

A SIMPLE THEORY OF DIFFERENTIAL CALCULUS IN LOCALLY CONVEX SPACES

BY

RICHARD A. GRAFF

ABSTRACT. A theory of differential calculus for nonlinear maps between general locally convex spaces is developed. All convergence notions are topological, and only familiarity with basic results from point set topology, differential calculus in Banach spaces, and locally convex space theory is assumed. The chain rule for continuous k th order differentiability, smoothness of inverse functions, and function space continuity properties of higher order derivatives are examined. It is shown that this theory extends the classical Fréchet theory of differential calculus for maps between Banach spaces.

Introduction. An open question for several decades has been whether it is possible to develop a “nice” theory of C^k differentiability for maps between general locally convex spaces which extends the standard theory of Fréchet differential calculus for maps between Banach spaces. Such a theory should have several properties:

- (1) The class of C^k maps should be closed under composition.
- (2) Continuous linear and multilinear maps between locally convex spaces should be C^∞ .
- (3) The theory should establish a framework for the study of inverse functions, local flows for C^k vector fields, and C^∞ partitions of unity.
- (4) Proofs and concepts should be simple, and close to those from Banach space Fréchet calculus.

We take an indirect approach to the solution of this problem. In [3], the author developed an alternative approach to Banach space differential calculus, in which derivatives are continuous with respect to the strong operator topology on spaces of linear and multilinear maps rather than the uniform operator topology. However, this theory (called \mathcal{S}^k calculus) was shown to be closely related to C^k calculus, and the notions of \mathcal{S}^k and C^k differentiability were shown to coincide in the case $k = \infty$.

In this article, the notion of \mathcal{S}^k differentiability is extended to the general locally convex setting. The chain rule is established for \mathcal{S}^k maps, continuous linear and multilinear maps are shown to be \mathcal{S}^∞ , and \mathcal{S}^1 inverses of \mathcal{S}^k maps are shown to be \mathcal{S}^k differentiable.

After the development of basic \mathcal{S}^k calculus, C^k differentiability is defined in terms of \mathcal{S}^k differentiability, and the corresponding results are established for C^k calculus. It is shown that this C^k calculus extends the notion of Banach space

Received by the editors July 27, 1984.

1980 *Mathematics Subject Classification*. Primary 58C20, 46G05; Secondary 46A05, 47H99.

©1986 American Mathematical Society
0002-9947/86 \$1.00 + \$.25 per page

Fréchet calculus, and that the generalized notions of \mathcal{S}^∞ and C^∞ coincide. In addition, the function space properties of \mathcal{S}^k and C^k derivatives are examined.

§§1 and 2 develop the basic properties of \mathcal{S}^k calculus (the reader will note that the results of §1 are valid for maps between general linear topological spaces, locally convex or not). §3 develops the function space properties of higher order derivatives for maps on compactly generated domains. §4 reviews the basic results of \mathcal{S}^k calculus for maps between Banach spaces as developed in [3], as well as some new results on the inverse function theorem for \mathcal{S}^k calculus. §5 defines C^k differentiability in the locally convex setting, develops the basic properties of C^k calculus, and establishes the relations between \mathcal{S}^k and C^k differentiability. It is shown that this notion of C^k differentiability not only extends the theory of C^k Banach space Fréchet calculus, but also the theory of C^k differentiability for maps between D -spaces which was developed in [1, 2]. We also note that our theory of \mathcal{S}^k differentiability extends the notion of continuous k th order differentiability for maps between Fréchet spaces developed as a framework for the Nash-Moser inverse function theorem by Hamilton in [6].

1. Basic properties of \mathcal{S}^k calculus. Assume throughout this section that V, Z, \tilde{Z} are linear topological spaces, W is open in V, U is open in $Z, f: W \rightarrow Z, g: U \rightarrow \tilde{Z}$. We will let $L(V, Z)$ denote the linear space of continuous linear maps from V to Z . More generally, if $r \geq 2$ and V_1, \dots, V_r are linear topological spaces, $L^r(V_1, \dots, V_r; Z)$ will denote the linear space of continuous r -linear maps from $V_1 \times \dots \times V_r$ to Z . For the special case of $V = V_1 = \dots = V_r$, we shorten the notation for the space of continuous r -linear maps to $L^r(V, Z)$. We do not assume the existence of any topology on these spaces of linear and multilinear maps in this section, although we will introduce topologies for these spaces in later sections.

1.1. DEFINITION. Let $\varepsilon > 0, r: ((-\varepsilon, 0) \cup (0, \varepsilon)) \times W \rightarrow Z, y \in W$. We will say that $r(\cdot, \cdot)$ approaches zero in the y -direction as $t \rightarrow 0$, if $\lim_{t \rightarrow 0; w \rightarrow y} r(t, w) = 0$.

1.2. DEFINITION. Let $x \in W, l \in L(V, Z)$. We will say that l is a (Hadamard) derivative for f at x if, for each $y \in V, r_y(\cdot, \cdot)$ approaches zero in the y -direction as $t \rightarrow 0$, where $r_y(t, w)$ is defined by

$$r_y(t, w) = \frac{f(x + tw) - f(x)}{t} - l(y).$$

1.3. REMARK. The map f has a derivative at $x \Leftrightarrow$ there exists a linear map $l \in L(V, Z)$ such that, for each $y \in V$, there exists a map $r_y(\cdot, \cdot)$ which approaches zero in the y -direction as $t \rightarrow 0$, such that

$$f(x + tw) = f(x) + tl(y) + t(r_y(t, w)).$$

It is immediate from the above remark that the derivative of f at x , if it exists, is unique. We will denote the derivative of f at x by $Df(x)$.

1.4. LEMMA. Assume that $f(W) \subset U, f: W \rightarrow Z$ is differentiable at x , and $g: U \rightarrow \tilde{Z}$ is differentiable at $f(x)$. Then $g \circ f$ is differentiable at x , and $D(g \circ f)(x) = Dg(f(x))Df(x)$.

PROOF. Let $y \in V$. Then $f(x + tw) = f(x) + tDf(x)(y) + tr_y(t, w)$, where $r_y(\cdot, \cdot)$ approaches zero in the y -direction as $t \rightarrow 0$. Thus

$$\begin{aligned} g \circ f(x + tw) &= g(f(x + tw)) = g(f(x) + tDf(x)(y) + tr_y(t, w)) \\ &= g(f(x) + t(Df(x)(y) + r_y(t, w))). \end{aligned}$$

Since g is differentiable at $f(x)$, there exists $s(\cdot, \cdot) = s_{Df(x)(y)}(\cdot, \cdot)$ such that

$$\begin{aligned} g(f(x) + t(Df(x)(y) + r_y(t, w))) &= g(f(x)) + tDg(f(x))Df(x)(y) \\ &\quad + ts(t, Df(x)(y) + r_y(t, w)), \end{aligned}$$

where $s(\cdot, \cdot)$ approaches zero in the $Df(x)(y)$ -direction as $t \rightarrow 0$. Then

$$\lim_{\substack{t \rightarrow 0 \\ w \rightarrow y}} s(t, Df(x)(y) + r_y(t, w)) = \lim_{\substack{t \rightarrow 0 \\ u \rightarrow Df(x)(y)}} s(t, u) = 0;$$

i.e. $s(\cdot, Df(x)(y) + r_y(\cdot, \cdot))$ approaches zero in the y -direction as $t \rightarrow 0$. Thus it follows from Remark 1.3 that $g \circ f$ is differentiable at x , and that $D(g \circ f)(x) = Dg(f(x))Df(x)$. \square

1.5. DEFINITION. We will say that f is of (differentiability) class \mathcal{S}^1 if $Df(x)$ exists for each $x \in W$ and if the map $Tf(\cdot, \cdot): U \times V \rightarrow Z \times Z$ defined by $Tf(x, y) = (f(x), Df(x)(y))$ is continuous.

1.6. DEFINITION. We define \mathcal{S}^k maps inductively for $k > 1$. For $k \in N, k \geq 2$, we will say that an \mathcal{S}^1 map f is \mathcal{S}^k if Tf is \mathcal{S}^{k-1} . We will call f an \mathcal{S}^∞ map if f is \mathcal{S}^k for every $k \in N$.

1.7. REMARK. If V and Z are finite-dimensional, then the definition of \mathcal{S}^k differentiability is equivalent to the standard definition of C^k Fréchet differentiability; i.e. f is $\mathcal{S}^k \Leftrightarrow f$ is C^k .

1.8. THEOREM (THE CHAIN RULE). Let $k \in N \cup \{\infty\}$, let $f: W \rightarrow Z$ and $g: U \rightarrow \tilde{Z}$ be \mathcal{S}^k maps, and assume that $f(W) \subset U$. Then $g \circ f$ is \mathcal{S}^k , and $T(g \circ f) = T(g) \circ T(f)$.

PROOF. It suffices to prove this theorem for the case $k = 1$, since the proof for the general case will follow immediately by induction. By Lemma 1.4, $D(g \circ f)$ exists for each $x \in V$, and $D(g \circ f)(x) = Dg(f(x))Df(x)$. Thus $T(g \circ f)$ exists, and

$$\begin{aligned} T(g \circ f)(x, y) &= (g \circ f(x), Dg(f(x))Df(x)(y)) \\ &= Tg(f(x), Df(x)(y)) = Tg \circ Tf(x, y), \end{aligned}$$

which implies that $T(g \circ f) = Tg \circ Tf$. Since $T(g \circ f) = Tg \circ Tf$ is the composition of continuous maps, it follows that $T(g \circ f)$ is continuous. Thus the composition of \mathcal{S}^1 maps is an \mathcal{S}^1 map. \square

We next turn our attention to showing that the class of \mathcal{S}^∞ maps is rich with examples. We begin by showing that continuous linear and multilinear maps are \mathcal{S}^∞ , and then show that it is possible to add \mathcal{S}^∞ maps together and multiply them by scalar-valued \mathcal{S}^∞ functions without leaving the class of \mathcal{S}^∞ maps.

1.9. LEMMA. Let $l \in L(V, Z)$. Then l is \mathcal{S}^∞ , and $Tl = l \times l$.

PROOF. Let $x \in V$. It is obvious that $Dl(x)$ exists, and that $Dl(x) = l$. Thus Tl exists, and

$$Tl = l \times l: V \times V \rightarrow Z \times Z, \quad (x, y) \rightarrow (l(x), l(y))$$

This formula implies that Tl is continuous, and hence that l is \mathcal{S}^1 . Furthermore, $l \times l \in L(V \times V, Z \times Z)$, which implies that Tl is again a linear map. Thus we can apply induction to conclude that l is \mathcal{S}^∞ . \square

1.10. LEMMA. Let $l_1 \in L(Z, \tilde{Z})$, $l_2 \in L(\tilde{V}, V)$, and assume that f is \mathcal{S}^k . Then $l_1 \circ f \circ l_2: l_2^{-1}(W) \rightarrow \tilde{Z}$ is \mathcal{S}^k , and $D(l_1 \circ f \circ l_2)(x) = l_1 \circ Df(x) \circ l_2$ for all $x \in l_2^{-1}(W)$.

PROOF. Since l_1, l_2 are \mathcal{S}^∞ , $l_1 \circ f \circ l_2$ is \mathcal{S}^k by the chain rule. The formula for the derivative of the composition follows immediately from Lemma 1.4. \square

1.11. REMARK. Let $Z = \prod_{\lambda \in \Lambda} Z_\lambda$, where each Z_λ is a linear topological space and the product is taken in the category of linear topological spaces. For each $\lambda \in \Lambda$, let $q_\lambda: Z \rightarrow Z_\lambda$ be projection onto the λ th factor. Let $\varepsilon > 0$, and assume that $r: ((-\varepsilon, 0) \cup (0, \varepsilon)) \times W \rightarrow Z$. Then $r(\cdot, \cdot)$ approaches zero in the y -direction as $t \rightarrow 0 \Leftrightarrow r_\lambda = q_\lambda \circ r$ approaches zero in the y -direction as $t \rightarrow 0$ for each $\lambda \in \Lambda$.

1.12. LEMMA. Let $Z = \prod_{\lambda \in \Lambda} Z_\lambda$. Then $f: W \rightarrow Z$ is $\mathcal{S}^k \Leftrightarrow f_\lambda = q_\lambda \circ f: W \rightarrow Z_\lambda$ is \mathcal{S}^k for each $\lambda \in \Lambda$.

PROOF. In one direction (\Rightarrow), this result is trivial: if f is \mathcal{S}^k , then each $f_\lambda = q_\lambda \circ f$ is the composition of \mathcal{S}^k maps, and hence is \mathcal{S}^k . We establish the converse result by induction on k .

To verify the result for \mathcal{S}^1 maps, let $x \in W$, and define $l_x \in L(V, Z)$ by $q_\lambda \circ l_x = Df_\lambda(x)$ for each $\lambda \in \Lambda$ (l_x is well defined by this representation, since Z is the linear topological product of the Z_λ 's). Let $y \in V$, and define $r(\cdot, \cdot)$ for $|t|$ sufficiently small, $t \neq 0$, by $r(t, w) = (f(x + tw) - f(x))/t - l_x(y)$. For each $\lambda \in \Lambda$, let $r_\lambda = q_\lambda \circ r$. Since each f_λ is differentiable at x , each $r_\lambda(\cdot, \cdot)$ approaches zero in the y -direction as $t \rightarrow 0$. Thus Remark 1.11 implies that $r(\cdot, \cdot)$ approaches zero in the y -direction as $t \rightarrow 0$, which implies that f is differentiable at x and that $Df(x) = l_x$. Since this reasoning holds for each $x \in W$, $Tf: W \times V \rightarrow Z \times Z$ exists. Furthermore, if we define the linear topological isomorphism $l: \prod_{\lambda \in \Lambda} (Z_\lambda \times Z_\lambda) \rightarrow Z \times Z$ by

$$(q_\lambda \times q_\lambda) \circ l \left(\prod_{\nu \in \Lambda} (x_\nu, y_\nu) \right) = (x_\lambda, y_\lambda),$$

then $Tf = l \circ (\prod_{\lambda \in \Lambda} Tf_\lambda)$, which implies that Tf is continuous. Thus f is \mathcal{S}^1 .

We complete the proof by a straightforward induction. Let $k \in N$, and assume the result has been established for k . Assume that each f_λ is \mathcal{S}^{k+1} . Then each Tf_λ is \mathcal{S}^k , which implies that $\prod_{\lambda \in \Lambda} Tf_\lambda$ is \mathcal{S}^k , and hence that $Tf = l \circ (\prod_{\lambda \in \Lambda} Tf_\lambda)$ is \mathcal{S}^k . Thus f is \mathcal{S}^{k+1} . \square

1.13. LEMMA. Assume that $f_1, \dots, f_r: W \rightarrow Z$ are \mathcal{S}^k maps, and let $f = f_1 + \dots + f_r$. Then $f: W \rightarrow Z$ is \mathcal{S}^k .

PROOF. It suffices to assume $r = 2$. We factor f as follows:

$$W \xrightarrow{(f_1, f_2)} Z \times Z \xrightarrow{\text{add}} Z,$$

where $\text{add}: Z \times Z \rightarrow Z$ is the addition map. By Lemma 1.12, (f_1, f_2) is \mathcal{S}^k . Since add is continuous and linear, it is \mathcal{S}^∞ . Thus the lemma follows from the chain rule.

□

1.14. PROPOSITION. Let $r \in \mathbb{N}$, $f \in L'(V_1, \dots, V_r; Z)$. Then f is \mathcal{S}^∞ .

PROOF. We already know this result for the case $r = 1$ from Lemma 1.9, so it suffices to assume $r \geq 2$. Let $x = (x_1, \dots, x_r)$, $y = (y_1, \dots, y_r)$, $w = (w_1, \dots, w_r)$. Then

$$\begin{aligned} f(x + tw) - f(x) &= f(x_1 + tw_1, x_2 + tw_2, \dots, x_r + tw_r) - f(x_1, \dots, x_r) \\ &= f(x_1 + tw_1, x_2 + tw_2, \dots, x_r + tw_r) - f(x_1 + tw_1, \dots, x_{r-1} + tw_{r-1}, x_r) \\ &\quad + f(x_1 + tw_1, \dots, x_{r-1} + tw_{r-1}, x_r) - f(x_1 + tw_1, \dots, x_{r-2} + tw_{r-2}, x_{r-1}, x_r) \\ &\quad + \dots + f(x_1 + tw_1, x_2, \dots, x_r) - f(x_1, \dots, x_r) \\ &= t(x_1 + tw_1, \dots, x_{r-1} + tw_{r-1}, w_r) \\ &\quad + tf(x_1 + tw_1, \dots, x_{r-2} + tw_{r-2}, w_{r-1}, x_r) + \dots + tf(w_1, x_2, \dots, x_r) \\ &= t(f(x_1, \dots, x_{r-1}, y_r) + f(x_1, \dots, x_{r-2}, y_{r-1}, x_r) + \dots + f(y_1, x_2, \dots, x_r)) \\ &\quad + tf((f(x_1 + tw_1, \dots, x_{r-1} + tw_{r-1}, w_r) - f(x_1, \dots, x_{r-1}, y_r)) \\ &\quad \quad + (f(x_1 + tw_1, \dots, x_{r-2} + tw_{r-2}, w_{r-1}, x_r) - f(x_1, \dots, x_{r-2}, y_{r-1}, x_r)) \\ &\quad \quad \quad + \dots + (f(w_1, x_2, \dots, x_r) - f(y_1, x_2, \dots, x_r))) \\ &= t(f(x_1, \dots, x_{r-1}, y_r) + f(x_1, \dots, x_{r-2}, y_{r-1}, x_r) + \dots + f(y_1, x_2, \dots, x_r)) \\ &\quad + t(r(t, w)). \end{aligned}$$

Since f is continuous, $\lim_{t \rightarrow 0; w \rightarrow y} r(t, w) = 0$, which implies that $Df(x)$ exists and that

$$\begin{aligned} Df(x)(y) &= f(x_1, \dots, x_{r-1}, y_r) + f(x_1, \dots, x_{r-2}, y_{r-1}, x_r) \\ &\quad + \dots + f(y_1, x_2, \dots, x_r). \end{aligned}$$

In addition, this formula implies that $Df(\cdot, \cdot)$ is continuous, which implies that Tf is continuous, and hence that f is \mathcal{S}^1 .

We finish the proof by induction. Let $k \in \mathbb{N}$, and assume we already know that continuous r -linear maps are \mathcal{S}^k . For each $1 \leq i \leq r$, define

$$l_i \in L(V_1 \times \dots \times V_r \times V_1 \times \dots \times V_r, V_1 \times \dots \times V_r)$$

by

$$l_i(x_1, \dots, x_r, y_1, \dots, y_r) = (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_r),$$

and let

$$p \in L(V_1 \times \dots \times V_r \times V_1 \times \dots \times V_r, V_1 \times \dots \times V_r)$$

be projection onto the first factor. From our formula for $Df(x)(y)$, we see that

$$Tf = \left(f \circ p, \sum_{i=1}^r f \circ l_i \right).$$

Thus it follows from Lemmas 1.12 and 1.13, together with the chain rule, that Tf is \mathcal{S}^k . The proposition now follows immediately by induction. \square

1.15. COROLLARY. *Let $k \in N \cup \{\infty\}$, and assume that $\alpha: W \rightarrow R$ and $f: W \rightarrow Z$ are \mathcal{S}^k maps. Then $\alpha f: W \rightarrow Z$ is \mathcal{S}^k , where αf is defined by $(\alpha f)(x) = \alpha(x)f(x)$ for each $x \in W$.*

PROOF. We factor αf as follows:

$$W \xrightarrow{(\alpha, f)} R \times Z \xrightarrow{m} Z,$$

where m is scalar multiplication. Since m is continuous and bilinear, it is \mathcal{S}^∞ by Proposition 1.14. Lemma 1.12 implies that (α, f) is \mathcal{S}^k . Thus the chain rule implies that αf is \mathcal{S}^k . \square

We conclude this section by proving that \mathcal{S}^k differentiability is a local property.

1.16. LEMMA. *Let $k \in N \cup \{\infty\}$, and let $f: W \rightarrow Z$. Let $W = \bigcup_{\lambda \in \Lambda} W_\lambda$, where each W_λ is open in V , and assume that each $f|_{W_\lambda}: W_\lambda \rightarrow Z$ is \mathcal{S}^k . Then $f: W \rightarrow Z$ is \mathcal{S}^k .*

PROOF. It suffices to prove this result for finite k . We proceed by induction. To see that the result holds for $k = 1$, assume that the restriction of f to each W_λ is \mathcal{S}^1 . This implies that f is differentiable at each point of W , so Tf is defined on $W \times V$. Since $W \times V = \bigcup_{\lambda \in \Lambda} (W_\lambda \times V)$, each $W_\lambda \times V$ is open in $V \times V$, and $Tf|_{W_\lambda \times V} = T(f|_{W_\lambda \times V})$ is continuous, it follows that $Tf: W \times V \rightarrow Z \times Z$ is continuous. Thus f is \mathcal{S}^1 .

Next, let $k \geq 2$, and assume that the result has been established for \mathcal{S}^i maps with $i < k$. If each $f|_{W_\lambda}: W_\lambda \rightarrow Z$ is \mathcal{S}^k , then we apply this lemma for the case $i = 1$ to see that $f: W \rightarrow Z$ is \mathcal{S}^1 . Thus Tf is defined on $W \times V$. Furthermore, each $Tf|_{W_\lambda \times V}: W_\lambda \times V \rightarrow Z \times Z$ is \mathcal{S}^{k-1} , so by this lemma for the case $i = k - 1$ it follows that $Tf: W \times V \rightarrow Z \times Z$ is \mathcal{S}^{k-1} , and hence that f is \mathcal{S}^k . By induction, we have the result for all $k \in N$. \square

2. Higher order derivatives and inverse functions. Assume for the remainder of this article that all linear topological spaces are locally convex. We will continue to use the notation introduced at the beginning of §1. In addition, if $r \in N$, $r \geq 2$, we will let $L'_r(V, Z)$ denote the subspace of $L'(V, Z)$ consisting of symmetric r -linear maps.

Let X be a locally convex space, and let $i: X \rightarrow \hat{X}$ be the embedding of X in its sequential completion. Let $\alpha: [0, 1] \rightarrow X$ be a continuous curve in X . Since X might not be sequentially complete, $\int_0^1 \alpha(t) dt$ may not exist. However, let $\hat{\alpha} = i \circ \alpha$. Since \hat{X} is sequentially complete, $\int_0^1 \hat{\alpha}(t) dt$ does exist. Suppose $\int_0^1 \hat{\alpha}(t) dt \in X$. Then, since the map i is an embedding, $\int_0^1 \alpha(t) dt$ exists, and $i(\int_0^1 \alpha(t) dt) = \int_0^1 \hat{\alpha}(t) dt$. Conversely, if $\int_0^1 \alpha(t) dt$ exists, then $\int_0^1 \hat{\alpha}(t) dt = i(\int_0^1 \alpha(t) dt) \in X$.

We will begin this section by establishing the Mean Value Theorem and its converse.

2.1. LEMMA. Assume that W is convex. Then f is $\mathcal{S}^1 \Leftrightarrow$ there exists a continuous map $g: W \times V \rightarrow Z$ such that $g(x, \cdot)$ is linear for each $x \in W$ and such that, for every pair $x, x + y \in W$, $\int_0^1 g(x + ty, y) dt$ exists and $f(y) - f(x) = \int_0^1 g(x + ty, y) dt$. Furthermore, if f is \mathcal{S}^1 , then $Df(x) = g(x, \cdot)$ for each $x \in W$.

PROOF. (\Rightarrow) Let $i: Z \rightarrow \hat{Z}$ be the embedding of Z in its completion, and let $\hat{f} = i \circ f$ be the induced \mathcal{S}^1 map. Note that $D\hat{f} = i \circ Df$. If we prove the result for \hat{f} , then

$$\int_0^1 D\hat{f}(x + ty)(y) dt = \hat{f}(x + y) - \hat{f}(x) = f(x + y) - f(x) \in Z,$$

so the integral in the statement of the lemma exists, and the desired equality follows from the corresponding equality for \hat{f} . Thus it suffices to prove the lemma in the case when Z is complete. But for this case, use Lemma 1.10 to reduce the problem to the case $V = Z = R$, where it is the Fundamental Theorem of Integral Calculus.

(\Leftarrow) Let $x \in W$. We begin by showing that $Df(x)$ exists and equals $g(x, \cdot)$. Let $y \in V$. Then, for w near y and $|t|$ sufficiently small, $x + tw \in W$, and

$$\begin{aligned} f(x + tw) &= f(x) + \int_0^1 g(x + stw, tw) ds \\ &= f(x) + tg(x, y) + t \int_0^1 (g(x + stw, w) - g(x, y)) ds. \end{aligned}$$

Define $r_y(\cdot, \cdot)$ by

$$r_y(t, w) = \int_0^1 (g(x + stw, w) - g(x, y)) ds.$$

Then it follows from the above expression for $f(x + tw)$ that it suffices to show that $r_y(\cdot, \cdot)$ approaches zero in the y -direction as $t \rightarrow 0$.

Define $h(\cdot, \cdot, \cdot)$ on an open subset of $R \times R \times V$ by $h(s, t, w) = g(x + stw, w) - g(x, y)$. Then $h(s, 0, y) = 0$ for all $s \in [0, 1]$. Let λ be a seminorm on Z , and let $\epsilon > 0$. Since $\lambda \circ h$ is continuous and $C = [0, 1] \times \{0\} \times \{y\}$ is compact, $\lambda \circ h$ is uniformly continuous on C . Thus there exists $\delta > 0$ and a convex neighborhood U of y such that $\lambda \circ h(s, t, w) < \epsilon$ for all $|t| < \delta$, $w \in W$, $s \in [0, 1]$. Since λ is a seminorm, this implies that, for $|t| < \delta$ and $w \in U$,

$$\lambda(r_y(t, w)) = \lambda\left(\int_0^1 h(s, t, w) ds\right) \leq \int_0^1 \lambda(h(s, t, w)) ds < \int_0^1 \epsilon ds = \epsilon,$$

which implies that $\lim_{t \rightarrow 0; w \rightarrow y} r_y(t, w) = 0$. Thus $Df(x)$ exists for each $x \in W$, and $Df(\cdot)(\cdot): W \times V \rightarrow Z$ is continuous. To complete the proof that Tf is continuous, it suffices to show that f is continuous.

Let $x \in W$. To see that f is continuous at x , let λ be a seminorm on Z and let $\epsilon > 0$. Choose a neighborhood \tilde{U} of $(x, 0)$ in $W \times V$ such that, for each $(v, w) \in \tilde{U}$, $\lambda(g(v, w)) < \epsilon$, and choose a convex neighborhood \tilde{W} of 0 such that $(x + \tilde{W}) \times \tilde{W} \subset \tilde{U}$. Then, for $y \in \tilde{W}$,

$$\begin{aligned} \lambda(f(x + y) - f(x)) &= \lambda\left(\int_0^1 g(x + ty, y) dt\right) \leq \int_0^1 \lambda(g(x + ty, y)) dt \\ &< \int_0^1 \epsilon dt = \epsilon. \end{aligned}$$

Thus f is continuous at x . Since x was an arbitrary point of W , we conclude that f is continuous. \square

We turn next to the definition of higher derivatives. We will need some additional notation, and then a formula (Lemma 2.2) which relates $T^k f$ and $D^k f$ for maps between Banach spaces which are k -times continuously Fréchet differentiable in the standard sense.

Let X be a locally convex space. For $n \geq 1$, let $p_n: X^n \rightarrow X$ be the projection onto the last factor, and let $i_n: X \rightarrow X^n$ be the injection into the first factor; i.e., $i_n(v) = (v, 0, 0, \dots, 0)$. Thus the map $i_1 \times i_{2^0} \times i_{2^1} \times \dots \times i_{2^{k-1}}$ is a linear injection of X^{k+1} into X^{2^k} .

For each $k \in \mathbb{N}$, let $\text{ev}: L^k(V, Z) \times V^k \rightarrow Z$ be the evaluation map, defined by $\text{ev}(g, w) = g(w)$ for each $g \in L^k(V, Z)$, $w \in V^k$.

2.2. LEMMA. *Assume that V and Z are Banach spaces, and that $f: W \rightarrow Z$ is C^k . Then*

$$p_{2^k} \circ T^k f \circ (i_1 \times i_{2^0} \times i_{2^1} \times \dots \times i_{2^{k-1}}) = \text{ev} \circ (D^k f \times \text{Id}_{V^k}): W \times V^k \rightarrow Z.$$

PROOF. See [3, Lemma 2.15]. \square

For each $x \in W$, let $j_x: V^k \rightarrow W \times V^k$ be given by $j_x(v_1, \dots, v_k) = (x, v_1, \dots, v_k)$ for each $v_1, \dots, v_k \in V$.

2.3. DEFINITION. Let $k > 1$. Assume that V and Z are locally convex spaces, and that $f: W \rightarrow Z$ is \mathcal{S}^k . For each $x \in W$, define $D^k f(x): V^k \rightarrow Z$ by

$$D^k f(x) = p_{2^k} \circ T^k f \circ (i_1 \times i_{2^0} \times \dots \times i_{2^{k-1}}) \circ j_x.$$

2.4. REMARK. It follows from Remark 1.7 and Lemma 2.2 that Definition 2.3 is equivalent to the standard definition of the k th order derivative if V and Z are finite-dimensional.

2.5. LEMMA. *Let $k > 1$, and assume that f is \mathcal{S}^k . Then, for each $x \in W$, $D^k f(x) \in L^k_s(V, Z)$. Furthermore, $\text{ev} \circ (D^k f \times \text{Id}_{V^k}): W \times V^k \rightarrow Z$ is continuous.*

PROOF. Let $x \in W$. Since $D^k f(x): V^k \rightarrow Z$ is defined as the composition of continuous maps, it clearly is continuous. Similarly, since $D^k f$ is defined so that the equation in Lemma 2.2 remains true for all locally convex spaces and all \mathcal{S}^k maps between them, continuity of the map $\text{ev} \circ (D^k f \times \text{Id}_{V^k})$ follows from continuity of the map $p_{2^k} \circ T^k f \circ (i_1 \times i_{2^0} \times \dots \times i_{2^{k-1}})$. So it only remains to show that $D^k f(x)$ is k -linear and symmetric. To see this, it suffices to show that $l \circ D^k f(x)|_{E^k}$ is k -linear and symmetric, where $E \subset V$ is a finite-dimensional subspace of V which contains x , and $l \in L(Z, R)$. So let $h: E \rightarrow V$ be the inclusion map. Then

$$\begin{aligned} l \circ p_{2^k} \circ T^k f \circ (i_1 \times i_{2^0} \times \dots \times i_{2^{k-1}}) \circ j_x \circ h^k & \\ &= p_{2^k} \circ l^{2^k} \circ T^k f \circ h^{2^k} \circ (i_1 \times i_{2^0} \times \dots \times i_{2^{k-1}}) \circ j_x \\ &= p_{2^k} \circ T^k (l \circ f \circ h) \circ (i_1 \times i_{2^0} \times \dots \times i_{2^{k-1}}) \circ j_x \\ &= \text{ev} \circ (D^k (l \circ f \circ h) \times \text{Id}_{E^k}) \circ j_x = D^k (l \circ f \circ h)(x). \end{aligned}$$

Since the domain and range of $l \circ f \circ h$ are finite-dimensional, $l \circ f \circ h$ is C^k by Remark 1.7, so Remark 2.4 implies that $D^k(l \circ f \circ h)(x)$ is k -linear and symmetric. Since $D^k(l \circ f \circ h)(x) = l \circ D^k f(x)|_E$, we are finished. \square

2.6. THEOREM. Assume that W is convex and $k \in \mathbb{N}$. Then the following are equivalent:

(i) f is \mathcal{S}^k .

(ii) For each $j = 1, \dots, k$, there exists $\theta^j f: W \rightarrow L^j(V, Z)$ such that $\text{ev} \circ (\theta^j f \times \text{Id}_{V^j}): W \times V^j \rightarrow Z$ is continuous and such that, for each $x, x + y \in W$ and $v_1, \dots, v_{j-1} \in V$, $\int_0^1 \theta^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt$ exists and

$$\begin{aligned} & \theta^{j-1} f(x + y)(v_1, \dots, v_{j-1}) - \theta^{j-1} f(x)(v_1, \dots, v_{j-1}) \\ &= \int_0^1 \theta^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt, \end{aligned}$$

where we let $\theta^0 f = f$.

Furthermore, if f satisfies (ii), then $\theta^j f = D^j f$ for $1 \leq j \leq k$.

PROOF. (i) \Rightarrow (ii) Assume that f is \mathcal{S}^k . Since each $\text{ev} \circ (D^j f \times \text{Id}_{V^j}): W \times V^j \rightarrow Z$ is continuous, the remarks on integration at the start of this section imply that it suffices to prove that (i) \Rightarrow (ii) for the map $\hat{f} = i \circ f$, where $i: Z \rightarrow \hat{Z}$ is the embedding of Z in its completion. Equivalently, we may simply assume that Z is complete.

So let us assume now that Z is complete. In this case, existence of the integral in (ii) is automatic, so we need only show that the equation in (ii) holds. To verify the equation, it suffices to show that

$$\begin{aligned} & l \circ D^{j-1} f(x + y)(v_1, \dots, v_{j-1}) - l \circ D^{j-1} f(x)(v_1, \dots, v_{j-1}) \\ &= \int_0^1 l \circ D^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt \end{aligned}$$

for each $l \in L(Z, R)$.

So let $l \in L(Z, R)$, and let E be a finite-dimensional subspace of V which contains $x, y, v_1, \dots, v_{j-1}$. Then

$$\begin{aligned} l \circ D^{j-1} f(x + y)(v_1, \dots, v_{j-1}) &= D^{j-1}(l \circ f \circ h)(x + y)(v_1, \dots, v_{j-1}), \\ l \circ D^{j-1} f(x)(v_1, \dots, v_{j-1}) &= D^{j-1}(l \circ f \circ h)(x)(v_1, \dots, v_{j-1}), \end{aligned}$$

and

$$l \circ D^j f(x + ty)(y, v_1, \dots, v_{j-1}) = D^j(l \circ f \circ h)(x + ty)(y, v_1, \dots, v_{j-1}).$$

Thus we are reduced to showing that

$$\begin{aligned} & D^{j-1}(l \circ f \circ h)(x + y)(v_1, \dots, v_{j-1}) - D^{j-1}(l \circ f \circ h)(x)(v_1, \dots, v_{j-1}) \\ (*) \quad &= \int_0^1 D^j(l \circ f \circ h)(x + ty)(y, v_1, \dots, v_{j-1}) dt. \end{aligned}$$

Now, the domain and range of $l \circ f \circ h$ are finite-dimensional, so $l \circ f \circ h$ is C^k by Remark 1.7. Thus (*) follows from Remark 2.4 together with the observation that, for a C^k map between Banach spaces, $D(D^{j-1} f) = D^j f$.

(ii) \Rightarrow (i) We prove this implication by induction on k . We already know it for $k = 1$ by Lemma 2.1. So let $k \in N$, and assume that the result has been established for \mathcal{S}^j maps with $j \leq k$. Assume that we are given $\theta^1, \dots, \theta^{k+1}$ as defined in statement (ii) of this theorem. By our induction hypothesis, it follows that f is \mathcal{S}^k and that $\theta^j f = D^j f$ for $j \leq k$.

For each $r = 0, \dots, k$, $x \in W$, $y \in V$, define $\phi^r f(x, y): (V \times V)^r \rightarrow Z \times Z$ by

$$\begin{aligned} \phi^r f(x, y)(v_1, w_1, \dots, v_r, w_r) = & \left(\theta^r f(x)(v_1, \dots, v_r), \theta^{r+1} f(x)(v_1, \dots, v_r, y) \right. \\ & \left. + \sum_{i=1}^r \theta^r f(x)(v_1, \dots, v_{i-1}, w_i, v_{i+1}, \dots, v_r) \right). \end{aligned}$$

From the definition, it follows that each $\phi^r(x, y)$ is continuous and r -linear from $(V \times V)^r$ to $Z \times Z$. In addition, the formula implies that each $\text{ev} \circ (\phi^r f \times \text{Id}_{(V \times V)^r}): (W \times V) \times (V \times V)^r \rightarrow Z \times Z$ is continuous. Note that, since $\phi^0 f(x, y) = (f(x), \theta^1 f(x)(y)) = (f(x), Df(x)(y))$, we have $\phi^0 f = Tf$. Thus, if we can show that, for each $1 \leq r \leq k$, $x, x + y \in W$, $u, z, v_1, w_1, \dots, v_{r-1}, w_{r-1} \in V$,

$$\int_0^1 \phi^r f(x + ty, u + tz)(y, z, v_1, w_1, \dots, v_{r-1}, w_{r-1}) dt$$

exists and

(**)

$$\begin{aligned} & \phi^{r-1} f(x + y, u + z)(v_1, w_1, \dots, v_{r-1}, w_{r-1}) - \phi^{r-1} f(x, u)(v_1, w_1, \dots, v_{r-1}, w_{r-1}) \\ & = \int_0^1 \phi^r f(x + ty, u + tz)(y, z, v_1, w_1, \dots, v_{r-1}, w_{r-1}) dt, \end{aligned}$$

then it will follow from our inductive hypothesis that Tf is \mathcal{S}^k . This will in turn imply that f is \mathcal{S}^{k+1} . From this, the result will follow for all $k \in N$ by an application of induction.

Let $i: Z \rightarrow \hat{Z}$ be the embedding of Z in its completion. Letting $\hat{f} = i \circ f$, $\theta^j \hat{f} = i \circ \theta^j f$, and $\phi^r \hat{f} = (i \times i) \circ \phi^r f$, it suffices by the remarks in the second paragraph of this section to verify (**) with $\phi^{r-1} f$ and $\phi^r f$ replaced by $\phi^{r-1} \hat{f}$ and $\phi^r \hat{f}$, respectively (since \hat{Z} is complete, it follows that

$$\int_0^1 \phi^r \hat{f}(x + ty, u + tz)(y, z, v_1, w_1, \dots, v_{r-1}, w_{r-1}) dt$$

automatically exists; then, it will follow from this modified version of (**) that the integral is actually a point in Z , from which it follows that the integral in the original version of (**) also exists).

To verify (**) for \hat{f} , it suffices to verify (**) with $\phi^{r-1} \hat{f}$ and $\phi^r \hat{f}$ replaced with $(l \times l) \circ \phi^{r-1} \hat{f} = \phi^{r-1}(l \circ \hat{f})$ and $(l \times l) \circ \phi^r \hat{f} = \phi^r(l \circ \hat{f})$ for each $l \in L(\hat{Z}, R) = L(Z, R)$. We verify this last equation by the now-familiar technique of reduction to finite-dimensional calculus.

Let $l \in L(Z, R)$, and let E be a finite-dimensional subspace of V which contains $x, y, u, z, v_1, w_1, \dots, v_{r-1}, w_{r-1}$. Then

$$\begin{aligned} &\phi^{r-1}(l \circ f)(x, y)(v_1, w_1, \dots, v_{r-1}, w_{r-1}) \\ &= \phi^{r-1}(l \circ f \circ h)(x, y)(v_1, w_1, \dots, v_{r-1}, w_{r-1}), \end{aligned}$$

and similarly for the other terms in (**). It follows that we are reduced to verifying that

$$\begin{aligned} &\phi^{r-1}(l \circ f \circ h)(x + y, u + z)(v_1, w_1, \dots, v_{r-1}, w_{r-1}) \\ (***) \quad &- \phi^{r-1}(l \circ f \circ h)(x, u)(v_1, w_1, \dots, v_{r-1}, w_{r-1}) \\ &= \int_0^1 \phi^r(l \circ f \circ h)(x + ty, u + tz)(y, z, v_1, w_1, \dots, v_{r-1}, w_{r-1}) dt, \end{aligned}$$

and this verification will be easy.

By Remark 1.7, $l \circ f \circ h$ is C^{k+1} , so $T(l \circ f \circ h)$ is C^k . It follows from the converse to the classical Mean Value Theorem applied to $D^k(l \circ f \circ h)$ and $D^{k+1}(l \circ f \circ h)$ that $D^{k+1}(l \circ f \circ h) = \theta^{k+1}(l \circ f \circ h)$. By [1, Lemma 6.9], $D^r(T(l \circ f \circ h)) = \phi^r(l \circ f \circ h)$ for $1 \leq r \leq k$. Since $D(D^{r-1}(T(l \circ f \circ h))) = D^r(T(l \circ f \circ h))$, (***) follows from the classical Mean Value Theorem for calculus in finite-dimensional spaces. Thus Tf is \mathcal{S}^k and f is \mathcal{S}^{k+1} . Furthermore, since $D^{k+1}(l \circ f \circ h) = \theta^{k+1}(l \circ f \circ h)$ for all $l \in L(Z, R)$ and all linear inclusions h of finite-dimensional spaces into V , we conclude that $D^{k+1}f = \theta^{k+1}f$. \square

We next examine the higher order differentiability of \mathcal{S}^1 inverse functions. We extend the definition of \mathcal{S}^k maps to the case $k = 0$ by calling f an \mathcal{S}^0 map if f is continuous. If X is a locally convex space, $n \in N$, then we let $\tilde{p}: X^n \rightarrow X$ denote the projection onto the first factor. Note that \tilde{p} is linear, hence \mathcal{S}^∞ .

2.7. LEMMA. *Let $k \in N \cup \{\infty\}$. Then f is $\mathcal{S}^k \Leftrightarrow Df(x)$ exists for each $x \in W$ and $\text{ev} \circ (Df \times \text{Id}_V)$ is \mathcal{S}^{k-1} .*

PROOF. (\Rightarrow) If f is \mathcal{S}^k then Tf is \mathcal{S}^{k-1} , which implies that $p_2 \circ Tf = \text{ev} \circ (Df \times \text{Id}_V)$ is \mathcal{S}^{k-1} .

(\Leftarrow) We prove this direction by induction on k . For $k = 1$, the result follows from Lemma 2.1. So assume that the result has been established for $k \in N$, and let f be a differentiable map such that $\text{ev} \circ (Df \times \text{Id}_V)$ is \mathcal{S}^k . Since $\text{ev} \circ (Df \times \text{Id}_V)$ is \mathcal{S}^{k-1} , it follows from our inductive assumption that f is \mathcal{S}^k , and hence that $Tf = (f \circ \tilde{p}, \text{ev} \circ (Df \times \text{Id}_V))$ is \mathcal{S}^k . Thus f is \mathcal{S}^{k+1} . By induction, the result is true for all $k \in N$. \square

2.8. LEMMA. *Let $g: W \rightarrow L(V, V)$. Assume that $g(x)$ is invertible for each $x \in W$, and that $\text{ev} \circ (g^{-1} \times \text{Id}_V): W \times V \rightarrow V$ is continuous. Assume in addition that $\text{ev} \circ (g \times \text{Id}_V): W \times V \rightarrow V$ is \mathcal{S}^1 . Then $\text{ev} \circ (g^{-1} \times \text{Id}_V): W \times V \rightarrow V$ is \mathcal{S}^1 .*

PROOF. Let $(x, y) \in W \times V, (u, z) \in V \times V$. Abbreviate $\text{ev} \circ (g \times \text{Id}_V)$ by h . Then, using the algebraic identity

$$g^{-1}(x + tv) - g^{-1}(x) = -g^{-1}(x + tv)(g(x + tv) - g(x))g^{-1}(x),$$

we see that

$$\begin{aligned}
 & \lim_{\substack{t \rightarrow 0 \\ (v,w) \rightarrow (u,z)}} \frac{g^{-1}(x+tw)(y+tw) - g^{-1}(x)(y)}{t} \\
 &= \lim_{\substack{t \rightarrow 0 \\ (v,w) \rightarrow (u,z)}} g^{-1}(x+tw)(w) + \lim_{\substack{t \rightarrow 0 \\ (v,w) \rightarrow (u,z)}} \frac{g^{-1}(x+tw)(y) - g^{-1}(x)(y)}{t} \\
 &= g^{-1}(x)(z) - \lim_{\substack{t \rightarrow 0 \\ (v,w) \rightarrow (u,z)}} g^{-1}(x+tw) \left(\frac{g(x+tw) - g(x)}{t} \right) g^{-1}(x)(y) \\
 &= g^{-1}(x)(z) - \lim_{\substack{t \rightarrow 0 \\ (v,w) \rightarrow (u,z)}} \left(\text{ev} \circ (g^{-1} \times \text{Id}_V) \right) \\
 &\quad \cdot \left(x+tw, \frac{h(x+tw, g^{-1}(x)(y)) - h(x, g^{-1}(x)(y))}{t} \right) \\
 &= g^{-1}(x)(z) - \left(\text{ev} \circ (g^{-1} \times \text{Id}_V) \right) (x, Dh(x, g^{-1}(x)(y))(u, 0)).
 \end{aligned}$$

This implies the $D(\text{ev} \circ (g^{-1} \times \text{Id}_V))(x, y)$ exists. Furthermore, the formula for $D(\text{ev} \circ (g^{-1} \times \text{Id}_V))(x, y)(u, z)$ implies that

$$\text{ev} \circ (D(\text{ev} \circ (g^{-1} \times \text{Id}_V)) \times \text{Id}_{V \times V}) : (W \times V) \times (V \times V) \rightarrow V$$

is continuous. Thus $\text{ev} \circ (g^{-1} \times \text{Id}_V)$ is \mathcal{S}^1 . \square

2.9. THEOREM. *Let $k \in N \cup \{\infty\}$, let W be open in V , and let $f: W \rightarrow V$ be an \mathcal{S}^k map. Assume that $U = f(W)$ is open in V , that f is injective, and that $f^{-1}: U \rightarrow V$ is \mathcal{S}^1 . Then f^{-1} is \mathcal{S}^k .*

PROOF. Let $k \in N, k \geq 2$, and assume that we have verified the theorem for \mathcal{S}^k maps. Let f be an \mathcal{S}^{k+1} map satisfying the hypothesis of the theorem. Then Tf is an \mathcal{S}^k map which satisfies the hypothesis of the theorem, so our inductive hypothesis implies that $(Tf)^{-1}$ is \mathcal{S}^k . Since $T(f^{-1}) = (Tf)^{-1}$, it follows that f^{-1} is \mathcal{S}^{k+1} . By induction, the theorem follows for all $k \in N$. So it only remains to verify the theorem for $k = 2$.

Let f be an \mathcal{S}^2 map which satisfies the hypothesis of the theorem. We need to show that f^{-1} is \mathcal{S}^2 , or equivalently that $T(f^{-1})$ is \mathcal{S}^1 . By Lemma 2.7, $T(f^{-1})$ is \mathcal{S}^1 if and only if $\text{ev} \circ (D(f^{-1}) \times \text{Id}_V)$ is \mathcal{S}^1 .

Since

$$\begin{aligned}
 D(f^{-1}) \times \text{Id}_V &= ((Df)^{-1} \circ f^{-1}) \times \text{Id}_V = ((Df)^{-1} \times \text{Id}_V) \circ (f^{-1} \times \text{Id}_V), \\
 \text{ev} \circ (D(f^{-1}) \times \text{Id}_V) &= \left(\text{ev} \circ ((Df)^{-1} \times \text{Id}_V) \right) \circ (f^{-1} \times \text{Id}_V).
 \end{aligned}$$

This factors $\text{ev} \circ (D(f^{-1}) \times \text{Id}_V)$ into the composition of two \mathcal{S}^1 maps, which implies that $\text{ev} \circ (D(f^{-1}) \times \text{Id}_V)$ is \mathcal{S}^1 . \square

3. Higher order derivatives and function spaces. We would like to introduce locally convex topologies on the spaces of linear and multilinear maps $L^r(V, Z)$ such that, for each \mathcal{S}^k map $f: W \rightarrow Z$ and each $1 \leq r \leq k$, $D^r f$ satisfies condition (ii) of Theorem 2.6 $\Leftrightarrow D^r f: W \rightarrow L^r(V, Z)$ is continuous. We will see in this section that it

is often possible to do this. Let $r \in N$. We will assume for the remainder of this article that $L^r(V, Z)$ has the compact-open topology, i.e. the topology of uniform convergence in Z on compact subsets of V^r .

The compact-open topology on spaces of linear and multilinear maps has two properties which interest us [7, Theorem 7.5]:

(i) $\text{ev}: L^r(V, Z) \times C \rightarrow Z$ is continuous for each compact subspace C of $L^r(V, Z)$.

(ii) If M is a topological space, $g: M \rightarrow L^r(V, Z)$ a map such that $\text{ev} \circ (g \times \text{Id}_{V^r}): M \times V^r \rightarrow Z$ is continuous, then $g: M \rightarrow L^r(V, Z)$ is continuous.

For $r \in N$, there is a natural linear inclusion of $L^{r+1}(V, Z)$ into $L(V, L^r(V, Z))$ which is defined as follows: for each $h \in L^{r+1}(V, Z)$, $v \in V$, let $h_v: V^r \rightarrow Z$ be the continuous r -linear map defined by $h_v(w) = h(v, w)$ for all $w \in V^r$. Define $\tilde{h}: V \rightarrow L^r(V, Z)$ by $\tilde{h}(v) = h_v$ for all $v \in V$. Note for future reference (Lemma 3.5) that h and \tilde{h} are related as follows: $h = \text{ev} \circ (\tilde{h} \times \text{Id}_{V^r})$. It follows as on [1, p. 47–48] (replacing the bounded sets in that argument with compact sets) that \tilde{h} is linear and continuous, and that the map which sends h to \tilde{h} is a linear topological embedding. Thus we have proved

3.1. LEMMA. *Let V and Z be locally convex spaces. Then $L^{r+1}(V, Z) \subset L(V, L^r(V, Z))$.*

Inductively extending the definition of this canonical linear embedding, we can extend Lemma 3.1 in a corresponding manner to prove the following

3.2. LEMMA. *Let V and Z be locally convex spaces, $r \in N, r \geq 2$. Then*

$$L^r(V, Z) \subset \underbrace{L(V, L(V, \dots, L(V, Z), \dots))}_{r \text{ times}}$$

It is convenient to work with locally convex spaces for which the canonical embedding in Lemma 3.2 is a linear topological isomorphism (i.e. for which the embedding is an onto map). One large class of spaces which have this property was introduced in [1, Definition 3.39]. Recall [9] that a *compactly generated* space M (also known as a Hausdorff k -space [7]) is a Hausdorff topological space such that F is closed in $M \Leftrightarrow F \cap C$ is closed for each compact set C in M . It is well known that open subsets of compactly generated regular spaces are compactly generated in the subspace topology. It is also straightforward to show that linear topological spaces are regular. Thus open subspaces of compactly generated linear topological spaces are compactly generated.

3.3. DEFINITION. An exponential space V is a locally convex space such that V^r is compactly generated for each $r \in N$.

3.4. REMARKS. Let V be a locally convex space, and Z a closed subspace of V .

(a) If V is an exponential space, then Z and V/Z are exponential spaces.

(b) If V is metrizable, then V is an exponential space.

(c) If V is an exponential space, then V^r is an exponential space for each $r \in N$.

(d) If W is an open subspace of V and V is an exponential space, then $W \times V^r$ is compactly generated for each $r \in N$.

The key property of compactly generated spaces is the following: if M, M_1 are topological spaces, and M is compactly generated, then a map $h: M \rightarrow M_1$ is continuous $\Leftrightarrow h|_C: C \rightarrow M_1$ is continuous for each compact subspace $C \subset M$.

3.5. LEMMA. *Let V be an exponential space, Z a locally convex space, W open in V , and $r \in \mathbb{N}$. Then a map $g: W \rightarrow L^r(V, Z)$ is continuous $\Leftrightarrow \text{ev} \circ (g \times \text{Id}_{V^r}): W \times V^r \rightarrow Z$ is continuous.*

PROOF. (\Rightarrow) Since $W \times V^r$ is compactly generated, it suffices to check the continuity of $\text{ev} \circ (g \times \text{Id}_{V^r})$ on compact subsets of $W \times V^r$ of the form $C_1 \times C$, where C_1 is compact in W and C is compact in V^r . So let C_1 be compact in W and let C be compact in V^r . It is clear that $g \times \text{Id}_{V^r}$ is continuous from $C_1 \times C$ to $L^r(V, Z) \times C$. Since $\text{ev}|_{L^r(V, Z) \times C}$ is continuous, the proof of this implication is concluded.

(\Leftarrow) This implication follows from property (ii) of the compact-open topology, and is true regardless of whether or not V is an exponential space. \square

3.6. LEMMA. *Let V be an exponential space, Z a locally convex space, $r \in \mathbb{N}, r \geq 2$. Then*

$$\underbrace{L(V, L(V, \dots, L(V, Z), \dots))}_{r \text{ times}} = L^r(V, Z).$$

PROOF. It is sufficient to show that $L(V, L^{r-1}(V, Z)) = L^r(V, Z)$, from which the result will follow by induction. To verify this equality, it suffices because of the remarks in the proof of Lemma 3.1 to verify that $\text{ev} \circ (\tilde{h} \times \text{Id}_{V^{r-1}}) \in L^r(V, Z)$ for each $\tilde{h} \in L(V, L^{r-1}(V, Z))$. Since Lemma 3.5 implies that $\text{ev} \circ (\tilde{h} \times \text{Id}_{V^{r-1}})$ is continuous, and since $\text{ev} \circ (\tilde{h} \times \text{Id}_{V^{r-1}})$ is clearly r -linear, we are done. \square

3.7. LEMMA. *Let V be an exponential space, Z a locally convex space, $r \in \mathbb{N}$. If Z is complete (resp. quasi-complete, sequentially complete), then $L^r(V, Z)$ is complete (resp. quasi-complete, sequentially complete).*

PROOF. Assume that Z is complete. We will show that $L^r(V, Z)$ is complete (the proofs of the other cases follow by essentially the same argument). Let $\{h_\lambda: \lambda \in \Lambda\}$ be a Cauchy net in $L^r(V, Z)$. Then $\{h_\lambda(v_1, \dots, v_r): \lambda \in \Lambda\}$ is a Cauchy net in Z for each $v_1, \dots, v_r \in V$. Since Z is complete, we can define $h: V^r \rightarrow Z$ by $h(v_1, \dots, v_r) = \lim_{\lambda \in \Lambda} h_\lambda(v_1, \dots, v_r)$. Since the convergence of $\{h_\lambda\}$ to h is at least as strong as pointwise convergence, it follows that h is r -linear. Since this convergence is uniform on compact sets, it follows that h is continuous on compact subsets of V^r . Since V^r is compactly generated, it follows that $h: V^r \rightarrow Z$ is continuous, and hence that $h \in L^r(V, Z)$. Finally, again since the convergence of $\{h_\lambda\}$ to h is uniform on compact subsets, $\lim_{\lambda \in \Lambda} (h_\lambda - h) = 0$, which implies that $L^r(V, Z)$ is complete. \square

3.8. REMARK. Let V, Z be locally convex spaces, $l \in L(V, Z)$. $\alpha: [0, 1] \rightarrow V$ a continuous curve in V . If $\int_0^1 \alpha(t) dt$ exists, then $\int_0^1 l \circ \alpha(t) dt$ exists, and $\int_0^1 l \circ \alpha(t) dt = l(\int_0^1 \alpha(t) dt)$.

3.9. THEOREM. Let $k \in N$, and assume that V is an exponential space. Then the following are equivalent:

(i) f is \mathcal{S}^k .

(ii) For each $j = 1, \dots, k$, there exists a continuous map $\theta^j f: W \rightarrow L^j(V, Z)$ such that $f = \theta^0 f, \theta^1 f, \dots, \theta^{k-1} f$ are \mathcal{S}^1 and $D(\theta^{j-1} f) = \theta^j f, 1 \leq j \leq k$.

Furthermore, if (ii) holds, then $D^j f = \theta^j f$ for $1 \leq j \leq k$.

PROOF. It suffices to prove this theorem for the case when W is convex, so we will make this additional assumption throughout the proof.

(\Rightarrow) Assume that f is \mathcal{S}^k , and let $j \in \{1, \dots, k\}$. By Theorem 2.6, we see that $\text{ev} \circ (D^j f \times \text{Id}_{V^j}): W \times V^j \rightarrow Z$ is continuous, $\int_0^1 D^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt$ exists for each $x, x + y \in W, v_1, \dots, v_{j-1} \in V$, and

$$\begin{aligned}
 & D^{j-1} f(x + y)(v_1, \dots, v_{j-1}) - D^{j-1} f(x)(v_1, \dots, v_{j-1}) \\
 (*) \quad & = \int_0^1 D^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt.
 \end{aligned}$$

Since $\text{ev} \circ (D^j f \times \text{Id}_{V^j})$ is continuous, $D^j f: W \rightarrow L^j(V, Z)$ is continuous. Let $i: Z \rightarrow \hat{Z}$ be the embedding of Z in its completion. Then $D^j \hat{f} = i \circ D^j f: W \rightarrow L^j(V, \hat{Z}), 1 \leq j \leq k$. Since $L^j(V, \hat{Z})$ is complete, it follows that $\int_0^1 D^j \hat{f}(x + ty)(y) dt$ exists for each $x, x + y \in W$. Furthermore, it follows from (*) that, for each $x, x + y \in W$,

$$D^{j-1} \hat{f}(x + y) - D^{j-1} \hat{f}(x) = \int_0^1 D^j \hat{f}(x + ty)(y) dt.$$

From the remarks at the beginning of §2, it follows that $\int_0^1 D^j \hat{f}(x + ty)(y) dt$ exists and that

$$D^{j-1} \hat{f}(x + y) - D^{j-1} \hat{f}(x) = \int_0^1 D^j f(x + ty)(y) dt.$$

Thus the converse to the Mean Value Theorem (Lemma 2.1) implies that each $D^{j-1} f$ is \mathcal{S}^1 and that $D(D^{j-1} f) = D^j f$.

(\Leftarrow) Assume that we are given $\theta^1 f, \dots, \theta^k f$ as defined in statement (ii) of this theorem. Let $j \in \{1, \dots, k\}$ and $v_1, \dots, v_{j-1} \in V$, and define $l: L^{j-1}(V, Z) \rightarrow Z$ by $l(h) = h(v_1, \dots, v_{j-1})$ for each $h \in L^{j-1}(V, Z)$. Clearly l is linear and continuous. Thus, for each $x, x + y \in W$,

$$\begin{aligned}
 & \theta^{j-1} f(x + y)(v_1, \dots, v_{j-1}) - \theta^{j-1} f(x)(v_1, \dots, v_{j-1}) \\
 & = l(\theta^{j-1} f(x + y)) - l(\theta^{j-1} f(x)) = l(\theta^{j-1} f(x + y) - \theta^{j-1} f(x)) \\
 & = l\left(\int_0^1 \theta^j f(x + ty)(y) dt\right) = \int_0^1 l \circ \theta^j f(x + ty)(y) dt \\
 & = \int_0^1 \theta^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt.
 \end{aligned}$$

Since each $\text{ev} \circ (\theta^j f \times \text{Id}_{V^j}): W \times V^j \rightarrow Z$ is continuous, Theorem 2.6 implies that f is \mathcal{S}^k , and that $D^j f = \theta^j f$ for $1 \leq j \leq k$. \square

3.10. COROLLARY. Assume that V is an exponential space, and let $j, k \in \mathbb{N}$, $j \leq k$. Then f is $\mathcal{S}^k \Leftrightarrow f$ is \mathcal{S}^j and $D^j f$ is \mathcal{S}^{k-j} . Furthermore, if f is \mathcal{S}^k , then $D^r(D^j f) = D^{r+j} f$ for $1 \leq r \leq k - j$.

PROOF. This is a triviality if $j = k$, so assume throughout this proof that $j < k$.

(\Rightarrow) Assume that f is \mathcal{S}^k . Since $j < k$, f is clearly \mathcal{S}^j . By Theorem 3.9, $D^j f, \dots, D^{k-1} f$ are \mathcal{S}^1 maps, and $D(D^r f) = D^{r+1} f$ for $j \leq r \leq k - 1$. Thus Theorem 3.9 implies that $D^j f$ is \mathcal{S}^{k-j} and that $D^r(D^j f) = D^{r+j} f$ for $1 \leq r \leq k - j$.

(\Leftarrow) Assume that f is \mathcal{S}^j and that $D^j f$ is \mathcal{S}^{k-j} . Define $\theta^r f: W \rightarrow L^r(V, Z)$ for $1 \leq j \leq k$ by $\theta^r f = D^r f$ for $1 \leq r \leq j$ and $\theta^r f = D^{r-j}(D^j f)$ for $j + 1 \leq r \leq k$. Let $\theta^0 f = f$. By Theorem 3.9, each $\theta^r f$ is \mathcal{S}^1 for $0 \leq r \leq k - 1$, and $D(\theta^r f) = \theta^{r+1} f$. Thus another application of Theorem 3.9 implies that f is \mathcal{S}^k , and that $D^r f = \theta^r f$ for $1 \leq r \leq k$. In particular, $D^r f = D^{r-j}(D^j f)$ for $j + 1 \leq r \leq k$. \square

We conclude this section with an observation about the compact-open topology on the spaces $L^r(V, Z)$, $r \in \mathbb{N}$. The only properties of this topology which we have used are properties (i) and (ii) from the second paragraph of this section. Thus any other topologies on these spaces which satisfy properties (i) and (ii) would be equally suitable for developing the material in this section.

We assert that an example of such an alternative function space topology is the precompact-open topology, i.e. the topology of uniform convergence in Z on precompact subsets of V^r . If V is complete, then precompact subsets of V^r are compact, and the precompact-open topology on $L^r(V, Z)$ coincides with the compact-open topology. However, in general, the precompact-open topology will be finer. Since every topology on $L^r(V, Z)$ which is at least as fine as the compact-open topology has property (i), the precompact-open topology clearly possesses this property. From [1, Remark 3.45] and the first half of the proof of [1, Theorem 3.46], it follows that the precompact-open topology also possesses property (ii).

When this section was being written, the author had difficulty deciding whether to give $L^r(V, Z)$ the compact-open topology or the precompact-open topology. The decision to go with the compact-open topology was made because this topology is more familiar to most mathematicians, and because it is discussed more frequently in the mathematical literature (e.g. [7, 9]). However, a final decision on which topology is more suitable for differential calculus will have to await the study of the function space properties of derivatives of functions on locally convex spaces other than exponential spaces.

4. Differential calculus in Banach spaces. The theory of \mathcal{S}^k differentiability for the case of maps between Banach spaces was developed in [3]. In this section we review the relevant conclusions from that article to provide motivation for the development in the next section of a theory of C^k differentiability in the general locally convex category. In addition, we include some results on inverse functions not covered in [3] ([3] was concerned exclusively with the study of local flows for partial differential operators). We assume throughout this section that V and Z are Banach spaces. As before, W is assumed to be an open subset of V . We also assume that the reader is familiar with the conventional theory of C^k differentiation for

maps between Banach spaces as developed in numerous standard references (e.g. [8]).

A look at [3] will reveal that the definition of \mathcal{S}^1 differentiability given in that reference is not the same as the one in §1 of the present article. However, we assert that the two definitions are equivalent. To see the equivalence, we refer to [3, Lemma 2.17]. This lemma states that, under the additional assumption that W is convex, $f: W \rightarrow Z$ is \mathcal{S}^1 in the sense of [3, Definition 2.2] \Leftrightarrow there exists a continuous map $g: W \times V \rightarrow Z$ such that $g(x, \cdot): V \rightarrow Z$ is linear for each $x \in W$ and such that, for each $x, x + y \in W$, $f(x + y) - f(x) = \int_0^1 g(x + ty, y) dt$. Since this is the same as the condition in Lemma 2.1 of the present article, we conclude that the two definitions are equivalent. Since both articles define \mathcal{S}^k differentiability for $k \geq 2$ in the same way in terms of \mathcal{S}^1 differentiability, it follows that the two definitions of \mathcal{S}^k differentiability are equivalent.

Letting $\mathcal{S}^k(W, Z)$ denote the set of \mathcal{S}^k maps from W to Z , and $C^k(W, Z)$ denote the set of C^k maps from W to Z , the following set of inclusions was established in [3]:

$$\begin{aligned} \mathcal{S}^0(W, Z) &= C^0(W, Z) \supset \mathcal{S}^1(W, Z) \supset C^1(W, Z) \supset \mathcal{S}^2(W, Z) \supset C^2(W, Z) \\ &\supset \dots \supset C^k(W, Z) \supset \mathcal{S}^{k+1}(W, Z) \supset C^{k+1}(W, Z) \supset \dots \supset \mathcal{S}^\infty(W, Z) \\ &= C^\infty(W, Z). \end{aligned}$$

This is the fundamental relation between \mathcal{S}^k and C^k differentiability, and in the next section it will be shown to hold for general locally convex spaces. The significance of this relation in the general locally convex case, and in particular of the equality $C^\infty(W, Z) = \mathcal{S}^\infty(W, Z)$, is that, when combined with the results of §1, it shows that C^k differentiability theory is rich with examples, a fact not immediately apparent from the definition.

We next examine the existence of local flows for vector fields and the inverse function theorem in terms of \mathcal{S}^k differentiability.

4.1. DEFINITION. Let $f: W \rightarrow V$ be continuous.

(i) A continuous curve $\alpha: (\delta, \epsilon) \rightarrow W$, $\delta, \epsilon \in R$, $\delta < 0 < \epsilon$, is called an integral curve for f if α is differentiable on (δ, ϵ) and $\dot{\alpha}(t) = f(\alpha(t))$ for each $t \in (\delta, \epsilon)$. The point $\alpha(0)$ is called the initial value of α .

(ii) f will be said to have unique integral curves if, for each $w \in W$, there exists at most one maximally defined integral curve for f with initial value x .

4.2. DEFINITION. Let $f: W \rightarrow V$ be continuous. A local flow for f is a continuous map $F: U \rightarrow W$, where U is open in $R \times W$, such that $U \cap (\{0\} \times W) = \{0\} \times W$ and such that, for each $x \in W$, $F(\cdot, x)$ is an integral curve for f with initial value x .

4.3. THEOREM. Let $k \in N \cup \{\infty\}$, and assume that $f: W \rightarrow V$ is an \mathcal{S}^k vector field. Then f has unique integral curves, and f generates an \mathcal{S}^k local flow.

PROOF. \mathcal{S}^k vector fields are locally uniformly Lipschitz, so the uniqueness of integral curves and existence of a continuous flow are immediate. That the flow is \mathcal{S}^1 was established in [2, Theorem 8.1]. That the flow is \mathcal{S}^k was observed in [3, Theorem 2.27]. \square

4.4. THEOREM. Let $k \in N \cup \{\infty\}$, $k \geq 2$, and let $f: W \rightarrow V$ be \mathcal{S}^k . Assume that $x \in W$ is a point for which $Df(x): V \rightarrow V$ is invertible. Then there exists an open neighborhood U of x , $U \subset W$, such that $f|_U$ is one-to-one, $W_1 = f(U)$ is open, and $(f|_U)^{-1}: W_1 \rightarrow V$ is \mathcal{S}^k .

PROOF. Since f is \mathcal{S}^2 , f is C^1 . By the standard inverse function theorem for C^1 maps, there exists a neighborhood U of x , $U \subset W$, such that $f|_U$ is one-to-one, $W_1 = f(U)$ is open, and $(f|_U)^{-1}: W_1 \rightarrow V$ is C^1 . Since $(f|_U)^{-1}$ is C^1 , it is \mathcal{S}^1 . Thus Theorem 2.9 implies that $(f|_U)^{-1}$ is \mathcal{S}^k . \square

The inverse function theorem is not valid for \mathcal{S}^1 maps. We demonstrate this by presenting an example of an \mathcal{S}^1 map f from a separable infinite-dimensional Hilbert space H to itself such that $f(0) = 0$ and such that $Df(0)$ is invertible, but such that $f(H)$ does not contain a neighborhood of zero.

4.5. EXAMPLE. Let H be a separable, infinite-dimensional Hilbert space, with orthonormal basis $\{e_j: j \in N\}$. Let $\rho: R \rightarrow [0, 1]$ be a smooth nondecreasing function such that $\rho(t) = t$ for all $t \leq 0$ and $\rho(t) = 1$ for all $t \geq 1$. For each $x = (x^1, x^2, \dots) \in H$, define $f(x) = \sum_{j=1}^\infty (1/j)\rho(jx^j)e_j$. Then f is \mathcal{S}^1 , $f(0) = 0$, and $Df(0) = Id_H$. However, $2e_j/j \notin f(H)$ for all $j \in N$, which implies that $f(H)$ does not contain a neighborhood of 0.

PROOF. It is clear that $f(0) = 0$. Furthermore, for each $j \in N$, $x \in H$,

$$\langle e_j, f(x) \rangle = \left\langle e_j, \sum_{k=1}^\infty \frac{\rho(kx^k)e_k}{k} \right\rangle = \frac{\rho(jx^j)}{j} \leq \frac{1}{j},$$

which implies that $2e_j/j \notin f(H)$. So it only remains to show that f is \mathcal{S}^1 and that $Df(0) = Id_H$.

To see that f is \mathcal{S}^1 , choose $c > 0$ such that $|\dot{\rho}(t)| \leq c$ for all $t \in R$. For each $x = (x^1, x^2, \dots) \in H$, $y = (y^1, y^2, \dots) \in H$, define $g(x, y) \in H$ by $\langle e_j, g(x, y) \rangle = \dot{\rho}(jx^j)y^j$. It is clear that each $g(x, \cdot)$ is a bounded linear operator on H such that $\|g(x, \cdot)\|_{H,H} \leq c$, and it is easily verified that $ev \circ (g \times Id_H): H \times H \rightarrow H$ is continuous. Furthermore, for each $j \in N$, $x, y \in H$,

$$\begin{aligned} \langle e_j, f(x + y) - f(x) \rangle &= (1/j)(\rho(jx^j + jy^j) - \rho(jx^j)) \\ &= \int_0^1 \dot{\rho}(jx^j + tjy^j)y^j dt = \int_0^1 \langle e_j, g(x + ty, y) \rangle dt, \end{aligned}$$

which implies that $f(x + y) - f(x) = \int_0^1 g(x + ty, y) dt$. By Lemma 2.1, f is \mathcal{S}^1 and $Df(x) = g(x, \cdot)$ for each $x \in H$.

Finally, note that, for each $j \in N$, $\langle e_j, g(0, y) \rangle = \dot{\rho}(0)y^j = y^j$, which implies that $g(0, y) = y$ for all $y \in H$, i.e. $Df(0) = Id_H$. \square

The basic reason the inverse function theorem is valid for \mathcal{S}^k maps between Banach spaces with $k \geq 2$ is that the existence of a continuous second derivative for f regulates the behavior of the first derivative enough to make the first derivative continuous into $L(V, Z)$ with the uniform operator topology. This notion of the $(j + 1)$ st derivative regulating the behavior of the j th derivative is central to the

relation between \mathcal{S}^k and C^k maps (see the proof of [3, Theorem 2.16]). It also provides us with a clue about how to define C^k differentiability in the next section so that the relation remains valid for maps between general locally convex spaces.

5. Basic properties of C^k calculus. In this section we again assume that $V, \tilde{V}, Z, \tilde{Z}, X, \tilde{X}$, are locally convex spaces, and that W is open in V . We will refer extensively to seminorms on these spaces. All seminorms which are introduced are assumed to be continuous on their respective domains of definition (we assume that the reader is familiar with the basic properties of continuous seminorms).

Let λ be a seminorm on X . We let $N_\lambda = \{x \in X: \lambda(x) = 0\}$. Note that N_λ is a closed linear subspace of X . We define X_λ to be the normed space whose underlying space is X/N_λ , equipped with the norm induced by λ . We let $p_\lambda: X \rightarrow X_\lambda$ denote the induced projection of X onto X_λ .

Let E be a normed space, $r \in N$, and consider the space $L^r(X, E)$. For each seminorm ν on X , define a subadditive function $\tilde{\nu}: L^r(X, E) \rightarrow R \cup \{\infty\}$ by

$$\tilde{\nu}(h) = \sup\{\|h(x_1, \dots, x_r)\|: \nu(x_i) \leq 1, 1 \leq i \leq r\}.$$

Note that $\tilde{\nu}(h) = 0 \Leftrightarrow h = 0$.

5.1. DEFINITION. For each $r \in N$, $L'_\nu(X, E)$ is the normed linear space $\{h \in L^r(X, E): \tilde{\nu}(h) < \infty\}$, with $\tilde{\nu}$ as the norm.

5.2. REMARKS. (a) The inclusion $L'_\nu(X, E) \rightarrow L^r(X, E)$ is continuous.

(b) For each $h \in L^r(X, E)$, there exists $\nu = \nu(h)$ such that $h \in L'_\nu(X, E)$.

For each seminorm ν on X , $r \in N$, let $L'_b(X_\nu, E)$ be the normed linear space of continuous r -linear maps from X_ν to E , with the norm of uniform convergence in the E -norm on $\{(x_1, \dots, x_r) \in (X_\nu)^r: \|x_j\| \leq 1\}$. Note that there exists a canonical linear map $L'_b(X_\nu, E) \rightarrow L'_\nu(X, E)$, $h \rightarrow \tilde{h}$ defined by $\tilde{h}(x_1, \dots, x_r) = h(p_\nu x_1, \dots, p_\nu x_r)$ for each $h \in L^r(X_\nu, E)$, $x_1, \dots, x_r \in X$. It is obvious that $\|\tilde{h}\| = \|h\|$.

5.3. LEMMA. *The canonical map $L'_b(X_\nu, E) \rightarrow L'_\nu(X, E)$ is an isomorphism of normed linear spaces.*

PROOF. See [1, Lemma 3.20]. \square

5.4. DEFINITION. Let $g: W \rightarrow L^r(X, \tilde{X})$. We will say that g is a D -map if, for each seminorm λ on \tilde{X} and $x \in W$, there exists a neighborhood $U = U(\lambda, x) \subset W$ of x and a seminorm ν on X such that $p_\lambda \circ g(U) \subset L'_\nu(X, \tilde{X}_\lambda)$ and such that the map $g_\lambda = p_\lambda \circ g: U \rightarrow L'_\nu(X, \tilde{X}_\lambda)$ is continuous.

5.5. REMARKS. Let $g, h: W \rightarrow L^r(X, \tilde{X})$ be D -maps, $\text{sym}: L^r(X, \tilde{X}) \rightarrow L'_s(X, \tilde{X})$ the (linear) symmetrization map, U open in \tilde{V} , and $\beta: U \rightarrow W$ a continuous map. Then

(a) $g + h: W \rightarrow L^r(X, \tilde{X})$ is a D -map,

(b) $g \circ \beta: U \rightarrow L^r(X, \tilde{X})$ is a D -map.

(c) $\text{sym} \circ g: W \rightarrow L'_s(X, \tilde{X})$ is a D -map.

(d) If $\psi: W \rightarrow L^r(X, \tilde{X})$ is a constant map (i.e. there exists $\sigma \in L^r(X, \tilde{X})$ such that $\psi(x) = \sigma$ for all $x \in W$), then ψ is a D -map.

5.6. LEMMA. Let $g: W \rightarrow L^r(X, \tilde{X})$ be a D -map. Assume that, for each $j = 1, \dots, r$, $h_j: W \rightarrow L^{k_j}(Z, X)$ is a D -map. Then the map

$$g \cdot (h_1 \times \dots \times h_r): W \rightarrow L^{k_1 + \dots + k_r}(Z, \tilde{X})$$

$$x \rightarrow g(x) \circ (h_1(x) \times \dots \times h_r(x))$$

is a D -map.

PROOF. Let λ be a seminorm on X , and let $x \in W$. Then there exists a neighborhood \tilde{W} of x in W and a seminorm ν on X such that $g_\lambda(\tilde{W}) \subset L'_\nu(X, \tilde{X}_\lambda) = L'_b(X_\nu, \tilde{X}_\lambda)$ and such that $g: \tilde{W} \rightarrow L'_b(X_\nu, \tilde{X}_\lambda)$ is continuous. For each $j \in 1, \dots, r$, choose a neighborhood U_j of x in W and a seminorm μ_j on Z such that $h_{j\nu}(U_j) \subset L^{k_j}_{\mu_j}(Z, X_\nu) = L^{k_j}_b(Z_{\mu_j}, X_\nu)$ and $h_{j\nu}: U_j \rightarrow L^{k_j}_{\mu_j}(Z, X_\nu)$ is continuous. Let $\mu = \mu_1 + \dots + \mu_r$. Note that, for each j , there is an induced continuous linear injection $L^{k_j}_{\mu_j}(Z, X_\nu) \rightarrow L^k_\mu(Z, X_\nu)$. Thus, for each j , $h_{j\nu}(U_j) \subset L^k_\mu(Z, X_\nu)$, and $h_{j\nu}: U_j \rightarrow L^k_\mu(Z, X_\nu)$ is continuous.

Let $U = \tilde{W} \cap U_1 \cap \dots \cap U_r$. Then, on U , we have the following factorization for $p_\lambda \circ (g \cdot (h_1 \times \dots \times h_r)) = g_\lambda \cdot (h_1 \times \dots \times h_r)$:

$$U \xrightarrow{g_\lambda \cdot (h_1 \times \dots \times h_r)} L'_\nu(X, \tilde{X}_\lambda) \times (L^{k_1}_\mu(Z, X_\nu) \times \dots \times L^{k_r}_\mu(Z, X_\nu))$$

$$= L'_b(X_\nu, \tilde{X}_\lambda) \times (L^{k_1}_b(Z_\mu, X_\nu) \times \dots \times L^{k_r}_b(Z_\mu, X_\nu)) \xrightarrow{\text{comp}} L^{k_1 + \dots + k_r}_b(Z_\mu, \tilde{X}_\lambda)$$

$$= L^{k_1 + \dots + k_r}_\mu(Z, \tilde{X}_\lambda).$$

Since Z_μ, X_ν, X_λ are normed spaces, the composition map in this factorization is continuous. Thus $g_\lambda \cdot (h_1 \times \dots \times h_r): U \rightarrow L^{k_1 + \dots + k_r}_\mu(Z, \tilde{X}_\lambda)$ is continuous. \square

The following proposition implies that D -map continuity is at least as strong as the continuity assumed for derivatives in the definition of \mathcal{S}^k differentiability. It can be proved via an obvious modification of the factorization used in the proof of Lemma 5.6. Details are left to the reader.

5.7. PROPOSITION. Let $g: W \rightarrow L^r(X, \tilde{X})$ be a D -map. Then $\text{ev} \circ (g \times \text{Id}_{X^r}): W \times X^r \rightarrow \tilde{X}$ is continuous.

5.8. DEFINITION. Let $k \in N \cup \{\infty\}$. We will say that $f: W \rightarrow Z$ is C^k if f is \mathcal{S}^k and $D^j f: W \rightarrow L^j(V, Z)$ is a D -map for each $1 \leq j \leq k$.

Note that, if f is C^k , then f is C^j for $0 \leq j \leq k$.

5.9. THEOREM (THE CHAIN RULE). Let $k \in N \cup \{\infty\}$, and assume that U is open in Z . Let $f: W \rightarrow Z$ and $g: U \rightarrow \tilde{Z}$ be C^k maps such that $f(W) \subset U$. Then $g \circ f$ is C^k .

PROOF. The chain rule for \mathcal{S}^k differentiability implies that $g \circ f$ is \mathcal{S}^k , so it only remains to verify that each $D^j(g \circ f)$ is a D -map for $1 \leq j \leq k$. Since

$$D^j(g \circ f) = \text{sym} \left(\sum_{i=1}^j \sum_{\substack{r_1 + \dots + r_i = j \\ 1 \leq r_1 \leq \dots \leq r_i}} \frac{i!}{r_1! \dots r_i!} ((D^i g) \circ f) \cdot (D^{r_1} f \times \dots \times D^{r_i} f) \right),$$

this verification follows immediately from Remark 5.5 and Lemma 5.6. \square

5.10. THEOREM. Let $k \in N \cup \{\infty\}$, let W be open in V , and let $f: W \rightarrow V$ be a C^k map. Assume that $U = f(W)$ is open in V , that f is injective, and that $f^{-1}: U \rightarrow V$ is C^1 . Then f^{-1} is C^k .

PROOF. The proof of this result is basically the same as the proof of Theorem 5.9. The theorem is a tautology for $k = 1$, so assume that $k \geq 2$. It follows from Theorem 2.9 that f^{-1} is \mathcal{S}^k , so we need only verify that each $D^j(f^{-1})$ is a D -map for $1 \leq j \leq k$. To see this, note first that, since f^{-1} is C^1 , $D(f^{-1})$ is a D -map. It follows from the chain rule applied to $f \circ f^{-1}$ that

$$D^2(f^{-1}) = -D(f^{-1}) \cdot ((D^2f) \circ f^{-1}) \cdot (D(f^{-1}) \times D(f^{-1})),$$

so Remark 5.5 and Lemma 5.6 imply that $D^2(f^{-1})$ is a D -map. For $j > 2$, $D^j(f^{-1})$ is the symmetrized sum of terms of the form

$$D(f^{-1}) \cdot ((D^qf) \circ f^{-1}) \cdot (D^{r_1}(f^{-1}) \times \cdots \times D^{r_q}(f^{-1})),$$

where $2 \leq q \leq j$, $1 \leq r_m$ for each $1 \leq m \leq q$, $r_1 + \cdots + r_q = j$, and terms of the form

$$D^i(f^{-1}) \cdot (\psi \times (((D^qf) \circ f^{-1}) \cdot (D^{r_1}(f^{-1}) \times \cdots \times D^{r_q}(f^{-1}))))),$$

where $2 \leq i \leq j - 1$, $2 \leq q \leq j$, $1 \leq r_m$ for each $1 \leq m \leq q$, $i + r_1 + \cdots + r_q = j + 1$, and $\psi: W \rightarrow (L(V, V))^{i-1}$ is given by $\psi(x) = (\text{Id}_V)^{i-1}$ for all $x \in W$. Since this implies that each $r_m \leq j - 1$, the result follows from Remark 5.5 and Lemma 5.6 by induction on j . \square

5.11. THEOREM. Let $k \in N$, and assume that $f: W \rightarrow Z$.

- (i) If f is C^k , then f is \mathcal{S}^k .
- (ii) If f is \mathcal{S}^k , then f is C^{k-1} .

PROOF. (i) follows from the definition of C^k differentiability. To see that (ii) holds, assume that f is \mathcal{S}^k . If $k = 1$, then f is \mathcal{S}^1 , and hence is $\mathcal{S}^0 = C^0$ (i.e. continuous). Thus we are reduced to considering the case $k \geq 2$, and to proving that each $D^j f$ is a D -map for $1 \leq j \leq k - 1$.

It suffices to prove this last statement under the additional assumption that W is convex. So assume that $j \in \{1, \dots, k - 1\}$, let $x \in W$, and let λ be a seminorm on Z . Since $h = p_\lambda \circ \text{ev} \circ (D^{j+1}f \times \text{Id}_{V^{j+1}}): W \times V^{j+1} \rightarrow Z_\lambda$ is continuous, there exists a convex neighborhood U of x , and a closed absolutely convex neighborhood F of 0 , such that $U \times F^{j+1} \subset h^{-1}(\{z \in Z_\lambda: \|z\|_\lambda \leq 1\})$ and such that $D^j f(x)(F^j) \subset \{z \in Z_\lambda: \|z\|_\lambda \leq 1\}$.

Let ν be the seminorm on V such that $F = \{w \in V: \nu(w) \leq 1\}$, let $x + y, x + y' \in W$, and let $v_1, \dots, v_j \in V$. By Theorem 2.6,

$$\begin{aligned}
 & \lambda(D^j f(x + y)(v_1, \dots, v_j) - D^j f(x)(v_1, \dots, v_j)) \\
 (*) \quad &= \lambda\left(\int_0^1 D^{j+1} f(x + ty)(y, v_1, \dots, v_j) dt\right) \\
 &\leq \int_0^1 \lambda(D^{j+1} f(x + ty)(y, v_1, \dots, v_j)) dt \\
 &\leq \nu(y)\nu(v_1) \cdots \nu(v_j).
 \end{aligned}$$

Thus

$$\begin{aligned} \lambda(D^j f(x+y)(v_1, \dots, v_j)) &\leq \lambda(D^j f(x)(v_1, \dots, v_j)) + \nu(y)\nu(v_1) \cdots \nu(v_j) \\ &\leq (\nu(y) + 1)\nu(v_1) \cdots \nu(v_j), \end{aligned}$$

which implies that $D^j f(x+y) \in L^j_\nu(V, Z_\lambda)$. Furthermore, (*) applied to $x+y, x+y'$, implies that

$$\tilde{\nu}(p_\lambda \circ D^j f(x+y) - p_\lambda \circ D^j f(x+y')) \leq \nu(y-y'),$$

where $\tilde{\nu}$ is the induced norm on $L^j_\nu(V, Z_\lambda)$. This in turn implies that $p_\lambda \circ D^j f: U \rightarrow L^j_\nu(V, Z_\lambda)$ is continuous. \square

5.12. COROLLARY. $\mathcal{S}^\infty(W, Z) = C^\infty(W, Z)$.

Up to this point, we have not given any examples of C^∞ maps. Now we obtain a large class of examples from Corollary 5.12—namely, all the examples of \mathcal{S}^∞ maps from §1.

Let $k \in N$, and assume that $f: W \rightarrow Z$ is C^k . It is natural to ask whether there is a topology on $L^k(V, Z)$ such that $D^k f$ is a D -map if and only if $D^k f$ is a continuous map from W to $L^k(V, Z)$ equipped with this topology. This question was examined in [1], and led to the definition of a D -space [1, Definition 3.21]. For each $r \in N$, let $L^r_b(V, Z)$ denote the linear space of continuous r -linear maps from V^r to Z , equipped with the topology of uniform convergence in Z on bounded subsets of V^r . If $g: W \rightarrow L^r_b(V, Z)$ is a D -map, it is immediate that g is continuous from W to $L^r_b(V, Z)$. We define D -spaces to be those locally convex spaces for which this is an if-and-only-if condition.

5.13. DEFINITION. A D -space V is a locally convex space such that, for each locally convex space $Z, r \in N$, and open subspace $W \subset V$, every continuous map $g: W \rightarrow L^r_b(V, Z)$ is a D -map.

5.14. LEMMA. Let V be a D -space, Z any locally convex space, $r \in N$. Then

$$\underbrace{L_b(V, L_b(V, \dots, L_b(V, Z), \dots))}_{r \text{ times}} = L^r_b(V, Z).$$

PROOF. See [1, Lemma 3.22]. \square

We will need to complete the space $L^r_b(V, Z)$ in the proof of Theorem 5.17. Unfortunately, it is not clear that this can always be done within the context of continuous function spaces. Therefore we introduce the following space of almost-continuous maps:

5.15. DEFINITION. Let V and Z be locally convex spaces. For each $r \in N, \tilde{L}^r_b(V, Z)$ will denote the locally convex space of r -linear maps from V^r to Z which are bounded and continuous on bounded subspaces of V^r , equipped with the topology of uniform convergence in Z on bounded subsets of V^r .

5.16. REMARKS. (a) $L^r_b(V, Z) \subset \tilde{L}^r_b(V, Z)$.

(b) $\text{ev}: \tilde{L}^r_b(V, Z) \times B \rightarrow Z$ is continuous for each bounded subspace $B \subset V^r$.

(c) $\tilde{L}_b(V, \tilde{L}^r_b(V, Z)) = \tilde{L}^{r+1}_b(V, Z)$.

(d) If Z is complete (resp. quasi-complete, sequentially complete), then $\tilde{L}'_b(V, Z)$ is complete (resp. quasi-complete, sequentially complete).

The next result is the analogue of Theorem 3.9 for C^k maps and D -spaces.

5.17. THEOREM. Let $k \in N$, $f: W \rightarrow Z$, and assume that V is a D -space. Then the following are equivalent:

(i) f is C^k .

(ii) For each $j = 1, \dots, k$, there exists a continuous map $\theta^j f: W \rightarrow L'_b(V, Z)$ such that $f = \theta^0 f$, $\theta^1 f, \dots, \theta^{k-1} f$ are C^1 and $D(\theta^{j-1} f) = \theta^j f$, $1 \leq j \leq k$.

Furthermore, if (ii) holds, then $D^j f = \theta^j f$ for $1 \leq j \leq k$.

PROOF. It suffices to prove this theorem for the case when W is convex, so we will make this additional assumption throughout the proof.

(\Rightarrow) Assume that f is C^k . Let $i: Z \rightarrow \hat{Z}$ be the embedding of Z in its completion, and let $\theta^0 f = i \circ f$. For each $j \in \{1, \dots, k\}$, let $\theta^j f = i_j \circ D^j f$, where i_j is the induced embedding of $L'_b(V, Z)$ in $\tilde{L}'_b(V, \hat{Z})$. Since $\tilde{L}'_b(V, \hat{Z})$ is complete, it follows that $\int_0^1 \theta^j f(x + ty)(y) dt$ exists for each $x, x + y \in W$. For each $v_1, \dots, v_{j-1} \in V$,

$$\begin{aligned} & \theta^{j-1} f(x + y)(v_1, \dots, v_{j-1}) - \theta^{j-1} f(x)(v_1, \dots, v_{j-1}) \\ &= D^{j-1} f(x + y)(v_1, \dots, v_{j-1}) - D^{j-1} f(x)(v_1, \dots, v_{j-1}) \\ &= \int_0^1 D^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt \\ &= \int_0^1 \theta^j f(x + ty)(y, v_1, \dots, v_{j-1}) dt. \end{aligned}$$

This implies that

$$D^{j-1} f(x + y) - D^{j-1} f(x) = \theta^{j-1} f(x + y) - \theta^{j-1} f(x) = \int_0^1 \theta^j f(x + ty)(y) dt.$$

From the remarks at the beginning of §2, it follows that $\int_0^1 D^j f(x + ty)(y) dt$ exists and that

$$D^{j-1} f(x + y) - D^{j-1} f(x) = \int_0^1 D^j f(x + ty)(y) dt.$$

Thus the converse to the Mean Value Theorem implies that $D^{j-1} f$ is \mathcal{S}^1 , and that $D(D^{j-1} f) = D^j f$. Since $D^j f: W \rightarrow L'_b(V, Z) = L_b(V, L'_b(V, Z))$ is a D -map, we conclude that $D^{j-1} f$ is C^1 .

(\Leftarrow) Assume that we are given $\theta^1 f, \dots, \theta^k f$ as defined in statement (ii) of this theorem. Since V is a D -space, each $\theta^j f$ is a D -map. Thus Proposition 5.7 implies that $ev \circ (\theta^j f \times \text{Id}_{V^j}): W \times V^j \rightarrow Z$ is continuous. With this observation, it is easily checked that the second half of the proof of Theorem 3.9 applies without modification to show that f is \mathcal{S}^k , and that $D^j f = \theta^j f$ for $1 \leq j \leq k$. Since each $D^j f$ is a D -map, we conclude that f is C^k . \square

5.18. COROLLARY. Assume that V is a D -space, and let $j, k \in N$, $j \leq k$. Then f is $C^k \Leftrightarrow f$ is C^j and $D^j f$ is C^{k-j} . Furthermore, if f is C^k , then $D^r(D^j f) = D^{r+j} f$ for $1 \leq r \leq k - j$.

PROOF. Replacing \mathcal{S}^k with C^k , $L'(V, Z)$ with $L'_b(V, Z)$, and references to Theorem 3.9 with references to Theorem 5.17, the proof is otherwise identical to the proof of Corollary 3.10. \square

We need to mention some examples of D -spaces. Clearly all normed spaces are D -spaces. In [1], two general classes of D -space examples are presented—normed spaces, and exponential Schwartz spaces. The latter class includes Fréchet Schwartz spaces, and the bounded weak-star topologies on dual Banach spaces (for which the D -space calculus in [1] was actually developed). In [4], the proof that exponential Schwartz spaces are D -spaces was modified and extended to show that exponential quasinormable spaces are D -spaces (quasinormability was introduced in [5], which is still one of the most readable references on this subject). The class of exponential quasinormable spaces includes all normed spaces and all exponential Schwartz spaces, and constitutes the entire class of examples of D -spaces known at the present time. Note that all results in §§3 and 5 on the function space properties of derivatives apply to maps defined on domains in exponential quasinormable spaces. Incidentally, [4] also shows that there exist separable reflexive Fréchet spaces which are not D -spaces.

Let $k \in N$. We know that $C^k(W, Z) \subset \mathcal{S}^k(W, Z)$, and that in general the inclusion is strict (e.g. if V and Z are French spaces and V is infinite-dimensional). We also know that $C^k(W, Z) = \mathcal{S}^k(W, Z)$ if V and Z are finite-dimensional spaces. We next discuss a more general condition which implies this equality.

5.19. THEOREM. *Assume that V is a D -space such that bounded sets in V are precompact, and let Z be any locally convex space. Then $C^k(W, Z) = \mathcal{S}^k(W, Z)$ for all $k \in N$.*

PROOF. Let $k \in N$, and let $f \in \mathcal{S}^k(W, Z)$. We must show that f is C^k . By Theorem 5.11, each $D^j f$ is a D -map for $1 \leq j \leq k - 1$, so it suffices to show that $D^k f$ is a D -map. Assume that $L^k(V, Z)$ has the precompact-open topology. By the comments at the end of §3, it follows that $D^k f: W \rightarrow L^k(V, Z)$ is continuous. Since bounded subsets of V are precompact, $L^k(V, Z) = L^k_b(V, Z)$. Thus $D^k f: W \rightarrow L^k_b(V, Z)$ is continuous. Since V is a D -space, this implies that $D^k f$ is a D -map. \square

5.20. COROLLARY. *Assume that V is an exponential Schwartz space, and let Z be any locally convex space. Then $C^k(W, Z) = \mathcal{S}^k(W, Z)$ for all $k \in N$.*

Let $f: W \rightarrow Z$. It only remains to show that the definition of C^k differentiability for f in this article is equivalent to the definition of C^k differentiability in [1] when V is a D -space. For the case $k = 1$, the converse to the Mean Value Theorem for D -space calculus [1, Lemma 2.8] implies that the two definitions of C^1 differentiability are equivalent. For $k > 1$, f is defined to be C^k in [1] if f is C^1 and $Df: W \rightarrow L^1_b(V, Z)$ is C^{k-1} . Letting $j = 1$ in Corollary 5.18, we see that f is C^k in the sense of this article if and only if f is C^1 and Df is C^{k-1} . Thus the two definitions of C^k differentiability are equivalent when V is a D -space. Since the definition of C^k differentiability in [1] extends the classical definition of C^k differentiability in the

Banach space category, it follows that the definition of C^k differentiability in this article is equivalent to the classical definition of C^k differentiability when V and Z are Banach spaces.

REFERENCES

1. R. A. Graff, *Elements of non-linear functional analysis*, Mem. Amer. Math. Soc., No. 206, (1978).
2. _____, *Flows for differentiable vector fields on conjugate Banach spaces*, J. Differential Geom. **15** (1980), 575–593.
3. _____, *Smoothness of semigroups for nonlinear equations*, J. Integral Equations. **4** (1982), 183–220.
4. R. A. Graff and W. M. Ruess, *Appropriate locally convex domains for differential calculus*, Proc. Amer. Math. Soc. **86** (1982), 331–335.
5. A. Grothendieck, *Sur les espaces (F) et (DF)* , Summa Brasil. Mat. **3** (1954), 57–123.
6. R. S. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. Amer. Math. Soc. (N.S.) **7** (1982), 65–222.
7. J. L. Kelley, *General topology*, Van Nostrand, New York, London and Toronto, 1955.
8. S. Lang, *Differential manifolds*, Addison-Wesley, Reading, Mass., 1972.
9. N. E. Steenrod, *A convenient category of topological spaces*, Michigan Math. J. **14** (1967), 133–152.

5125 SOUTH KENWOOD AVENUE, APARTMENT 201, CHICAGO, ILLINOIS 60615