*(x_i, y_i) = (\lambda_i x_i, \mu_i y_i),$$

where $0 < \mu_1 < 1 < \lambda_1$, $0 < \lambda_2 < 1 < \mu_2$. We write $w = w_1 + w_2$ where $\mathrm{supp}\,(w_i) \subset M_i$; the arguments for $i = 1$ and $i = 2$ are the same (replace f by f^{-1}) so we suppose without loss of generality that $\mathrm{supp}\,(w) \subset M_1$.

We first try to solve (1) by the Neumann series

$$\zeta = \sum_{n=0}^{\infty} f_{\#}^{n} w. \tag{3.s}$$

Note that if $\mathrm{supp}\,(w) \subset \{|y_1| < L\}$ then $\mathrm{supp}\,(f_{\#}^n w) \subset \{|y_1| < \mu^n L\}$, so that (3.s) converges on M if it converges on M_1. To investigate this write

$$z_{i\#} w = (u_i, v_i)$$

$$z_{i\#} \zeta = (\xi_i, \eta_i)$$

so that

$$\xi_1(x, y) = \sum_{n=0}^{\infty} \lambda_1^n u_1(\lambda_1^{-n} x, \mu_1^{-n} y),$$

$$\eta_1(x, y) = \sum_{n=0}^{\infty} \mu_1^n v_1^n(\lambda_1^{-n} x, \mu_1^{-n} y).$$

The series for η_1 converges as $\mu_1 < 1$. The series for ξ_1 apparently diverges as $\lambda_1 > 1$ however if we differentiate twice with respect to x we obtain

$$(\partial_1^2 \xi_1)(x, y) = \sum_{n=0}^{\infty} \lambda_1^{-n}(\partial_1^2 u)(\lambda_1^{-n} x, \mu_1^{-n} y),$$

which also converges.

(14.3) We have reduced to the case

$$u_1(x, y) = u_{10}(y) + u_{11}(y)x$$
$$v_1(x, y) = 0.$$

In this case we try the other Neumann series

$$\zeta = -\sum_{n=1}^{\infty} f_{\#}^{-n}w. \tag{3.u}$$

This gives

$$\xi_1(x, y) = \xi_{10}(y) + \xi_{11}(y)x$$
$$\eta_1(x, y) = 0$$

$$\xi_{10}(y) = -\sum_{n=1}^{\infty} \lambda_1^{-n} u_{10}(\mu_1^n y) \tag{4.0}$$

$$\xi_{11}(y) = -\sum_{n=1}^{\infty} u_{11}(\mu_1^n y). \tag{4.1}$$

This series for ξ_{10} converges as $\lambda_1 > 1$; as $\mu_1 < 1$ the series for ξ_{11} also converges provided that

$$u_{11}(0) = 0 \tag{5}$$

and u_{11} satisfies a Holder estimate.

We have solved (1) on M_1. We cannot stop here however for the series (3.u) blows up on M_2. Indeed since supp $(w) \subset M_1$, ζ will be invariant by $f_{\#}$ near Ω_2 and Ω_3. Near Ω_3 this causes no problem, for any invariant vector field which is continuous on $W^u(\Omega_3) \setminus \Omega_3$ and vanishes at Ω_3 is automatically continuous at Ω_3. Near Ω_2 however we are only assured of continuity if $\xi_2 = 0$.

(14.4) To get a feeling for how to attack this problem, assume the W^{us} consists of a single orbit $W^{us} = \{f^n(p): n \in \mathbb{Z}\}$ and that S is a small neighbourhood of p. We assume further that the coordinate systems z_1 and z_2 are related by formulae

$$x_2 = x_1 + b(y_1)$$
$$y_2 = c(y_1) \tag{6}$$

on S. (We shall relax these assumptions somewhat in § 22.) Instead of seeking the solution of (1) in the form $z_{1\#}\zeta = (\xi_1, 0)$ we allow $z_{1\#}\zeta = (\xi_1, \eta_1)$ but impose the condition that $z_{2\#}\zeta = (0, \eta_2)$ so that ζ will be continuous on M_2. Thus we seek ζ such that $z_{i\#}\zeta = (\xi_i, \eta_i)$, $i = 1, 2$, where

$$\xi_1(x, y) = \xi_{10}(y) + \xi_{11}(y)x$$
$$\eta_1(x, y) = \eta_{10}(y) + \eta_{11}(y)x$$
$$\xi_2(x, y) = 0$$
$$\eta_2(x, y) = \eta_{20}(y) + \eta_{11}(y)x \tag{7}$$

where ξ_{10}, ξ_{11} are given by (4.0) and (4.1) as before and the η_{ij} are to be determined. (6) and (7) give the overlap relations

$$\xi_2(x + b(y), c(y)) = \xi_1(x, y) + b'(y)\eta_1(x, y)$$
$$\eta_2(x + b(y), c(y)) = c'(y)\eta_1(x, y)$$

on S which reduce to:

$$0 = \xi_{10}(y) + b'(y)\eta_{10}(y)$$
$$0 = \xi_{11}(y) + b'(y)\eta_{11}(y)$$
$$\eta_{20}(c(y)) = c'(y)\eta_{10}(y)$$
$$\eta_{21}(c(y)) = c'(y)\eta_{11}(y).$$

Now $c'(y) \neq 0$ so it is enough to solve the first two equations to obtain a solution on S. Once we have the solution on S, the invariance

$$\eta_{i0}(\mu_i y) = \mu_i \eta_{i0}(y)$$
$$\eta_{i1}(\mu_i y)\lambda_i = \mu_i \eta_{i1}(y)$$

assures that η_i is continuous on M_i and hence that ζ is continuous on M. Now we see that we can solve by dividing through by $b'(y)$; this requires that at any zero of b', ξ_{10} and ξ_{11} vanish to the same order as b'.

Now $W^u(\Omega_2) = \{x_2 = 0\}$, $W^s(\Omega_1) = \{x_1 = 0\}$, so by (6)

$$\{p\} = W^{us} \cap S = \{x_1 = 0, b(y_1) = 0\}.$$

Thus the order of the zero of b' at $y_1(p)$ is the degree of tangency of $W^u(\Omega_2)$ and $W^s(\Omega_1)$ at p. If this degree is r we require the $2r$ conditions

$$d^k \xi_{10}(y_1(p)) = 0$$
$$d^k \xi_{11}(y_1(p)) = 0$$

$k = 0, 1, \ldots, r - 1$ for solving (1). This in addition to (5) and the analogous condition $u_{21}(0) = 0$ (to handle the case supp $(w) \subset M_2$) gives $2r + 2$ sufficient conditions for solving (1). (We have solved (1) on a neighbourhood of cl (W^{us}); we have already argued that this is enough.)

(14.5) We now argue that these $2r + 2$ conditions are necessary as well as sufficient; i.e. that dim coker $(1 - f_\#) = 2r + 2$. For this we construct $2r + 2$ $f_\#$-invariant linear functionals $\delta_1, \delta_2, \omega_1, \ldots, \omega_{2r}$. They are given by:

$$\langle \delta_i, w \rangle = \partial_1 u_i(0, 0)$$
$$\langle \omega_j, w \rangle = \sum_{n = -\infty}^{\infty} \langle \varepsilon_j, f_\#^n w \rangle$$

for $j = 1, \ldots, 2r$ where $z_{i\#}w = (u_i, v_i)$, $i = 1, 2$ and ε_j is defined by:

$$\langle \varepsilon_j, w \rangle = \partial_2^{j-1} u_i(0, y_i(p)),$$
$$\langle \varepsilon_{r+j}, w \rangle = (\partial_2^{j-1} \partial_1 u_i)(0, y_i(p)),$$

for $j = 1, \ldots, r$. We are assuming $d^j b(y_1(p)) = 0$ for $j = 0, 1, \ldots, r$ and the overlap relations give

$$u_2(x + b(y), c(y)) = u_1(x, y) + b'(y)v_1(x, y),$$

so that the definition of ε_j is independent of the choice of $i = 1, 2$.

The convergence of the bi-infinite series is easy; in fact

$$\sum_{n=0}^{\infty} \langle \varepsilon_j, f_{\#}^n w \rangle = \sum_{n=0}^{\infty} \mu_2^{1-j} \lambda_2^j (\partial_2^{j-1} u_2)(0, c_n)$$

$$\sum_{n=1}^{\infty} \langle \varepsilon_j, f_{\#}^{-n} w \rangle = \sum_{n=1}^{\infty} \mu_1^{j-1} \lambda_1^{-j} (\partial_2^{j-1} u_1)(0, c_{-n})$$

$$\sum_{n=0}^{\infty} \langle \varepsilon_{r+j}, f_{\#}^n w \rangle = \sum_{n=0}^{\infty} \mu_2^{1-j} (\partial_1 \partial_2^{j-1} u_2)(0, c_n)$$

$$\sum_{n=1}^{\infty} \langle \varepsilon_{r+j}, f_{\#}^{-n} w \rangle = \sum_{n=1}^{\infty} \mu_1^{j-1} (\partial_1 \partial_2^{j-1} u_1)(0, c_{-n})$$

where $c_n = \mu_2^{-n} y_2(p)$, $c_{-n} = \mu_1^n y_1(p)$. These series clearly converge (the convergence in the case of ω_{r+1} requires that $\langle \delta_i, w \rangle = 0$ for $i = 1, 2$). The required invariance

$$\langle \delta_i, f_{\#} w \rangle = \langle \delta_i, w \rangle$$

$$\langle \omega_j, f_{\#} w \rangle = \langle \omega_j, w \rangle$$

is also immediate.

(14.6) Finally let us ask what must be the nature of the $\mathbb{N}$-normed space $\mathcal{F}(M)$ in order that these constructions converge. We have required that $\partial_1^j u_i$ exist on M_i and that $\partial_2^j u_i$ exist near p. As we must allow perturbations of $f_{\#}$ and we solved near p by dividing by b' some form of the C^{∞} preparation theorem will be required. Hence we insist that elements of $\mathcal{F}(M)$ be C^{∞} near p. These requirements together with the requirement that $\mathcal{F}(M)$ be adapted to the unfolding, essentially determine the definition of $\mathcal{F}(M)$ given below.

15. *The tame division theorem*

For each open interval $I \subset \mathbb{R}$ let $\mathcal{B}^r(I)$ denote the Banach space of C^r-functions $f: I \to \mathbb{R}$ with bounded derivatives of order $\leq r$:

$$f \in \mathcal{B}^r(I) \Leftrightarrow |f|_r < \infty$$

where $|f|_r = \sum_{k=0}^{r} \sup \{|d^k f(t)| : t \in I\}$. Thus the $\mathbb{N}$-normed space $\mathcal{B}(I)$ is defined by

$$\mathcal{B}(I) = \bigcap_{r \in \mathbb{N}} \mathcal{B}^r(I).$$

Denote by S_p the p-dimensional vector space of all (real) polynomials $u: \mathbb{R} \to \mathbb{R}$ of degree $< p$:

$$u \in S_p \Leftrightarrow u(t) = \sum_{j=1}^{p} u_j t^{p-j},$$

for $t \in \mathbb{R}$ and suitable $u_1, \ldots, u_p \in \mathbb{R}$.

THEOREM. *Let $I \subset \mathbb{R}$ be an open interval about 0 and $g_0 \in \mathcal{B}(I)$ be regular of order p at 0:*

$$d^j g_0(0) = 0 \qquad j = 0, 1, \ldots, p-1,$$

$$d^p g_0(0) \neq 0.$$

Then there exists an open interval $J \subset I$ about 0, a neighbourhood $\mathcal{N}$ of g_0 in $\mathcal{B}(I)$ a tame family $q: \mathcal{N} \to L(\mathcal{B}(I), \mathcal{B}(J))$ and a continuous family $h: \mathcal{N} \to L(\mathcal{B}(I), S_p)$ such that

$$f = g \cdot q + h \tag{1}$$

(on J) for $g \in \mathcal{N}$, $f \in \mathcal{B}(I)$, $q = q(g)f$, and $h = h(g)f$.

Proof. Each cut-off function $\theta \in C_0^\infty(I)$ gives a linear map $\mathcal{B}(I) \to \mathcal{B}(\mathbb{R}): g \to \theta g$ so there is no loss of generality in assuming $I = \mathbb{R}$.

We shall use the C^∞-division theorem as formulated by Sergeraert [**19**, p. 636] (See also [**9**].) This asserts the existence of mappings

$$Q: C^\infty(\mathbb{R}) \times \mathbb{R} \times S_p \to \mathbb{R}$$
$$H: C^\infty(\mathbb{R}) \times S_p \to S_p$$

satisfying the equation:

$$f(t) = (t^p + u(t))Q(f, t, u) + H(f, u)(t) \tag{2}$$

for $f \in C^\infty(\mathbb{R})$, $u \in S_p$, and $t \in \mathbb{R}$. Moreover $Q(f, t, u)$ and $H(f, u)$ are $\mathbb{R}$-linear in f for fixed (t, u) and C^∞ in (t, u) for fixed f and we have estimates:

$$\left|D_2^k H(f, u)\right| \le C|f|_{(4p-1)k+p+4} \tag{3}$$

for $k \in \mathbb{N}$, $f \in \mathcal{B}(\mathbb{R})$, $u \in X$ where $X \subset S_p$ is compact and $C = C(k, X)$ and:

$$\left|D_2^r D_3^k Q(f, t, u)\right| \le C|f|_{r+(4p-1)k+p+4} \tag{4}$$

for $r, k \in \mathbb{N}$, $f \in \mathcal{B}(\mathbb{R})$, $(t, u) \in Y$ where $Y \subset \mathbb{R} \times S_p$ is compact and $C = C(r, k, Y)$. (We require the inequalities (3) and (4) only for $k \le 2$).

We will now imitate the proof of the preparation theorem from the division theorem as in [**24**]. (Wall's formulation involves a finite dimensional parameter z; we take g itself as the parameter.) First note that (3) together with the Lipschitz estimates

$$\left|D_2^j H(f, v) - D_2^j H(f, u)\right| \le L|v - u|,$$

with $L = \sup \{|D_2^{j+1} H(f, u + t(v - u))|: 0 \le t \le 1\}$ give that the map $u \to (f \to H(f, u))$ is (essentially) a map $S_p \to L(\mathcal{B}^l(\mathbb{R}), S_p)$ of class C^{k-1} for $l = (4p-1)k+p+4$. (One uses uniform continuity to extend the map $f \to H(f, u)$ to $\mathcal{B}^l(\mathbb{R})$.) The evaluation map

$$\mathcal{B}^l(\mathbb{R}) \times L(\mathcal{B}^l(\mathbb{R}), S_p) \to S_p$$

is a continuous bilinear map so we have extended H to a map $H: \mathcal{B}^l(\mathbb{R}) \times S_p \to S_p$ of class C^{k-1} (as a map between Banach spaces). Note that exactly the same reasoning shows that Q extends to a map

$$Q: \mathcal{B}^l(\mathbb{R}) \times \mathbb{R} \times S_p \to \mathbb{R}$$

of class C^{k-1}. We now take $k = 2$ so that $l = 9p+2$. Thus $H: \mathcal{B}^{9p+2}(\mathbb{R}) \times S_p \to S_p$ is class C^1. Similarly (with $r = k = 1$) $Q: \mathcal{B}^{5p+3}(\mathbb{R}) \times \mathbb{R} \times S_p \to \mathbb{R}$ is continuous.

Next take $f = g_0$ and $u = 0$ in (2) to obtain

$$g_0(t) = t^p Q(g_0, t, 0) + H(g_0, 0)(t).$$

Since g_0 is regular of order p we conclude that

$$H(g_0, 0) = 0 \qquad (5)$$

$$Q(g_0, 0, 0) \neq 0. \qquad (6)$$

Now let $\{e_j\}, j = 0, 1, \ldots, p-1$, be the basis of S_p determined by $e_j(t) = t^j$ and differentiate (2) in the direction e_j and evaluate at $u = 0$, $f = g_0$. The result is the identity:

$$0 = e_j Q_0 + e_p Q_1 + D_2 H(g_0, 0) e_j \qquad (7)$$

where $Q_0(t) = Q(g_0, t, 0)$, $Q_1(t) = D_3 Q(g_0, t, 0) e_j$. Differentiate (7) (repeatedly) in t and set $t = 0$. The result shows that the matrix representing the linear transformation $D_2 H(g_0, 0): S_p \to S_p$ with respect to the basis $\{e_j\}$ is triangular, and by (6) the diagonal entries are not zero. We conclude

$$\det(D_2 H(g_0, 0)) \neq 0. \qquad (8)$$

Now (5) and (8) are the hypotheses of the (Banach space) implicit function theorem. This gives a neighbourhood $\mathcal{N}$ of g_0 and a C^1 map $P: \mathcal{N} \subset \mathcal{B}^{9p+2}(\mathbb{R}) \to S_p$ satisfying

$$P(g_0) = 0 \qquad (9)$$

$$H(g, P(g)) = 0 \qquad (10)$$

for $g \in U$. By (6) and the continuity of Q we may shrink $\mathcal{N}$ and find an interval J about 0 and a positive real number δ so that

$$0 < \delta \leq Q(g, t, P(g)) \qquad (11)$$

for $g \in \mathcal{N}$ and $t \in \text{closure}(J)$. We now define:

$$h(g)f = H(f, P(g))$$

$$(q(g)f)(t) = Q(f, t, P(g))/Q(g, t, P(g))$$

for $g \in \mathcal{N} \cap \mathcal{B}(\mathbb{R}) \subset \mathcal{B}^{9p+2}(\mathbb{R})$, $f \in \mathcal{B}(\mathbb{R})$, and $t \in J$. The tame estimates follow from (11) and (4) (with $k = 0$). Read $P(g)$ for u in (2) to obtain

$$f(t) = (t^p + P(g)(t))Q(f, t, P(g)) + P(f, g)(t)$$

and read g for f and $P(g)$ for u in (2) and use (10) to obtain:

$$g(t) = (t^p + P(g)(t))Q(g, t, P(g)).$$

These equations imply (1).

16. *The tame sub-module perturbation theorem*

We continue the notation of § 15. Our aim is to present a tame version of a well-known variant of the C^∞-preparation theorem of Malgrange and Mather.

THEOREM. *Let I be an interval about 0 in $\mathbb{R}$ and suppose*

$$A_0 \in \mathcal{B}(I) \otimes L(\mathbb{R}^l, \mathbb{R}^n),$$

$$B_0 \in \mathcal{B}(I) \otimes L(\mathbb{R}^m, \mathbb{R}^n).$$

Assume that there is a tame linear map $\mathcal{B}(I) \otimes \mathbb{R}^n \to \mathcal{B}(I) \otimes \mathbb{R}^m: F \to G_0$ and a continuous linear map $\mathcal{B}(I) \otimes \mathbb{R}^n \to \mathbb{R}^l: F \to H_0$, such that

$$F(t) = A_0(t)H_0 + B_0(t)G_0(t) \qquad (1)$$

for $t \in I$. Then there exists an interval $J \subset I$ about 0, a neighbourhood $\mathcal{N} \subset \mathcal{B}(I) \otimes \{L(\mathbb{R}^l, \mathbb{R}^n) \times L(\mathbb{R}^m, \mathbb{R}^n)\}$ of (A_0, B_0), a tame map $\mathcal{N} \times \mathcal{B}(I) \otimes \mathbb{R}^n \to \mathcal{B}(J) \otimes \mathbb{R}^m$, $(A, B, F) \to G$, which is linear in F and a continuous map $\mathcal{N} \times \mathcal{B}(I) \otimes \mathbb{R}^n \to \mathbb{R}^l$, $(A, B, F) \to H$, also linear in F such that

$$F(t) = A(t)H + B(t)G(t) \tag{2}$$

for $(A, B) \in \mathcal{N}$ and $F \in \mathcal{B}(I) \otimes \mathbb{R}^n$.

Proof. The analogous statement for germs is corollary 5.6 of theorem 5.5 (entitled the 'Preparation Theorem') of [25]. We shall mimic the proof given there. Our proof will be more wordy since it contains implicitly a proof of the preparation theorem and since we have not developed a theory of quotient modules in the tame category. Note also that in [25] the variable t can be multi-dimensional; whereas here it is one-dimensional. The reason is that the inductive argument in the preparation theorem is not valid in the tame category. (See (13.3.3).)

We shall view elements of $\mathbb{R}^n$ as column vectors so that an element of $L(\mathbb{R}^m, \mathbb{R}^n)$ is a matrix with m columns. This explains the title of the theorem: equation (1) says that the columns of the matrix A_0 span an l-dimensional complement in the module $\mathcal{B}(I) \otimes \mathbb{R}^n$ to the submodule generated by the columns of B_0 (the ring is $\mathcal{B}(I)$); equation (2) says that this remains true when A_0 and B_0 are perturbed to A and B. If there were no loss of differentiability in the map $F \to G_0$ this would follow trivially by Banach space arguments.

We now proceed with the proof. We assume $\mathcal{N}$ is sufficiently small that the constructions we give are well defined. We use Greek letters to denote quantities (which depend on A and B and) which vanish when $A = A_0$ and $B = B_0$.

Note that (1) implies an equation

$$F(t) = A(t)H_0 + B(t)G_0(t) + \Omega(t).$$

Also the columns of $A_0(0)$ and $B_0(0)$ span $\mathbb{R}^m$ so (shrinking I if necessary) we have a linear map

$$\mathcal{B}(I) \otimes \mathbb{R}^n \to \mathcal{B}(I) \otimes \{L(\mathbb{R}^l, \mathbb{R}^n) \times L(\mathbb{R}^m, \mathbb{R}^n)\}$$

$$\Omega \to (\Theta, \Gamma)$$

such that

$$\Omega(t) = A(t)\Theta(t) + B(t)\Gamma(t)$$

for $t \in I$. This gives:

$$F(t) = A(t)\{H_0 + \Theta(t)\} + B(t)\{G_0(t) + \Gamma(t)\}. \tag{3}$$

Apply (3) column by column to obtain a matrix L_0 and a map $(A, B) \to (\Lambda, X)$ such that

$$tA(t) = A(t)\{L_0 + \Lambda(t)\} + B(t)X(t). \tag{4}$$

Thus $L_0 \in L(\mathbb{R}^l)$, $\Lambda \in \mathcal{B}(I) \otimes L(\mathbb{R}^l)$ and $X \in \mathcal{B}(I) \otimes L(\mathbb{R}^l, \mathbb{R}^n)$. This implies

$$d(t)A(t) = B(t)Y(t) \tag{5}$$

where $d(t) = \det(tE - L - \Lambda(t))$ and

$$Y(t) = X(t)\{tE - L_0 - \Lambda(t)\}^{\ddagger}$$

with E the $l \times l$ identity matrix and $\ddagger$ denoting the transposed matrix of cofactors.

When $\Lambda \equiv 0$, d is a monic polynomial of degree l; hence d is regular of order $p \leq l$. Thus by the tame division theorem there is a tame map $\mathcal{B}(I) \times \mathcal{N} \to \mathcal{B}(J)$, $(f, A, B) \to f_p$, linear in f and continuous maps $\mathcal{B}(I) \times \mathcal{N} \to \mathbb{R}$, $(f, A, B) \to f_i$, $i = 0, 1, \ldots, p-1$, such that

$$f(t) = d(t)f_p(t) + \sum_{i=0}^{p-1} t^i f_i \tag{6}$$

for $t \in J$. As in (4) we have

$$t^j A(t) = A(t)\{L_j + \Lambda_j(t)\} + BX_j(t) \tag{7}$$

for $j = 0, 1, \ldots, p-1$ and applying (6) component by component gives

$$\Lambda_j(t) = d(t)\Lambda_{pj}(t) + \sum_{i=0}^{p-1} t^i \Lambda_{ij}. \tag{8}$$

By (5), (7), and (8):

$$\sum_{i=0}^{p-1} t^i A(t)\{\delta_{ij}E - \Lambda_{ij}\} = A(t)L_j + B(t)\{Y(t)\Lambda_{pj}(t) + X_j(t)\}, \tag{9}$$

where δ_{ij} is the Kronecker δ. Define $\Delta_{jk} \in L(\mathbb{R}^l)$ by

$$\sum_{j=0}^{p-1} \{\delta_{ij}E - \Lambda_{ij}\}\{\delta_{jk}E + \Delta_{jk}\} = \delta_{ik}E, \tag{10}$$

so that (9) and (10) give

$$t^k A(t) = A(t)P_k + B(t)Q_k(t) \tag{11}$$

for $k = 0, 1, \ldots, p-1$ where:

$$P_k = \sum_{j=0}^{p-1} L_j\{\delta_{jk}E + \Delta_{jk}\}$$

$$Q_k(t) = \sum_{j=0}^{p-1} \{Y(t)\Lambda_{pj}(t) + X_j(t)\}\{\delta_{jk}E + \Delta_{jk}\}.$$

Apply (6) to the Θ from (3) to obtain

$$\Theta(t) = d(t)\Theta_p(t) + \sum_{k=0}^{p-1} t^k \Theta_k \tag{12}$$

and substitute (12) in (3) and use (5) to obtain:

$$F(t) = A(t)\left\{ H_0 + \sum_{k=0}^{p-1} t^k \Theta_k \right\} + B(t)\{Y(t)\Theta_p(t) + G_0(t) + \Gamma(t)\}. \tag{13}$$

Finally (13) and (11) give (2) with

$$H = H_0 + \sum_{k=0}^{p-1} P_k \Theta_k$$

$$G(t) = G_0(t) + \Gamma(t) + Y(t)\Theta_p(t) + \sum_{k=0}^{p-1} Q_k(t)\Theta_k.$$

The required estimates are easy since all our constructions were either explicit C^∞ functions of other constructions or defined by the tame division theorem.

17. $\mathbb{N}$-*jet semi-finslers on* $\mathbb{R}^m$

(17.1) For each multi-index $\alpha \in \mathbb{N}^m$ let $\Theta_\alpha : \mathbb{R}^m \to [0, \infty)$ be a continuous function. This determines an $\mathbb{N}$-jet semi-Finsler Θ on $\mathbb{R}^m$ via the formula

$$\Theta((j^r u)(x)) = \|u\|_{r,x}$$

for $u : \mathbb{R}^m \to \mathbb{R}$ and $x \in \mathbb{R}^m$ where $\|u\|_{r,x} = \sum_{|\alpha| \le r} \Theta_\alpha(x) |\partial^\alpha u(x)|$ and the sum is over all multi-indices $\alpha = (\alpha_1, \cdots, \alpha_m)$ of degree $|\alpha| = \alpha_1 + \cdots + \alpha_m \le r$. As usual put

$$\|u\|_r = \sup \{\|u\|_{r,x} : x \in \mathbb{R}^m\}$$

and recall the $\mathbb{N}$-normed space $\mathcal{B}(\Theta) = \mathcal{B}(\mathbb{R}^m, \Theta)$ of (5.3)

$$u \in \mathcal{B}(\Theta) \Leftrightarrow \|u\|_r < \infty \qquad \text{for all } r \in \mathbb{N}.$$

We generally use the abbreviations $\mathcal{X}(\Theta) = \mathcal{B}(\Theta) \otimes \mathbb{R}^m$, $\mathcal{L}(\Theta) = \mathcal{B}(\Theta) \otimes L(\mathbb{R}^m)$.

PROPOSITION. *Assume that for each pair* (α, β) *of multi-indices we have a 'pointwise multiplication inequality'*

$$\Theta_{\alpha+\beta}(x) \le c_{\alpha,\beta}(x) \Theta_\alpha(x) \Theta_\beta(x) \tag{1}$$

for $x \in \mathbb{R}^m$ *where* $c_{\alpha,\beta} : \mathbb{R}^m \to [0, \infty)$ *is continuous. Then given* $k, r \in \mathbb{N}$ *there is a continuous function* $c = c_{r,k} : \mathbb{R}^m \to \mathbb{R}$ *such that we have the 'pointwise composition inequality'*

$$\|v\|_{r,x} \le c(x) \sum |q|_{\gamma,x,u} \|u\|_{i_1,x} \cdots \|u\|_{i_j,x} \tag{2}$$

where $v(x) = q(x, u(x))$, $|q|_{\gamma,x,u} = |(\partial^\gamma q)(x, u(x))|$, *and the sum is over all multi-indices* $\gamma = (\beta, \delta) \in \mathbb{N}^m \times \mathbb{N}^k$ *and* $i_1, \ldots, i_j$ *satisfying* $|\gamma| = |\beta| + |\delta| \le r$, $i_1 + \cdots + i_j = |\delta|$.

Proof. This is immediate from Faa di-Bruno's formula

$$\partial^\beta v(x) = \sum (\partial^\gamma q)(x, u(x)) \, \partial^{\alpha_1} u_{k_1}(x) \cdots \partial^{\alpha_j} u_{k_j}(x).$$

Remark. If all the $c_{\alpha,\beta}$ are constants so are the $c_{r,k}$.

COROLLARY. *As in corollary 1 of § 6 this implies that* $\mathcal{B}(\Theta)$ *is a 'tame algebra'*

$$\|uv\|_r \le c \|u : v\|_r \tag{2}$$

(17.2) We give some examples.

Example 1. Take $\Theta_\alpha \equiv 1$ for all α. Then $\mathcal{B}(\Theta) = \mathcal{B}(\mathbb{R}^m)$ where $\mathcal{B}(\mathbb{R}^m)$ is the space of all functions having bounded derivatives of all orders. This is a tame algebra.

Example 2. Take $m = 1$,

$$\Theta_0(y) = 1$$
$$\Theta_r(y) = \min(1, |y|^{r-\rho}) \qquad r \ge 1,$$

where ρ is a real number with $0 < \rho < 1$. Then $\mathcal{B}(\Theta)$ is a tame algebra.

Example 3. Let $m = 1$,

$$\Theta_0(x) = (1 + x^2)^{-1}$$
$$\Theta_1(x) = (1 + |x|)^{-1}$$
$$\Theta_r(x) = 1 \qquad r \ge 2.$$

Then elements of $\mathcal{B}(\Theta)$ may grow no faster than quadratically (but $u(x) = x^2$ is an element of $\mathcal{B}(\Theta)$) so it is not an algebra.

PROPOSITION. *All three of the examples admit a family of smoothing operators with offset $d = 0$ and hence (by the proposition of (2.6)) satisfy the interpolation inequalities*:

$$\|u\|_q \leq C \|u\|_p^\alpha \|u\|_r^\beta \tag{3}$$

for $C = C(q, p, r)$, $p \leq q \leq r$ where

$$\alpha = (r-q)/(r-p), \qquad \beta = (q-p)/(r-p).$$

Proof. For example 1 use convolutions. For example 3 note the isomorphism

$$\mathbb{R}^2 \times \mathscr{B}(\mathbb{R}) \to \mathscr{B}(\Theta) : (a, b, v) \to u$$

$$u(x) = a + bx + \int_0^x \int_0^t v(s) \, ds \, dt.$$

For example 2 use a cut-off function to express any element in the form $u_1 + u_2$ with supp $(u_1) \subset [-2, 2]$, supp $(u_2) \subset \mathbb{R} \setminus [-1, 1]$. Clearly u_2 lies in $\mathscr{B}(\mathbb{R})$ so we may smooth it with convolutions. For u_1 note the map

$$\mathscr{B}(\Theta) \to \mathscr{B}(\mathbb{R}) \oplus \mathbb{R} \oplus \mathscr{B}(\mathbb{R}) : u \to (v_+, v_0, v_-)$$

where $v_\pm(t) = (u(\pm e^t) - u(0)) e^{-\rho t}$, $v_0 = u(0)$. This map is injective and in fact is an isomorphism if Θ is replaced by a locally equivalent $\mathbb{N}$-jet semi-Finsler

$$\Theta_0(y) = 1 + |y|^\rho$$

$$\Theta_r(y) = |y|^{r-\rho}, \qquad r \geq 1.$$

This follows from the formula

$$(d^r v)(t) = \sum_{k=1}^r c_k (d^k u)(e^t) e^{t(k-\rho)}$$

with $c_r = 1$.

(17.3) PROPOSITION. *Suppose $f: \mathbb{R}^m \to \mathbb{R}^m$ is a linear automorphism of form*

$$f(x) = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_m x_m)$$

for $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$ with $\lambda_1, \ldots, \lambda_m > 0$. Suppose the $\mathbb{N}$-jet semi-Finsler Θ satisfies inequalities:

$$\lambda^\alpha \Theta_\alpha(x) \leq \nu_r \Theta_\alpha(f(x)) \tag{4}$$

for each multi-index α with $|\alpha| \leq r$. Then the automorphism $f^*: \mathscr{B}(\Theta) \to \mathscr{B}(\Theta)$ satisfies:

$$\|f^* u\|_r \leq \nu_r \|u\|_r. \tag{5}$$

Proof. $\partial^\alpha (f^* u)(x) \Theta_\alpha(x) = \lambda^\alpha (\partial^\alpha u)(f(x)) \Theta_\alpha(x)$.

COROLLARY 1. *For Θ as in example 2 and $f: \mathbb{R} \to \mathbb{R}$ given by $f(x) = \lambda x$ with $\lambda < 1$ we have (after a suitable renorming)*

$$\|f^* u\|_r \leq \lambda^2 \|u\|_r$$

for all $u \in \mathscr{B}(\Theta)$ with $u(0) = du(0) = 0$.

(17.4) THEOREM (Dynamical product estimate). *Assume Θ satisfies (1) so that $\mathscr{B}(\Theta)$ is a tame algebra. Let f as in proposition (2) satisfy (4), and $A \in \mathscr{B}(\Theta) \otimes L(\mathbb{R}^k)$. Define*

a linear operator

$$T: \mathcal{B}(\Theta) \otimes \mathbb{R}^k \to \mathcal{B}(\Theta) \otimes \mathbb{R}^k$$
$$Tu = f^*(Au)$$

for $u \in \mathcal{B}(\Theta) \otimes \mathbb{R}^k$. *Assume* $\Theta_0 \equiv 1$, *the inequalities* (3), *and that* T *satisfies the estimate*

$$\|Tu\|_0 \le \sigma \|u\|_0 \tag{6}$$

for some σ *with* $0 < \sigma < 1$. *Then for each* r *there is a constant* $C = C(r)$ *such that for* $n \ge 1$:

$$\|T^n u\|_r \le C n^r \sigma^n \|A : u\|_r. \tag{7}$$

Proof. Since f^* is an algebra homomorphism we have the formula

$$T^n u = (f^*A)(f^{*2}A) \cdots (f^{*n}A)(f^{*n}u).$$

Applying ∂^α with $|\alpha| \le r$ gives a formula

$$\partial^\alpha (T^n u) = \sum M$$

where each term M is of form $M = \{\partial^{\beta(1)} f^*A\} \cdots \{\partial^{\beta(n)} f^{*n}A\}\{\partial^\gamma f^{*n}u\}$, $\alpha = \beta(1) + \cdots + \beta(n) + \gamma$, and where there are $(n+1)^r \le C n^r$ terms in all.

We first show that each M satisfies the pointwise estimate;

$$|M(x)| \Theta_\alpha(x) \le C \sigma^n |A_1| \cdots |A_l| |U_\gamma|, \tag{8}$$

where $U_\gamma = \lambda^\gamma (\partial^\gamma u)(f^n(x)) \Theta_\gamma(x)$, $A_k = (\lambda^{\beta(j)})^r (\partial^{\beta(j)} A)(f^r(x)) \Theta_{\beta(j)}(x)$, and $j = j(k)$, $k = 1, \ldots, l$, are precisely those integers with $\beta(j(k)) \neq 0$. This follows by double induction on n and $|r|$ using (6) in the case where M has form $M = TM'$, and the crude estimate

$$|\partial^\beta (f^*A)(y)| \le c |(\partial^\beta A)(f(y))|$$

in the case where M has form $M = \{\partial^\beta (f^*A)\} f^* M'$ with $\beta = \beta(1) + 0$.

Now by (4) and (8) we obtain

$$M(x) \gamma_\alpha(x) \le C \sigma^n \|A\|_{b(1)} \cdots \|A\|_{b(l)} \|u\|_c,$$

with $b(k) = |\beta(j(k))|$, $c = |\gamma|$ so that $|\alpha| = b(1) + \cdots + b(l) + c$. Now take the sup on x and use interpolation.

Remark. By enlarging σ (slightly) and C (a lot) we may replace $C n^r \sigma^n$ by $C \sigma^n$ in (7).

18. *An $\mathbb{N}$-jet semi-finsler on $\mathbb{R}^2$*

(18.1) Define as in § 17 an $\mathbb{N}$-jet semi-Finsler Θ on $\mathbb{R}^2$ by the formulae

$$\Theta_{i0}(x, y) = (1 + |x|^{2^{-i}})^{-1},$$
$$\Theta_{ij}(x, y) = \min(1, |y|^{j-\rho}) \Theta_{i0}(x, y),$$
$$\Theta_{k0}(x, y) = 1,$$
$$\Theta_{kj}(x, y) = \min(1, |y|^{j-\rho}),$$

for $i = 0, 1$; $k \ge 1$. (Here $\rho > 0$ is a small positive number specified more precisely below.) As in § 17 this gives an $\mathbb{N}$-normed space:

$$u \in \mathcal{B}(\Theta) \Leftrightarrow \|u\|_r < \infty \qquad \text{for all } r \in \mathbb{N},$$

where

$$\|u\|_r = \text{supp } \{\|u\|_{r,z} : z \in \mathbb{R}^2\}$$
$$\|u\|_{r,z} = \sum_{i+j \leq r} |(\partial_1^i \partial_2^j u)(x, y)| \Theta_{ij}(x, y)$$

for $z = (x, y) \in \mathbb{R}^2$.

PROPOSITION 1. *If* $u \in \mathcal{B}(\Theta)$ *then* u *is continuous on* $\mathbb{R}^2$ *and* C^∞ *on* $y \neq 0$. *Moreover the partial derivatives* $\partial_1^k u$ *are* C^∞ *for* $k \geq 0$ *and bounded for* $k \geq 2$. *For* $j \geq 1$ *the mixed partials* $\partial_1^i \partial_2^j u$ *blow up at worst like* $|y|^{p-j}$ *as* $y \to 0$.

PROPOSITION 2. *The functions* Θ_{ij} *satisfy the pointwise inequalities*:
$$\Theta_{ij}(x, y) \leq (4 + 4x^2)\Theta_{i'j'}(x, y)\Theta_{i''j''}(x, y),$$
for $i = i' + i''$, $j = j' + j''$. (*Hence proposition 1 of* § 17 *applies.*)

PROPOSITION 3. *The* $\mathbb{N}$-*normed space* $\mathcal{B}(\Theta)$ *admits a family of smoothing operators with offset* 0 *and hence satisfies the interpolation inequalities.*

Proof. As in examples 2 and 3 of § 17.

(18.2) Now fix $\mu, \lambda_1, \lambda_2 \in \mathbb{R}$ satisfying $0 < \mu < 1 < \lambda_1 < \lambda_2$ and define an unfolding $F \subset \text{Diff}(\mathbb{R}^2)$ by

$$f \in F \Leftrightarrow f(x, y) = (\lambda x, \mu y)$$

for some $\lambda \in (\lambda_1, \lambda_2)$.

PROPOSITION 4. *Let* $\psi \in C_0^\infty(\mathbb{R}^2)$ *be a cut-off function. Then the map*
$$F \times \mathcal{B}(\Theta) \to \mathcal{B}(\Theta)$$
$$(f, u) \to \psi f^* u$$

is tame-C^r of degree r *for each* $r \in \mathbb{N}$.

Proof. For $\lambda \in (\lambda_1, \lambda_2)$ and $k \leq r$ let

$$v = \left(\frac{d}{dk}\right)^k (\psi f^* u)\bigg|_{\lambda = \lambda}.$$

Then

$$v(x, y) = \psi(x, y)x^k(\partial_1^k u)(\lambda x, \mu y)$$

so the estimate $\|v\|_p \leq C\|u\|_{p+k}$ is obvious. As v is linear in u this gives the desired tame estimates. For continuity note the inequalities

$$|(\partial_1^k u)(\lambda x, \mu y) - (\partial_1^k u)(\lambda_0 x, \mu y)| \leq c|x|\,|\lambda - \lambda_0|\,\|u\|_r,$$
$$|y|^{j-p}|(\partial_1^i \partial_2^j u)(\lambda x, \mu y) - (\partial_1^i \partial_2^j(\lambda_0 x, \mu y)| \leq c|x|\,|\lambda - \lambda_0|\,\|u\|_r,$$

for $k < r$, $1 \leq j$, $k+j < r$, $(x, y) \in \text{supp}(\psi)$, and $\lambda, \lambda_0 \in (\lambda_1, \lambda_2)$. Take the sup over (x, y), note that

$$\partial_1^i \partial_2^j f^* u = \lambda^i \mu^j f^* \partial_1^i \partial_2^j u$$

and use the triangle inequality to obtain the Lipschitz estimate

$$\|f^* u - f_0^* u\|_{r-1} \leq c\|u\|_r |\lambda - \lambda_0|$$

which (with the tame estimate) easily implies the continuity of the map $(f, u) \to v$.

(18.3) Now fix λ and hence f and recall the abbreviation $\mathscr{X}(\Theta) = \mathscr{B}(\Theta) \otimes \mathbb{R}^2$. Define $f^{\#}, f_{\#} \colon \mathscr{X}(\Theta) \to \mathscr{X}(\Theta)$ by:

$$f_{\#} w = (\lambda f_* u, \mu f_* v),$$
$$f^{\#} w = (\lambda^{-1} f^* u, \mu^{-1} f^* v),$$

for $w = (u, v) \in \mathscr{X}(\Theta)$. (Thus $f_{\#}^{-1} = f^{\#}$.)

We shall define an $f_{\#}$-invariant splitting

$$\mathscr{X}(\Theta) = \mathscr{E}^u(f_{\#}) \oplus \mathscr{E}^s(f_{\#}). \tag{1}$$

To this end define projection operators e_0, e_1, e_2 by:

$$(e_0 v)(x, y) = v(0, y)$$
$$(e_1 v)(x, y) = (\partial_1 v)(0, y)x$$
$$(e_2 v)(x, y) = \int_0^x \int_0^t (\partial_1^2 v)(s, y) \, ds \, dt.$$

where v is any vector valued function defined on $\mathbb{R}^2$. Note the formulae:

$$v = e_0 v + e_1 v + e_2 v;$$
$$e_i e_j = 0, \qquad i \neq j;$$
$$e_i^2 = e_i;$$
$$f_*(e_i v) = e_i(f_* v);$$

and

$$\begin{bmatrix} e_0(au) \\ e_1(au) \\ e_2(au) \end{bmatrix} = \begin{bmatrix} e_0 a & 0 & 0 \\ e_1 a & e_0 a & 0 \\ e_2 a & a - e_0 a & a \end{bmatrix} \begin{bmatrix} e_0 u \\ e_1 u \\ e_2 u \end{bmatrix}.$$

Define the splitting (1) by

$$\mathscr{E}^u(f_{\#}) = \{(u, 0) \in \mathscr{X}(\Theta) \colon e_2 u = 0\},$$
$$\mathscr{E}^s(f_{\#}) = \{(u, v) \in \mathscr{X}(\Theta) \colon e_2 u = u\}.$$

Also define a linear functional $\delta(f_{\#}) \colon \mathscr{X}(\Theta) \to \mathbb{R}$ by:

$$\langle \delta(f_{\#}), w \rangle = (\partial_1 u)(0, 0).$$

PROPOSITION. *For $\sigma = \max \{\lambda^{-1} \mu^{-\rho}, \mu^{1-\rho}, \mu^{\rho}\}$ we have for $r \geq 2$:*

$$w \in \mathscr{E}^s(f_{\#}) \Rightarrow \|f_{\#} w\|_r \leq \sigma \|w\|_r$$
$$w \in \mathscr{E}^u(f_{\#}) \Rightarrow \|f^{\#} w\| \leq \|w\|_r$$

while if $\langle \delta(f_{\#}), w \rangle = 0$ this last inequality can be improved:

$$w \in \mathscr{E}^u(f_{\#}) \cap \operatorname{Ker}(\delta(f_{\#})) \Rightarrow \|f^{\#} w\|_r \leq \sigma \|w\|_r.$$

Hence if ρ is sufficiently small that $\sigma < 1$, the operator $f_{\#}$ is (except for an invariant one dimensional space on which it is the identity) hyperbolic in each of the norms $\|\cdot\|_r$.

Proof. This all follows from (17.3). We include some inequalities for future reference:

$$e_0 u = u \Rightarrow \begin{cases} \|f_* u\|_0 = \|u\|_0, \\ \|f_* u\|_r \leq \mu^{-\rho} \|u\|_r, \end{cases}$$

$$e_1 v = v \Rightarrow \begin{cases} \|f^* v\|_0 = \|v\|_0 \\ \|f^* v\|_r \leq \lambda \mu^{\rho} \|v\|_r, \end{cases}$$

$$e_2 w = w \Rightarrow \ \|f_* w\|_r \leq \lambda^{-2} \mu^{-\rho} \|w\|_r,$$

(up to equivalent norms). Note also the sharper inequality

$$\left. \begin{array}{l} e_2 u = 0 \\ (\partial_1 u)(0,0) = 0 \end{array} \right\} \Rightarrow \|f^* u\|_r \leq \lambda \mu^{\rho} \|u\|_r$$

provided that $\rho > 0$ is sufficiently small that $1 < \lambda \mu^{\rho}$.

19. *The splitting mod e_2.*

In § 20 we study a perturbation $Tw = f_*(au + bv, cu + dv)$ of the operator $f_\#$ of § 18; here we study it modulo higher order terms in x.

We take Θ to be the $\mathbb{N}$-jet semi-Finsler (on $\mathbb{R}$) of § 17 example 2 and denote by $\mathscr{P}$ the algebra

$$\mathscr{P} = \mathscr{B}(\Theta) \otimes \mathbb{R}[x]/(x^2).$$

We write elements of $\mathscr{P}$ as matrix valued functions of y:

$$U \in \mathscr{P} \Leftrightarrow u = \begin{bmatrix} u_0 & 0 \\ u_1 & u_0 \end{bmatrix}$$

where $u_0, u_1 : \mathbb{R} \to \mathbb{R}$ with $\|u_0\|_r, \|u_1\|_r < \infty$.

Thus $U = U(y)$ is identified with the map $u : \mathbb{R}^2 \to \mathbb{R}$ given by $u(x, y) = u_0(y) + u_1(y) x$ so we define $f^* : \mathscr{P} \to \mathscr{P}$ and $f_* : \mathscr{P} \to \mathscr{P}$ by:

$$(f^* U)(y) = \begin{bmatrix} u_0(\mu y) & 0 \\ \lambda u_1(\mu y) & u_0(\mu y) \end{bmatrix}$$

$$(f_* U)(y) = \begin{bmatrix} u_0(\mu^{-1} y) & 0 \\ \lambda^{-1} u_1(\mu^{-1} y) & u_0(\mu^{-1} y) \end{bmatrix}$$

where f is as in § 18. Let

$$\Lambda_0 = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix},$$

$$f_\# : \mathscr{P} \otimes \mathbb{R}^2 \to \mathscr{P} \otimes \mathbb{R}^2,$$

$$f_\# W = f_*(\Lambda_0 W),$$

$$\mathscr{R}_\varepsilon = \{ \Lambda \in \mathscr{P} \otimes L(\mathbb{R}^2) : \|\Lambda - \Lambda_0\|_0 < \varepsilon \}.$$

THEOREM. *For $\varepsilon > 0$ sufficiently small there are tame maps*

$$\mathscr{R}_\varepsilon \subset \mathscr{P} \otimes L(\mathbb{R}^2) \to \mathscr{P} : \Lambda \to G, H$$

such that

$$(f_* \Lambda) \mathscr{E}^u(\Lambda) = \mathscr{E}^u(\Lambda)$$

$$(f_* \Lambda) \mathscr{E}^s(\Lambda) = \mathscr{E}^s(\Lambda)$$

where $\mathscr{E}^u(\Lambda) = \{(U, GU) : U \in \mathscr{P}\}$, $\mathscr{E}^s(\Lambda) = \{(HV, V) : V \in \mathscr{P}\}$.

Proof. Let $\Lambda \in \mathscr{P} \otimes L(\mathbb{R}^2)$ be given by

$$\Lambda = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Since $\mathscr{P}$ is a commutative algebra:

$$\Lambda^{-1} = \begin{bmatrix} D & -B \\ -C & A \end{bmatrix} \begin{bmatrix} AD - BC & 0 \\ 0 & AD - BC \end{bmatrix}.$$

Now define

$$\Gamma(\Lambda, G) = (C + DG)(A + BG)^{-1}$$

$$\Delta(\Lambda, H) = -(DH - B)(CH - A)^{-1}$$

so that the invariance conditions become fixed point equations:

$$G = f_* \Gamma(\Lambda, G), \tag{1.u}$$

$$H = \Delta(\Lambda, f^* H). \tag{1.s}$$

For $\Lambda = \Lambda_0$ we have

$$\Gamma(\Lambda_0, G) = \lambda^{-1} \mu G, \tag{2.u}$$

$$\Delta(\Lambda_0, H) = \lambda^{-1} \mu H, \tag{2.s}$$

so that the operators $f_* \Gamma$ and Δf^* are (linear) contractions on $\mathscr{P}_0$:

$$\|f_* \Gamma(\Lambda_0, G)\|_0 \le \sigma_0 \|G\|_0,$$

$$\|f^* \Delta(f_* \Lambda_0, H)\|_0 \le \sigma_0 \|G\|_0,$$

where $\sigma_0 < 1$. (In fact we can take $\sigma_0 = \lambda^{-1} \mu$ in the former inequality and $\sigma_0 = \mu$ in the latter.) Hence by the tame contraction principle of § 4 we only need estimates

$$\|D_2(f_* \Gamma)(\Lambda, G)^n \hat{G}\|_r \le C n^r \sigma^n [\![\Lambda, G : \hat{G}]\!]_r \tag{3.u}$$

$$\|D_2(f^* \Delta)(f_* \Lambda, H)^n \hat{H}\|_r \le C n^r \sigma^n [\![\Lambda, H : \hat{H}]\!]_r \tag{3.s}$$

for $n \ge 1$, $\|\Lambda - \Lambda_0\|_0 < \varepsilon$, $\|G\|_0$, $\|H\|_0 < 1$, where

$$[\![\Lambda, G : \hat{G}]\!]_r = (1 + \|\Lambda\|_r + \|G\|_r) \|\hat{G}\|_0 + \|\hat{G}\|_r,$$

$0 < \sigma < 1$ with $\sigma = \sigma(r)$, $C = C(r)$ independent of n. (Note that by increasing σ slightly and enlarging C we may replace $C n^r \sigma^n$ by $C \sigma^n$.) For $r \ge 1$ the fact that these inequalities hold for $\Lambda = \Lambda_0$ is of no help since we are *not* assuming that $\|\Lambda - \Lambda_0\|_r$ is small.

Explicit calculation gives

$$D_2(f_* \Gamma)(\Lambda, G) \hat{G} = f_*(L \hat{G}),$$

$$D_2(f^* \Delta)(f_* \Lambda, H) \hat{H} = f^*(M \hat{H}),$$

where $L(y) = D_2 \Gamma(\Lambda(y), G(y))$, $M(y) = D_2 \Delta((f_* \Lambda)(y), H(y))$. The inequalities (3.u) and (3.s) follow by continuity, (1), (2), the dynamic product estimate of (17.4) and the tame estimates

$$\|L\|_r \le C [\![\Lambda, G : \hat{G}]\!]_r,$$

$$\|M\|_r \le C [\![\Lambda, H : \hat{H}]\!]_r.$$

PROPOSITION. *Let* $\mathbb{E} \subset (\mathscr{P} \otimes \mathbb{R}^2)^*$ *denote the 4-dimensional space of all linear functionals* γ *of form*:

$$\langle \gamma, W \rangle = \langle \alpha, U(0) \rangle + \langle \beta, V(0) \rangle$$

for $W = (U, V) \in \mathscr{P} \otimes \mathbb{R}^2$ *where* $\langle \alpha, U(0) \rangle = \alpha_0 u_0(0) + \alpha_1 u_1(0)$ *and* $\langle \beta, V(0) \rangle = \beta_0 v(0) + \beta_1 v_1(0)$. *Then there is a continuous map* $\delta : \mathscr{R}_\varepsilon \subset \mathscr{P} \otimes L(\mathbb{R}^2) \to \mathbb{E}$ *such that we have estimates*:

$$\| (f_* \Lambda)^n W \|_r \leq C\sigma^n [\![\Lambda : W]\!]_r$$

for $W \in \mathscr{E}^s(\Lambda)$, $n \in \mathbb{N}$ *and*:

$$\| (f_* \Lambda)^{-n} W \|_r \leq C\sigma^n [\![\Lambda : W]\!]_r$$

for $W \in \mathscr{E}^u(\Lambda) \cap \mathrm{Ker}\, (\delta(\Lambda))$ *where* $[\![\Lambda : W]\!]_r = (1 + \| \Lambda \|_r) \| W \|_0 + \| W \|_r$.

Proof. The space $\mathbb{E}$ is invariant under $f_* \Lambda$ since

$$\mathbb{E}^\perp = \{ W \in \mathscr{P} \otimes \mathbb{R}^2 : W(0) = 0 \}.$$

When $\Lambda = \Lambda_0$ the eigenvalues of $(f_* \Lambda)^*$ on $\mathbb{E}$ are $\lambda, 1, \mu, \lambda^{-1}\mu$ so let $\delta(\Lambda)$ denote the eigenvector corresponding to the eigenvalue nearest 1 normalized so that its value on $W = (u_0, u_1, v_0, v_1) = (0, 1, 0, 0)$ is 1. The desired estimate follows from the corresponding estimates for $f_* \Lambda_0$ in (18.3) and the dynamic product estimate of (17.4).

20. The splitting $\mathscr{X}(\Theta) = \mathscr{E}^u(\Phi) \oplus \mathscr{E}^s(\Phi)$

We revert to the notation of § 18. Put $Z = \{ (x, y) \in \mathbb{R}^2 : |x|, |y| < l \}$ where l is any (large) positive real number.

THEOREM. *There is neighbourhood* $\mathscr{N}$ *of* 0 *in* $\mathscr{L}(\Theta)$, *a tame linear family* $T : \mathscr{N} \to L(\mathscr{X}(\Theta))$, *a tame family of splittings* $\mathscr{X}(\Theta) = \mathscr{E}^u(\Phi) \oplus \mathscr{E}^s(\Phi)$, *indexed by* $\Phi \in \mathscr{N}$, *and a continuous family of linear maps* $\delta : \mathscr{N} \to L(\mathscr{X}(\Theta), \mathbb{R})$, *such that the following conditions hold*:

(*invariance*):

$$T(\Phi) \mathscr{E}^e(\Phi) = \mathscr{E}^e(\Phi) \qquad e \in \{s, u\}, \Phi \in \mathscr{N}.$$

(*local form*):

$$\{ T(\Phi) w \} | Z = \{ (1 + \Phi) f_\# w \} | Z \qquad \Phi \in \mathscr{N}, w \in \mathscr{X}(\Theta);$$

(*tame asymptotic estimates*): *there exists* $\sigma \in (0, 1)$ *such that*

$$\| T(\Phi)^n w^s \|_r \leq c\sigma^n [\![\Phi : w^s]\!]_r,$$

$$\| T(\Phi)^{-n} w^u \|_r \leq c\sigma^n [\![\Phi : w^u]\!]_r,$$

for $\Phi \in \mathscr{N}$, $n \in \mathbb{N}$, $w^s \in \mathscr{E}^s(\Phi)$, $w^u \in \mathscr{E}^u(\Phi)$. *Here* $[\![\Phi : w]\!]_r = (1 + \| \Phi \|_r) \| w \|_0 + \| w \|_r$ *and the constant* C *is independent of* Φ, w^s, w^n *and* n (*but may depend on* r).

Proof. Fix $\psi \in C_0^\infty(\mathbb{R}^2)$ satisfying $\psi | Z \equiv 1$. Throughout we use the notation $w = (u, v) \in \mathscr{X}(\Theta)$ and $m = f^*(1 + \psi\Phi) m_0 \in \mathscr{L}(\Theta)$ where

$$m_0 = \begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix} \in L(\mathbb{R}^2),$$

so that

$$(1 + \psi\Phi) f_\# w = f_*(mw).$$

Let a, b, c, d denote the components of m:

$$m = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

We identify w with the 6 by 1 matrix

$$W = \begin{bmatrix} U \\ V \\ \Omega \end{bmatrix}$$

where

$$U = \begin{bmatrix} e_0 u \\ e_1 u \end{bmatrix}, \quad V = \begin{bmatrix} e_0 v \\ e_1 v \end{bmatrix}, \quad \Omega = \begin{bmatrix} e_2 u \\ e_2 v \end{bmatrix}.$$

The multiplication operator $w \to mw$ then takes the form $W \to MW$ where:

$$M = \begin{bmatrix} A & B & 0 \\ C & D & 0 \\ P & Q & R \end{bmatrix}$$

with

$$A = \begin{bmatrix} e_0 a & 0 \\ e_1 a & e_0 a \end{bmatrix}, \quad B = \begin{bmatrix} e_0 b & 0 \\ e_1 b & e_0 b \end{bmatrix}, \quad C = \begin{bmatrix} e_0 c & 0 \\ e_1 c & e_0 c \end{bmatrix},$$

$$D = \begin{bmatrix} e_0 d & 0 \\ e_1 d & e_0 d \end{bmatrix}, \quad P = \Theta \begin{bmatrix} e_2 a & a - e_0 a \\ e_2 c & c - e_0 c \end{bmatrix}, \quad Q = \Theta \begin{bmatrix} e_2 b & b - e_0 b \\ e_2 d & c - e_0 c \end{bmatrix},$$

and

$$R = \begin{bmatrix} (1 - \theta)\lambda + \theta a & \theta b \\ \theta c & (1 - \theta)\mu + \theta d \end{bmatrix}$$

with $\theta = 1$. To overcome the difficulty that $\mathscr{B}(\Theta)$ is not an algebra we replace $\theta \equiv 1$ in these formulae by a cut-off function $\theta = \theta(x)$ where $[-l, l] \subset \theta^{-1}(1)$, supp $(\theta) \subset [-2l, 2l]$. This gives a linear T, $TW = f_*(MW)$, which agrees with $(1 + \Phi)f_\#$ on Z. We shall construct an invariant splitting for T. Thus we construct T-invariant spaces

$$\mathscr{E}^u(T) = \{(U, GU, \Gamma U) : U \in \mathscr{P}\},$$
$$\mathscr{E}^s(T) = \{(HV, V, \Omega) : V \in \mathscr{P}, \Omega \in \mathscr{W}\},$$

where $\mathscr{P}$ (resp. $\mathscr{W}$) is the set of all matrices of form (1) (resp. (3)).

The matrices G and H are constructed as in § 16. The condition that $\mathscr{E}^u(\Phi)$ be invariant under T takes the form

$$\Gamma f_*(A + BG) = f_*(P + QG + R\Gamma)$$

which has the explicit solution:

$$\Gamma = \sum_{n=0}^{\infty} (f_* L)^n (f_*^{n+1} \Gamma_0)$$

where $\Gamma_0 = (P + QG)(A + BG)^{-1}$, $L(\Gamma) = R\Gamma(A + BG)^{-1}$. The convergence of this series follows from the estimates:

$$\| f_* \Omega \|_r \leq \lambda^{-2} \| \Omega \|_r$$

(when $e_2\Omega = 0$) and the dynamical product estimate of (17.4). The tame hyperbolic estimates follow from (18.3) and the dynamical product estimate of (17.4). The other estimates are easy.

COROLLARY. *Let F denote the space of all diffeomorphisms f; i.e. $f \in F$ iff for some $\lambda \in (1, \infty)$ we have $f(x, y) = (\lambda x, \mu y)$, ($\mu$ is fixed). Fix $f_0 \in \Gamma$ and let $N \subset \mathbb{R}^2$ satisfy*

$$\mathrm{cl}\,(N \cup f_0^{-1}(N)) \subset Z.$$

Then the unfolding F is preversal at N, i.e. there is a neighbourhood $\mathcal{N}$ of $(0, f_0)$ in $\mathcal{L}(\Theta) \times F$, a tame family J and a continuous family δ, $J: \mathcal{N} \to L(\mathcal{X}(\Theta))$, $\delta: \mathcal{N} \to L(\mathcal{X}(\Theta), \mathbb{R})$, such that

$$(1 - (1 + \Phi)f_\#)\zeta \,|\, N = w \,|\, N$$

if $\zeta = J(\Phi, f)w$, $\delta(\Phi, f)w = 0$, $(\Phi, f) \in \mathcal{N}$.

Proof. Take $w = w^s + w^u \in \mathscr{C}^u(\Phi) \oplus \mathscr{C}^s(\Phi)$, $J(\Phi, f)w = \zeta = \zeta^u + \zeta^s$, where

$$\zeta^u = -\sum_{n=1}^{\infty} T(\Phi)^{-n} w^u,$$

$$\zeta^s = \sum_n T(\Phi)^n w^s.$$

(The constructions of the theorem depend continuously on λ.)

21. Another $\mathbb{N}$-jet semi-Finsler

The $\mathbb{N}$-jet semi-Finsler of § 15 was designed to be used at a saddle point. Now we present one for use at a sink or source.

(21.1) In the notation of § 14 the $\mathbb{N}$-jet semi-Finsler on $\mathbb{R}^2$ is determined by

$$\Theta_{00}(x, y) = 1 + (|x| + |y|)^\sigma |y|^\rho$$
$$\Theta_{i0}(x, y) = |x|^{i-\sigma} + \theta(x/y)|y|^{i-\sigma}$$
$$\Theta_{ij}(x, y) = |x|^{i+\rho-\sigma}|y|^{j-\rho} + \theta(x/y)|y|^{i+j-\sigma}$$

for $i, j \geq 1$ and $(x, y) \in \mathbb{R}^2$. Here $0 < \rho < \sigma < 1$ and $\theta: \mathbb{R} \to [0, 1]$ is a C^∞ cut-off function with $\theta(0) > 0$, $\mathrm{supp}\,(\theta) \subset [-e^{-1}, e^{-1}]$, $|a| \leq b \Rightarrow \theta(b) \leq \theta(a)$, (and $e = 2.71 \ldots$). Note the homogeneity property

$$\Theta_{ij}(\tau x, \tau y) = \tau^{i+j-\sigma} \Theta_{ij}(x, y)$$

for $(i, j) \neq (0, 0)$ and $\tau \in (0, \infty)$.

PROPOSITION 1. *If $u \in \mathscr{B}(\Theta)$, then u is continuous on $\mathbb{R}^2$, and C^∞ on $y \neq 0$. Its restriction to $y = 0$, $x \neq 0$ is also C^∞, and for $j \geq 1$ the mixed partials $\partial_1^i \partial_2^j u$ blow up at worst like $|y|^{\sigma-j}$ as $(x, y) \to (x_0, 0)$ (where $x_0 \neq 0$).*

PROPOSITION 2. *Given i, j there is a continuous function $c = c_{ij}: \mathbb{R}^2 \to \mathbb{R}$ such that*

$$\Theta_{ij}(x, y) \leq c(x, y)\Theta_{i'j'}(x, y)\Theta_{i''j''}(x, y),$$

for $i = i' + i''$, $j = j' + j''$. (Hence the proposition of (17.1) applies.)

Proof. The left side is homogeneous of degree $i + j - \sigma$ and the right side is homogeneous of degree $i + j - 2\sigma$ so it is enough to define c on $x^2 + y^2 = 1$. The

functions Θ_{ij} are continuous on this circle and vanish only at $(x, y) = (\pm 1, 0)$; hence it is enough to construct c near those two points. But $\Theta_{ij}(\pm 1, y) = |y|^{j-\rho}$ and $\Theta_{i0}(1, y) = 1$ so the construction is obvious.

PROPOSITION 3. *The $\mathbb{N}$-normed space admits a family of smoothing operators with offset 0 and hence satisfies the interpolation inequalities.*

Proof. This is much like the analogous result for example 2 of § 17 but more complicated. Let

$$H_{\pm} = \{(x, y) : \pm y \geq 0\}$$
$$\mathcal{B}(H_{\pm}, \Theta) = \{u \in \mathcal{B}(\Theta) : \text{supp}\,(u) \subset H_{\pm}\}$$
$$\mathcal{B}(\mathbb{R}, \Theta) = \{u \in \mathcal{B}(\Theta) : u(x, y) = u(x, 0) \quad \text{for all } x, y\}.$$

Then we have a direct sum decomposition

$$\mathcal{B}(\Theta) = \mathcal{B}(H_{+}, \Theta) \oplus \mathcal{B}(\mathbb{R}, \Theta) \oplus \mathcal{B}(H_{-}, \Theta),$$

so we show that each summand admits smoothing operators. The middle summand is example 2 of § 17 and the other two are isomorphic so it suffices to construct an isomorphism $\mathcal{B}(H, \Theta) \to \mathcal{B}(\mathbb{R}^2) : u \to v$, (where we have abbreviated $H = H_{+}$). For this let $\psi : \mathbb{R}^2 \to H$ be a diffeomorphism of form:

$$\psi(s, t) = (e^s \xi(t), e^s \eta(t)) = (x, y)$$

where $\mathbb{R} \to H : t \to (\xi(t), \eta(t))$ is a C^{∞} embedding parameterizing a smooth convex curve from $(-1, 0)$ to $(1, 0)$. We assume

$$\xi(t) = \text{sgn}\,(t), \qquad \eta(t) = e^{-|t|},$$

for $|t| \geq 1$ where $\text{sgn}\,(t) = \pm 1$ for $\pm t > 0$.

The desired isomorphism $\mathcal{B}(H, \Theta) \to \mathcal{B}(\mathbb{R}^2) : u \to v$ is given by:

$$v(s, t) = u(e^s \xi(t), e^s \eta(t)) e^{-\sigma s + \rho |t|}.$$

We must prove:

$$c^{-1} |v|_r \leq \|u\|_r \leq c |v|_r$$

for suitable constant c. Differentiating gives

$$\partial_1^a \partial_2^b v(s, t) = \sum (\partial_1^i \partial_2^j u)(x, y)\, e^{(i+j-\sigma)s + \rho|t|} A(t) B(t), \tag{1}$$

where $1 \leq i + j \leq a + b$ for each term in the sum, and

$$A(t) = \sum \alpha \xi(t)^{i_0} (d\xi(t))^{i_1} \ldots,$$
$$B(t) = \sum \beta \eta(t)^{j_0} (d\xi(t))^{j_1} \ldots,$$

where the coefficients α, β are expressions in $a, b, i, j, \rho, \sigma, i_0, i_1, \ldots, j_0, j_1, \ldots$, and for each term which actually appears we have $1 \leq i + j \leq a + b$, $i_0 + (i_1 + 2i_2 + 3i_3 + \cdots) \leq i$, $j_0 + (j_1 + 2j_2 + 3j_3 + \cdots) \leq j$. For $|t| \leq 1$ (1) gives:

$$|\partial_1^a \partial_2^b v(s, t)| \leq \sum \|u\|_{ij} S_{ij},$$

where $S_{ij} = \text{supp}_{|t| \leq 1} |e^{\rho|t|} A(t) B(t)|$, while for $|t| \geq 1$ (1) simplifies to:

$$\partial_1^a \partial_2^b v(s, t) = \sum \gamma\, \partial_1^i \partial_2^j u(x, y) |x|^{i+j-\sigma} |y|^{j-\rho}, \tag{2}$$

where $\gamma = \gamma(a, b, i, j, \sigma, \rho)$. In either case we obtain

$$|\partial_1^a \partial_2^b v(s, t)| \le c\|u\|_r$$

for $a + b \le r$: i.e. $|v|_r \le c\|u\|_r$.

For the reverse inequality note that we have an inequality

$$|(\Theta_{ij} \partial_1^i \partial_2^j u)(x, y)| \le c \sum_{a+b \le i+j} |\partial_1^a \partial_2^b v(s, t)|, \tag{3}$$

for $|t| \le 1$ and $s = 0$ by continuity and hence for $|t| \le 1$ and all s by homogeneity. For $|t| \ge 1$ we have:

$$u(x, y) = v(s, t)|x|^{i+\rho-\sigma}|y|^{j-\rho}$$

with $s = \log|x|$, $|t| = \log|x| - \log y$, so that we again obtain (3) by differentiation. This completes the proof.

(21.2) Now fix $\mu, \lambda_1, \lambda_2 \in \mathbb{R}$ satisfying $0 < \mu < \lambda_1 < \lambda_2 < 1$ and define an unfolding $F \subset \text{Diff}(\mathbb{R}^2)$ by

$$f \in F \Leftrightarrow f(x, y) = (\lambda x, \mu y)$$

for some $\lambda \in (\lambda_1, \lambda_2)$.

PROPOSITION. *Let $\psi \in C_0^\infty(\mathbb{R}^2)$ be a cut-off function. Then the map*

$$F \times \mathcal{B}(\Theta) \to \mathcal{B}(\Theta)$$

$$(f, u) \to \psi f^* u$$

is tame-C^r of degree r for each $r \in \mathbb{N}$.

Proof. As in proposition 4 of (18.2) put

$$v(x, y) = \psi(x, y)x^k(\partial_1^k u)(\lambda x, \mu y)$$

for $\lambda \in (\lambda_1, \lambda_2)$, $k \le r$. We must prove

$$\|v\|_l \le c\|u\|_{r+l}.$$

On the set $y \ne 0$, Θ is an $\mathbb{N}$-jet Finsler and near the set $y = 0$, $x \ne 0$, Θ is locally equivalent to the $\mathbb{N}$-jet semi-Finsler of § 18 so we shall assume $\psi \equiv 1$ and prove an estimate

$$\Theta_{ij}(x, y)|\partial_1^i \partial_2^j v(x, y)| \le c\|u\|_{r+l}$$

for (x, y) near $(0, 0)$, $i + j \le l$. We have the explicit formula

$$\partial_1^i \partial_2^j v(x, y) = \sum_{p+q=i} bx^{k-p}(\partial_1^{k+q} \partial_2^j u)(\lambda x, \mu y)$$

where $b = b(i, k, p)\lambda^q \mu^j$ so it suffices to prove

$$|\Theta_{ij}(x, y)x^{k-p}(\partial_1^{k+q} \partial_2^j u)(\lambda x, \mu y)| \le c\|u\|_{k+i+j}$$

for (x, y) near $(0, 0)$ and $p + q = i$. This in turn follows from the inequality

$$\Theta_{ij}(x, y)x^{k-p} \le c\Theta_{k+q,j}(\lambda x, \mu y).$$

Both sides are homogeneous of degree $(i+j-\sigma) + (k-p) = k+q+j-\sigma$ so it is enough to prove the inequality on the unit circle. There $\Theta_{k+q,j}$ vanishes only when $y = 0$ so it is enough to check it for $x = \pm 1$. But then the left side is $y^{j-\rho}$ and the right side is $(\mu y)^{j-\rho}$ as required.

For continuity note that for a suitable positive constant $c = c(i, j)$ we have

$$x\Theta_{ij}(x, y) \le c\Theta_{i+1,j}(\lambda x, \mu y)$$

for all $(x, y) \in \mathbb{R}^2$ and $\lambda \in (\lambda_1, \lambda_2)$. (This is because both sides are homogeneous of degree $i + 1 + j - \sigma$ and on the unit circle, the right side vanishes only if $j \ge 1$ and $(x, y) = (\pm 1, 0)$. But for $x = \pm 1$, $j \ge 1$, and $|y|$ small the left side is $|y|^{j-\rho}$ and the right side is $|\mu y|^{j-\rho}$.) We then have

$$\Theta_{ij}(x, y)|(\partial_1^i \partial_2^j u)(\lambda x, \mu y) - (\partial_1^i \partial_2^j u)(\lambda_0 x, \mu y)|$$

$$= \left| \int_{\lambda_0}^{\lambda} \Theta_{ij}(x, y) x (\partial_1^{i+1} \partial_2^j u)(\tilde{\lambda} x, \mu y) \, d\tilde{\lambda} \right|$$

$$\le \int_{\lambda_0}^{\lambda} c\Theta_{i+1,j}(\tilde{\lambda} x, \mu y)|(\partial_1^{i+1} \partial_2^j u)(\tilde{\lambda} x, \mu g)| \, d\tilde{\lambda}$$

$$\le \int_{\lambda_0}^{\lambda} c\|u\|_{i+1+j} \, d\tilde{\lambda}$$

$$= c\|u\|_{i+1+j}|\lambda - \lambda_0|.$$

As in proposition 4 of § 14 this gives the Lipschitz estimate

$$\|f^* u - f_0^* u\|_r \le c|\lambda - \lambda_0| \, \|u\|_{r+1}$$

for suitable $c = c(r)$ and hence the required continuity.

(21.3) PROPOSITION. *Let N be a bounded open set in $\mathbb{R}^2$. Then for suitable ρ, σ the unfolding F is preversal at N in the following sense: for $f_0 \in F$ there is a neighbourhood $\mathcal{N} \subset \mathcal{L}(\Theta) \times F$ of $(0, f_0)$ and a tame family $J : \mathcal{N} \to L(\mathcal{X}(\Theta))$ such that*

$$\{\zeta - (1 + \Phi)f_\# \zeta\}|N = w|N$$

for $(\Phi, f) \in \mathcal{N}$, $\zeta = J(\Phi, f)w$, $w \in \mathcal{X}(\Theta)$. Here as usual

$$(f_\# w)(x, y) = (\lambda u(\lambda^{-1} x, \mu^{-1} y), \mu v(\lambda^{-1} x, \mu^{-1} y))$$

for $w = (u, v)$ and $(x, y) \in \mathbb{R}^2$.

Proof. Choose a bounded open set $Z \subset \mathbb{R}^2$ satisfying closure$(N \cup f^{-1}(N)) \subset Z$ for $f \in F$ and choose $\psi \in C_0^\infty(Z)$ with $\psi|N \cup f^{-1}(N) \equiv 1$. By the remark in § 11 we may replace Φ by $\psi\Phi$. Hence we may replace Θ by a locally equivalent (i.e. equivalent on Z) $\mathbb{N}$-jet semi-Finsler. We do this by redefining Θ_{00} so that $\Theta_{00} \equiv 1$.

LEMMA. *This Θ satisfies (4) of (17.3), in fact:*

$$\lambda^{-i}\mu^{-j}\Theta_{ij}(x, y) \le \mu^{-\sigma}\Theta_{ij}(\lambda^{-1} x, \mu^{-1} y).$$

Proof. This follows from

$$\lambda^{-i}\mu^{-j} \le \mu^{-\sigma}\lambda^{-i-\rho+\sigma}\mu^{-j+\rho},$$

(i.e. $1 \le (\mu/\lambda)^{\rho-\sigma}$ which follows from $\mu \le \lambda$ and $\rho \le \sigma$), and

$$\lambda^{-i}\mu^{-j}\theta(x/y) \le \mu^{-\sigma}\mu^{-i-j+\sigma}\theta(\lambda^{-1}x/\mu^{-1}y),$$

which follows from $|\lambda^{-1}x/\mu^{-1}y| \le |x/y|$ and the monotonicity property of the cut-off function θ.

By proposition 2 and (17.1) $\mathcal{B}(\Theta)$ is a 'tame algebra'. By (17.3) $\|f_* u\|_r \le \mu^{-\sigma} \|u\|_r$ for $u \in \mathcal{B}(\Theta)$ and $r \in \mathbb{N}$. Hence

$$\|f_\# w\|_r \le \lambda \mu^{-\sigma} \|w\|_r$$

for $w \in \mathscr{X}(\Theta)$. For σ sufficiently small the coefficient $\lambda\mu^{-\sigma}$ is less than one; hence we may choose $\tau < 1$ and $\mathcal{N}$ so that

$$\|T(\Phi, f)w\|_0 \le \tau \|w\|_0$$

for $(\Phi, f) \in \mathcal{N}$ where

$$T(\Phi, f) = (1 + \psi\Phi)f_\#.$$

By the dynamical product estimate of (17.4) we have

$$\|T(\Phi, f)^n w\|_r \le c\tau^n \|\Phi : w\|_r,$$

so that we may define J by the Neuman series

$$J(\Phi, f)w = \sum_{n=0}^{\infty} T(\Phi, f)^n w.$$

22. *Admissible diffeomorphisms*

(22.1) Let M denote a torus of dimension two. More precisely, we suppose $M = M_0 \cup M_1 \cup M_2 \cup M_3$ and that there are four diffeomorphisms (coordinate systems)

$$z_i = (x_i, y_i) : M_i \to \mathbb{R}^2 \qquad i = 0, \ldots, 3,$$

related by

$$y_0 = y_1, \quad y_2 = y_3, \quad y_1 = y_2^{-\tau},$$
$$x_0 = x_1^{-1}, \quad x_2 = x_3^{-1}, \quad x_1 = x_2,$$

where $\tau \in (0, \infty)$. Fix positive numbers $\mu_i, \alpha_i, \beta_i, i = 1, 2$, with:

$0 < \mu_1 < 1 < \mu_2$;

$\mu_1 = \mu_2^{-\tau}, \quad \alpha_1 = \alpha_2^{-\tau}, \quad \beta_1 = \beta_2^{-\tau}$;

$\alpha_1 < \mu_1^{-1}\alpha_1 < \beta_1 < \mu_1^{-2}\alpha_1$;

$\alpha_2 < \mu_2\alpha_2 < \beta_2 < \mu_2^2\alpha_2$.

(The last line is redundant.) Call $f \in \mathrm{Diff}(M)$ *admissible* iff there exist smooth functions $a_i = a_{if} : \mathbb{R} \to (0, \infty)$, $b_i = b_{if} : \mathbb{R} \to \mathbb{R}$, $i = 1, 2$, such that:

$0 < a_1(0)^{-1} < \mu_1 < 1$;

$0 < a_2(0) < \mu_2^{-1} < 1$;

$f^*x_i = a_i(y_i)x_i + b_i(y_i)$;

$f^*y_i = \mu_i y_i$;

$|y_i| < \alpha_i \Rightarrow \begin{cases} a_i(y_i) = a_i(0), \\ b_i(y_i) = 0. \end{cases}$

Call an unfolding $F \subset \mathrm{Diff}(M)$ *admissible* iff each $f \in F$ is admissible and if the maps

$$F \times \mathbb{R} \to (0, \infty) : (f, y_i) \to a_{if}(y_i)$$
$$F \times \mathbb{R} \to \mathbb{R} : (f, y_i) \to b_{if}(y_i)$$

are smooth.

We introduce the following notation:

$$\mu_0 = \mu_1, \quad \mu_3 = \mu_2;$$
$$\lambda_i(f) = a_{if}(0), \quad i = 1, 2;$$
$$\lambda_0(f) = \lambda_1(f)^{-1}, \quad \lambda_3(f) = \lambda_2(f)^{-1};$$

so that for $|x_i| + |y_i|$ sufficiently small and $i = 0, \ldots, 3$ we have

$$f^*(x_i, y_i) = (\lambda_i(f)x_i, \mu_i y_i).$$

(22.2) Now let $\Theta^1 = \Theta^2$ (resp. $\Theta^0 = \Theta^3$) be the $\mathbb{N}$-jet semi-Finsler on $\mathbb{R}^2$ defined in § 18 (resp. § 21). From the formulae one easily checks

$$z_i^*\Theta^i | M_i \cap M_j \sim z_j^*\Theta^j | M_i \cap M_j,$$

where $\sim$ denotes local equivalences of $\mathbb{N}$-jet semi-Finslers so by the patching construction we have an $\mathbb{N}$-jet semi-Finsler Θ on M with

$$\Theta | M_i \sim z_i^*\Theta^i | M_i.$$

Write $\mathscr{F}(M) = \mathscr{F}(M, \Theta)$, $\mathscr{X}(M) = \mathscr{F}(M) \otimes TM$, $\mathscr{L}(M) = \mathscr{F}(M) \otimes L(TM)$, and note that by § 18 and § 21 the space $\mathscr{F}(M)$ is adapted to any admissible unfolding.

(22.3) Given an admissible diffeomorphism f we shall now define a level decomposition

$$R_0, S_1, R_1, S_2, R_2, S_3, R_3$$

which respects f. Take

$$S_2 = \{\alpha_1 < |y_1| < \beta_1\} \cup \{\alpha_1 < |y_0| < \beta_1\}$$

so that

$$S_2 = \{\alpha_2 < |y_2| < \beta_2\} \cup \{\alpha_2 < |y_3| < \beta_2\},$$

(note that $y_1 | M_1 \cap M_0 = y_0 | M_1 \cap M_0$ etc.). Next choose $\alpha_{ki}, \beta_{ki} > 0$, $i = 0, 3$, $k = 1, 2$, satisfying

$$\mu_0^{-2}\beta_{20} < \alpha_1;$$
$$\mu_3^2\beta_{30} < \alpha_2;$$
$$\alpha_{10} < \lambda_0(f)^{-1}\alpha_{10} < \beta_{10} < \lambda_0(f)^{-2}\alpha_{10};$$
$$\alpha_{20} < \mu_0^{-1}\alpha_{20} < \beta_{20} < \mu_0^{-2}\alpha_{20};$$
$$\alpha_{13} < \lambda_3(f)\alpha_{13} < \beta_{13} < \lambda_3(f)^2\alpha_{13};$$
$$\alpha_{23} < \mu_2\alpha_{23} < \beta_{23} < \mu_2^2\alpha_{23};$$

and define:

$$S_1 = \{\alpha_{10} < |x_0| < \beta_{10}\} \cap \{\alpha_{20} < |y_0| < \beta_{20}\},$$
$$S_3 = \{\alpha_{13} < |x_3| < \beta_{13}\} \cap \{\alpha_{23} < |y_3| < \beta_{23}\}.$$

Thus

$$f^n(\text{cl}(S_i)) \cap \text{cl}(S_i) = \varnothing$$

for $i = 1, 2, 3$ and $|n| > 1$. Now choose $\varepsilon > 0$ and put

$$R_i = \{|x_i| < \alpha_{1i} + \varepsilon, |y_i| < \alpha_{2i} + \varepsilon\}$$

for $i = 0, 3$. For $i = 1, 2$ we define R_i by:

$$R_1 = \{\alpha_{20} - \varepsilon < |y_0| < \alpha_1 + \varepsilon\} \cup \{|x_1| < \alpha_{10}^{-1} + \varepsilon, |y_1| < \alpha_1 + \varepsilon\}$$

$$R_2 = \{\alpha_2 - \varepsilon < |y_3| < \alpha_2 + \varepsilon\} \cup \{|x_2| < \alpha_{13}^{-1} + \varepsilon, |y_2| < \alpha_2 + \varepsilon\}.$$

When $\varepsilon = 0$ the sets $R_0, S_1, \ldots, R_3$ are pairwise disjoint, while for $\varepsilon > 0$ they cover M. Hence for sufficiently small ε they give a level decomposition. Note that for this level decomposition, $\Omega_i(f)$ is defined by

$$\Omega_i(f) = \{z_i^{-1}(0, 0)\}.$$

Thus $\Omega_0(f)$ is a sink, $\Omega_3(f)$ is a source and $\Omega_1(f), \Omega_2(f)$ are saddles. Also note that

$W^s(\Omega_0, f) = M_0;$

$W^u(\Omega_1, f) = \{y_1 = 0\};$

$W^s(\Omega_2, f) = \{y_2 = 0\};$

$W^u(\Omega_3, f) = M_3;$

$W^s(\Omega_1, f) \cap \{|y_1| < \alpha\} = \{x_1 = 0, |y_1| < \alpha_1\};$

$W^u(\Omega_2, f) \cap \{|y_2| < \alpha_2\} = \{x_2 = 0, |y_2| < \alpha_2\};$

so that these sets are independent of f. By § 20 and (21.3) we have:

PROPOSITION. *Let F be an admissible unfolding. Then this level decomposition (for $f = f_0 = $ the centre of F) is locally preversal. The linear functionals $\delta_i: \mathcal{N} \to L(\mathcal{X}(M), \mathbb{E}_i)$ of the definition are determined by*:

$$\mathbb{E}_0 = \mathbb{E}_3 = \{0\}, \quad \mathbb{E}_1 = \mathbb{E}_2 = \mathbb{R}, \quad \delta_i(0, f)w = \partial_1 u_i(0, 0),$$

for $w \in \mathcal{X}(M)$, $i = 1, 2$ where $z_{i\#} w = (u_i, v_i)$.

Remark. The definitions have been arranged so that

$$\left. \begin{array}{r} w \in f_* T_f F \\ \delta_1 \cdot w = \delta_2 \cdot w = 0 \end{array} \right\} \Rightarrow \mathrm{supp}\,(w) \subset S_2$$

for $(\Phi, f) \in \mathcal{N} \subset \mathcal{L}(M) \times F$, where $\delta_i w = \delta_i(\Phi, f)w$.

23. *Universal admissible unfoldings*

(23.1) We continue the notation of § 22 but we denote the four coordinate systems z_i, $i = 0, \ldots, 3$, defined there by $z_{0i} = (x_{0i}, y_{0i})$. For $i = 1, 2$ we call a coordinate system $z_i = (x_i, y_i): M_i \to \mathbb{R}^2$ admissible if

$$z_i(p) = z_{i0}(p) \qquad \text{for } |y_{i0}(p)| < \alpha_i,$$

and

$$f^*(x_i, y_i) = (\lambda_i(f)x_i, \mu_i y_i)$$

(on all of M_i) for some admissible diffeomorphism f. This gives a bijective correspondence $f \leftrightarrow (z_1, z_2)$ between admissible diffeomorphisms and certain pairs of admissible coordinate systems as z_1, z_2 are determined uniquely by f:

$$x_1(p) = \lim_{n \to \infty} \lambda_1(f)^{-n} x_{10}(f^n(p)),$$

$$x_2(p) = \lim_{n \to \infty} \lambda_2(f)^n x_{20}(f^{-n}(p)),$$

$$y_i = y_{i0}, \qquad i = 1, 2,$$

where the limits exist trivially as the argument of each limit is independent of $n > n(p)$. Note that the coordinate systems z_1, z_2 are related by the formulae

$$x_2 = a(y_1)x_1 + b(y_1)$$
$$y_2 = c(y_1) \tag{1}$$

where $c(y) = y^{-\tau}$ and $a: \mathbb{R} \setminus \{0\} \to (0, \infty)$ and $b: \mathbb{R} \setminus \{0\} \to \mathbb{R}$ are smooth and satisfy $a(\mu_1 y)\lambda_1 = \lambda_2 a(y)$, $b(\mu_1 y) = \lambda_2 b(y)$. (These equations determine the set of pairs (z_1, z_2) which correspond to some f.)

(23.2) In these coordinates we have $W^s(\Omega_1, f) = \{x_1 = 0\}$, $W^u(\Omega_2, f) = \{x_2 = 0\}$, so that if we define

$$W^{us}(f) = W^s(\Omega_1, f) \cap W^u(\Omega_2, f),$$

we obtain

$$W^{us}(f) = \{x_1 = 0, b = 0\}.$$

We therefore define for $p \in W^{us}(f)$ the *order of intersection* at p by

$$(d^k b)(y_1(p)) = 0 \qquad k \le \text{ord}\,(p, f),$$
$$(d^k b)(y_1(p)) \ne 0 \qquad k = \text{ord}\,(p, f) + 1.$$

Thus p is a point of transversal intersection if and only if $(b(y_1(p)) = 0$ and) $\text{ord}\,(p, f) = 0$.

We say that f has *finite codimension* iff

$$p \in W^{us}(f) \Rightarrow \text{ord}\,(p, f) < \infty.$$

As $\text{ord}\,(p, f) < \infty$ implies that the zero of b at $y_1(p)$ is isolated, it follows that (when f has finite codimension) the set $W^{us}(f)$ consists of finitely many orbits of f. We may adjust the interface set S_2 so that each orbit of $W^{us}(f)$ intersects S_2 in exactly one point and enumerate:

$$W^{us}(f) \cap S_2 = \{p_1, p_2, \dots, p_l\}.$$

We then define the *codimension* of f by

$$\text{codim}\,(f) = 2 + 2 \sum_{j=1}^{l} \text{ord}\,(p_j, f).$$

(23.3) THEOREM. *Let F be an admissible unfolding whose centre f_0 has finite codimension. Then F is preversal. In fact, one can choose $\omega: \mathcal{N} \to L(\mathcal{X}(M), \mathbb{F})$ as in the definition so that*

$$\dim\,(\mathbb{F}) = \text{codim}\,(f_0).$$

Proof. We have already constructed a locally preversal level decomposition so it remains only to check the interface condition. Thus we must construct

$$\omega_i: \mathcal{N} \to L(\mathcal{X}(S_i)_0, \mathbb{F}_i)$$
$$\pi_i^u, \pi_i^s: \mathcal{N} \to L(\mathcal{X}(S_i)_0)$$

as in the definition for $i = 1, 2, 3$. For $i = 1, 3$ there is nothing to do; we take $\mathbb{F}_1 = \mathbb{F}_3 = \{0\}$, $w = \pi_1^s(\Phi, f)w$, $w = \pi_3^u(\Phi, f)w$. Hence we drop the subscript 2: $S = S_2$, $\mathbb{F} = \mathbb{F}_2$, and construct $\omega: \mathcal{N} \to L(\mathcal{X}(S)_0, \mathbb{F})$ $\pi^s, \pi^u: \mathcal{N} \to L(\mathcal{X}(S)_0)$.

Now let

$$W^{us}(f_0) \cap S = \{p_1, p_2, \ldots, p_l\}$$

and write

$$S = P_1 \cup P_2 \cup \cdots \cup P_l \cup P^u \cup P^s$$

as a union of open sets with

$W^{us}(f) \cap S \subset P_1 \cup P_2 \cup \cdots \cup P_l;$

$W^s(\Omega_1, f) \cap \mathrm{cl}\,(P^s) = \varnothing;$

$W^u(\Omega_2, f) \cap \mathrm{cl}\,(P^u) = \varnothing;$

$\mathrm{cl}\,(P_j) \cap \mathrm{cl}\,(P_k) = \varnothing, \quad j \neq k;$

for $(\Phi, f) \in \mathcal{N}$. We will define

$$\varepsilon_j \colon \mathcal{N} \to [(\mathcal{X}(S)_0, \mathbb{E}_j),$$

$$\pi_j^u, \pi_j^s \colon \mathcal{N} \to L(\mathcal{X}(S)_0),$$

for $j = 1, \ldots, l$, such that

$$w|P_j = \{\pi_j^s(\Phi, f)w + \pi_j^u(\Phi, f)w\}|P_j$$

when $\varepsilon_j(\Phi, f)w = 0$ and the appropriate tame asymptotic estimates hold. We then write

$$w - \sum_{j=1}^{l} w_j = w^u + w^s$$

where:

$w_j = \psi_j\{\pi_j^s(\Phi, f)w + \pi_j^u(\Phi, f)w\};$

$\mathrm{supp}\,(w^u) \subset P^u;$

$\mathrm{supp}\,(w^s) \subset P^s;$

$\psi_j|P_j \equiv 1;$

$\mathrm{supp}\,(\psi_j) \cap \mathrm{cl}\,(P_k) = \varnothing, \quad j \neq k,$

and define

$$\pi^s(\Phi, f) = w^s + \sum_{j=1}^{l} \psi_j \pi_j^s(\Phi, f)w,$$

$$\pi^u(\Phi, f) = w^u + \sum_{j=1}^{l} \psi_j \pi_j^u(\Phi, f)w.$$

Since $f^n(\mathrm{cl}\,(P^s)) \subset R_0, f^{-n}(\mathrm{cl}\,(P^u)) \subset R_3$, for sufficiently large n we have the tame asymptotic estimates on w^s and w^u by § 21. Hence it suffices to construct $\varepsilon_j, \pi_j^s, \pi_j^u$. Thus we drop the subscript j and construct $\varepsilon \colon \mathcal{N} \to L(\mathcal{X}(S)_0, \mathbb{F}), \pi^u, \pi^s \colon \mathcal{N} \to L(\mathcal{X}(S)_0)$, satisfying

$$w|P = \{\pi^s(\Phi, f)w + \pi^u(\Phi, f)w\}|P$$

when $\varepsilon(\Phi, f)w = 0$, where P is a small neighbourhood of $p \in W^{us}(f_0) \cap S$ in $S = S_2$.

By § 20 we may construct (using admissible coordinates and appropriate cut-off functions, an $\mathbb{N}$-normed space $\mathcal{X}(M_i)$, $i = 1, 2$, of vector fields on M_i and for

$(\Phi, f) \in \mathcal{N} \subset \mathcal{L}(M) \times F$ an operator

$$T_i(\Phi, f): \mathcal{X}(M_i) \to \mathcal{X}(M_i)$$

and a $T_i(\Phi, f)$-invariant splitting

$$\mathcal{X}(M_i): \mathcal{E}_i^u(\Phi, f) \oplus \mathcal{E}_i^s(\Phi, f)$$

such that for $w \in \mathcal{X}(M_i)$

$$T_i(\Phi, f) w | Z_i = T(\Phi, f) w | Z_i.$$

Here $Z_i \subset M_i$ is open and we may assume $\mathrm{cl}\,(P) \subset Z_1 \cup Z_2$ and that $P^{f+} \subset Z_1^{f+}$, $P^{f-} \subset Z_2^{f-}$. Now choose $w \in \mathcal{X}(P)_0$ and define (u_i, v_i), $i = 1, 2$ by $z_{i\#} w = (u_i, v_i)$. Write

$$u_i(x, y) = u_{i0}(y) + u_{i1}(y)x + u_{i2}(x, y),$$

$$v_i(x, y) = v_{i0}(y) + v_{i1}(y)x + v_{i2}(x, y),$$

where $e_2 u_{i2} = u_{i2}$, $e_2 v_{i2} = v_{i2}$ and note that $\tilde{w} \in \mathcal{X}(S)_0$ defined by $z_{2\#} \tilde{w} = (u_{22}, v_{22})$ satisfies $\tilde{w} \in \mathcal{E}_2^u(\Phi, f)$ and hence satisfies appropriate backwards tame estimates. Hence put

$$\pi^u(\Phi, f) w = \pi^u(\Phi, f)(w - \tilde{w}) + \tilde{w}$$

and assume without loss of generality that $\tilde{w} = 0$; i.e. that

$$u_2(x, y) = u_{20}(y) + u_{21}(y)x$$
$$v_2(x, y) = v_{20}(y) + v_{21}(y)x. \tag{2}$$

It follows that

$$u_1(x, y) = u_{10}(y) + u_{11}(y)x + u_{12}(y)x^2,$$
$$v_1(x, y) = v_{10}(y) + v_{11}(y)x. \tag{3}$$

Now by (1) (u_1, v_1) and (u_2, v_2) are related by formulae

$$u_2(ax + b) = au_1(x) + (a'x + b')v_1(x)$$
$$v_2(ax + b) = c'v_1(x)$$

where we have suppressed the dependence on y. Formulae (2) and (3) yield the relations:

$$au_{10} = u_{20} + bu_{21} - b'v_{20}$$
$$au_{11} = \quad\quad au_{21} - a'v_{20} - b'v_{21}$$
$$au_{12} = \quad\quad a'v_{21}$$
$$c'v_{10} = v_{20} + bv_{21}$$
$$c'v_{11} = \quad\quad av_{21}$$
$$c'v_{12} = 0$$

Now we shall attempt to solve:

$$w = \zeta^s + \zeta^u, \tag{4}$$

$$z_{1\#} \zeta^s \in \mathcal{E}_1^s(\Phi, f), \tag{5}$$

$$z_{2\#} \zeta^u \in \mathcal{E}_2^u(\Phi, f). \tag{6}$$

We assume ζ^u has form

$$z_{2\#} \zeta^u = (\xi_2^u, \eta_2^u),$$

with

$$\xi_2^u(x) = \xi_{20}^u + \xi_{21}^u x, \qquad \eta_2^u(x) = \eta_{20}^u + \eta_{21}^u x,$$

so that for $i = 1$ we have:

$$z_{1\#}\zeta^u = (\xi_1^u, \eta_1^u)$$

$$\xi_1^u(x) = \xi_{10}^u + \xi_{11}^u x + \xi_{12}^u x^2 \qquad \eta_1^u(x) = \eta_{10}^u + \eta_{12}^u x,$$

as above. We assume ζ^s has form:

$$z_{1\#}\zeta^s = (\xi_1^s, \eta_1^s)$$

with

$$\xi_1^s(x) = \xi_{10}^s + \xi_{11}^s x + \xi_{12}^s x^2 \qquad \eta_1^s(x) = \eta_{10}^u + \eta_{11}^u x.$$

By § 19 (5) and (6) take the form

$$\xi_{10}^s = h_{10}\eta_{10}^s \qquad \xi_{11}^s = h_{11}\eta_{10}^s + h_{10}\eta_{11}^s,$$

$$\xi_{20}^u = h_{20}\eta_{20}^u \qquad \xi_{21}^u = h_{21}\eta_{20}^u + h_{20}\eta_{21},$$

for certain functions h_{i0}, h_{i1}, $i = 1, 2$, which depend tamely on (Φ, f) and vanish when $\Phi = 0$. To express (4) we rewrite our equations in matrix form:

$$U_1 = X_1^s + X_1^u \qquad V_1 = Y_1^s + Y_1^u,$$

$$X_1^s = H_1 Y_1^s \qquad X_2^u = H_2 Y_2^u$$

$$X_1^u = AX_2^u - BY_1^u \qquad Y_1^u = CY_2^u$$

where for $i = 1, 2$, $\sigma = s, u$:

$$U_1 = \begin{bmatrix} u_{10} \\ u_{11} \end{bmatrix}, \quad V_1 = \begin{bmatrix} v_{10} \\ v_{11} \end{bmatrix}, \quad X_i^\sigma = \begin{bmatrix} \xi_{i0}^\sigma \\ \xi_{i1}^\sigma \end{bmatrix}, \quad Y_i^\sigma = \begin{bmatrix} \eta_{i0}^\sigma \\ \eta_{i1}^\sigma \end{bmatrix},$$

$$H_i = \begin{bmatrix} h_{i0} & 0 \\ h_{i1} & h_{i0} \end{bmatrix}, \quad A = a^{-1}\begin{bmatrix} 1 & b \\ 0 & a \end{bmatrix},$$

$$B = -a^{-1}\begin{bmatrix} b' & 0 \\ a' & b' \end{bmatrix}, \quad C = (c')^{-1}\begin{bmatrix} 1 & b \\ 0 & a \end{bmatrix}.$$

Now if

$$Q = (B + R)Y \tag{7}$$

where $Q = a(U_1 - H_1 V_1)$, $R = H_1 - AH_2C^{-1}$, our equations are solved by $X_1^s = H_1(V_1 - Y)$, $Y_1^s = V_1 - Y$, $X_1^u = (AH_2C^{-1} - B)Y$, $Y_1^u = Y$, so we must solve (7) for Y. When $\Phi = 0$ we have $H_1 = 0$, $H_2 = 0$ so (7) takes the form

$$q_0 = b'\eta_0,$$

$$q_1 = a'\eta_0 + b'\eta_1.$$

All quantities are functions of y and we must only solve in a neighbourhood of y_0 where b' is regular of order r at y_0. (y_0 is defined by the condition that $z_1(p) = (0, y_0)$, $p \in W^{us}(f_0)$.) Define matrices $E_1, \ldots, E_{2r}$ by

$$E_i(y) = \begin{bmatrix} (y - y_0)^{r-i} \\ 0 \end{bmatrix}, \qquad E_{r+i}(y) = \begin{bmatrix} 0 \\ (y - y_0)^{r-i} \end{bmatrix},$$

for $i = 1, \ldots, r$ so that the equation

$$Q = BY + \sum_{i=1}^{2r} \varepsilon_i E_i$$

has a unique solution $Y \in C^{\infty}(\mathbb{R}) \otimes \mathbb{R}^2$, $\varepsilon_1, \ldots, \varepsilon_{2r} \in \mathbb{R}$ for any $U \in C^{\infty}(\mathbb{R}) \otimes \mathbb{R}^2$. (The loss of smoothness is $2r$.) Thus by the tame submodule perturbation theorem there is an interval I about Y_0, a tame map

$$\mathscr{B}(I) \otimes (\mathbb{R}^2 \times L(\mathbb{R}^2)) \to \mathscr{B}(I) \otimes \mathbb{R}^2$$

$$(Q, R) \to Y$$

of degree $2r$ and continuous maps

$$\mathscr{B}(I) \otimes (\mathbb{R}^2 \times L(\mathbb{R}^2)) \to \mathbb{R}$$

$$(Q, R) \to \varepsilon_i$$

such that

$$Q = (B + R) Y + \sum_{i=1}^{2r} \varepsilon_i E_i. \tag{8}$$

Our constructions give tame maps $(\Phi, f) \to R$, $(\Phi, f, w) \to Q$, $(\Phi, f, Y) \to (\zeta^s, \zeta^u)$, so that we have constructed the desired maps $(\Phi, f, w) \to (\zeta^s, \zeta^u)$ and functionals $(\Phi, f, w) \to \varepsilon_i(\Phi, f)w$. Note that (8) and $\varepsilon_i = 0$, $i = 1, \ldots, 2r$, imply (7). This completes the proof.

COROLLARY. *There exists an admissible universal unfolding F of dimension equal to the codimension of f_0. In fact, an admissible unfolding is universal provided that for any system of real numbers $c_1, c_2, a_{kj}, b_{kj} \in \mathbb{R}$, where $j = 0, 1, \ldots, r_k - 1 = \text{ord}(f_0, p_k) - 1$ and $k = 1, \ldots, l$ the equations*

$$\partial_1 u_1(0, 0) = c_1$$

$$\partial_1 u_2(0, 0) = c_2$$

$$\partial_2^j u_1(0, y_1(m_k)) = b_{kj}$$

$$\partial_2^j \partial_1 u_1(0, y_1(m_k)) = a_{kj}$$

have a solution $w \in f_* T_f F$ (where $z_{i \neq} w = (u_i, v_i)$ for $i = 1, 2$).

To compare this result with those of Palis [14] and De Melo [3] assume that $\text{ord}(f_0, p_k) = 1$ for $k = 1, \ldots, l$. A universal unfolding obviously requires l parameters to break these non-transversal intersections. When these parameters vanish Palis and De Melo give $l = 1 + (l - 1)$ additional moduli of stability. Thus apparently a universal unfolding requires at least $2l$ parameters. Our univeral unfolding has $2l + 2$ parameters in this case.

Added in proof. A recent preprint of W. de Melo and S. J. van Strien entitled 'Diffeomorphisms on surfaces with a finite number of moduli' shows that conjecture C of § 13 cannot hold in the generality stated there. In particular any diffeomorphism having a certain configuration of saddle connections has infinite codimension. Nonetheless, the techniques described in § 13 can probably still be used to construct universal unfoldings when they exist. Also the arguments in S. J. van Strien, Saddle

connections of arcs of diffeomorphisms, Dynamical systems and turbulence, Warwick 1980 Springer lecture notes in math. **898** (1981) seem to prove that a universal unfolding must have $2\ell + 2$ parameters.

REFERENCES

[1] R. Abraham & J. Robbin. *Transversal Mappings and Flows.* Benjamin: New York, 1967.

[2] C. C. Conley. *Isolated Invariant Sets and the Morse Index*; CBMS monographs **38** Amer. Math. Soc: Providence, R.I., 1978.

[3] W. DeMelo. Moduli of stability of diffeomorphisms. *Topology* **19** (1978), 9–21.

[4] M. Golubitsky & V. Guillemin. *Stable Mappings and Their Singularities.* Springer Graduate Texts in Math. # 14. Springer-Verlag: Berlin–Heidelberg–New York, 1974.

[5] R. S. Hamilton. The inverse function theorem of Nash and Moser. *Bull. Amer. Math. Soc.* **7** (1982), 65–222.

[6] M. C. Irwin. On the smoothness of the composition map. *Quart. J. Math. Oxford* (2) **23** (1972), 113–33.

[7] D. Knuth. *The Art of Computer Programming* Vol. I. Addison-Wesley: Reading, Mass., 1973.

[8] N. Kopell. Commuting diffeomorphisms. In *Proc. of Symposia in Pure Math* (Vol. XIV: Global Analysis) 165–184. Amer. Math. Soc: Providence, R.I. (1970).

[9] M. G. Lassalle. Une demonstration du theoreme de division. *Topology* **12** (1973) 41–62.

[10] R. Mañe. Absolute stability and bifurcation theory. *Invent. Math.* **53** (1979), 59–75.

[11] J. Moser. A new technique for the construction of solutions of non-linear differentiation equations. *Proc. Nat. Acad. Sci. U.S.A.*, **47** (1961), 1824–1831.

[12] J. Moser. A rapidly convergent iteration method and non-linear partial differential equations. *Ann. Scuola Norm. Sup. Pisa*, Ser. III, Vol. XXI (1966), (pp. 265–315) II (pp. 499–535).

[13] J. Nash. The embedding problem for Riemannian manifolds. *Ann. of Math.* **63** (1956), 20–63.

[14] J. Palis. A differentiable invariant of topological conjugacies and moduli of stability. *Asterisque* **51** (1978).

[15] M. Reed & B. Simon. *Methods of Modern Mathematical Physics* II, *Fourier Analysis, Self-Adjointness.* Academic Press: New York, 1975.

[16] J. Robbin. A structural stability theorem. *Ann. of Math.* **94** (1971), 447–493.

[17] C. Robinson. Structural stability of C^1 diffeomorphisms. *J. Diff. Eq.* **22** (1976), 28–73.

[18] J. Schwartz. On Nash's Implicit Function Theorem, *Comm. Pure Appl. Math.* **13** (1960), 509–530. See also his lecture notes 'Nonlinear Functional Analysis' NYU, 1964, [**18a**].

[19] F. Sergeraert. Un theoreme de fonctions implicites sur certains espace de Frechet et quelques applications. *Ann. Sci. École Norm. Sup.* (4), **5** (1972), 599–660.

[20] S. Smale. Differentiable dynamical systems. *Bull. Amer. Math. Soc.* **73** (1967), 747–817; (reprinted in Smale [**22**]).

[21] S. Smale. Notes on differentiable dynamical systems. In *Proc. of Symposia in Pure Math* (Global Analysis) XIV, pp. 277–287. Amer. Math. Soc: Providence, RI, 1970.

[22] S. Smale. *The Mathematics of Time.* Springer-Verlag: New York, 1980.

[23] S. Sternberg. Infinite Lie groups and formal aspects of dynamical systems. *J. of Math. & Mech.* **10** (1961), 451–474.

[24] C. T. C. Wall. Introduction to the preparation theorem. In *Proc. of Liverpool Singularities Symposium* I, Springer lecture notes 192, pp. 90–95. Springer-Verlag: Berlin–Heidelberg–New York, 1971.

[25] E. C. Zeeman. The classification of elementary catastrophes of codimension ≤5. (Lecture notes of a course given at Warwick in 1973). Published in: *Structural Stability, the Theory of Catastrophes, and Applications in Science*, Springer lecture notes in math. 525 (1976), pp. 263–327; reprinted in *Catastrophe Theory, Selected papers* 1972–1977, Addison-Wesley, 1977, pp. 497–561.

[26] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems, I. *Comm. Pure Appl. Math* **28** (1975), 91–140.