Mathematical Proceedings of the Cambridge Philosophical Society

VOL. 83 MARCH 1978 PART 2

*Math. Proc. Camb. Phil. Soc.* (1978), **83**, 159
*Printed in Great Britain*



# Formulae for high derivatives of composite functions

BY L. E. FRAENKEL
*Mathematics Division, University of Sussex*

(*Received* 12 *September* 1977)

1. *Introduction.* This note concerns a question of elementary calculus. Given a smooth composite function $u = g \circ f$ [with values $u(x) = g(f(x))$], we write explicit formulae for its derivatives, of arbitrary order, in terms of derivatives of $f$ and $g$. We consider (A) the general case,

$$u = g \circ f, \quad U \xrightarrow{f} V \xrightarrow{g} G, \quad U \subset E \quad \text{and} \quad V \subset F, \tag{1}$$

in which $E$, $F$ and $G$ are Banach spaces, and $U$, $V$ are open sets; (B) the finite-dimensional case $E = \mathbb{R}^M$ and $F = \mathbb{R}^N$, where $\mathbb{R}^M$ denotes real $M$-dimensional Euclidean space; and (C) the particular case of (B) (due to restricting $g$ to part of an $M$-dimensional surface in $\mathbb{R}^{M+1}$) in which $N = M + 1$ and $u(x) = g(x, \phi(x))$, so that $\phi$ denotes a real-valued (scalar-valued) function of $x = (x_1, \ldots, x_M)$.

Case (B) deserves separate treatment because it involves *bona fide* partial derivatives (in (A), only multi-linear operators are available to us), and it is easier to derive this result *ab initio* than to treat it as a particular form of the result for (A). Case (C) is included because it arises frequently in its own right, and because the formula is less monstrous than that for (B). Of course, functions with values $g(x_1, \phi(x_1, x_2))$ or $g(x_1, x_2, \phi(x_1))$, say, are covered by (C) because these values can be regarded as $\hat{g}(x_1, x_2, \hat{\phi}(x_1, x_2))$ with $D_2\hat{g} = 0$ or $D_2\hat{\phi} = 0$, respectively.

The progenitor of the formulae in this note is the result of de Bruno (2) for the one-dimensional case $E = F = \mathbb{R}$; it is quoted by Goursat ((7), p. 34) and Chaundy ((4), p. 131). Cartan ((3), p. 86) describes the method of constructing what we shall call formula A, and Dieudonné ((6), pp. 149, 429) does virtually the same for formula B in the context of Taylor's formula in a Lie group. However, these authors do not display the fruits of their work in an explicit form, and I have not found such a display in the literature. Of course, the formulae are unwieldy and contrary to the style of much modern mathematics.



https://doi.org/10.1017/S0305004100054402 Published online by Cambridge University Press

Nevertheless, there are circumstances in which results are proved more easily by means of such formulae than by a less precise description of the derivatives in question. Abraham and Robbin (1) evidently share this view in the context of jets and of preparations for Sard's theorem, but the formula that they give ((1), p. 3) is incorrect.† Another application, set out at the end of this note for the simplest case, is to functions constructed by means of the implicit function theorem; for estimates of high derivatives of such a function, it is helpful to have a formula of the present kind (although sharp estimates are difficult even then). A third example is provided by Stein's beautiful extension theorem ((8), pp. 180–192) for functions in Sobolev spaces, defined on a domain with minimally smooth boundary. With the present notation, this extension involves a function $g(.,\phi(.))$ in which $D^k\phi(x)\to\infty$ like $d^{1-k}$ as

$$d = \operatorname{dist}(x, \partial U)\to 0$$

(here $D^k\phi$ denotes any partial derivative of order $k \geqslant 2$, and $\partial U$ is the boundary of $U$). Stein remarks ((8), p. 185) that 'the argument is entirely typical' for a particular derivative of order two, and this is quite true, but I was able to see why, for derivatives of arbitrary order, only with the help of a formula.

2. *Notation and terminology.* An *n-index*, usually called a *multi-index*, is an ordered $n$-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of integers $\alpha_i \geqslant 0$, and $|\alpha| = \alpha_1 + \ldots + \alpha_n$ is called the *order* of $\alpha$ because partial derivatives can be written in terms of the symbol

$$D^\alpha = D_1^{\alpha_1} \ldots D_n^{\alpha_n} = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \ldots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}.$$

Addition, and multiplication by integers $k \geqslant 0$, are defined respectively by

$$(\alpha+\beta)_i = \alpha_i + \beta_i$$

and by $(k\alpha)_i = k\alpha_i$; also, $x^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ and $\alpha! = \alpha_1! \ldots \alpha_n!$. (For $x_i^{\alpha_i}$, the convention $0^0 = 1$ is adopted!) The notation $\alpha \leqslant \beta$ means that $\alpha_i \leqslant \beta_i$ for each $i$; the meaning of $\alpha < \beta$ is that $\alpha \leqslant \beta$ and $|\alpha| < |\beta|$; and, where no confusion is possible, the zero $n$-index $(0, \ldots, 0)$ will be written 0.

The symbol $\langle n\rangle$ will denote the set of $n$-indices endowed with these properties, and $\langle n\rangle_+$ the subset of those with uniformly positive entries: $\beta \in \langle n\rangle_+$ if $\beta_i \geqslant 1$ for each $i$.

With the notation in (1), $C^m(U; F)$ denotes the set of functions $U \to F$ that have continuous derivatives to order $m$, $\mathscr{L}(E; F)$ the Banach space of bounded linear operators $E \to F$, and, for any positive integer $n$, $\mathscr{L}_n(E; F) = \mathscr{L}(E, \ldots, E; F)$ the Banach space of bounded $n$-linear operators $E^n \to F$. As is usual ((3), p. 59; (5), p. 179), we identify successively for $n = 1, 2, \ldots$ the spaces $\mathscr{L}(E; \mathscr{L}_n(E; F))$ and $\mathscr{L}_{n+1}(E; F)$, there being a natural linear isometry between each pair (the former then houses derivatives of order $n+1$ in the first instance). An operator $S \in \mathscr{L}_n(E; F)$ is *symmetric* if

$$S(h_1, \ldots, h_n) = S(h_{\sigma(1)}, \ldots, h_{\sigma(n)})$$

for every permutation $\sigma$ of $\{1, \ldots, n\}$.

† I am indebted to J. Marsden and to J. Robbin for helpful letters about the statement in (1). Only case (A) is considered there. The result can be corrected by symmetrization, but the use of Leibniz's rule seems (to me) an inappropriate detour, even then.

https://doi.org/10.1017/S0305004100054402 Published online by Cambridge University Press

3. *Composite functions between Banach spaces.* We follow the instructions of Cartan ((3), p. 86). The $n$th derivative of $f$, evaluated at $x \in U$, is $f^{(n)}(x) \in \mathscr{L}_n(E; F)$; a dot will be used between such an operator and its operand; and, when we write $f^{(n)}(x) \cdot (h, \ldots, h)$, the label $(n)$ implies that $h$ appears $n$ times.

FORMULA A. *Let $u$ be as in* (1), *and assume that $f \in C^m(U; F)$ and $g \in C^m(V; G)$. Then, for $k \in \{1, \ldots, m\}$, $x \in U$ and $(h_1, \ldots, h_k) \in E^k$,*

$$u^{(k)}(x) \cdot (h_1, \ldots, h_k) = \sum_{n=1}^{k} \sum_{\beta \in \langle n \rangle_+, |\beta| = k} \sum_{\sigma} \frac{g^{(n)}(f(x))}{n!\,\beta!} \cdot$$

$$(f^{(\beta_1)}(x) \cdot (h_{\sigma(1)}, \ldots, h_{\sigma(\beta_1)}), \ldots, f^{(\beta_n)}(x) \cdot (h_{\sigma(k-\beta_n+1)}, \ldots, h_{\sigma(k)})), \tag{2}$$

*where $\sum_{\sigma}$ denotes summation over the $k!$ permutations $\sigma$ of $\{1, \ldots, k\}$.*

*Derivation.* To compute the value $u^{(k)}(a)$ at any given point $a \in U$, we expand $u(a+h)$ in Taylor series, extract the homogeneous polynomial of degree $k$, and symmetrize the operator associated with that polynomial. Define

$$\hat{f}(h) = f(a+h) - f(a), \quad \hat{g}(y) = g(f(a)+y) - g(f(a)), \quad \hat{u} = \hat{g} \circ \hat{f}; \tag{3a}$$

then

$$\hat{u}(h) = u(a+h) - u(a). \tag{3b}$$

Clearly, $\hat{u}(0) = 0$ and $\hat{f}(0) = 0$ at the origin in $E$, $\hat{g}(0) = 0$ at the origin in $F$, and $\hat{u}^{(n)}(0) = u^{(n)}(a)$, $\hat{f}^{(n)}(0) = f^{(n)}(a)$ and $\hat{g}^{(n)}(0) = g^{(n)}(f(a))$ for $n \geqslant 1$. We can now *omit the labels* $\hat{(\cdot)}$ without confusion. Moreover ((3), pp. 84–86), it is sufficient for the task in hand to suppose that $f$ and $g$ are polynomials (finite expansions, truncated Taylor series) of degree $k$ or more. Therefore, the terms that we shall sum *are zero from some point onwards.* Accordingly,

$$\begin{aligned} u(h) &= \sum_{n=1}^{\infty} \frac{g^{(n)}(0)}{n!} \cdot (f(h), \ldots, f(h)), \\ &= \sum_{n=1}^{\infty} \frac{g^{(n)}(0)}{n!} \cdot \left( \sum_{\beta_1=1}^{\infty} \frac{f^{(\beta_1)}(0)}{\beta_1!} \cdot (h, \ldots, h), \ldots, \sum_{\beta_n=1}^{\infty} \frac{f^{(\beta_n)}(0)}{\beta_n!} \cdot (h, \ldots, h) \right), \\ &= \sum_{n=1}^{\infty} \frac{g^{(n)}(0)}{n!} \cdot \sum_{\beta \in \langle n \rangle_+} \frac{1}{\beta!} (f^{(\beta_1)}(0) \cdot (h, \ldots, h), \ldots, f^{(\beta_n)}(0) \cdot (h, \ldots, h)). \end{aligned}$$

We split the summation over $\beta$ (in which $|\beta| \geqslant n$) into ($a$) summation over all $\beta \in \langle n \rangle_+$ of order $|\beta| = r$, ($b$) summation over $r$ from $n$ to $\infty$. Changing the order of summation, we obtain

$$u(h) = \sum_{r=1}^{\infty} \frac{1}{r!} T_r(h, \ldots, h), \quad T_r \in \mathscr{L}_r(E; G),$$

where

$$T_r(h_1, \ldots, h_r) = r! \sum_{n=1}^{r} \sum_{\beta \in \langle n \rangle_+, |\beta| = r} \frac{g^{(n)}(0)}{n!\,\beta!} \cdot (f^{(\beta_1)}(0) \cdot (h_1, \ldots, h_{\beta_1}), \ldots, f^{(\beta_n)}(0) \cdot (h_{r-\beta_n+1}, \ldots, h_r)). \tag{4}$$

Now the operator $S_r$ defined by

$$S_r(h_1, \ldots, h_r) = \frac{1}{r!} \sum_{\sigma} T_r(h_{\sigma(1)}, \ldots, h_{\sigma(r)}) \tag{5}$$

is the only symmetric operator in $\mathscr{L}_r(E; G)$ such that $S_r(h, \ldots, h) = T_r(h, \ldots, h)$, and then $u(h) = \Sigma S_r(h, \ldots, h)/r!$. But such an expansion (with $S_r$ symmetric and in $\mathscr{L}_r(E; G)$) holds only if $S_r = u^{(r)}(0)$. Thus (4) and (5) imply (2).



https://doi.org/10.1017/S0305004100054402 Published online by Cambridge University Press

4. *The finite-dimensional case.* We now have at our disposal (*a*) the *multinomial theorem*: for any integer $n \geqslant 0$,

$$(x_1 + \ldots + x_M)^n = n! \Sigma_{|\alpha|=n} \frac{x^\alpha}{\alpha!} \quad (x \in \mathbb{R}^M, \alpha \in \langle M \rangle); \tag{6}$$

and (*b*) the following more detailed form of *Taylor's theorem*. Let $v \in C^m(U; B)$, where $U$ is an open set in $\mathbb{R}^M$ and $B$ is a Banach space; let $a \in U$ and let $|h|$ be so small that the closed ball with centre $a$ and radius $|h|$ lies in $U$. Then

$$v(a+h) = \Sigma_{|\beta| \leqslant m} v_\beta(a)\, h^\beta + o(|h|^m) \quad (\beta \in \langle M \rangle), \tag{7}$$

if and only if

$$v_\beta(a) = (D^\beta v)(a)/\beta!. \tag{8}$$

(The multinomial theorem is proved easily by induction on $M$. To prove the existence part of this Taylor theorem one defines $w(t) = v(a+th)$, $0 \leqslant t \leqslant 1$, notes the result

$$w(1) = \sum_{k=0}^{m} \frac{w^{(k)}(0)}{k!} + \int_0^1 \frac{(1-t)^{m-1}}{(m-1)!} \{w^{(m)}(t) - w^{(m)}(0)\}\, dt$$

of repeated integration by parts, and uses (6) for $(d/dt)^k = (h_1 D_1 + \ldots + h_M D_M)^k$. For a proof of uniqueness independent of that cited earlier, although stated only for $B = \mathbb{R}$, see (6), p. 149.)

*Notation.* The use of a multi-index as subscript to denote Taylor coefficients is convenient; *henceforth we regard* (8) *as defining notation.* Observe also that $h_1, \ldots, h_M$ are now the components of a vector $h \in \mathbb{R}^M$ (whereas, in (2), $h_1, \ldots, h_k$ were distinct vectors in $E$).

LEMMA (powers of truncated Taylor series). *Let v, a and h be as in* (7), *but with v real-valued* ($B = \mathbb{R}$). *Then, for any integer* $n \geqslant 0$ *and M-index* $\alpha$ *with* $|\alpha| \leqslant m$,

$$\{v(a+h) - v(a)\}^n = \Sigma_{0 \leqslant \gamma \leqslant \alpha} P_\gamma(n, v; a)\, h^\gamma + \mathrm{nc}(h^\alpha) \quad (\gamma \in \langle M \rangle), \tag{9}$$

*where* nc($h^\alpha$) *denotes terms that make* no contribution *to the* $h^\alpha$ *term in a product of such expansions, and* $P_\gamma(n, v;.)$ *is a homogeneous polynomial of degree n in derivatives of v* (*up to* $D^\gamma v$) *as follows.*

*Let* $\beta(1), \ldots, \beta(r)$ *be an enumeration of the set* $\langle M \rangle_\gamma = \{\beta \in \langle M \rangle | 0 < \beta \leqslant \gamma\}$ *if* $\gamma > 0$; *choose* $r = 1$, $\beta(1) > 0$ *if* $\gamma = 0$. *Then*

$$P_\gamma(n, v; a) = \Sigma_{\rho \in R(\gamma, n)} \frac{n!}{\rho!} v_{\beta(1)}(a)^{\rho_1} \ldots v_{\beta(r)}(a)^{\rho_r}, \tag{10a}$$

*where*

$$R(\gamma, n) = \left\{ \rho \in \langle r \rangle \,\middle|\, \sum_{j=1}^{r} \rho_j \beta(j) = \gamma, \quad |\rho| = n \right\}, \tag{10b}$$

*and* $P_\gamma$ *is defined to be the zero polynomial when* $R(\gamma, n)$ *is empty.*

*Remark.* Note that, having used the *o*-symbol (that is, the limiting process $|h| \to 0$) in (7) to associate a finite power series with *v* and to prove uniqueness of its coefficients, we can now pursue the coefficient of $h^\alpha$ in a product of such series by purely algebraic means. For it is an obvious algebraic reason that justifies the truncation $\gamma \leqslant \alpha$ in (9), and the set nc($h^\alpha$) is far larger, in general, than the set $o(|h|^{|\alpha|})$. Dieudonné ((6),

https://doi.org/10.1017/S0305004100054402 Published online by Cambridge University Press

p. 148 *et seq.*), using $C^\infty$ functions, builds such facts into a complete structure, of which only a very small part is needed here.

*Proof of the lemma.* Assume that $\alpha > 0$. We may rewrite (7) as

$$v(a+h)-v(a) = \Sigma_{0<\delta\leqslant\alpha} v_\delta(a)\, h^\delta + \mathrm{nc}(h^\alpha) \quad (\delta\in\langle M\rangle).$$

Let $\delta(1), \ldots, \delta(s)$ be an enumeration of $\langle M\rangle_\alpha$; by the multinomial theorem

$$\{v(a+h)-v(a)\}^n = \Sigma_{|\sigma|=n} \frac{n!}{\sigma!} (v_{\delta(1)}(a)\, h^{\delta(1)})^{\sigma_1} \ldots (v_{\delta(s)}(a)\, h^{\delta(s)})^{\sigma_s} + \mathrm{nc}(h^\alpha),$$

where $\sigma\in\langle s\rangle$. The coefficient of $h^\gamma$ in this series comes from those $\sigma$ (of order $n$) for which $\Sigma\delta(j)\,\sigma_j = \gamma$; therefore, it is

$$\Sigma_{\sigma\in S(\gamma,n,\alpha)} \frac{n!}{\sigma!} v_{\delta(1)}(a)^{\sigma_1} \ldots v_{\delta(s)}(a)^{\sigma_s}, \tag{11a}$$

where

$$S(\gamma,n,\alpha) = \left\{\sigma\in\langle s\rangle \,\middle|\, \sum_{j=1}^{s} \sigma_j\,\delta(j) = \gamma, \quad |\sigma| = n\right\}. \tag{11b}$$

It remains only to show that in these expressions $\{\delta(j)\}$ may be reduced to $\{\beta(j)\}$ (hence $S(\gamma,n,\alpha)$ to $R(\gamma,n)$) when $\gamma < \alpha$. Now if $\delta(k)\in\langle M\rangle_\alpha \setminus \langle M\rangle_\gamma$, then $\delta_i(k) > \gamma_i$ for some $i\in\{1,\ldots,M\}$; hence $\sigma_k = 0$ in (11*b*), and $\delta(k)$ does not appear in (11*a*, *b*). Therefore the enumeration $\{\delta(j)\}$ of $\langle M\rangle_\alpha$ can be reduced to an enumeration of $\langle M\rangle_\gamma$ for the definition of $P_\gamma$, except when $\gamma = 0$ and $\langle M\rangle_\gamma$ is empty. For that case, one easily checks that (10) implies the right answer:

$$P_0(n,v;a) = 1 \quad \text{if} \quad n = 0, \quad \text{or} \quad 0 \quad \text{if} \quad n \geqslant 1. \tag{12}$$

FORMULA B. *Let $u$ be as in* (1), *with $E = \mathbb{R}^M$ and $F = \mathbb{R}^N$, and assume that*

$$f\in C^m(U;\mathbb{R}^N) \quad \textit{and} \quad g\in C^m(V;G).$$

*Then, for $1 \leqslant |\alpha| \leqslant m$ and $x\in U$,*

$$u_\alpha(x) = \Sigma_{1\leqslant|\lambda|\leqslant|\alpha|} g_\lambda(f(x))\, Q_{\alpha,\lambda}(f;x) \quad (\alpha\in\langle M\rangle, \lambda\in\langle N\rangle), \tag{13}$$

*where $Q_{\alpha,\lambda}(f;.)$ is a homogeneous polynomial of degree $|\lambda|$ in derivatives of $f_1,\ldots,f_N$. In fact,*

$$Q_{\alpha,\lambda}(f;x) = \sum_{\gamma^1+\ldots+\gamma^N=\alpha} P_{\gamma^1}(\lambda_1,f_1;x)\ldots P_{\gamma^N}(\lambda_N,f_N;x), \tag{14}$$

*where the polynomials $P_{\gamma^i}$ are defined by* (10).

*Derivation.* To compute the value $u_\alpha(a)$ at any given point $a\in U$, we expand $u(a+h)$ in Taylor series and extract the coefficient of $h^\alpha$. Again we make the transformation (3), note the simple facts stated immediately after (3), and then omit the label $\hat{(.)}$. By (7),

$$u(h) = \Sigma_{1\leqslant|\lambda|\leqslant|\alpha|} g_\lambda(0) f(h)^\lambda + o(|h|^{|\alpha|}) \quad (\lambda\in\langle N\rangle), \tag{15}$$

because $|f(h)| = O(|h|)$. Now $f(h)^\lambda$ is a product of $N$ factors, any one of which can be written

$$f_i(h)^{\lambda_i} = \Sigma_{0\leqslant\gamma\leqslant\alpha} P_\gamma(\lambda_i,f_i;\,0)\, h^\gamma + \mathrm{nc}(h^\alpha) \quad (\gamma\in\langle M\rangle),$$

https://doi.org/10.1017/S0305004100054402 Published online by Cambridge University Press

in view of (9). We form the product of these $N$ factors, changing the summation index to $\gamma^i$ in the $i$th to avoid chaos, and observe that the coefficient of $h^\alpha$ in the expansion of $f(h)^\lambda$ is the polynomial (14). Together with (15), this implies the result.

FORMULA C. *Let $u$ be as in* (1), *with $E = \mathbb{R}^M$, $F = \mathbb{R}^{M+1}$ and $f(x) = (x, \phi(x))$, and assume that $\phi \in C^m(U; \mathbb{R})$ and $g \in C^m(V; G)$. Then for $1 \leqslant |\alpha| \leqslant m$ and $x \in U$,*

$$\left.\begin{aligned} u_\alpha(x) &= \textstyle\sum_{1 \leqslant |\lambda| \leqslant |\alpha|} g_\lambda(x, \phi(x))\, P_{\alpha-\lambda'}(\lambda_{M+1}, \phi; x), \\ &\alpha \in \langle M \rangle, \lambda \in \langle M+1 \rangle, \quad \lambda' = (\lambda_1, \ldots, \lambda_M), \end{aligned}\right\} \tag{16}$$

*where $P_{\alpha-\lambda'}(\lambda_{M+1}, \phi; .)$ is a homogeneous polynomial of degree $\lambda_{M+1}$ in derivatives of $\phi$, and is defined by* (10).

*Derivation.* It is not difficult to reduce formula B to the present form, but it is easier still to begin again. Making the transformation (3) and then omitting $\hat{(.)}$, we have, by Taylor's theorem and because $\phi(h) = O(|h|)$,

$$\begin{aligned} u(h) &= \textstyle\sum_{1 \leqslant |\lambda| \leqslant |\alpha|} g_\lambda(0)\, h^{\lambda'}\, \phi(h)^{\lambda_{M+1}} + o(|h|^{|\alpha|}) \\ &= \textstyle\sum_{1 \leqslant |\lambda| \leqslant |\alpha|} g_\lambda(0)\, h^{\lambda'} \sum_{0 \leqslant \gamma \leqslant \alpha} P_\gamma(\lambda_{M+1}, \phi; 0)\, h^\gamma + \mathrm{nc}(h^\alpha), \end{aligned} \tag{17}$$

in view of (9). The terms containing $h^\alpha$ are those with $\gamma = \alpha - \lambda'$; thus (17) implies (16).

*Remark.* It is useful to separate from the sum in (16) terms which are exceptional in some way. The relevant definitions are

$$P_{\alpha-\lambda'}(\lambda_{M+1}, \phi; x) = \textstyle\sum_{\rho \in R} \frac{\lambda_{M+1}!}{\rho!} \phi_{\beta(1)}(x)^{\rho_1} \ldots \phi_{\beta(r)}(x)^{\rho_r}, \tag{18a}$$

where

$$R = \left\{ \rho \in \langle r \rangle \,\middle|\, \sum_{j=1}^{r} \rho_j \beta(j) = \alpha - \lambda', \quad |\rho| = \lambda_{M+1} \right\}, \tag{18b}$$

and where $\beta(1), \ldots, \beta(r)$ are described before (10); obviously we assume $R$ to be not empty.

($a$) If $\lambda' = \alpha$, then $\rho = 0$ because each $\beta(j) > 0$; hence $\lambda_{M+1} = 0$. Thus $\lambda = (\alpha, 0)$ and, by (12) or (18$a$), the polynomial is $P_0(0, \phi; x) = 1$.

($b$) The polynomial $P_{\alpha-\lambda'}$ is independent of $\phi$ only if $\rho = 0$; this is again case ($a$).

($c$) If $P_{\alpha-\lambda'}$ contains a highest derivative $\phi_{\beta(k)}$, $|\beta(k)| = |\alpha|$, for some $k \in \{1, \ldots, r\}$, then $\lambda' \neq \alpha$, by ($a$), and so each $\beta(j) \leqslant \alpha - \lambda'$ because each is in $\langle M \rangle_{\alpha-\lambda'}$. Hence $\lambda' = 0$, and, if the $\beta(j)$ are so labelled that $\beta(r) = \alpha$, we must have $k = r$ and (by (18$a$) and the hypothesis) $\rho_r \geqslant 1$. The first condition in (18$b$) now requires first $\rho_r = 1$ and then (since each $\beta(j) > 0$) $\rho = (0, \ldots, 0, 1)$; the second gives $\lambda_{M+1} = 1$. Therefore $P_{\alpha-\lambda'} = \phi_\alpha$ and $g_\lambda = D_{M+1} g$.

($d$) If a lowest derivative $g_\lambda$, $|\lambda| = 1$, of $g$ occurs, we have either $\lambda_{M+1} = 0$, which makes $\rho = 0$ and hence is case ($a$), or $\lambda' = 0$ and $\lambda_{M+1} = 1$, which is case ($c$).

Accordingly, we have

FORMULA C′.

$$\begin{aligned} u_\alpha(x) = (D_{M+1} g)(x, \phi(x))\, \phi_\alpha(x) + \sum_{2 \leqslant |\lambda| \leqslant |\alpha|,\, \lambda' \neq \alpha} & g_\lambda(x, \phi(x))\, P_{\alpha-\lambda'}(\lambda_{M+1}, \phi; x) \\ & + g_{(\alpha,0)}(x, \phi(x)). \end{aligned} \tag{19}$$

https://doi.org/10.1017/S0305004100054402 Published online by Cambridge University Press

Suppose now that existence of $\phi$ has been proved from the equation $g(x, \phi(x)) = 0$ by means of the implicit function theorem. Then $u_\alpha = 0$ in the neighbourhood in question, a positive lower bound for $\|D_{M+1} g\|$ in that neighbourhood has been established, and, in principle, (19) can be used recursively to estimate all derivatives $\phi_\alpha$ that exist. (We have seen in (*c*) that the sum over $\lambda$ involves only lower derivatives of $\phi$.) Of course, this principle is well known; all that may be new is the avoidance of 'and so on for higher derivatives'.

## REFERENCES

(1) Abraham, R. and Robbin, J. *Transversal mappings and flows* (New York, Benjamin, 1967).
(2) Bruno, F. de. Note sur une nouvelle formule de calcul différentiel. *Quart. J. Math.* **1** (1856), 359–360.
(3) Cartan, H. *Differential calculus* (Paris, Hermann, and Boston, Houghton Mifflin, 1971).
(4) Chaundy, T. *The differential calculus* (Oxford, Clarendon Press, 1935).
(5) Dieudonné, J. *Foundations of modern analysis* (New York, Academic Press, 1969).
(6) Dieudonné, J. *Treatise on analysis*, vol. iv (New York, Academic Press, 1974).
(7) Goursat, E. *Mathematical analysis*, vol. I (Boston, Ginn, 1904).
(8) Stein, E. M. *Singular integrals and differentiability properties of functions* (Princeton, University Press, 1970).

https://doi.org/10.1017/S0305004100054402 Published online by Cambridge University Press