

JOINT CONTINUITY OF DIVISION OF SMOOTH FUNCTIONS. I: UNIFORM LOJASIEWICZ ESTIMATES

BY

MARK ALAN MOSTOW AND STEVEN SHNIDER¹

ABSTRACT. In this paper we study the question of the existence of a continuous inverse to the multiplication mapping $(f, g) \rightarrow (fg, g)$ defined on pairs of C^∞ functions on a manifold M . Obviously, restrictions must be imposed on the domain of such an inverse. This leads us to the study of a modified problem: Find an appropriate domain for the inverse of $(f, G) \rightarrow (f(p \circ G), G)$, where G is a C^∞ mapping of the manifold M into an analytic manifold N and p is a fixed analytic function on N . We prove a theorem adequate for application to the study of inverting the mapping $(A, X) \rightarrow (A, AX)$, where X is a vector valued C^∞ function and A is a square matrix valued C^∞ function on M whose determinant may vanish on a nowhere dense set.

1. Introduction. In this paper we shall consider triples of smooth functions (f, g, h) satisfying $f = g \cdot h$, and study the continuity of the quotient $h = f/g$ as a function f for fixed g and as a function of the pair (f, g) .

Let M be a smooth (C^∞), finite-dimensional, second countable, Hausdorff manifold. For $p = 0, 1, \dots, \infty$, let $E^p = E^p(M)$ be the Fréchet space of C^p real-valued functions on M , with the Fréchet C^p topology of uniform convergence of derivatives of order $\leq p$ ($< p$ if $p = \infty$) on compact sets [Tre1, p. 94]. Let $E = E^\infty$.

DEFINITION 1. If, for some fixed $g \in E$, the multiplication map $m_g: E \ni h \mapsto gh \in E$ is one-to-one and its inverse map

$$m_g^{-1}: gE \ni f = gh \mapsto h = f/g \in E$$

is continuous, we say that division by g is continuous in the numerator.

Observe that m_g is one-to-one if and only if

$$g \in E_{n.d.} = (\text{def}) \{ g \in E \mid g^{-1}(0) \subset M \text{ is nowhere dense} \}.$$

DEFINITION 2. Let m be the multiplication map

$$m: E \times E_{n.d.} \ni (h, g) \mapsto (gh, g) \in E \times E_{n.d.}.$$

If its inverse map m^{-1} : image $m \mapsto E \times E_{n.d.}$ is continuous at a point $(f_0 = g_0 h_0, g_0) \in \text{image } m$, we say that division is jointly continuous at (f_0, g_0) .

Received by the editors December 31, 1984.

1980 *Mathematics Subject Classification.* Primary 46E25; Secondary 26E10, 46E10, 46F10, 46H10.

Key words and phrases. Division of smooth functions, continuity of division, Mather Division Theorem, division of distributions, continuous dependence of solutions on parameters, matrix equations, function spaces.

¹The second author was supported in part by an NSERCC operating grant.

©1985 American Mathematical Society
0002-9947/85 \$1.00 + \$.25 per page

A simple example shows that the two types of continuity are quite different. Let $M = \mathbf{R}$ with coordinate x , and let t be a real parameter. For each fixed value of t , including 0, division by $t^2 + x^2$ is continuous (by [Hörm] or by direct calculation). But division is not jointly continuous at $(0, x^2)$, since if it were, the quotient $t/(t^2 + x^2)$ would approach $0/x^2 = 0$ in E as $t \rightarrow 0$, contradicting the fact that at $x = 0$, $t/t^2 \rightarrow \infty \neq 0$ as $t \rightarrow 0$.

The preceding example $t/(t^2 + x^2)$ shows that joint continuity can only be proven for a restricted class of denominators, which is not large enough to include the cases of interest to us. However, if we require the denominator g to be of the form $p \circ G$ where G is a C^∞ mapping of the manifold M into an analytic manifold N and p is a fixed analytic function on N , and require the quotient $h = f/p \circ G$ to depend continuously on the pair (f, G) , then we can prove a result which does cover the examples that led us to this problem.

We were led to the study of joint continuity by a particular case of the linear equation $AX = B$, where X and B are smooth vector valued functions on \mathbf{R}^m and A is a smooth $n \times n$ matrix valued function on \mathbf{R}^m . Generically, A is invertible on an open dense set U and on that set we can solve X as $A^{-1}B$. Since X is assumed continuous, its values on all of \mathbf{R}^m are determined by its values on U . But the following question arises: Can the uniform norms on X and its derivatives be estimated on all of \mathbf{R}^m from the uniform norms of A , B and their derivatives? In other words does the solution depend continuously in the Fréchet topology of function spaces on the data A , B ?

This problem could arise in many contexts. We came across it in what is called the field copy problem in gauge theory on \mathbf{R}^4 . In differential geometric terms, the problem is to determine a generic set of connections on a principal bundle over \mathbf{R}^4 which depend uniquely and continuously on their curvatures. One approach is to use the Bianchi identity $d\Omega = [\Omega, \omega]$ where $\Omega = d\omega + \frac{1}{2}[\omega, \omega]$ is the curvature of the connection ω . We can solve for ω in terms of Ω if the map $\text{ad}_\Omega: \eta \rightarrow [\Omega, \eta]$, as a map from Lie algebra valued one-forms to Lie algebra valued three-forms, is invertible. For a principal bundle with semisimple structure group over a four-dimensional base space one can find a generic set of connections for which ad_Ω is invertible on an open dense subset of the base space. Thus we are in the situation of the linear equation $AX = B$. Moreover, by Kramer's rule, when $\det A \neq 0$, solving for X involves multiplication by the transpose of the matrix of cofactors of A and division by $\det A$. Thus continuity of X as a function of A and B would follow from joint continuity in (f, G) of division with denominators of the form $p \circ G$.

In §2 we review some facts about continuity in the numerator such as equivalence to the closedness of an ideal in the ring of smooth functions and divisibility of distributions. The basic results here are due to Schwarz, Whitney, Hörmander, Łojasiewicz, Malgrange and Tougeron. On the other hand we have not found any explicit discussion of joint continuity of division.

In §3 we formulate and prove a modified joint continuity theorem along the lines of the previous remarks. The proof involves a particular choice of Whitney stratification of the zero set of p and an adaptation of Hörmander's proof [Hörm] to the case

of a varying denominator. We show that the constants appearing in his estimates vary continuously with G if the denominator is of the form $p \circ G$ with G transverse to the chosen Whitney stratification of the zero set of p . Most careful attention must be paid to the Łojasiewicz inequality estimating a function by the distance to its zero set. Our proof depends in part on a theorem of independent interest about the behavior of this inequality under pull back which the first author has proven, which will appear in part II of this paper [Mos].

2. Continuity in the numerator of division by a fixed function. The following proposition summarizes some known facts about division [Whi, Mal], distributions [Schw], and Fréchet spaces [Tre1, Tre2].

PROPOSITION 1. *For any smooth function $g: M \rightarrow \mathbf{R}$ with $g^{-1}(0)$ nowhere dense, the following eight conditions on g are equivalent.*

- (1) (*Separate continuity of division by g .*) $(m_g)^{-1}: gE \rightarrow E$ is continuous.
- (2) (*Closed ideals.*) gE is a closed ideal in E .
- (3) (*Division of distributions.*) The transpose $m'_g: E' \rightarrow E'$ ($E' =$ continuous dual of E) is onto. That is, every distribution $T \in E'$ can be written in the form gS for some (nonunique) $S \in E'$.

(4) (*Whitney's criterion.*) Given $f \in E$, if for each point $x \in M$ and each integer $k \geq 0$ there exists a function $F \in gE$ (depending on x, k) whose k -jet $j^k F(x)$ at x equals $j^k f(x)$, then $f \in gE$.

(1') to (4'): Same as (1) to (4), but with E replaced by D [Schw, p. 65], the LF space [Tre2, Chapter 13] of smooth functions on M of compact support.

REMARK. These equivalences, and in particular the equivalence with Whitney's criterion, show that the continuity of division by g is a purely local question.

PROPOSITION 2. *Assume the same hypotheses as in Proposition 1. Then the division operator m_g^{-1} is continuous relative to the Whitney C^∞ topology on E and its subspace gE if and only if for each nonnegative integer p there exists an integer $k = k(p)$ (of necessity $\geq p$) for which the set map $m_g^{-1}: gE \rightarrow E$ is continuous with respect to the Fréchet $C^{k(p)}$ topology on gE and the Fréchet C^p topology on E .*

COROLLARY. *If division by g is continuous with respect to the Whitney C^∞ topology, then it is with respect to the Fréchet C^∞ topology.*

REMARKS. Proposition 2 is not hard to prove using the description of the Whitney C^∞ topology given in [Hirs, p. 35]. The converse of the Corollary is false. For example, if division by g is Fréchet-continuous but g has zeros of arbitrarily high finite order, then division by g is not Whitney-continuous. The reason is that the numbers $k(p)$ in Proposition 2 can be shown to satisfy $k(0) \geq r$ if g and its derivatives of order $< r$ vanish at some point. On the other hand, if $g^{-1}(0)$ is compact, then one can show that Fréchet continuity of division by g implies Whitney continuity.

Hörmander [Hörm] proved continuity in the numerator of division by a fixed polynomial. Independently Łojasiewicz [Łoj] showed that division by a fixed real-analytic function is continuous in the numerator. Tougeron [Tou, p. 160, Corollary

1.9] extended this to the following result. Let M be a C^∞ manifold, P a real-analytic manifold, and $p: P \rightarrow \mathbf{R}$ a fixed real-analytic function. Then division by $p \circ G$ is continuous in the numerator for “general” $G: M \rightarrow P$, that is, for each C^∞ map G whose ∞ -jet $j^\infty G$ avoids a certain “algebraic variety of infinite codimension” in the ∞ -jet bundle $J^\infty(M, P)$.

Since the division problem is local, we see that division by Morse functions (which near their zero set are linear or quadratic functions of smooth local coordinates) is continuous in the numerator. This shows that the set T of functions g for which division by g is continuous contains an open dense set in the Whitney C^∞ topology [GG]. On the other hand, if $\dim M \geq 2$, then T is not open in either topology. This is clear locally from the family of functions $g_c(x) = x_1^2 + \cdots + x_{n-1}^2 + c \exp(-1/x_n^2)$, since the polynomial g_0 is in T but g_c is not in T for positive c [Mal, p. 89]. The example can be made global. Finally, we note that division by a smooth complex-valued function g is continuous relative to the Fréchet or Whitney C^∞ topology if and only if division by the real valued function $|g|^2$ is continuous in the same topology.

3. Joint continuity of division. Our main result is the following.

THEOREM. *Let $p: P \rightarrow \mathbf{R}$ be a fixed analytic function on a real-analytic manifold P . Let $E(M)$ (respectively $E(M, P)$) denote the Fréchet space of C^∞ functions from a smooth manifold M to \mathbf{R} (resp. from M to P), with the Fréchet C^∞ topology. Let M_p be the operator*

$$M_p: E(M) \times E(M, P) \ni (h, G) \mapsto (f, G) \in E(M) \times E(M, P),$$

where $f(x) = h(x)p(G(x))$. There exists a Whitney stratification of $p^{-1}(0)$ for which the vanishing order of p is constant on each stratum. Let W be the set of all $G \in E(M, P)$ which are transversal to every stratum. (W is open and dense in the Whitney C^1 topology.) Then $M_p|_{E(M) \times W}$ is one-to-one, and its inverse operator (defined on $M_p(E(M) \times W)$ and mapping (f, G) to $f/(p \circ G)$) is continuous (in the Fréchet C^∞ topology).

Before proceeding with the proof we make some remarks and state two corollaries. To prove the theorem it suffices to find inequalities of the form

$$(1) \quad \|h\|_{r,K} \leq C \|f\|_{r',K'}$$

(K, K' compact, $K \subset K'$, $f = h \cdot (p \circ G)$, $\|h\|_{r,K} = \sup_{x \in K} \|j^r h(x)\|$ over $x \in K$, using some norm on the fibers of $J^r(M)$) in which C is independent of G locally in $E(M, P)$. Hörmander [Hörm], in his proof that division by a polynomial is continuous, obtained bounds of this type for division of f by a fixed polynomial g (i.e. $f = gh$, $h = f/g$). The transversality hypothesis on G guarantees that the zero set of $g = p \circ G$ varies “continuously” as G varies in $E(M, P)$. Also, one can show [Mos] that as x varies over a compact set in M , the distances $\text{dist}(x, g^{-1}(0))$ and $\text{dist}(G(x), p^{-1}(0))$ (using Riemannian metrics) are bounded by constant multiples of each other, and that these constants can be chosen to work for all maps G whose

1-jets are close enough to the 1-jet of a given map G_0 on a neighborhood of the compact set. Using these facts, one can adapt Hörmander's proof and prove that his constants can also be chosen to work for all G close enough to G_0 in $E(M, P)$.

COROLLARY 1. *For any smooth manifold M , the division map $(f = gh, g) \mapsto h = f/g$ is jointly continuous (in the Fréchet C^∞ topology) at all pairs $(f = gh, g) \in \text{image } m \subset E \times E_{n.d.}$ for which the differential of g vanishes nowhere on $g^{-1}(0)$.*

PROOF. Take for p the polynomial x on \mathbf{R} . One can also give a direct proof as follows. By the Implicit Function Theorem we can find local coordinate systems using g as the first coordinate. Write

$$\begin{aligned} f(g, x_2, \dots, x_n) &= \int_0^1 (d/dt)(f(tg, x_2, \dots, x_n)) dt \\ &= g \int_0^1 (d_1 f)(tg, x_2, \dots, x_n) dt. \end{aligned}$$

From this, we get uniform bounds for $h = f/g$ locally in terms of bounds on $d_1 f$, and for the r -jet $j^r h$ of h in terms of bounds on $j^{r+1} f$, for each r , relative to the (g, x_2, \dots, x_n) coordinate system. When we change to a coordinate system independent of g , we get uniform bounds on compact sets for $j^r h$ in terms of similar bounds on $j^{r+1} f$ and $j^{r+1} g$, locally in $E \times E_{n.d.}$. This implies the result.

COROLLARY 2. *Let \mathcal{X} be the space of smooth (C^∞) \mathbf{R}^n -valued functions on M , \mathcal{A} the space of smooth $(n \times n)$ matrix-valued functions on a smooth m -manifold M , and*

$$\mathcal{A}_0 = \{A \in \mathcal{A} \mid \det A \neq 0 \text{ on a dense set in } M\}.$$

Endow them with the Fréchet C^∞ topology. Then the map

$$M: \mathcal{A}_0 \times \mathcal{X} \ni (A, X) \mapsto (A, AX) \in \mathcal{A}_0 \times \mathcal{X}$$

is one-to-one, and the inverse map $M^{-1}: \text{image } M \rightarrow \mathcal{A}_0 \times \mathcal{X}$ is continuous at all pairs (A, AX) for which A is transversal to a Whitney stratification of $\det^{-1}(0) \subset \mathbf{R}^{n^2}$, provided that the vanishing order of \det is constant on each stratum.

PROOF. At each $x \in M$, $(\det A)X = (\text{Cof}'_A)(AX)$ (Cramer's Rule), where Cof'_A is the transpose of the matrix of cofactors of A . Now apply Theorem 2, taking $h = i$ th entry of $\text{Cof}'_A(AX)$, $G = A$, and $p = \det A$.

PROOF OF THE THEOREM. Hörmander [Hörm] has proven that division by a polynomial is continuous in the numerator. In his proof he points out that one could replace the polynomial by any function satisfying two conditions. The first condition, numbered 4.2, states that the order of vanishing of $p(x)$ is bounded on all of \mathbf{R}^n . The second condition, numbered 4.10, is a "Lojasiewicz inequality" of the following type: For arbitrary j, k , there exist constants c, s, u such that

$$(2) \quad \sum_{|q|=jk} |D^q(p(x)^j)|^2 \geq c(1 + \|x\|)^{-s} d(x, N_{k+1})^u$$

for $x \in N_k$ = the set of points where p vanishes to order k . Lojasiewicz [Loj] independently and at about the same time proved that such estimates are true for an arbitrary analytic function if we restrict to compact sets. With this restriction, the

Schwarz space multiplier becomes irrelevant. For any function satisfying his conditions 4.2 and 4.10, Hörmander proves that for all m, r there exist a constant C and integers m', r' such that

$$(3) \quad \sup_{x \in \mathbf{R}^n} (1 + \|x\|)^m \sum_{|a| \leq r} |D^a h(x)|^2 \leq C \sup_{x \in \mathbf{R}^n} (1 + \|x\|)^{m'} \sum_{|b| \leq r'} |D^b(p(x)h(x))|^2.$$

Since we are interested only in continuity with respect to seminorms defined with respect to compact sets in \mathbf{R}^n , we can restrict x to a closed ball B in \mathbf{R}^n and drop the Schwarz multiplicative factors $(1 + \|x\|)^m, (1 + \|x\|)^{m'}$. Hörmander's proof shows that if p is a smooth function defined on B whose order of vanishing is bounded on B (condition 4.2) and satisfies estimates (condition 4.10)

$$(4) \quad \sum_{|q|=jk} |D^q(p(x)^j)|^2 \geq cd(x, N_{k+1})^u$$

for $x \in N_k$, then for any r there exist a constant C and an integer $r' \geq r$ such that

$$(5) \quad \sup_{x \in B} \left[\sum_{|a| \leq r} |D^a h(x)|^2 \right] \leq C' \sup_{x \in B} \left[\sum_{|b| \leq r'} |D^b(p(x)h(x))|^2 \right]$$

provided the suprema exist.

To prove joint continuity it is sufficient to show first that (4) with fixed c, u holds uniformly for all $p \circ G$ in place of p when G is restricted to a suitable open set in the function space of mappings from the open ball A into B and is further required to be transverse to a particular stratification of $p^{-1}(0)$ and second that the constant C' and integer r' in (5) can be chosen uniformly for all such $p \circ G$.

Assuming we have demonstrated the first fact, to demonstrate the second we follow the proof given in the book by Treves [Tre]. It is an inverse induction on propositions P_k and P'_k with P'_0 directly implying (5). Conditions P_k and P'_k are trivially satisfied for k greater than the maximal order of vanishing of $p(x)$. We claim that within the class of functions satisfying (4) with fixed exponent u and constant c , the constant C' and integer r' can be chosen uniformly for all p' close to p . The claim is a consequence of the nature of the induction steps

- (a) $P_k + P'_{k+1} \Rightarrow P'_k$,
- (b) $P_{k+1} \Rightarrow P_k$.

Induction (a) depends on three types of estimates: (1) estimates on the derivatives of the Whitney extension of a function f from the set N_{k+1} in terms of its derivatives on N_{k+1} , (2) estimates on the sup norm of certain universal polynomials in p and its derivatives arising from application of the Leibnitz formula, and (3) estimates on the remainder in the Taylor formula. Induction (b) depends on sup norm estimates on universal polynomials in p and its derivatives (the operators labelled $L_k(x, F)$), estimates on the remainder in Taylor's formula and finally the Lojasiewicz inequality (4). Thus when we restrict to those functions satisfying (4) with the same exponent and constant, all the other estimates introduce constants which can be bounded uniformly for all p' near p , or do not depend on p at all.

To establish the claim we must show that (4) is satisfied by $p \circ G$ for appropriately restricted G . Let \tilde{N}_k be the set of points where $p \circ G$ vanishes together with all its derivatives up to order k . We will show that transversality of G to a certain stratification of $p^{-1}(0)$ guarantees that $G^{-1}(N_k) = \tilde{N}_k$ for all k , thus proving that the order of vanishing of $p \circ G$ at x equals the order of vanishing of p at $G(x)$. Then we prove that (4) is true for $p \circ G$ with the same exponent u unchanged and the constant chosen uniformly for all G near G_0 . Recall the definition of N_k . Let $N = N_1$ be the set of points where $p(x) = 0$. N_k is the set of points where p vanishes to order k ; that is, the total j th derivative, $d^j p(x)$, which considered as a j -fold symmetric tensor on \mathbf{R}^n vanishes, for all $j < k$. Let $d^j p(x)[v]$ stand for the evaluation of $d^j p(x)$ on the k -tuple (v, v, \dots, v) . Then

$$\begin{aligned} x \in N_k \quad & \text{iff } d^j p(x)[v] = 0 && \text{for all } j < k \text{ and } v \in \mathbf{R}^n, \\ & \text{iff } \frac{\partial^j p}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_j}}(x) = 0 && \begin{aligned} & \text{for all } (i_1 \cdots i_j), j \text{ a tuple of integers} \\ & \text{from 1 to } n \text{ and } j \text{ an integer} \\ & \text{from 0 to } k - 1. \end{aligned} \end{aligned}$$

For $x \in N_k - N_{k+1}$ the vectors v such that $d^k p(x)[v] = 0$ are called nongeneric directions.

LEMMA 1. *Suppose $x \in N_k - N_{k+1}$ and there exists a submanifold S of \mathbf{R}^n through x contained in N_k . Let V be a vector subspace of \mathbf{R}^n contained in the cone of nongeneric directions. Then $V + T_x S \neq \mathbf{R}^n$.*

PROOF. We claim that $d^k p(x)[v] = 0$ for any $v \in V + T_x S$; thus $V + T_x S \neq \mathbf{R}^n$ since by assumption $x \notin N_{k+1}$. The claim is clearly true for $v \in T_x S$ or $v \in V$ so let $v = v_1 + u_1$ with $0 \neq v_1 \in V \setminus T_x S$ and $0 \neq u_1 \in T_x S$. Find a coordinate system near x such that $\partial/\partial x_1 = u_1$ and $\partial/\partial x_n = v_1$ at x , and the submanifold S is defined by $x_{k+1} = x_{k+2} = \cdots = x_n = 0$. Then

$$\begin{aligned} d^k p(x)[u_1 + v_1] &= d^k p(x) \left[\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_n} \right] \\ &= \sum_{l=0}^k \binom{k}{l} \frac{\partial^k p(x)}{\partial x_1^l \partial x_n^{k-l}} = \frac{\partial^k p(x)}{\partial x_n^k} = 0 \end{aligned}$$

because $\partial^{k-l} p / \partial x_n^{k-l}$ vanishes identically along $S \subset N_k$ for any $l > 0$, so any $\partial/\partial x_1$ derivative of such a term vanishes and $\partial^k p(x)/\partial x_n^k = d^k p(x)[v_1] = 0$ since v_1 is nongeneric.

LEMMA 2. *Let $p: \mathbf{R}^m \rightarrow \mathbf{R}$ be a real analytic function. Then there exists a Whitney stratification of $p^{-1}(0)$ for which the vanishing order of p is constant on each stratum.*

PROOF. The case $p \equiv 0$ is trivial so assume p is not identically zero and for each positive integer k let N_k be defined as above. On each compact set K the nested sequence

$$K \cap N_1 \supset K \cap N_2 \supset K \cap N_k \supset \cdots$$

must end in the empty set after finitely many steps, since otherwise we could find a limit point $x \in K$ in which p vanishes to infinite order, implying $p \equiv 0$. Furthermore, each set N_k is an analytic set. A standard result about stratifications (see [Ver]) implies the existence of a Whitney stratification for which each N_k is a union of strata. Thus $N_k - N_{k+1}$ is a union of strata and the vanishing order of p at every point of each stratum in this union is k .

LEMMA 3. *Let $\mathcal{S} = \{S_\alpha\}$ be a stratification of the zero set of p such that the vanishing order is constant on each stratum. Let $G: U \subset \mathbf{R}^n \rightarrow \mathbf{R}^m$ be transverse to each stratum in \mathcal{S} (transverse to \mathcal{S}). Then $G^{-1}(N_k) = \tilde{N}_k = \{x \in U \mid p \circ G \text{ vanishes to order } k \text{ at } x\}$.*

PROOF. It is always true that $G^{-1}(N_k) \subseteq \tilde{N}_k$. Suppose there exists $x \in \tilde{N}_k$ such that $G(x) \notin N_k$. Then p vanishes to order $l < k$ at $G(x)$. Since $p \circ G$ vanishes to order k at x , the subspace $dG_x(T_x \mathbf{R}^n)$ consists entirely of nongeneric directions. By transversality, $dG_x(T_x \mathbf{R}^n) + T_{G(x)}S_\alpha = T_{G(x)}\mathbf{R}^m$ if S_α is the stratum containing $G(x)$. This contradicts Lemma 1.

There are two equivalent norms on the space of k -fold symmetric tensors on \mathbf{R}^n :

$$\|F\| = \sup\{ |F[v]| \mid \|v\| = 1, v \in \mathbf{R}^n \}$$

and

$$\|F\|_1 = \left(\sum_{1 \leq i_1 \leq \dots \leq i_k \leq n} |F(e_{i_1} \cdots e_{i_k})|^2 \right)^{1/2}.$$

Recall the notation $F[v] = F(v, \dots, v)$. The inequality (4) uses the norm $\|\cdot\|_1$ but the proof is simplified by using $\|\cdot\|$ so we write (4) as

$$(4') \quad \|d^k p(y)\| \geq C_1 d(y, N_{k+1})^u \quad \text{for } y \in N_k \cap K_1.$$

We want to prove

$$(6) \quad \|d^k(p \circ G)(x)\| \geq C_2 d(x, \tilde{N}_{k+1})^u \quad \text{for } x \in \tilde{N}_k \cap K_2,$$

where K_1 is a compact set in \mathbf{R}^m and K_2 is a compact set in \mathbf{R}^n such that $G(K_2) \subset K_1$ and the constant C_2 can be chosen uniformly for all G close to an initial G_0 and transverse to \mathcal{S} .

Substituting $y = G(x)$ in (4') gives

$$\|d^k p(G(x))\| \geq C_1 d(G(x), N_{k+1})^u.$$

We estimate the right side from below using the theorem on behavior under pullback of the distance to a Whitney stratification which is proven in part II of this paper:

$$(*) \quad d(G(x), N_{k+1}) \geq C_3 d(x, G^{-1}(N_{k+1})) = C_3 d(x, \tilde{N}_{k+1}).$$

Therefore

$$\|d^k p(G(x))\| \geq C_4 d(x, \tilde{N}_{k+1})^u.$$

The desired inequality for $p \circ G$ will follow once we have shown

$$(7) \quad \|d^k(p \circ G)(x)\| \geq C_5 \|d^k p(G(x))\|.$$

That the constant C_3 can be chosen uniformly in G is proven in Part II, and the uniformity in G of C_5 follows from the proof of (7) given below.

Before proving (7) we note that if F is a k -fold symmetric tensor on \mathbf{R}^n and W is a subspace of \mathbf{R}^n such that $F(v_1, \dots, v_{k-1}, w) = 0$ for all $w \in W$ and $v_1, \dots, v_{k-1} \in \mathbf{R}^n$, then $F[v] = F[\pi(v)] \stackrel{\text{def}}{=} F \circ \pi[v]$ where π is the orthogonal projection of \mathbf{R}^n on W^\perp .

If $G(x)$ is in the stratum $S_\alpha \subset N_k$, then $d^k p(G(x))$ vanishes on a k -tuple which includes a direction tangent to S_α . Thus $d^k p(G(x)) = d^k p(G(x)) \circ \pi_{G(x)}$, where $\pi_{G(x)}$ is the orthogonal projection of $T_{G(x)}\mathbf{R}^n$ on $(T_{G(x)}S_\alpha)^\perp$, and

$$\begin{aligned} \|d^k(p \circ G)(x)\| &= \|d^k p(G(x)) \circ dG(x)\| \\ &= \|d^k p(G(x)) \circ \pi_{G(x)} \circ dG(x)\|. \end{aligned}$$

By the transversality assumption, $\pi_{G(x)} \circ dG(x)$ is a surjection of $T_x\mathbf{R}^n$ onto $(T_{G(x)}S_\alpha)^\perp$. Let $R_\alpha = G^{-1}(S_\alpha)$ be the pull-back stratification of $\tilde{N} = G^{-1}(N)$ [GWPL, p. 14]. If we restrict the linear transformation $\pi_{G(x)} \circ dG(x)$ to $(T_x R_\alpha)^\perp$, it gives an isomorphism onto $(T_{G(x)}S_\alpha)^\perp$. Thus for $v \in (T_x R_\alpha)^\perp$

$$\|\pi_{G(x)} dG(x)[v]\| \geq \lambda_x |v|$$

and therefore if we pick $w_0 \in (T_{G(x)}S_\alpha)^\perp$, $|w_0| = 1$, such that $\|d^k p(G(x))\| = |d^k p(G(x))[w_0]|$, then there is a $v_0 \in (T_x R_\alpha)^\perp$ such that $\pi_{G(x)} dG(x)[v_0] = w_0$ and $1 = |w_0| \geq \lambda_x |v_0|$. Then

$$\begin{aligned} \|d^k(p \circ G)(x)\| &= \|d^k p(G(x)) \circ \pi_{G(x)} dG(x)\| \\ &\geq |d^k p(G(x)) \pi_{G(x)} dG(x)[v_0/v_0]| \\ &\geq |d^k p(G(x))[w_0]|/|v_0|^k \geq \|d^k p(G(x))\| \lambda_x^k. \end{aligned}$$

We want a uniform lower bound on λ_x as x varies over a compact set but is not necessarily restricted to one stratum. The existence of such a bound follows from the next lemma.

LEMMA 4. *Let $G: X \rightarrow Y$ be a smooth mapping. Let $\mathcal{S} = \{S_\alpha\}$ be a stratification of $N \subset Y$ and assume G is transverse to \mathcal{S} . Let $\{R_\alpha = G^{-1}(S_\alpha)\}$ be the pull-back stratification of $G^{-1}N$. Let $W_y = (T_y S_\alpha)^\perp$ and $V_x = (T_x R_\alpha)^\perp$ and let $\pi_y: T_y Y \rightarrow W_y$ be orthogonal projection. Then*

$$\inf_{\substack{|v|=1 \\ (v \in V_x)}} |\pi_{G(x)} dG(x)[v]| = \lambda(x)$$

is lower semicontinuous on $G^{-1}N$.

PROOF. Pick $x \in G^{-1}(N)$. If there exists a neighborhood of x in $G^{-1}(N)$ which is contained in one stratum then $\lambda(x)$ is continuous at x by standard results for transversal maps of manifolds. Let $x \in R_1 \cap \bar{R}_2$ and assume the conclusion is not true. There exists a sequence of points $x_n \in R_2$ converging to x with $\lim \lambda(x_n) < \lambda(x)$. For each x_n there exists a $v_n \in T_{x_n} R_2^\perp$ such that $|v_n| = 1$ and $\lambda(x_n) = |\pi_{G(x_n)} dG(x_n)[v_n]|$. Assume the sequence is chosen such that the tangent spaces $T_{x_n} R_2$ converge to a subspace $T_0 \subset T_x X$ in the Grassmann q -plane bundle over X ($q = \dim R_2$) and such that the v_n converge to v in the sphere bundle. Then

$T_0 \supset T_x R_1$ by Whitney condition (a) and $v \in V_0 = T_0^\perp \subset T_x R_1^\perp = V_x$. In fact we can write $\lim \pi_{G(x_n)} dG(x_n)$ as $\pi_0 \circ \pi_{G(x)} dG(x) \circ j_0$, where $\pi_0: W_y \rightarrow W_0$ is orthogonal projection on W_0 and $W_0 = (T_0^1)^\perp$ for $T_0^1 = \lim T_{G(x_n)} S_2$ and $j_0: V_0 \rightarrow V_x$ is inclusion. The kernel of π_0 is $W_y \cap T_0^1$. By transversality at x we know that $\pi_{G(x)} dG(x): V_x \rightarrow W_y$ is an isomorphism. By continuity $dG(x)$ maps T_0 into T_0^1 . Furthermore $\pi_{G(x)}$ maps T_0^1 into itself since it is orthogonal projection along $T_y S_1 \subset T_0^1$. Thus $\pi_{G(x)} dG(x)(T_0 \cap V_x) \subseteq T_0^1 \cap W_y$. But

$$\begin{aligned} \dim T_0 \cap V_x &= \dim R_2 - \dim R_1 = \text{codim } R_1 - \text{codim } R_2 \\ &= \text{codim } S_1 - \text{codim } S_2 = \dim S_2 - \dim S_1 \\ &= \dim T_0^1 \cap W_y. \end{aligned}$$

Thus $\pi_{G(x)} dG(x): T_0 \cap V_x \rightarrow T_0^1 \cap W_y$ is an isomorphism. Let w be the orthogonal projection along W_0 of $\pi_{G(x)} dG(x)[v]$ in $T_0^1 \cap W_y$. Then $\pi_0 \pi_{G(x)} dG(x)[v] = \pi_{G(x)} dG(x)[v] - w$. As we saw above w can be written $\pi_{G(x)} dG(x)[u]$ for $u \in T_0 \cap V_x$. Then

$$\pi_0 \pi_{G(x)} dG(x)[v] = \pi_{G(x)} dG(x)[v - u]$$

and

$$\begin{aligned} \lim \lambda(x_n) &= |\pi_0 \pi_{G(x)} dG(x)[v]| = |\pi_{G(x)} dG(x)[v - u]| \\ &\geq \lambda(x) |v - u| \geq \lambda(x) (|v|^2 + |u|^2)^{1/2} \geq \lambda(x) |v|, \end{aligned}$$

contradicting the assumption that $\lim \lambda(x_n) < \lambda(x)$. From the lower semicontinuity of $\lambda(x)$ we conclude that $\lambda(x)$ has a positive lower bound on $K \cap \tilde{N}_k$ for any compact set K . Therefore $\|d^k(p \circ G)(x)\| \geq c \|d^k p(G(x))\|$ for $x \in K \cap \tilde{N}_k$. Together with the estimate * established in part II this proves the Theorem.

REFERENCES

- [GG] M. Golubitsky and V. Guillemin, *Stable mappings and their singularities*, Springer-Verlag, New York, 1973. Second, corrected, printing 1980.
- [GWPL] C. G. Gibson et al., *Topological stability of smooth mappings*, Lecture Notes in Math., vol. 552, Springer-Verlag, New York, 1976.
- [Hirs] M. W. Hirsch, *Differential topology*, Springer, New York, 1976.
- [Hörm] L. Hörmander, *On the division of distributions by polynomials*, Ark. Mat. **3** (1958), 555–568.
- [Loj] S. Lojasiewicz, *Sur le problème de la division*, Studia Math. **18** (1959), 87–136.
- [Mal] B. Malgrange, *Ideals of differentiable functions*, Oxford Univ. Press, London, 1966.
- [Mat] J. N. Mather, *Stability of C^∞ mappings: II. Infinitesimal stability implies stability*, Ann. of Math. (2) **89** (1969), 254–291.
- [Mos] M. A. Mostow, *Joint continuity of division II. The distance to a Whitney stratified set from a transversal submanifold*, Trans. Amer. Math. Soc. **292** (1985), 585–594.
- [MS1] M. A. Mostow and S. Shnider, *The continuity of computing connections from curvatures, and of dividing smooth functions*, Proc. Internat. Conf. Differential Geometric Methods in Theoretical Physics (Jerusalem, Israel, August 5–11, 1982), Mathematical Physics Studies Series, Reidel, Dordrecht, Holland.
- [MS2] ———, *An application of a division theorem to the continuous determination of connections from curvatures* (in preparation).

[Schw] L. Schwartz, *Théorie des distributions*, Hermann, Paris, 1966.

[Tou] J.-C. Tougeron, *Idéaux de fonctions différentiables*, Springer, New York, 1972.

[Tre1] F. Trèves, *Linear partial differential equations with constant coefficients*, Gordon and Breach, New York, 1966.

[Tre2] _____, *Topological vector spaces, distributions and kernels*, Academic Press, New York, 1967.

[Ver] J.-L. Verdier, *Stratifications de Whitney et théorème de Bertini-Sard*, *Invent. Math.* **36** (1976), 295–312.

[Whi] H. Whitney, *On ideals of differentiable functions*, *Amer. J. Math.* **70** (1948), 635–658.

DEPARTMENT OF MATHEMATICS, NORTH CAROLINA STATE UNIVERSITY, RALEIGH, NORTH CAROLINA 27650

DEPARTMENT OF MATHEMATICS, BEN-GURION UNIVERSITY OF THE NEGEV, B'ER-SHEVA, ISRAEL; On leave from Mathematics Department, McGill University, Montreal, P.Q., Canada

Current address (Mark Mostow): Clarity, Ltd., P. O. Box 3112, Herzelia 46103, Israel