J. Austral. Math. Soc. 19 (Series A), (1975), 121–128.

DIFFERENTIATION OF COMPOSITES WITH RESPECT TO A PARAMETER

ALISTAIR GRAY[1]

(Received 21 April 1973; revised 29 November 1973)

Communicated by E. Strzelecki

A formula is established for differentiating the composite $F(\xi) \circ G(\xi)$ with respect to the parameter ξ where F and G are maps which assume their values in function spaces. This composite is denoted by $F \underset{-}{\circ} G(\xi)$. Use of this notation, together with the tangent functor T, enables the formula to be written in the variable-free form

$$T(F \underset{-}{\circ} G) = \underline{T}(F \circ \Pi_1) \underset{-}{\circ} TG + TF \underset{-}{\circ} (G \circ \Pi_1)$$

where Π_1 is a projection map and $\underline{T}F(\xi) = T(F(\xi))$. Formal verification of this formula is straightforward. It has application to many small divisor problems such as those which occur in celestial mechanics.

1. Notation

Let X be an open set containing 0 in a normed vector space V_X and let A, B and C be open sets in V_A, V_B and V_C respectively, where V_A, V_B and V_C are vector spaces with norm denoted by $|\ \ |$. For each such A and B we put

$$\mathscr{F}(A,B) = \{\xi : A \to B \,|\, \xi \text{ is bounded and continuously differentiable}\}$$

This vector space will be assumed to have the sup norm $\|\ \ \|$.

Now for each $\xi \in X$, we may form the composite of the two functions $F(\xi): B \to C$ and $G(\xi): A \to B$ to give

$$F(\xi) \circ G(\xi): A \to C.$$

We wish to differentiate this composite with respect to the parameter ξ, that is, we wish to differentiate the operator

$$F \underset{-}{\circ} G: X \to \mathscr{F}(A,C): \xi \mapsto F(\xi) \circ G(\xi)$$

[1] The author wishes to thank Dr Arthur Jones for his help in the preparation of this paper.

We shall use the symbol "D" to denote *Fréchet differentiation*, as defined for instance in Dieudonné (1960), page 143. For differentiating composites, the chain rule is available Dieudonné (1960), page 145. The aim of this paper is to prove an analogous formula for differentiating the operator $F \circ G$.

It is important to note that Fréchet differentiation may occur at two levels. Thus the derivative at $\xi \in X$ of a *map* $F: X \to \mathscr{F}(B, C)$ is a map

$$DF(\xi): V_X \to \mathscr{F}(B, C)$$

while the derivative of the *function* $F(\xi): B \to C$ at $b \in B$ is a function

$$D(F(\xi))(b): V_B \to V_C.$$

In the present context, however, the Fréchet derivative D is less natural than the *tangent functor* T, Abraham (1967) page 8, which is defined in terms of D by putting

$$TF: Y \times V_Y \to W \times W: (\xi, \eta) \mapsto (F(\xi), DF(\xi) \cdot (\eta))$$

where $F: Y \to Z$ maps an open subset Y of a normed vector space V_Y into another normed vector space W.

The introduction of the tangent functor makes it possible to express the usual chain rule in variable-free notation, Abraham (1967) page 8, and it will play a similar role here.

2. Preliminary results

The symbol "o_i" ($i = 1, 2, 3 \cdots$) will be used to denote a continuous map between normed vector spaces with the property

$$\lim_{v \to 0} \frac{\| o_i(v) \|}{\| v \|} = 0$$

LEMMA 1. *Suppose that the function $\mu \in \mathscr{F}(B, C)$ has derivative $D\mu$ which is bounded in the sense that for some positive M and all $b \in B$, $x \in V_B$*

$$|D\mu(b) \cdot (x)| < M|x|.$$

For each map $o_1: \mathscr{F}(A, B) \to \mathscr{F}(A, B)$ and function $\beta: A \to B$ there is then a map $o_2: \mathscr{F}(A, B) \to \mathscr{F}(A, C)$ such that, for all $a \in A$,

$$D\mu(\beta(a)) \cdot (o_1(\eta)(a)) = o_2(\eta)(a).$$

PROOF. We define o_2 by the above equation and then get

$$0 \leqq \lim_{\eta \to 0} \| o_2(\eta) \| / \| \eta \| = \lim_{\eta \to 0} (\sup_{a \in A} |D\mu(\beta(a)) \cdot (o_1(\eta)(a))|) / \| \eta \|$$

$$\leqq \lim_{\eta \to 0} M \| o_1(\eta) \| / \| \eta \| = 0, \text{ as required.}$$

LEMMA 2. *Let* $G: X \to \mathscr{F}(A,B)$ *be a differentiable map. Then for each map* $o_1: B \to C$ *and for each map* $o_2: X \to \mathscr{F}(A,B)$ *and for each* ξ, X *there is a map* $o_3: X \to \mathscr{F}(A,C)$ *such that for all* $\eta \in X$

$$o_1(DG(\xi)\cdot(\eta)(a) + o_2(\eta)(a)) = o_3(\eta)(a)$$

PROOF. Because G is differentiable for each $\xi \in X$ there is a positive constant K such that for all $\eta \in X$

$$\text{(I)} \qquad \| DG(\xi)\cdot(\eta)\| \leqq K \|\eta\|.$$

By definition of o_1 for any positive ε there is a positive δ such that for all $a \in A$

$$\text{(II)} \qquad |DG(\xi)\cdot(\eta)(a) + o_2(\eta)(a)| < \delta \Rightarrow$$
$$|o_1(DG(\xi)\cdot(\eta)(a) + o_2(\eta)(a))| \leqq \varepsilon |DG(\xi)\cdot(\eta)(a) + o_2(\eta)(a)|.$$

But using the triangle inequality, (I) above together with the definition of o_2 we get

$$\begin{aligned}|DG(\xi)\cdot(\eta)(a) + o_2(\eta)(a)| &\leqq |DG(\xi)\cdot(\eta)(a)| + |o_2(\eta)(a)| \\ &\leqq K\|\eta\| + L\|\eta\| \text{ for some } L > 0 \\ &\to 0 \text{ as } \eta \to 0.\end{aligned}$$

Combining this with (II) gives the required result.

LEMMA 3. *Let* $H: X \to \mathscr{F}(B,C)$ *and assume that there is a positive constant* M *such that for each* $b \in B$, *each* $y \in B$ *and each sufficiently small* $\eta \in X$

$$|D(H(\eta))(b)\cdot(y)| \leqq M\|\eta\| |y|.$$

Then for each $\xi \in X$ *and* $\beta: A \to B$ *there is a map* $o_1: X \to \mathscr{F}(A,C)$ *such that*

$$D(H(\eta))(\beta(a))\cdot(DG(\xi)\cdot(\eta)(a)) = o_1(\eta)(a)$$

provided $G: X \to \mathscr{F}(A,B)$ *is differentiable.*

PROOF.
$$\begin{aligned}0 &\leqq \lim_{\eta\to 0} |D(H(\eta))(\beta(a))\cdot DG(\xi)\cdot(\eta)(a)| / \|\eta\| \\ &\leqq \lim_{\eta\to 0} M\|\eta\| |DG(\xi)\cdot(\eta)(a)| / \|\eta\| \\ &\leqq \lim_{\eta\to 0} MN\|\eta\| \qquad \text{for some positive constant } N \\ &= 0\end{aligned}$$

which gives the result.

3. The differentiation formula

THEOREM. *Let the maps* $F: X \to \mathscr{F}(B,C)$ *and* $G: X \to \mathscr{F}(A,B)$ *be differentiable and assume that they have the following properties.*

(I) *For each* $\xi, \eta \in X$, *the functions* $F(\xi): B \to C$ *and* $DF(\xi)\cdot(\eta): B \to C$

https://doi.org/10.1017/S1446788700023594 Published online by Cambridge University Press

are differentiable and have derivatives bounded over the domain B.

(II) *For each $\xi \in X$ there are positive constants M and K such that for each $b \in B$ and every $y \in B$*

$$\left| D(DF(\xi)\cdot(\eta))(b)\cdot(y) \right| \leqq M \|\eta\| \, |y|$$

provided $\|\eta\| < K$.

The operator $F \underset{-}{\circ} G \colon X \to \mathscr{F}(A, C)$ *is then differentiable and*

$$T(F \underset{-}{\circ} G) = \underline{T}(F \circ \Pi_1) \underset{-}{\circ} TG + TF \underset{-}{\circ} (G \circ \Pi_1)$$

where $\Pi_1 \colon X \times X \to X \colon (\xi, \eta) \times \xi$ *and where* $\underline{T}F(\xi) = T(F(\xi))$.

Proof. $D(F \underset{-}{\circ} G)$ is defined by the condition

$$D(F \underset{-}{\circ} G)(\xi)\cdot(\eta) = F \underset{-}{\circ} G(\xi+\eta) - F \underset{-}{\circ} G(\xi) + o_1(\eta)$$

together with the requirement that $D(F \underset{-}{\circ} G)(\xi)\cdot(\eta)$ be continuous and linear in η. So consider

$$\begin{aligned} F \underset{-}{\circ} G(\xi+\eta) &= F(\xi+\eta) \circ G(\xi+\eta) \\ &= (F(\xi) + DF(\xi)\cdot(\eta) + o_2(\eta)) \circ G(\xi+\eta) \\ &= F(\xi) \circ G(\xi+\eta) + DF(\xi)\cdot(\eta) \circ G(\xi+\eta) + o_2(\eta) \circ G(\xi+\eta) \end{aligned}$$

Now consider the right hand side of the above equation term by term, evaluated at $a \in A$.

The *first term* is

$$\begin{aligned} F(\xi) \circ G(\xi+\eta)(a) &= F(\xi)(G(\xi)(a) + DG(\xi)\cdot(\eta)(a) + o_3(\eta)(a)) \\ &= F(\xi) \circ G(\xi)(a) + D(F(\xi))(G(\xi)(a))\cdot(DG(\xi)\cdot(\eta)(a) \\ &\qquad + o_3(\eta)(a)) \\ &\qquad + o_4(DG(\xi)\cdot(\eta)(a) + o_3(\eta)(a) \\ &= F(\xi) \circ G(\xi)(a) + D(F(\xi))(G(\xi)(a))\cdot(DG(\xi)\cdot(\eta)(a)) \\ &\qquad + o_5(\eta)(a) \text{ by Lemmas 1 and 2.} \end{aligned}$$

The *second term* is

$$\begin{aligned} DF(\xi)\cdot(\eta) \circ G(\xi+\eta)(a) &= DF(\xi)\cdot(\eta)(G(\xi)(a) + DG(\xi)\cdot(\eta)(a) + o_6(\eta)(a)) \\ &= DF(\xi)\cdot(\eta)(G(\xi)(a)) \\ &\quad + D(DF(\xi)\cdot(\xi)(a))\cdot(DG(\xi)\cdot(\eta)(a) + o_6(\eta)(a)) \\ &\quad + o_7(DG(\xi)\cdot(\eta)(a) + o_6(\eta)(a)) \\ &= DF(\xi)\cdot(\eta)(G(\xi)(a)) + o_8(\eta)(a) \text{ by hypothesis (II)} \end{aligned}$$

of the Theorem and Lemmas 1, 2 and 3.

https://doi.org/10.1017/S1446788700023594 Published online by Cambridge University Press

The *third term* is easily seen to be $o_9(\eta)(a)$.
Thus

$$D(F \underset{-}{\circ} G)(\xi)\cdot(\eta)(a) = D(F(\xi))(G(\xi)(a))\cdot(DG(\xi)\cdot(\eta)(a)) + DF(\xi)\cdot(\eta)(G(\xi)(a)).$$

Each of the terms on the right hand side will now be simplified by the use of the tangent functor.

Thus

$$\begin{aligned}(F \underset{-}{\circ} G)(\xi)(a), D(F(\xi))(G(\xi)(a))\cdot(DG(\xi)\cdot(\eta)(a))) &= T(F(\xi))(G(\xi)(a),\\ &\qquad DG(\xi)\cdot(\eta)(a))\\ &= \underline{T}(F\circ\Pi_1)\underset{-}{\circ} TG(\xi,\eta)(a)\end{aligned}$$

and

$$\begin{aligned}(F \underset{-}{\circ} G(\xi)(a), DF(\xi)\cdot(\eta)(G(\xi)(a))) &= TF(\xi,\eta)(G(\xi)(a))\\ &= TF \underset{-}{\circ} (G\circ\Pi_1)(\xi,\eta)(a))\end{aligned}$$

which completes the proof.

4. Applications of the formula

Sternberg (1969) gives a treatment of a number of small divisor problems including the Siegel centre theorem, the Moser twist theorem and the Kolmogorov-Arnold theorem. All of these problems involve differentiation of a composite with respect to a parameter. Sternberg gives the details of the differentiations only for the relatively simple case of a circle mapping problem. But even for this case his treatment is incomplete. Furthermore he uses a notation which confuses functions with their values.

Our formula can be used in all of of the above small divisor problems. We illustrate its use by redoing the differentiation which arises for the circle problem, Sternberg (1969) pages 88 and 89.

For $h>0$ the set of complex numbers $S[h]$ is defined by $S[h]=\{z:\operatorname{im} z|<h\}$. The set of periodic functions bounded and analytic on $S[h]$ with period 1 is denoted by $\mathscr{S}[h]$. The derivative, in the traditional sense, of a function $\mu\in\mathscr{S}[h]$, for some $h>0$, will be denoted by μ'.

Now let $X=\mathscr{S}[h+2k]\times\mathscr{S}[h+2k]$, $A=S[h]$, $B=S[h+k]$ and $C=S[h+2k]$. The operators P_2 and $I+P_1$ are defined by

$$P_2: X\to\mathscr{F}(B,C): (\xi,\alpha)\mapsto\alpha$$

$$I+P_1: X\to\mathscr{F}(A,B): (\xi,\alpha)\mapsto\iota+\xi$$

where ι denotes the identity function. Sternberg's (1969) composite "$\alpha\circ(\iota+\xi)$" may now be written in variable-free notation as H: $X\to\mathscr{F}(A,C)$ where $H=P_2\underset{-}{\circ}(I+P_1)$. It is required to find the second Fréchet derivative of H.

The following result will be found useful for the estimation of derivatives. Let $\xi \in \mathscr{S}[h+k]$ where $h > 0$ and $k > 0$. Then

$$\text{(1)} \qquad \sup_{z \in S[h]} \left| \xi^{(n)}(z) \right| \leqq n! \pi^{-1} k^{-n-1} \sup_{z \in S[h+k]} \left| \xi(z) \right|$$

where $n = 1, 2, 3, \cdots$. This may be proved by using Cauchy's formula and integrating around the contour $\{z : \operatorname{im} z| = h, 0 \leqq \operatorname{re} z \leqq 1\}$. The conditions of the theorem are now checked for the variable-free version of Sternberg's composite. The condition (II) of the theorem in this case is for each $(\xi, \alpha), (\eta, \mu) \in X \times X$

$$\left| D(DP_2(\xi, \alpha) \cdot (\eta, \mu))(b) \cdot (y) \right| \leqq M \left\| (\eta, \mu) \right\| \left| y \right|$$

i.e. $\left| \mu'(b) y \right| \leqq M \left\| (\eta, \mu) \right\| \left| y \right|$ for each $b \in B$ and every $y \in B$ for some positive constant M. This condition will be satisfied since by (1) above

$$\sup_{b \in S[h+k]} \left| \mu'(b) \right| \leqq (\pi k^2)^{-1} \sup_{b \in S[h+2k]} \left| \mu(b) \right|.$$

The other hypotheses of the theorem are easily verified.

Now

$$\begin{aligned} \underline{T}(P_2 \circ \Pi_1) \underline{\circ}\, T(I + P_1)((\xi, \alpha), (\eta, \mu))(a) & \\ = T\alpha \circ ((\iota + \xi)(a), \eta(a)) & \\ = (\alpha \circ (\iota + \xi)(a), \alpha' \circ (\iota + \xi)(a) \eta(a)) & \end{aligned}$$

and

$$\begin{aligned} TP_2 \underline{\circ}\, ((I + P_1) \circ \Pi_1)((\xi, \alpha), (\eta, \mu))(a) & \\ = TP_2((\xi, \alpha), (\eta, \mu)) \circ (\iota + \xi)(a) & \\ = (\alpha \circ (\iota + \xi)(a), \mu \circ (\iota + \xi)(a)) & \end{aligned}$$

thus

$$\text{(2)} \qquad DH(\xi, \alpha) \cdot (\eta, \mu) = \alpha' \circ (\iota + \xi)\eta + \mu \circ (\iota + \xi).$$

The maps P_2', K_μ and K_η are defined by

$$P_2' : X \to \mathscr{F}(B, C) : (\xi, \alpha) \to \alpha'$$

$$K_\mu : X \to \mathscr{F}(B, C) : (\xi, \alpha) \to \mu$$

$$K_\eta : X \to \mathscr{F}(A, B) : (\xi, \alpha) \to \eta.$$

Note that in this context K_μ and K_η are constant maps and that $\alpha' \in \mathscr{S}[h+k]$ since $\alpha \in \mathscr{S}[h+2k]$. Differentiation of (2) involves the differentiation of a product of an operator with a composite and differentiation of another composite.

Consider first $J : X \to \mathscr{F}(A, C)$ where $J = P_2 \underline{\circ} (I + P_1)$. Condition (II) of the theorem becomes

$$|D(DP_2'(\xi,\alpha)\cdot(\gamma,\delta))(b)\cdot(y)| \leqq M\|(\gamma,\delta)\|\,|y|$$

i.e. $|\delta''(b)y| \leqq M\|(\gamma,\delta)\|\,|y|$

for each $b \in B$ and every $y \in B$ for some positive constant M. This will be satisfied since by (1) above

$$\sup_{b \in S[h]} |\delta''(b)| \leqq (\pi k^3)^{-1} \sup_{b \in S[h+k]} |\delta(b)|.$$

The other hypotheses are again easily verified.

Now

$$\underline{T}(P_2' \circ \Pi_1) \underline{\circ}\, T(I+P_1)((\xi,\alpha),(\gamma,\delta))(a) = T\alpha'((\iota+\xi)(a),\gamma(a))$$
$$= (\alpha' \circ (\iota+\xi)(a), \alpha'' \circ (\iota+\xi)(a)\delta(a))$$

and

$$TP_2' \underline{\circ}\, ((I+P_1)\circ \Pi_1)((\xi,\alpha),(\gamma,\delta))(a) = TP_2((\xi,\alpha)',(\gamma,\delta))\circ(\iota+\xi)(a)$$
$$= (\alpha'\circ(\iota+\xi)(a), \gamma'\circ(\iota+\xi)(a))$$

thus

$$DJ(\xi,\alpha)\cdot(\gamma,\delta) = \alpha''\circ(\iota+\xi)\delta + \gamma'\circ(\iota+\xi).$$

The derivative of $Q: X \to \mathscr{F}(A,C)$ where $Q = K_\mu \underline{\circ}\, (I+P_1)$ is now found. The conditions (II) of the theorem is satisfied since

$$DK_\mu(\xi,\alpha)\cdot(\gamma,\delta) = 0.$$

Now

$$\underline{T}(K_\mu \circ \Pi_1)\underline{\circ}\, T(I+P_1)((\xi,\alpha),(\gamma,\delta))(a) = T\mu((\iota+\xi)(a),\gamma(a))$$
$$= (\mu\circ(\iota+\xi)(a), \mu'\circ(\iota+\xi)(a)\gamma(a))$$

and

$$TK_\mu \underline{\circ}\, ((I+P_1)\circ\Pi_1)((\xi,\alpha),(\gamma,\delta))(a) = (\mu\circ(\iota+\xi),0).$$

These results together with the usual product formula yield the following.

$$D^2H(\xi,\alpha)\,.\,((\mu,\eta),(\gamma,\delta))$$
$$= \alpha''\circ(\iota+\xi)\gamma\mu + \delta'\circ(\iota+\xi)\mu + \eta'\circ(\iota+\xi)\gamma$$

and so the answer to Sternberg's problem (1969), page 89, is then

$$D^2H(\xi,\alpha)\cdot((\mu,\eta)\,.\,(\mu,\eta)) = \alpha''\circ(\iota+\xi)\mu^2 + 2\eta'\circ(\iota+\xi)\mu.$$

https://doi.org/10.1017/S1446788700023594 Published online by Cambridge University Press

References

R. Abraham (1967), *Foundations of Mechanics* (W. A. Benjamin, Inc.; New York, 1967).

J. Dieudonné (1967), *Foundations of Modern Analysis* (Academic Press; New York and London; 1960).

Shlomo Sternberg (1969), *Celestial Mechanics, Part II* (W. A. Benjamin, Inc.; New York; 1969).

Department of Pure Mathematics
La Trobe University
Victoria, 3083
Australia

https://doi.org/10.1017/S1446788700023594 Published online by Cambridge University Press