J. Austral. Math. Soc. **25** (Series A) (1978), 348–361.

# HIGHER ORDER DERIVATIVES IN TOPOLOGICAL LINEAR SPACES

J. W. LLOYD

(Received 20 December 1976)

Communicated by J. Virsik

**Abstract**

The higher order chain rule for Fréchet and Hadamard differentiable mappings on topological linear spaces is proved and various formulae for $(g \circ f)^{(n)}(x)$ are given. Leibniz' theorem (in a very general form) is also proved.

*Subject Classification* (*Amer. Math. Soc.* (*MOS*) 1970): primary 58C20, 46A99, 47H99.

*Keywords*: topological linear space, Fréchet derivative, Hadamard derivative, higher order chain rule.

## 1. Discussion and preliminaries

Let $f: U \to V$ and $g: V \to W$ be $n$-times Fréchet differentiable, where $U$, $V$ and $W$ are open subsets of normed linear spaces $E$, $F$ and $G$. Then it is well known that $g \circ f$ is $n$-times Fréchet differentiable. The proof depends on the formula $(g \circ f)'(x) = g'(f(x)) \circ f'(x)$ and the fact that the composition mapping comp: $\mathscr{L}(E, F) \times \mathscr{L}(F, G) \to \mathscr{L}(E, G)$ defined by $\mathrm{comp}(u, v) = v \circ u$ is continuous bilinear and thus infinitely Fréchet differentiable.

If $E$, $F$ and $G$ are only supposed to be topological linear spaces, then comp may only be (once) Fréchet differentiable and this proof breaks down. Consequently, we are forced to use the formula for $(g \circ f)^{(n)}(x)$ and this leads to a rather more complicated induction argument. This paper contains the details of this proof as well as a similar proof of Leibniz' theorem (in a very general form) and various formulae for $(g \circ f)^{(n)}(x)$.

This paper arose from an attempt to reconcile three versions of the formula for $(g \circ f)^{(n)}(x)$. One appears (without proof) in Averbuh and Smoljanov (1967), p. 234, the second is due to Penot (1973), p. 8 (proof outline for locally convex spaces) and the third is given by Abraham and Robbin (1967), p. 3 for normed linear spaces. We show that Abraham and Robbin's formula is incorrect. Penot's formula has an omission in the coefficients. The formula of Averbuh and Smoljanov is correct and the most general. We derive several simpler formulae which arise in special cases.

The higher order chain rule is used constantly in the theory of smoothness properties of topological linear spaces (Lloyd (1972, 1973, 1974, 1975)).

TLS denotes the class of all (Hausdorff) topological linear spaces over the real field $\mathbf{R}$. Throughout $E, F, G, E_1, \cdots, E_n, F_1, F_2 \in$ TLS. TLS* denotes those topological linear spaces over $\mathbf{R}$, which are separated by their dual. $\mathcal{O}(E)$ denotes the class of all open subsets of $E$.

Certain classes of multilinear mappings arise naturally in the calculus in topological linear spaces and one must be careful to distinguish between them.

$\mathcal{L}H(E_1, \cdots, E_n; F)$ denotes the linear space of hypocontinuous $n$-linear maps from $E_1 \times \cdots \times E_n$ into $F$. That is, $n$-linear $u \in \mathcal{L}H(E_1, \cdots, E_n; F)$ if, given a 0-neighbourhood $V$ in $F$, given $k \in \{1, 2, \cdots, n\}$, given bounded $B_i$ in $E_i (i \neq k)$, there exists a 0-neighbourhood $U_k$ in $E_k$ such that

$$u(B_1 \times \cdots \times B_{k-1} \times U_k \times B_{k+1} \times \cdots \times B_n) \subset V.$$

$\mathcal{L}(E_1, \mathcal{L}(E_2, \cdots, \mathcal{L}(E_n, F) \cdots)$ is defined by induction. $\mathcal{L}(E_1, F)$ is the linear space of continuous linear maps from $E_1$ into $F$, given the topology of uniform convergence on bounded subsets of $E_1$. Then, having defined $\mathcal{L}(E_2, \cdots, \mathcal{L}(E_n, F) \cdots)$, we define $\mathcal{L}(E_1, \mathcal{L}(E_2, \cdots, \mathcal{L}(E_n, F) \cdots)$ to be the linear space of continuous linear maps from $E_1$ into $\mathcal{L}(E_2, \cdots, \mathcal{L}(E_n, F) \cdots)$, given the topology of uniform convergence on bounded subsets of $E_1$.

We identify $\mathcal{L}(E_1, \mathcal{L}(E_2, \cdots, \mathcal{L}(E_n, F) \cdots)$ (as a linear space) with the linear space $\mathcal{L}(E_1, \cdots, E_n; F)$ consisting of those $n$-linear maps $u$ from $E_1 \times \cdots \times E_n$ into $F$, which satisfy the following "continuity" condition:

Given $k \in \{1, 2, \cdots, n\}$, given $(x_1, \cdots, x_{k-1}) \in E_1 \times \cdots \times E_{k-1}$, given bounded $B_{k+1}$ in $E_{k+1}, \cdots$, bounded $B_n$ in $E_n$, given a 0-neighbourhood $V$ in $F$, there exists a 0-neighbourhood $U_k$ in $E_k$ such that $u(\{x_1\} \times \cdots \times \{x_{k-1}\} \times U_k \times B_{k+1} \times \cdots \times B_n) \subset V$.

Using this identification, we can transfer the topology on $\mathcal{L}(E_1, \cdots, \mathcal{L}(E_n, F) \cdots)$ over to $\mathcal{L}(E_1, \cdots, E_n; F)$. A basic 0-neighbourhood in this topology on $\mathcal{L}(E_1, \cdots, E_n; F)$ is then a set of the form $(B_1, \cdots, B_n, W) = \{u : u(B_1 \times \cdots \times B_n) \subset W\}$, where the $B_i$ are bounded subsets of the $E_i$ and $W$ is a 0-neighbourhood in $F$. Note that $\mathcal{B} \subset \mathcal{L}(E_1, \cdots, E_n; F)$ is bounded in this topology if and only if $\bigcup_{u \in \mathcal{B}} u(B_1 \times \cdots \times B_n)$ is bounded in $F$, for each bounded $B_i$ in $E_i$ $(i = 1, 2, \cdots, n)$.

$\mathcal{L}B(E_1, \cdots, E_n; F)$ denotes the linear space of all separately continuous

$n$-linear maps from $E_1 \times \cdots \times E_n$ into $F$, which also map bounded subsets of $E_1 \times \cdots \times E_n$ into bounded subsets of $F$.

Clearly

$$\mathscr{L}H(E_1, \cdots, E_n; F) \subset \mathscr{L}(E_1, \cdots, E_n; F) \subset \mathscr{L}B(E_1, \cdots, E_n; F).$$

Note that if $\rho$ is a permutation of $\{1,2,\cdots,n\}$, then $\mathscr{L}H(E_1, \cdots, E_n; F)$ can be naturally identified with $\mathscr{L}H(E_{\rho(1)}, \cdots, E_{\rho(n)}; F)$. The same is true for $\mathscr{L}B(E_1, \cdots, E_n; F)$, but it is not true for $\mathscr{L}(E_1, \cdots, E_n; F)$. (See example 1.2 below.)

The *composition mapping* $\text{comp}: \mathscr{L}(E,F) \times \mathscr{L}(F,G) \to \mathscr{L}(E,G)$ is defined by $\text{comp}(u,v) = v \circ u$. A special case is the *evaluation mapping* $\text{ev}: E \times E' \to \mathbf{R}$ defined by $\text{ev}(x,x') = \langle x, x' \rangle$.

EXAMPLE 1.1. $\mathscr{L}H(E_1, \cdots, E_n; F) \subsetneqq \mathscr{L}(E_1, \cdots, E_n; F)$.

Consider $\text{ev}: l^2 \times \sigma(l^2, l^2) \to \mathbf{R}$, where $\sigma(l^2, l^2)$ means $l^2$ with the $\sigma(l^2, l^2)$-topology. Then $\text{ev} \in \mathscr{L}(l^2, \sigma(l^2, l^2); \mathbf{R})$, using the facts that $\text{ev}: l^2 \times l^2 \to \mathbf{R}$ is continuous and $l^2$ and $\sigma(l^2, l^2)$ have the same bounded sets. However, $\text{ev} \notin \mathscr{L}H(l^2, \sigma(l^2, l^2); \mathbf{R})$. For let $B$ be the unit ball in $l^2$ and $W = \{\xi \in \mathbf{R}: |\xi| \leqq 1\}$. Suppose there exists a 0-neighbourhood $U$ in $\sigma(l^2, l^2)$ such that $\text{ev}(B \times U) \subset W$. Then $U \subset B^0 = B$, which is a contradiction.

EXAMPLE 1.2. $\mathscr{L}(E_1, \cdots, E_n; F) \subsetneqq \mathscr{L}B(E_1, \cdots, E_n; F)$.

Clearly comp has the property that given a bounded set $\mathscr{B}$ in $\mathscr{L}(E,F)$ and a 0-neighbourhood $\mathscr{W}$ in $\mathscr{L}(E,G)$, there exists a 0-neighbourhood $\mathscr{V}$ in $\mathscr{L}(F,G)$ such that $\text{comp}(\mathscr{B} \times \mathscr{V}) \subset \mathscr{W}$. This implies that comp maps bounded sets into bounded sets. Also comp is clearly separately continuous, and so,

$$\text{comp} \in \mathscr{L}B(\mathscr{L}(E,F), \mathscr{L}(F,G); \mathscr{L}(E,G))$$

and

$$\text{comp} \in \mathscr{L}(\mathscr{L}(F,G), \mathscr{L}(E,F); \mathscr{L}(E,G)).$$

Now consider $\text{ev}: E \times E' \to \mathbf{R}$, where $E$ is a non-quasi-barrelled locally convex space and $E'$ is the strong dual of $E$. We show that $\text{ev} \notin \mathscr{L}(E, E'; \mathbf{R})$. In fact, suppose the contrary. Let $B$ be a bounded set in $E'$. Then, for each $\varepsilon > 0$, there exists a 0-neighbourhood $U$ in $E$ such that $|\langle x, y \rangle| \leqq \varepsilon$, whenever $x \in U$ and $y \in B$. That is, $B$ is equicontinuous, and so $E$ is quasi-barrelled.

We emphasize that while $\text{comp} \in \mathscr{L}(\mathscr{L}(F,G), \mathscr{L}(E,F); \mathscr{L}(E,G))$, generally $\text{comp} \notin \mathscr{L}(\mathscr{L}(E,F), \mathscr{L}(F,G); \mathscr{L}(E,G))$.

In case $E_1 = E_2 = \cdots = E_n = E$, we write $\mathscr{L}H_n(E,F)$ instead of $\mathscr{L}H(E, \cdots, E; F)$, $\mathscr{L}_n(E,F)$ instead of $\mathscr{L}(E, \cdots, E; F)$ and $\mathscr{L}B_n(E,F)$ instead of $\mathscr{L}B(E, \cdots, E; F)$.

Let $\sigma$ denote a collection of bounded subsets of $E$, which includes all

https://doi.org/10.1017/S1446788700021091 Published online by Cambridge University Press

single point subsets of $E$. Let $f: U \to V$, where $U \in \mathcal{O}(E)$ and $V \in \mathcal{O}(F)$. Then we say $f$ is *$\sigma$-differentiable at $x \in U$* if there exists $u \in \mathscr{L}(E, F)$ such that $t_n \to 0$ $(t_n \neq 0)$ and $\{h_n\} \subset B \in \sigma$ imply $t_n^{-1} \cdot [f(x + t_n h_n) - f(x)] - u \cdot h_n \to 0$.

The mapping $u$ is then uniquely determined and is denoted by $f'(x)$. We say $f'(x)$ is the *$\sigma$-derivative of $f$ at $x$*. If $f'(x)$ exists for each $x \in U$, then we can define a map $f': U \to \mathscr{L}(E, F)$ by $x \to f'(x)$. We say $f'$ is the *$\sigma$-derivative* of $f$ and $f$ is *$\sigma$-differentiable*.

By induction, we define an *$n$-times $\sigma$-differentiable* map $f: U \to V$ as an $(n-1)$-times $\sigma$-differentiable map, whose $(n-1)$th $\sigma$-derivative is $\sigma$-differentiable. The $n$th $\sigma$-derivative is denoted by $f^{(n)}$ and is a map from $U$ into $\mathscr{L}_n(E, F)$.

$f: U \to V$ is *$n$-times $\sigma$-differentiable at $x \in U$*, if $f$ is $(n-1)$-times $\sigma$-differentiable and $f^{(n-1)}: U \to \mathscr{L}_{n-1}(E, F)$ is $\sigma$-differentiable at $x$.

When $\sigma$ is the class of all bounded (resp. sequentially compact) subsets of $E$, the $\sigma$-derivative is called the *Fréchet* (resp. *Hadamard*) derivative.

PROPOSITION 1.3. (Symmetry of the higher derivative). *Let $f: U \to V$, where $V \in \mathcal{O}(F)$ and $F \in$ TLS*. Suppose $f$ is $n$-times Hadamard differentiable at $x \in U$. Then $f^{(n)}(x)$ is symmetric. Consequently, $f^{(n)}(x) \in \mathscr{L}H_m(E, F)$.*

PROOF. Lloyd (1973), p. 16, Penot (1973), p. 7.

It would be interesting to know if it is possible to prove that $f^{(n)}(x) \in \mathscr{L}H_n(E, F)$ without the assumption that $F$ be separated by its dual. This is of some interest since there is one step in the proof of the higher order chain rule in which hypocontinuity seems to be essential.

## 2. Leibniz' theorem

LEMMA 2.1. *Let $f: U \to V$ be $\sigma$-differentiable at $x \in U$. If $t_n \to 0$ and $\{h_n\} \subset B \in \sigma$, then $\{t_n^{-1} \cdot [f(x + t_n h_n) - f(x)]\}$ is bounded in $F$.*

PROOF. We know $t_n^{-1} \cdot [f(x + t_n h_n) - f(x)] - f'(x) \cdot h_n \to 0$. Since $\{h_n\}$ is bounded, $\{f'(x) \cdot h_n\}$ is bounded and the result follows.

LEMMA 2.2. *Let $U \in \mathcal{O}(E)$ and $u \in \mathscr{L}B(F_1, F_2; G)$. Let $f_i: U \to F_i$ be $\sigma$-differentiable at $x \in U$ $(i = 1, 2)$. Then $f: U \to G$ defined by $f(y) = u(f_1(y), f_2(y))$ is $\sigma$-differentiable at $x$ and*

$$f'(x) \cdot h = u(f_1(x), f_2'(x) \cdot h) + u(f_1'(x) \cdot h, f_2(x)).$$

PROOF. Certainly, since $u$ is separately continuous, $f'(x)$ is a continuous linear map from $E$ into $G$. Let $t_n \to 0$ and $\{h_n\} \subset B \in \sigma$. Then

$$\begin{aligned}
&t_n^{-1}\cdot[u(f_1(x+t_nh_n), f_2(x+t_nh_n)) - u(f_1(x), f_2(x))] \\
&\qquad - u(f_1(x), f_2'(x)\cdot h_n) - u(f_1'(x)\cdot h_n, f_2(x)) \\
&\quad = u(f_1(x), t_n^{-1}\cdot[f_2(x+t_nh_n) - f_2(x)] - f_2'(x)\cdot h_n) \\
&\qquad + u(t_n^{-1}\cdot[f_1(x+t_nh_n) - f_1(x)] - f_1'(x)\cdot h_n, f_2(x)) \\
&\qquad + t_n\cdot u(t_n^{-1}\cdot[f_1(x+t_nh_n) - f_1(x)], t_n^{-1}\cdot[f_2(x+t_nh_n) - f_2(x)]) \\
&\quad \to 0,
\end{aligned}$$

using the separate continuity of $u$, 2.1 and the fact that $u$ maps bounded sets to bounded sets.

LEMMA 2.3. *Let $\nu$ be a partition of $\{1,2,\cdots,n\}$ into an ordered pair of disjoint subsets $\{i_1,\cdots,i_k\}$ and $\{j_1,\cdots,j_l\}$ such that $k+l=n, k\geqq 0, l\geqq 0, n\geqq 1, i_1<i_2<\cdots<i_k$ and $j_1<j_2<\cdots<j_l$. Let $u\in\mathscr{L}H(F_1,F_2;G)$. Define*

$$u_\nu:\mathscr{L}_k(E,F_1)\times\mathscr{L}_l(E,F_2)\to\mathscr{L}_n(E,G)$$

*by*

$$u_\nu(\alpha,\beta)(h_1,\cdots,h_n)=u(\alpha(h_{i_1},\cdots,h_{i_k}),\beta(h_{j_1},\cdots,h_{j_l})),$$

*where $\alpha\in\mathscr{L}_k(E,F_1)$, $\beta\in\mathscr{L}_l(E,F_2)$ and $h_1,\cdots,h_n\in E$. Then $u_\nu$ is well-defined and*

$$u_\nu\in\mathscr{L}H(\mathscr{L}_k(E,F_1),\mathscr{L}_l(E,F_2);\mathscr{L}_n(E,G)).$$

PROOF. To show $u_\nu$ is well-defined, we have to show that, in fact, $u_\nu(\alpha,\beta)$ does belong to $\mathscr{L}_n(E,G)$. Let $(x_1,\cdots,x_{p-1})\in E^{p-1}$, $B_{p+1},\cdots,B_n$ be bounded subsets of $E$ and $V$ be a 0-neighbourhood in $G$. Suppose $p=i_s$. (The proof is similar if $p$ is one of the $j$'s.) Then $B=\beta(\{x_{j_1}\}\times\cdots\times\{x_{j_t}\}\times B_{j_{t+1}}\times\cdots\times B_{j_l})$ is a bounded subset of $F_2$. Since $u$ is hypocontinuous, there exists a 0-neighbourhood $W$ in $F_1$ such that $u(W\times B)\subset V$. Since $\alpha\in\mathscr{L}_k(E,F_1)$, there exists a 0-neighbourhood $U$ in $E$ such that

$$\alpha(\{x_{i_1}\}\times\cdots\times\{x_{i_{s-1}}\}\times U\times B_{i_{s+1}}\times\cdots\times B_{i_k})\subset W.$$

Then

$$\begin{aligned}
&u_\nu(\alpha,\beta)(\{x_1\}\times\cdots\times\{x_{p-1}\}\times U\times B_{p+1}\times\cdots\times B_n) \\
&\quad = u(\alpha(\{x_{i_1}\}\times\cdots\times\{x_{i_{s-1}}\}\times U\times B_{i_{s+1}}\times\cdots\times B_{i_k}), \\
&\qquad\quad \beta(\{x_{j_1}\}\times\cdots\times\{x_{j_t}\}\times B_{j_{t+1}}\times\cdots\times B_{j_l})) \\
&\quad \subset u(W\times B)\subset V.
\end{aligned}$$

That is, $u_\nu(\alpha,\beta)\in\mathscr{L}_n(E,G)$, as required.

Clearly $u_\nu$ is bilinear. So we have only to show it is hypocontinuous. First let $\mathscr{B}$ be a bounded subset of $\mathscr{L}_k(E,F_1)$ and $\mathscr{U}$ be a 0-neighbourhood in $\mathscr{L}_n(E,G)$. We have to show there exists a 0-neighbourhood $\mathscr{V}$ in $\mathscr{L}_l(E,F_2)$ such that $u_\nu(\mathscr{B}\times\mathscr{V})\subset\mathscr{U}$.

Suppose $\mathscr{U}=(B_1,\cdots,B_n,U)$ where the $B_i$ are bounded subsets of $E$ and $U$ is a 0-neighbourhood in $G$. Put $B=\bigcup_{\alpha\in\mathscr{B}}\alpha(B_{i_1}\times\cdots\times B_{i_k})$. Then $B$ is a bounded subset of $F_1$. Since $u$ is hypocontinuous, there exists a 0-neighbourhood $V$ in $F_2$ such that $u(B\times V)\subset U$. Put $\mathscr{V}=(B_{j_1},\cdots,B_{j_l},V)$. Then $\mathscr{V}$ is a 0-neighbourhood in $\mathscr{L}_l(E,F_2)$. Also $\alpha\in\mathscr{B}$ and $\beta\in\mathscr{V}$ implies $u_\nu(\alpha,\beta)(h_1,\cdots,h_n)=u(\alpha(h_{i_1},\cdots,h_{i_k}),\beta(h_{j_1},\cdots,h_{j_l}))\in U$, whenever $h_i\in B_i$ $(i=1,\cdots,n)$. That is, $u_\nu(\alpha,\beta)\in\mathscr{U}$.

Similarly, if $\mathscr{B}$ is a bounded subset of $\mathscr{L}_l(E,F_2)$ and $\mathscr{U}$ is a 0-neighbourhood in $\mathscr{L}_n(E,G)$, we can find a 0-neighbourhood $\mathscr{W}$ in $\mathscr{L}_n(E,F_1)$ such that $u_\nu(\mathscr{W}\times\mathscr{B})\subset\mathscr{U}$. Thus $u_\nu$ is hypocontinuous.

THEOREM 2.4. (Leibniz' theorem, compare with Averbuh and Smoljanov (1967), p. 233). *Let $u\in\mathscr{L}H(F_1,F_2;G)$ and $U\in\mathscr{O}(E)$. Let $f_i:U\to F_i$ be $n$-times $\sigma$-differentiable at $x\in U$ $(i=1,2)$. Then $f:U\to G$ defined by $f(y)=u(f_1(y),f_2(y))$ is $n$-times $\sigma$-differentiable at $x$ and*

$$f^{(n)}(x)=\sum_\nu u_\nu(f_1^{(k)}(x),f_2^{(l)}(x)),$$

*where the sum is over all partitions $\nu$ of $\{1,2,\cdots,n\}$ into ordered pairs of disjoint subsets $\{i_1,\cdots,i_k\}$ and $\{j_1,\cdots,j_l\}$ such that $k+l=n$, $k\geqq 0$, $l\geqq 0$, $i_1<i_2<\cdots<i_k$ and $j_1<j_2<\cdots<j_l$.*

*In other words*

$$f^{(n)}(x)(h_1,\cdots,h_n)=\sum_\nu u(f_1^{(k)}(x)(h_{i_1},\cdots,h_{i_k}),f_2^{(l)}(x)(h_{j_1},\cdots,h_{j_l})).$$

PROOF. The proof is by induction. The case $n=1$ is covered by Lemma 2.2. So suppose the theorem is true for some $n$. Let the $f_i$ be $(n+1)$-times $\sigma$-differentiable at $x$. By the induction hypothesis, $f$ is $n$-times $\sigma$-differentiable at each $y\in U$ and

$$f^{(n)}(y)=\sum_\nu u_\nu(f_1^{(k)}(y),f_2^{(l)}(y)).$$

By 2.3, $u_\nu$ is hypocontinuous. Hence, by 2.2, $f^{(n)}$ is $\sigma$-differentiable at $x$ and

$$f^{(n+1)}(x)\cdot h_0=\sum_\nu\left[u_\nu(f_1^{(k+1)}(x)\cdot h_0,f_2^{(l)}(x))+u_\nu(f_1^{(k)}(x),f_2^{(l+1)}(x)\cdot h_0)\right].$$

Thus

$$f^{(n+1)}(x)(h_0, h_1, \cdots, h_n) = \sum_{\nu} [u(f_1^{(k+1)}(x)(h_0, h_{i_1}, \cdots, h_{i_k}), f_2^{(l)}(x)(h_{j_1}, \cdots, h_{j_l})) + u(f_1^{(k)}(x)(h_{i_1}, \cdots, h_{i_k}), f_2^{(l+1)}(x)(h_0, h_{j_1}, \cdots, h_{j_l}))].$$

Hence

$$f^{(n+1)}(x) = \sum_{\mu} u_\mu(f_1^{(m)}(x), f_2^{(p)}(x)),$$

where the sum is over all partitions $\mu$ of $\{0, 1, \cdots, n\}$ into ordered pairs of disjoint subsets $\{i_1, \cdots, i_m\}$ and $\{j_1, \cdots, j_p\}$ such that $m + p = n + 1, m \geqq 0, p \geqq 0, i_1 < i_2 < \cdots < i_m$ and $j_1 < j_2 < \cdots < j_p$.

This completes the proof of Leibniz' theorem.

If it is known that the derivatives $f_1^{(k)}(x)$ and $f_2^{(l)}(x)$ are all symmetric, then we can write the formula more concisely (as J. P. Penot has pointed out to me):

$$f^{(n)}(x)(h_1, \cdots, h_n) = \sum_{k=0}^{n} \sum_{\rho \in S_n} \frac{1}{k!(n-k)!} u(f_1^{(k)}(x)(h_{\rho(1)}, \cdots, h_{\rho(k)}), f_2^{(n-k)}(x)(h_{\rho(k+1)}, \cdots, h_{\rho(n)})),$$

where $S_n$ is the group of permutations on $\{1, 2, \cdots, n\}$.

In particular, if $h^{(n)}$ denotes the $n$-tuple $(h, \cdots, h)$, then

$$f^{(n)}(x) \cdot h^{(n)} = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} u(f_1^{(k)}(x) \cdot h^{(k)}, f_2^{(n-k)}(x) \cdot h^{(n-k)}).$$

Next we give a counterexample to Abraham and Robbin's version of Leibniz' theorem. Let $E$ be a normed space. Put $\mathscr{L}(E) = \mathscr{L}(E, E)$. Supose there exist $h_1, h_2 \in \mathscr{L}(E)$ such that $h_2 \circ h_1 \neq h_1 \circ h_2$. Consider $f : \mathscr{L}(E) \to \mathscr{L}(E)$ defined by $f(x) = x \circ x$, where $x \in \mathscr{L}(E)$. Now $f(x) = u(f_1(x), f_2(x))$, where $f_1 = f_2 =$ identity map on $\mathscr{L}(E)$ and $u$ is the composition map from $\mathscr{L}(E) \times \mathscr{L}(E)$ into $\mathscr{L}(E)$ defined by $u(x, y) = y \circ x$. Let $x \in \mathscr{L}(E)$. According to Abraham and Robbin's formula (Abraham and Robbin (1967), p. 3), $f^{(2)}(x)(h_1, h_2) = 2h_2 \circ h_1$. Hence by our choice of $h_1$ and $h_2$, $f^{(2)}(x)(h_1, h_2) \neq f^{(2)}(x)(h_2, h_1)$. That is, $f^{(2)}(x)$ is not symmetric! The correct formula is $f^{(2)}(x)(h_1, h_2) = h_1 \circ h_2 + h_2 \circ h_1$.

Note that Leibniz' theorem is no longer true (for $n = 2$) if we replace "$u \in \mathscr{L}H(F_1, F_2; G)$" by "$u \in \mathscr{L}(F_1, F_2; G)$". In fact, suppose $u \in \mathscr{L}(F_1, F_2; G)$. Then according to 3.1 below, $u$ is Fréchet differentiable and $u'(x) \cdot y = u(x_1, y_2) + u(y_1, x_2)$. The map $u' : F_1 \times F_2 \to \mathscr{L}(F_1 \times F_2, G)$ is linear. Consequently, if $u$ is twice $\sigma$-differentiable, $u'$ must be continuous. Let $B$ be a bounded subset of $F_1$ and $W$ be a 0-neighbourhood in $G$. Then there exists a 0-neighbourhood $U \times V$ in $F_1 \times F_2$ such that $u'(U \times V) \subset (B \times \{0\}, W)$. Thus by the formula for $u', u(B \times V) \subset W$. Hence $u \in \mathscr{L}H(F_1, F_2; G)$. But in Example 1.1, we showed $\mathscr{L}H(F_1, F_2; G) \subsetneqq \mathscr{L}(F_1, F_2; G)$.

## 3. Higher order chain rule

LEMMA 3.1. *Let $u \in \mathscr{L}B(E_1, \cdots, E_n; F)$. Then $u$ is Fréchet differentiable and*

$$u'(x) \cdot y = \sum_{i=1}^{n} u(x_1, \cdots, x_{i-1}, y_i, x_{i+1}, \cdots, x_n).$$

PROOF. Since $u$ is separately continuous, $u'(x)$ is certainly a continuous linear map from $E_1 \times \cdots \times E_n$ into $F$. Let $B = B_1 \times \cdots \times B_n$ be a bounded subset of $E_1 \times \cdots \times E_n$ and $V$ be a 0-neighbourhood in $F$. Choose a balanced 0-neighbourhood $W$ in $F$ such that

$$\sum_{j=1}^{2^n-(n+1)} W \subset V.$$

Put $\bar{B}_i = \{x_i\} \cup B_i$ and $\bar{B} = \bar{B}_1 \times \cdots \times \bar{B}_n$. Since $u$ maps bounded sets into bounded sets, $u(\bar{B})$ is a bounded subset of $F$. Hence there exists $\delta \in (0,1]$ such that $|t| \leqq \delta$ implies $t \cdot u(\bar{B}) \subset W$.

Then $|t| \leqq \delta$ and $h \in B$ implies

$$
\begin{aligned}
u(x+th) - u(x) - &\sum_{i=1}^{n} u(x_1, \cdots, x_{i-1}, th_i, x_{i+1}, \cdots, x_n) \\
= t^2 &\sum_{\substack{i,j \\ (i \neq j)}} u(x_1, \cdots, x_{i-1}, h_i, x_{i+1}, \cdots, x_{j-1}, h_j, x_{j+1}, \cdots, x_n) \\
&+ \cdots \\
&+ t^n u(h_1, \cdots, h_n) \\
&\in t \sum_{j=1}^{2^n-(n+1)} W \\
&\subset tV.
\end{aligned}
$$

Thus $u$ is Fréchet differentiable.

LEMMA 3.2. *Let $\nu$ be a partition of $\{1, 2, \cdots, n\}$ into $k$ disjoint non-empty sets $\{i_1^1, \cdots, i_{l_1}^1\}, \cdots, \{i_1^k, \cdots, i_{l_k}^k\}$ with $l_1 + \cdots + l_k = n$, $l_j > 0$, $i_1^j < i_2^j < \cdots < i_{l_j}^j$ $(j = 1, 2, \cdots, k)$ and $i_{l_1}^1 < i_{l_2}^2 < \cdots < i_{l_k}^k$. Consider the map*

$$w_\nu : \mathscr{L}_{l_1}(E, F) \times \cdots \times \mathscr{L}_{l_k}(E, F) \times \mathscr{L}H_k(F, G) \to \mathscr{L}_n(E, G)$$

*defined by*

https://doi.org/10.1017/S1446788700021091 Published online by Cambridge University Press

$$w_\nu(u_1,\cdots,u_k,v)(h_1,\cdots,h_n)$$
$$= v(u_1(h_{i_1^1},\cdots,h_{i_{l_1}^1}),\cdots,u_k(h_{i_1^k},\cdots,h_{i_{l_k}^k})).$$

*Then $w_\nu$ is well-defined and*

$$w_\nu \in \mathscr{L}B(\mathscr{L}_{l_1}(E,F),\cdots,\mathscr{L}_{l_k}(E,F),\mathscr{L}H_k(F,G);\mathscr{L}_n(E,G)).$$

PROOF. To show $w_\nu$ is well defined, we have to show that $w_\nu(u_1,\cdots,u_k,v)$ does indeed belong to $\mathscr{L}_n(E,G)$. To this end, let $p\in\{1,2,\cdots,n\}$, $x_1,\cdots,x_{p-1}\in E$ and $B_{p+1},\cdots,B_n$ be bounded subsets of $E$. Put $D_1=\{x_1\},\cdots,D_{p-1}=\{x_{p-1}\}$, $D_{p+1}=B_{p+1},\cdots,D_n=B_n$. Suppose $p=i_s^m$. Put $\bar{B}_j=u_j(D_{i_1^j}\times\cdots\times D_{i_{l_j}^j})$ $(j\in\{1,2,\cdots,k\}, j\neq m)$. Then each $\bar{B}_j$ is a bounded subset of $F$. Let $W$ be a 0-neighbourhood in $G$. Choose a 0-neighbourhood $V$ in $F$ such that

$$v(\bar{B}_1\times\cdots\times\bar{B}_{m-1}\times V\times\bar{B}_{m+1}\times\cdots\times\bar{B}_k)\subset W.$$

Then choose a 0-neighbourhood $U$ in $E$ such that

$$u_m(\{x_{i_1^m}\}\times\cdots\times\{x_{i_{s-1}^m}\}\times U\times B_{i_{s+1}^m}\times\cdots\times B_{i_{l_m}^m})\subset V.$$

Hence

$$w_\nu(u_1,\cdots,u_k,v)(\{x_1\}\times\cdots\times\{x_{p-1}\}\times U\times B_{p+1}\times\cdots\times B_n)$$
$$\subset v(\bar{B}_1\times\cdots\times\bar{B}_{m-1}\times V\times\bar{B}_{m+1}\times\cdots\times\bar{B}_k)$$
$$\subset W.$$

This shows $w_\nu$ is well-defined. Clearly $w_\nu$ is multilinear and maps bounded sets into bounded sets. So it remains to show that $w_\nu$ is separately continuous.

*Case* (i). $u_1,\cdots,u_k$ are fixed and $v^\alpha\to 0$. Let $(B_1,\cdots,B_n,W)$ be a 0-neighbourhood in $\mathscr{L}_n(E,G)$. Put $\bar{B}_j=u_j(B_{i_1^j}\times\cdots\times B_{i_{l_j}^j})$ $(j=1,\cdots,k)$. Each $\bar{B}_j$ is a bounded subset of $F$. Now $v^\alpha\in(\bar{B}_1,\cdots,\bar{B}_k,W)$ eventually. Hence $w_\nu(u_1,\cdots,u_k,v^\alpha)\in(B_1,\cdots,B_n,W)$ eventually.

*Case* (ii). $v,u_1,\cdots,u_{m-1},u_{m+1},\cdots,u_k$ are fixed and $u_m^\alpha\to 0$. Put $\bar{B}_j=u_j(B_{i_1^j}\times\cdots\times B_{i_{l_j}^j})$ $(j\in\{1,2,\cdots,k\}, j\neq m)$. Since $v\in\mathscr{L}H(F,G)$, there exists a 0-neighbourhood $U$ in $F$ such that $v(\bar{B}_1\times\cdots\times\bar{B}_{m-1}\times U\times\bar{B}_{m+1}\times\cdots\times\bar{B}_k)\subset W$. Now $u_m^\alpha\in(B_{i_1^m},\cdots,B_{i_{l_m}^m},U)$ eventually. Hence $w_\nu(u_1,\cdots,u_{m-1},u_m^\alpha,u_{m+1},\cdots,u_k,v)\in(B_1,\cdots,B_n,W)$ eventually.

This completes the proof of 3.2. Note that $\mathscr{L}H_k(F,G)$ was needed, rather than just $\mathscr{L}_k(F,G)$.

THEOREM 3.3. (Higher Order Chain Rule, compare with Averbuh and Smoljanov (1967), p. 234). *Let $f:U\to V$ be $n$-times Fréchet (resp. Hadamard) differentiable at $x\in U$ and $g:V\to W$ be $n$-times Fréchet (resp.*

https://doi.org/10.1017/S1446788700021091 Published online by Cambridge University Press

*Hadamard) differentiable at* $f(x)$, *where* $U \in \mathcal{O}(E)$, $V \in \mathcal{O}(F)$, $W \in \mathcal{O}(G)$ *and* $G \in \mathrm{TLS}^*$. *Then* $g \circ f$ *is* $n$*-times Fréchet (resp. Hadamard) differentiable at* $x$ *and*

$$(g \circ f)^{(n)}(x) = \sum_{\nu} w_\nu(f^{(l_1)}(x), \cdots, f^{(l_k)}(x), g^{(k)}(f(x))),$$

*where the sum is over all partitions* $\nu$ *of* $\{1,2,\cdots,n\}$ *into* $k$ *disjoint non-empty sets* $\{i_1^1, \cdots, i_{l_1}^1\}, \cdots, \{i_1^k, \cdots, i_{l_k}^k\}$ *with* $l_1 + \cdots + l_k = n$, $l_j > 0$, $i_1^j < i_2^j < \cdots < i_{l_j}^j$ $(j = 1,2,\cdots,k)$, $i_{l_1}^1 < i_{l_2}^2 < \cdots < i_{l_k}^k$ *and with* $k$ *varying from* 1 *to* $n$. *In other words,*

$$(g \circ f)^{(n)}(x)(h_1, \cdots, h_n) = \sum_{\nu} g^{(k)}(f(x))(f^{(l_1)}(x)(h_{i_1^1}, \cdots, h_{i_{l_1}^1}),$$
$$\cdots, f^{(l_k)}(x)(h_{i_1^k}, \cdots, h_{i_{l_k}^k})).$$

*In particular,*

$$(g \circ f)^{(n)}(x) \cdot h^{(n)} = \sum_{\mu} \binom{n-1}{l_k - 1} \binom{n - l_k - 1}{l_{k-1} - 1}$$
$$\cdots \binom{l_1 - 1}{l_1 - 1} g^{(k)}(f(x))(f^{(l_1)}(x) \cdot h^{(l_1)}, \cdots, f^{(l_k)}(x) \cdot h^{(l_k)}),$$

*where the sum is over all (ordered)* $k$*-tuples* $\mu = (l_1, \cdots, l_k)$ *of positive integers such that* $l_1 + \cdots + l_k = n$ *and with* $k$ *varying from* 1 *to* $n$.

Proof. The proof is by induction, the case $n = 1$ being well known with $(g \circ f)'(x) = g'(f(x)) \circ f'(x)$.

So suppose the theorem is true for some $n$. Let $f$ be $(n+1)$-times differentiable (Fréchet or Hadamard) at $x$ and $g$ be $(n+1)$-times differentiable at $f(x)$. Hence by the induction hypothesis, $(g \circ f)^{(n)}(y)$ exists for each $y \in U$ and

$$(g \circ f)^{(n)}(y) = \sum_{\nu} w_\nu(f^{(l_1)}(y), \cdots, f^{(l_k)}(y), g^{(k)}(f(y))).$$

Now $w_\nu$ is Fréchet differentiable by 3.1 and 3.2. Consequently, $(g \circ f)^{(n)}$ is differentiable at $x$ and

$$(g \circ f)^{(n+1)}(x) \cdot h_0 = \sum_{\nu} w'_\nu(f^{(l_1)}(x), \cdots, f^{(l_k)}(x), g^{(k)}(f(x))) \cdot f^{(l_1+1)}(x) \cdot h_0,$$
$$\cdots, f^{(l_k+1)}(x) \cdot h_0, g^{(k+1)}(f(x)) \cdot (f'(x) \cdot h_0))$$
$$= \sum_{\nu} [w_\nu(f^{(l_1+1)}(x) \cdot h_0, f^{(l_2)}(x), \cdots, f^{(l_k)}(x), g^{(k)}(f(x)))$$

$$+ w_\nu(f^{(l_1)}(x), f^{(l_2+1)}(x)\cdot h_0, \cdots, f^{(l_k)}(x), g^{(k)}(f(x)))$$
$$+ \cdots$$
$$+ w_\nu(f^{(l_1)}(x), \cdots, f^{(l_{k-1})}(x), f^{(l_k+1)}(x)\cdot h_0, g^{(k)}(f(x)))$$
$$+ w_\nu(f^{(l_1)}(x), \cdots, f^{(l_k)}(x), g^{(k+1)}(f(x))\cdot(f'(x)\cdot h_0))].$$

Thus

$$(g\circ f)^{(n+1)}(x)(h_0, h_1, \cdots, h_n)$$
$$= \sum_\nu [g^{(k)}(f(x))(f^{(l_1+1)}(x)(h_0, h_{i_1^1}, \cdots, h_{i_{l_1}^1}),$$
$$f^{(l_2)}(x)(h_{i_1^2}, \cdots, h_{i_{l_2}^2}), \cdots, f^{(l_k)}(x)(h_{i_1^k}, \cdots, h_{i_{l_k}^k}))$$
$$+ \cdots$$
$$+ g^{(k)}(f(x))(f^{(l_1)}(x)(h_{i_1^1}, \cdots, h_{i_{l_1}^1}), \cdots, f^{(l_k+1)}(x)(h_0, h_{i_1^k}, \cdots, h_{i_{l_k}^k}))$$
$$+ g^{(k+1)}(f(x))(f'(x)\cdot h_0, f^{(l_1)}(x)(h_{i_1^1}, \cdots, h_{i_{l_1}^1}),$$
$$\cdots, f^{(l_k)}(x)(h_{i_1^k}, \cdots, h_{i_{l_k}^k}))]$$
$$= \sum_\mu g^{(p)}(f(x))(f^{(s_1)}(x)(h_{j_1^1}, \cdots, h_{j_{s_1}^1}), \cdots, f^{(s_p)}(x)(h_{j_1^p}, \cdots, h_{j_{s_p}^p})),$$

where the sum is over all partitions $\mu$ of $\{0, 1, \cdots, n\}$ into $p$ non-empty disjoint sets $\{j_1^1, \cdots, j_{s_1}^1\}, \cdots, \{j_1^p, \cdots, j_{s_p}^p\}$ with $s_1 + \cdots + s_p = n + 1$, $s_i > 0$, $j_1^i < j_2^i < \cdots < j_{s_i}^i$ $(i = 1, \cdots, p)$, $j_{s_1}^1 < j_{s_2}^2 < \cdots < j_{s_p}^p$ and $p$ varies from 1 to $n + 1$.

Now, suppose $h_1 = h_2 = \cdots h_n = h$. Then many of the terms in $(g\circ f)^{(n)}(x)(h_1, \cdots, h_n)$ will be equal and can be collected together. Let $(l_1, \cdots, l_k)$ be an (ordered) $k$-tuple of positive integers such that $l_1 + \cdots + l_k = n$. How many terms of the form $g^{(k)}(f(x))(f^{(l_1)}(x)\cdot h^{(l_1)}, \cdots, f^{(l_k)}(x)\cdot h^{(l_k)})$ will there be? This is the same as asking: how many ways can we partition $\{1, 2, \cdots, n\}$ into (ordered) $k$-tuples $(\{i_1^1, \cdots, i_{l_1}^1\}, \cdots, \{i_1^k, \cdots, i_{l_k}^k\})$ with $l_1 + \cdots + l_k = n$, $l_j > 0$, $i_1^j < \cdots < i_{l_j}^j$ $(j = 1, 2, \cdots, k)$ and $i_{l_1}^1 < i_{l_2}^2 < \cdots < i_{l_k}^k$?

Consider the last set $\{i_1^k, \cdots, i_{l_k}^k\}$ in the $k$-tuple. Because of the ordering restrictions, $i_{l_k}^k$ is fixed. It must be $n$. With this fixed, the other $l_k - 1$ elements can be chosen arbitrarily and there are $\binom{n-1}{l_k-1}$ ways of doing this.

Now look at the second last set $\{i_1^{k-1}, \cdots, i_{l_{k-1}}^{k-1}\}$. This time $i_{l_{k-1}}^{k-1}$ is fixed by the ordering restrictions. It must be the largest integer in $\{1, 2, \cdots, n\}$, which is not in $\{i_1^k, \cdots, i_{l_k}^k\}$. The other $l_{k-1} - 1$ elements can be chosen arbitrarily and there are $\binom{n - l_k - 1}{l_{k-1} - 1}$ ways of doing this. This process leads to the coefficient

$\binom{n-1}{l_k-1}\binom{n-l_k-1}{l_{k-1}-1}\cdots\binom{l_1-1}{l_1-1}$, as required.

This completes the proof of 3.3.

When the $\sigma$-derivative is either the Hadamard or Fréchet derivative, one can deduce Leibniz' theorem (2.4) from the higher order chain rule.

If it is known that all derivatives $g^{(k)}(f(x))$ and $f^{(l)}(x)$ are symmetric, then

$$(g\circ f)^{(n)}(x)(h_1,\cdots,h_n) = \sum_{\mu}\sum_{\rho\in S_n}\frac{1}{l_1!\cdots l_k!k!}g^{(k)}(f(x))(f^{(l_1)}(x)(h_{\rho(1)},\cdots,h_{\rho(l_1)}),$$
$$\cdots,f^{(l_k)}(x)(h_{\rho(n-l_k+1)},\cdots,h_{\rho(n)})),$$

where the first sum is over all (ordered) $k$-tuples $\mu=(l_1,\cdots,l_k)$ of positive integers such that $l_1+\cdots+l_k=n$ and with $k$ varying from 1 to $n$.

In particular,

$$(g\circ f)^{(n)}(x)\cdot h^{(n)} = \sum_{\mu}\frac{n!}{l_1!\cdots l_k!k!}g^{(k)}(f(x))(f^{(l_1)}(x)\cdot h^{(l_1)},$$
$$\cdots,f^{(l_k)}(x)\cdot h^{(l_k)}).$$

This formula was given in Penot (1973), p. 8. Still under the hypothesis of symmetry, this last formula can also be written:

$$(g\circ f)^{(n)}(x)\cdot h^{(n)} = \sum\frac{n!}{l_k!\cdots l_1!m_n!\cdots m_1!}g^{(k)}(f(x))(f^{(l_1)}(x)\cdot h^{(l_1)},$$
$$\cdots,f^{(l_k)}(x)\cdot h^{(l_k)}),$$

where the sum is over all (ordered) $k$-tuples $(l_1,\cdots,l_k)$ of positive integers such that $l_1\leqq l_2\leqq\cdots\leqq l_k$ and $l_1+\cdots+l_k=n$, with $k$ varying from 1 to $n$. $m_j$ $(j=1,\cdots,n)$ is the number of numbers $l_1,\cdots,l_k$ equal to $j$.

This last formula is given in Averbuh and Smoljanov (1967), p. 234 and Schwartz (1967), p. 262.

Under mild restrictions, there is a rather simpler proof of the higher order chain rule. This depends on the following lemma.

LEMMA 3.4. (Lloyd (1973), p. 16). *Let $u\in\mathscr{L}H(E_1,E_2;F)$. Then $u$ is infinitely Fréchet differentiable.*

Now suppose, with the same notation as Theorem 3.3, comp $\in\mathscr{L}H(\mathscr{L}(E,F),\mathscr{L}(F,G);\mathscr{L}(E,G))$. Then comp is infinitely Fréchet differentiable. A sufficient condition for comp to have this property is for $F$ and $G$ to be locally convex with $F$ quasi-barrelled. Alternatively, it suffices that $F$ be a Baire space.

Then a simple proof that the composite of $n$-times (Fréchet or Hadamard) differentiable maps is $n$-times differentiable can be obtained by slightly modifying the standard argument in Dieudonné (1969), Theorem 8.12.10. One can then obtain the formula for $(g \circ f)^{(n)}(x)$ by using the Leibniz formula to differentiate $n$-times the formula $(g \circ f)'(x) = \operatorname{comp}(f'(x), g'(f(x)))$.

Finally, we give a counterexample to Abraham and Robbin's version of the formula for $(g \circ f)^{(n)}(x)$ (Abraham and Robin (1967), p. 3). Let $E$ be a normed space. Put $\mathscr{L}(E) = \mathscr{L}(E, E)$. Suppose there exist $x, h_1, h_2 \in \mathscr{L}(E)$ such that

$$x \circ h_1 \circ h_2 + h_2 \circ h_1 \circ x \neq x \circ h_2 \circ h_1 + h_1 \circ h_2 \circ x.$$

Consider the map $f : \mathscr{L}(E) \to \mathscr{L}(E)$ defined by $f(\underline{x}) = x \circ x$. Put $g = f$. Then $g \circ f : \mathscr{L}(E) \to \mathscr{L}(E)$ is given by $(g \circ f)(x) = x \circ x \circ x \circ x$.

According to Abraham and Robbin's (1967), p. 3,

$$\begin{aligned}(g \circ f)^{(3)}(x)(h_1, h_2, h_3) &= g'(f(x)) \cdot f^{(3)}(x)(h_1, h_2, h_3) \\ &\quad + g^{(2)}(f(x))(f^{(2)}(x)(h_1, h_2), f'(x) \cdot h_3) \\ &\quad + 2g^{(2)}(f(x))(f'(x) \cdot h_1, f^{(2)}(x)(h_2, h_3)) \\ &\quad + g^{(3)}(f(x))(f'(x) \cdot h_1, f'(x) \cdot h_2, f'(x) \cdot h_3).\end{aligned}$$

We show this formula for $(g \circ f)^{(3)}(x)$ is not symmetric. Let $h_3$ = identity on $E$. Then we show

$$(g \circ f)^{(3)}(x)(h_1, h_2, h_3) \neq (g \circ f)^{(3)}(x)(h_2, h_1, h_3).$$

For this, we need to show that

$$g^{(2)}(f(x))(f'(x) \cdot h_1, f^{(2)}(x)(h_2, h_3)) \neq g^{(2)}(f(x))(f'(x) \cdot h_2, f^{(2)}(x)(h_1, h_3)).$$

Now

$$\begin{aligned}&g^{(2)}(f(x))(f'(x) \cdot h_1, f^{(2)}(x)(h_2, h_3)) \\ &\qquad = 2x \circ h_1 \circ h_2 + 2h_1 \circ x \circ h_2 + 2h_2 \circ h_1 \circ x + 2h_2 \circ x \circ h_1\end{aligned}$$

and

$$\begin{aligned}&g^{(2)}(f(x))(f'(x) \cdot h_2, f^{(2)}(x)(h_1, h_3)) \\ &\qquad = 2x \circ h_2 \circ h_1 + 2h_2 \circ x \circ h_1 + 2h_1 \circ h_2 \circ x + 2h_1 \circ x \circ h_2.\end{aligned}$$

By our choice of $x$, $h_1$ and $h_2$, these two expressions differ.

However, the coefficients $\sigma_k$ in Abraham and Robbin's formula turn out to be exactly the same as the coefficients $\binom{n-1}{l_k-1}\binom{n-l_k-1}{l_{k-1}-1}\cdots\binom{l_1-1}{l_1-1}$ in Theorem 3.3, and so their formula is correct if $h_1 = h_2 = \cdots = h_n = h$.

## REFERENCES

R. Abraham and J. Robbin (1967), ***Transversal mappings and flows*** (W. A. Benjamin, New York).

V. I. Averbuh and O. G. Smoljanov (1967), 'Differentiation theory in linear topological spaces', *Uspehi Mat. Nauk* **6**, 201–260 = *Russian Math. Surveys* **6**, 201–258.

V. I. Averbuh and O. G. Smoljanov (1968), 'Different definitions of derivative in linear topological spaces', *Uspehi Mat. Nauk*, **4**, 67–116 = *Russian Math. Surveys* **4**, 67–113.

J. Dieudonné (1969), *Foundations of modern analysis* (Academic Press, New York, London).

J. W. Lloyd (1973), Ph.D. Dissertation, Australian National University, Canberra.

J. W. Lloyd (1972), 'Inductive and projective limits of smooth topological vector spaces', *Bull. Austral. Math. Soc.* **6**, 227–240.

J. W. Lloyd (1974), 'Smooth partitions of unity on manifolds', *Trans. Amer. Math. Soc.* **187**, 249–259.

J. W. Lloyd (1975), 'Smooth locally convex spaces', *Trans. Amer. Math. Soc.* **212**, 383–392.

J.-P. Penot (1973), 'Calcul différentiel dans les espaces vectoriels topologiques', *Studia Math.* **47**, 1–23.

L. Schwartz (1967), *Cours d'analyse* (Hermann, Paris).

Department of Mathematics,
University of Melbourne,
Parkville, Victoria 3052, Australia.