$$= \frac{M}{24n^3}x(1-x)[3(n-2)x(1-x)+1]$$

from eqn. (5), and the required result follows.

The class of functions to which the proof applies is sufficiently general to show that for most practical purposes the Bernstein polynomials do not provide useful approximations.

Acknowledgment: I am grateful to the referee for pointing out an error in an earlier version of this paper.

References

1. J. Todd, Introduction to the Constructive Theory of Functions, Birkhauser Verlag, 1963.
2. A. Ralston, A First Course in Numerical Analysis, McGraw Hill, New York, 1965.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DURHAM, SOUTH ROAD, DURHAM, ENGLAND.

DIRECTIONAL DIFFERENTIATION IN THE PLANE AND TANGENT VECTORS ON C^r MANIFOLDS

H. G. ELLIS

If at a certain point each of the functions f and g is differentiable, then so is their product fg. The product can, however, be differentiable at that point without both, or even either of f and g being so, as the example $f(x) = x^{1/3}$, $g(x) = x^{2/3}$ shows. This simple observation suggests the possibility of refining Chevalley's widely adopted "algebraic" or "intrinsic" definition of a tangent vector at a point P of a C^ω(that is to say, an analytic) manifold [1, p. 76] to make it applicable also to every C^r manifold with $r \geqq 1$. That definition, abstracted from the notion of directional differentiation, says that a tangent vector at P is a real-valued linear operator on the real-valued functions analytic at P which obeys the algebraic rule for differentiation at P of products of functions differentiable at P. The analyticity assumption ensures that the tangent vectors so defined are in fact directional differentiations at P, and that the dimensionality of the tangent space made up of these tangent vectors is the same as that of the manifold. Here I shall show how to arrive at the same goals, relaxing the requirement from analyticity at P to that of mere differentiability at P of the functions operated upon, but compensating for this loss by incorporating into the definition an additional algebraic condition. The revised definition will work for all C^r manifolds with $r \geqq 1$; it will also be "intrinsic" in that it will make no direct reference to a coordinate system.

The basic ideas can be explained more readily for the vector space $\mathbb{R}^2$ than for an abstract manifold, so I shall begin there, saving the generalities for later. It will be convenient to have at our disposal the coordinate functions x and y, such that

$x(\mathbf{u}) = u_x$ and $y(\mathbf{u}) = u_y$ if $\mathbf{u} = [\![u_x, u_y]\!] \in \mathbb{R}^2$. The Euclidean norm function $(x^2 + y^2)^{1/2}$ will be denoted by $|\cdot|$. As the point in question let us take the vector $\mathbf{0}$ in $\mathbb{R}^2$.

The definition of differentiability most appropriate in this context is that the real-valued function f defined on a neighborhood of $\mathbf{0}$ in $\mathbb{R}^2$ is differentiable at $\mathbf{0}$ if and only if there exist a (homogeneous) linear mapping H: $\mathbb{R}^2 \mapsto \mathbb{R}$ and a real-valued function η on dmn f such that

(a) $f(\mathbf{u}) = f(\mathbf{0}) + H\mathbf{u} + |\mathbf{u}|\eta(\mathbf{u})$ if $\mathbf{u} \in \operatorname{dmn} f$,

(b) $\eta(\mathbf{0}) = 0$, and

(c) η is continuous at $\mathbf{0}$.

A familiar consequence of this definition is that if f is differentiable at $\mathbf{0}$, then (i) the partial derivatives $(\partial f/\partial x)(\mathbf{0})$ and $(\partial f/\partial y)(\mathbf{0})$ both exist, (ii) the mappings H and η are uniquely determined by the conditions (a), (b), and (c), and (iii) $H = \alpha x + \beta y$, where $\alpha = (\partial f/\partial x)(\mathbf{0})$ and $\beta = (\partial f/\partial y)(\mathbf{0})$. This linear mapping H: $\mathbb{R}^2 \mapsto \mathbb{R}$, which is called the differential of f at $\mathbf{0}$ and denoted by $df(\mathbf{0})$, is an alter ego of the gradient vector $\nabla f(\mathbf{0})$; precisely,

$$df(\mathbf{0}) = (\partial f/\partial x)(\mathbf{0})x + (\partial f/\partial y)(\mathbf{0})y = \nabla f(\mathbf{0}) \cdot [\![x, y]\!],$$

where the dot stands for the usual inner product. Because of this relationship the differential operator d and the gradient operator ∇ obey the same rules of operation, which are, in addition to linearity, those peculiar to differentiation, such as the chain rule and the product rule.

For a fixed vector $\mathbf{u}$ we can define an operator $L_{\mathbf{u}}$ by the phrase "$L_{\mathbf{u}}f$ is the unnormalized directional derivative of f at $\mathbf{0}$ along $\mathbf{u}$," which is equivalent to the formula $L_{\mathbf{u}}f = df(\mathbf{0})\mathbf{u}$. The domain of this real-valued operator $L_{\mathbf{u}}$ is the set $\mathscr{D}^1(\mathbf{0})$, consisting of all real-valued functions defined on neighborhoods of $\mathbf{0}$ in $\mathbb{R}^2$ and differentiable at $\mathbf{0}$. It is easy to see that $L_{\mathbf{u}}$, because it is a linear combination of differential operators, is *linear* in the sense that $L_{\mathbf{u}}(af + bg) = aL_{\mathbf{u}}f + bL_{\mathbf{u}}g$ if $a, b \in \mathbb{R}$ and $f, g \in \mathscr{D}^1(\mathbf{0})$, and is *semi-Leibnizian* in the sense that $L_{\mathbf{u}}(fg) = (L_{\mathbf{u}}f)g(\mathbf{0}) + f(\mathbf{0})(L_{\mathbf{u}}g)$ if $f, g \in \mathscr{D}^1(\mathbf{0})$ (the reference to Leibniz is in respect of his rule for differentiating products; the reason for including the prefix "semi-" will become apparent). What Chevalley did in 1946 was to seize upon these two algebraic properties of $L_{\mathbf{u}}$ and use them, in effect, to identify $\mathbf{u}$ with $L_{\mathbf{u}}$ without mentioning $\mathbf{u}$. He could do this because he could show that the two properties serve to characterize $\{L_{\mathbf{u}} \mid \mathbf{u} \in \mathbb{R}^2\}$. He proved that not only is each $L_{\mathbf{u}}$ linear and semi-Leibnizian, but also, if L is a linear and semi-Leibnizian real-valued operator whose domain is a subset $\mathscr{A}$ of $\mathscr{D}^1(\mathbf{0})$, then $L = L_{\mathbf{u}|\mathscr{A}}$ for some vector $\mathbf{u}$ in $\mathbb{R}^2$—*provided* that $\mathscr{A}$ consists of just those functions in $\mathscr{D}^1(\mathbf{0})$ that are real-analytic on a neighborhood of $\mathbf{0}$. His argument makes full use of this restriction.

In 1952 Flanders, putting to work an observation that Bohnenblust had made, established the same characterization when L is only restricted to operate on C^∞

functions [2]. Then in 1955 it was found that the characterization does not go through if L is allowed to operate on functions that are less than infinitely differentiable [3, 4]. What we shall see is that in this latter case the characterization can be reestablished if the two algebraic properties are supplemented by a third, to make L, so to speak, fully Leibnizian. To do this we need a couple of simple lemmas. The space in question is still $\mathbb{R}^2$, but it will not be mentioned in the statements of these lemmas, as the propositions will then hold true in much wider contexts.

LEMMA 1. *If each of f and g is a real-valued function on a neighborhood of $\mathbf{0}$, and f is differentiable at $\mathbf{0}$ with $f(\mathbf{0}) = 0$, and g is continuous at $\mathbf{0}$, then fg is differentiable at $\mathbf{0}$, and $d(fg)(\mathbf{0}) = g(\mathbf{0})df(\mathbf{0})$.*

Proof. If $\mathbf{u} \in \operatorname{dmn} f \cap \operatorname{dmn} g$, and η_f is the function η that enters into the determination that f is differentiable at $\mathbf{0}$ according to the definition adopted, then

$$(fg)(\mathbf{u}) = (fg)(\mathbf{0}) + [g(\mathbf{0})df(\mathbf{0})]\mathbf{u} + |\mathbf{u}|\eta_{fg}(\mathbf{u}),$$

where

$$\eta_{fg}(\mathbf{u}) = \begin{cases} [g(\mathbf{u}) - g(\mathbf{0})]df(\mathbf{0})(\mathbf{u}/|\mathbf{u}|) + g(\mathbf{u})\eta_f(\mathbf{u}) & \text{if } \mathbf{u} \neq \mathbf{0}, \\ 0 \text{ if } \mathbf{u} = \mathbf{0}. \end{cases}$$

Because $\mathbf{u}/|\mathbf{u}|$ is a unit vector, $|df(\mathbf{0})(\mathbf{u}/|\mathbf{u}|)| = |\nabla f(\mathbf{0}) \cdot (\mathbf{u}/|\mathbf{u}|)| \leqq |\nabla f(\mathbf{0})|$, and it follows readily that η_{fg} is continuous at $\mathbf{0}$. The rest is straightforward.

LEMMA 2. *If h is a real-valued function on a neighborhood of $\mathbf{0}$, and h is continuous at $\mathbf{0}$ with $h(\mathbf{0}) = 0$, then $|\cdot|h$ is differentiable at $\mathbf{0}$, and $d(|\cdot|h)(\mathbf{0}) = 0$.*

Finding a proof for Lemma 2 is not hard to do and may serve as a test of students' understandings of the definition of differentiability.

From Lemma 1 it follows that if $f \in \mathscr{D}^1(\mathbf{0})$ with $f(\mathbf{0}) = 0$, and $g \in C^0(\mathbf{0})$ (the set of all real-valued functions on neighborhoods of $\mathbf{0}$ in $\mathbb{R}^2$ that are continuous at $\mathbf{0}$), then $fg \in \mathscr{D}^1(\mathbf{0})$, and if $\mathbf{u} \in \mathbb{R}^2$, then $L_{\mathbf{u}}(fg) \equiv d(fg)(\mathbf{0})\mathbf{u} = [g(\mathbf{0})df(\mathbf{0})]\mathbf{u} = (L_{\mathbf{u}}f)g(\mathbf{0})$. This is the missing property, that $L_{\mathbf{u}}(fg) = (L_{\mathbf{u}}f)g(\mathbf{0})$ if $f \in \mathscr{D}^1(\mathbf{0})$, $g \in C^0(\mathbf{0})$, and $f(\mathbf{0}) = 0$. When an operator $L: \mathscr{D}^1(\mathbf{0}) \mapsto \mathbb{R}$ is semi-Leibnizian and has in addition this property, then it may be termed *fully Leibnizian.* The missing property thus identified, the extended characterization can now be established.

THEOREM 1. *If $L: \mathscr{D}^1(\mathbf{0}) \mapsto \mathbb{R}$, then L is linear and fully Leibnizian if and only if there is a vector $\mathbf{u}$ such that $Lf = L_{\mathbf{u}}f \equiv df(\mathbf{0})\mathbf{u}$ if $f \in \mathscr{D}^1(\mathbf{0})$; specifically, $\mathbf{u} = [\![Lx, Ly]\!]$, so that $Lf = (Lx)(\partial f/\partial x)(\mathbf{0}) - (Ly)(\partial f/\partial y)(\mathbf{0})$ if $f \in \mathscr{D}^1(\mathbf{0})$.*

Proof. If $L = L_{\mathbf{u}}$, then, as we have seen, L is linear and fully Leibnizian. Going the other way, notice first that if c_U denotes the constant function defined on the neighborhood U of $\mathbf{0}$ and having the real number c as its value, then $Lc_U = L(1_U c_U)$

$= (L1_U)c_U(\mathbf{0}) + 1_U(\mathbf{0})(Lc_U) = cL1_U + Lc_U = 2Lc_U$, hence $Lc_U = 0$. Next, observe that if $g \in \mathscr{D}^1(\mathbf{0})$, and U is a neighborhood of $\mathbf{0}$, then $g\,|_U \in \mathscr{D}^1(\mathbf{0})$, and $L(g\,|_U) = L(1_U g) = (L1_U)g(\mathbf{0}) + 1_U(\mathbf{0})(Lg) = Lg$.

Now suppose that $f \in \mathscr{D}^1(\mathbf{0})$, and let $U = \operatorname{dmn} f$. Then

$$\begin{aligned} f &= f(\mathbf{0})_U + df(\mathbf{0})\,|_U + |\cdot|\,\eta \\ &= f(\mathbf{0})_U + (\partial f/\partial x)(\mathbf{0})x\,|_U + (\partial f/\partial y)(\mathbf{0})y\,|_U + |\cdot|\,\eta, \end{aligned}$$

where $\operatorname{dmn}\eta = U$, $\eta(\mathbf{0}) = 0$, and η is continuous at $\mathbf{0}$. According to Lemma 2 the function $|\cdot|\,\eta^{1/3}$ is in $\mathscr{D}^1(\mathbf{0})$; moreover, its value at $\mathbf{0}$ is 0. The function $\eta^{2/3}$, on the other hand, is in $C^0(\mathbf{0})$. Hence, because L is fully Leibnizian, $L(|\cdot|\,\eta) = L(|\cdot|\,\eta^{1/3})\eta^{2/3}(\mathbf{0}) = 0$. Therefore, because L is linear,

$$\begin{aligned} Lf &= L(f(\mathbf{0})_U) + (\partial f/\partial x)(\mathbf{0})L(x\,|_U) + (\partial f/\partial y)(\mathbf{0})L(y\,|_U) + L(|\cdot|\,\eta) \\ &= 0 + (\partial f/\partial x)(\mathbf{0})(Lx) + (\partial f/\partial y)(\mathbf{0})(Ly) + 0 \\ &= \nabla f(\mathbf{0}) \cdot \mathbf{u} = df(\mathbf{0})\mathbf{u} = L_{\mathbf{u}} f, \end{aligned}$$

where $\mathbf{u} = [\![Lx, Ly]\!]$. This completes the proof and thereby establishes the extended version of Chevalley's characterization.

Now it is time to state the generalities. In place of $\mathbb{R}^2$ we have a finite-dimensional differentiable manifold $\mathscr{M}$, "differentiable" meaning here C^ω, or C^∞, or C^r with $r \geqq 1$. The required definition of differentiability at a point reads as follows.

DEFINITION. The real-valued function f defined on a neighborhood of the point P of the M-dimensional differentiable manifold $\mathscr{M}$ is differentiable at P with respect to the coordinate system X of $\mathscr{M}$ at P if and only if there exist a (homogeneous) linear mapping $H\colon \mathbb{R}^M \mapsto \mathbb{R}$ and a real-valued function η on $\operatorname{dmn} f \cap \operatorname{dmn} X$ such that

(a) $f(Q) = f(P) + H[X(Q) - X(P)] + |X(Q) - X(P)|\,\eta(Q)$ if $Q \in \operatorname{dmn} f \cap \operatorname{dmn} X$,
(b) $\eta(P) = 0$, and
(c) η is continuous at P;

f is differentiable at P if and only if f is differentiable at P with respect to each coordinate system of $\mathscr{M}$ at P. (Here $|\cdot|$ can be any norm function on $\mathbb{R}^M$, the Euclidean in particular.)

Chevalley's original definition of tangent vector is now replaced by the following one, in which $\mathscr{D}^1(P)$, respectively $C^0(P)$, is the set of all real-valued functions defined on neighborhoods of P and differentiable, respectively continuous, at P.

DEFINITION. By a tangent vector of $\mathscr{M}$ at P is meant a mapping $L\colon \mathscr{D}^1(P) \mapsto \mathbb{R}$ such that

(1) $L(af + bg) = aLf + bLg$ if $a, b \in \mathbb{R}$ and $f, g \in \mathscr{D}^1(P)$,
(2) $L(fg) = (Lf)g(P) + f(P)(Lg)$ if $f, g \in \mathscr{D}^1(P)$, and
(3) $L(fg) = (Lf)g(P)$ if $f \in \mathscr{D}^1(P)$, $g \in C^0(P)$, and $f(P) = 0$.

To recover Chevalley's definition from this one, just leave out condition(3) and use, instead of $\mathscr{D}^1(P)$, the set of all real-valued functions analytic at P (this requires that $\mathscr{M}$ be a C^ω manifold).

The characterizing theorem, analogous to Theorem 1, that justifies this definition is the following one.

THEOREM 2. *L is a tangent vector of $\mathscr{M}$ at P if and only if $L: \mathscr{D}^1(P) \mapsto \mathbb{R}$ and, for some one or, equivalently, for every coordinate system X of $\mathscr{M}$ at P, one has $L = \sum_{m=1}^{m=M}(Lx^m)(\partial/\partial x^m)(P)$, where $X = [\![x^1, \cdots, x^M]\!]$.*

An immediate corollary of that theorem is this one.

THEOREM 3. *The tangent space of $\mathscr{M}$ at P has dimensionality M, and a basis for it is $\{(\partial/\partial x^m)(P) \mid 1 \leqq m \leqq M\}$.*

These theorems, which Chevalley established in the C^ω case and Flanders in the C^∞ case, under the old definition break down in the C^r cases with $1 \leqq r < \infty$, where it can be shown that the tangent spaces are infinite-dimensional [**3**, **4**]. Proofs of them are omitted, but are not difficult to construct once one knows the earlier development in $\mathbb{R}^2$.

It is worth mentioning in passing that the new definition of tangent vector can, without alteration of its content, be modified to have in place of condition (2) the weaker condition (2′) $L(1) = 0$. Yet other equivalent definitions are obtained when one uses in conjunction with (1) and (2) and with (1) and (2′) the condition (3) $L(fg) = 0$ if $f \in \mathscr{D}^1(P)$, $g \in C^0(P)$, and $f(P) = g(P) = 0$. An interesting exercise is to find substitutes for Theorems 2 and 3 when linearity of L is replaced by mere additivity: $L(f + g) = Lf + Lg$ if $f, g \in \mathscr{D}^1(P)$.

I take this opportunity to express my gratitude to the referee, whose suggestions of ways to make these ideas accessible to a wider readership I was happy to follow.

References

1. C. Chevalley, Theory of Lie Groups I, Princeton Univ. Press, Princeton, N. J., 1946.

2. H. Flanders, Development of an extended exterior differential calculus, Trans. Amer. Math. Soc., 75 (1953) 311–326.

3. W. F. Newns and A. G. Walker, Tangent planes to a differentiable manifold, J. London Math. Soc., 31 (1956) 400–407 (MR 18 (1957), 821).

4. G. Papy, Sur la définition intrinsèque des vecteurs tangents à une variété de classe C^r lorsque $1 \leqq r < \infty$, C. R. Acad. Sci. Paris, 242 (1956) 1573–1575 (MR 17 (1956), 892).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COLORADO, BOULDER, CO 80302.