M. Schreiber

Differential Forms

A Heuristic Introduction

M. Schreiber

Differential Forms

A Heuristic Introduction

Springer-Verlag
New York Heidelberg Berlin

M. Schreiber

Department of Mathematics
The Rockefeller University
1230 York Ave.
New York, N.Y. 10021

AMS Subject Classifications: 26-01, 26A57, 26A60, 26A66

Library of Congress Cataloging in Publication Data

Schreiber, Morris, 1926–
Differential forms.

(Universitext)
Bibliography: p.
1. Differential forms. I. Title.
QA381.S4 515'.37 77-14392

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90287-2 Springer-Verlag New York
ISBN 3-540-90287-2 Springer-Verlag Berlin Heidelberg

Preface

A working knowledge of differential forms so strongly illuminates the calculus and its developments that it ought not be too long delayed in the curriculum. On the other hand, the systematic treatment of differential forms requires an apparatus of topology and algebra which is heavy for beginning undergraduates. Several texts on advanced calculus using differential forms have appeared in recent years. We may cite as representative of the variety of approaches the books of Fleming [2],(1) Nickerson-Spencer-Steenrod [3], and Spivak [6]. Despite their accommodation to the innocence of their readers, these texts cannot lighten the burden of apparatus exactly because they offer a more or less full measure of the truth at some level of generality in a formally precise exposition. There is consequently a gap between texts of this type and the traditional advanced calculus. Recently, on the occasion of offering a beginning course of advanced calculus, we undertook the experiment of attempting to present the technique of differential forms with minimal apparatus and very few prerequisites. These notes are the result of that experiment.

Our exposition is intended to be heuristic and concrete. Roughly speaking, we take a differential form to be a multi-dimensional integrand, such a thing being subject to rules making change-of-variable calculations automatic. The domains of integration (manifolds) are explicitly given "surfaces" in Euclidean space. The differentiation of forms (exterior

(1) Numbers in brackets refer to the Bibliography at the end.

differentiation) is the obvious extension of the differential of functions, and this completes the apparatus. To avoid the geometric and not quite elementary subtleties of a correct proof of the general Stokes formula we offer instead a short plausibility argument which we hope will be found attractive as well as convincing. This is one of several abbreviations we have made in the interests of maintaining an elementary level of exposition.

The prerequisite for this text is a standard first course of calculus and a bit more. The latter, though not very specific, may be described as some familiarity with Euclidean space of k dimensions, with k-by-k matrices and the row-by-column rule for multiplying them, and with the simpler facts about k-by-k determinants. Serious beginning undergraduates seem generally to possess this equipment at the present time. Linear algebra proper is not required, except at one place in Chapter 6, where we must diagonalize a real symmetric matrix. For this theorem, and for several other facts of algebra (such as those mentioned above), we offer references to the text [5] of Schreier and Sperner. There the matters in question are well presented without prerequisites. For analytical matters we provide citations to Courant [1]. We have tried to design the text so that, with the books of Courant and Schreier-Sperner as his only other equipment, the industrious reader working alone will find here an essentially self-contained course of study. However, the better use of this text is probably its obvious one as part of a modern sophomore or junior course of advanced calculus.

The content of each Chapter is clear from the Table of Contents, with two exceptions: in 6.2 we give the theorem on the geometric and arithmetic means, and in 7.3 we prove the isoperimetric inequality. Our notations, all standard, are listed on page (x). The symbol n.m(k) means Formula (k) in Section n.m.

It is a pleasure to acknowledge several debts of gratitude: to P.A. Griffiths, who encouraged the project and suggested the inclusion of "something on integral geometry"; to Mary Ellen O'Brien, who gave the manuscript its format in the course of typing it; and to the students, who were willing to participate in an experiment.

M. Schreiber
21 July 1977

Notations

$a \in A$	a is a member of the set A
$\vec{x}$	a vector
$\Vert\vec{x}\Vert$	the length of $\vec{x}$
$\vec{x} \cdot \vec{y}$	the scalar product of $\vec{x}$ and $\vec{y}$.
$\vec{x} \times \vec{y}$	the vector product of $\vec{x}$ and $\vec{y}$
$\binom{n}{k}$	the binomial coefficient $\binom{n}{k} = \frac{n!}{(n-k)!\,k!}$
$\mathbb{R}$	the set of real numbers
$\mathbb{R}^k$	Euclidean space of dimension k
$f : \mathbb{R}^n \to \mathbb{R}^m$	m-vector-valued function of an n-vector argument
$Z(\phi)$	the set of zeros of $\phi : \mathbb{R}^k \to \mathbb{R}^1$
$f \circ g$	the composition $f \circ g(x) = (f(g(x))$ of functions f and g
$\omega \wedge \tau$	the wedge product of differential forms ω and τ
Λ^r	the space of r-forms
$\mathbb{W}$	the set of vector fields
$\mathbb{S}$	the set of scalar fields
$\vert T \vert$	the determinant of the matrix T
${}^tT$	the transpose of the matrix T
$\operatorname{tr} T$	the trace of the matrix T

Table of contents

Chapter 1
Partial differentiation

1.1 Partial Derivatives

We denote by $\mathbb{R}^k$ the set of real ordered k-tuples $\vec{x} = (x_1, x_2, \ldots, x_k)$. Such a k-tuple is called a k-vector, the numbers $x_1, x_2, \ldots$ being its components. k-vectors are added and multiplied by scalars in the component-wise function familiar in the plane and in three-space. In this notation the plane and three-space are denoted $\mathbb{R}^2$ and $\mathbb{R}^3$ respectively.

The inner (or scalar) product $\vec{x} \cdot \vec{y} = \sum x_i y_i$ of k-vectors $\vec{x}$ and $\vec{y}$ determines length and angle in $\mathbb{R}^k$ as follows. The length $\|\vec{x}\|$ of $\vec{x}$ is $\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}}$, and the angle θ between $\vec{x}$ and $\vec{y}$ is defined by the relation (law of Cosines) $\cos\theta = \vec{x} \cdot \vec{y} / \|\vec{x}\| \cdot \|\vec{y}\|$. These are exact analogs for $\mathbb{R}^k$ of the corresponding constructions in $\mathbb{R}^2$ and $\mathbb{R}^3$. In particular, $\vec{x} \in \mathbb{R}^k$ is called a unit vector if $\|\vec{x}\| = 1$, and k-vectors $\vec{x}, \vec{y}$ are orthogonal if $\vec{x} \cdot \vec{y} = 0$.

A function of n variables may be viewed as a function of an n-vector argument or variable. We shall be concerned also with functions taking vector values. The notation $f : \mathbb{R}^n \to \mathbb{R}^m$ signifies that f is a function of an n-vector argument taking m-vector values. Since a 1-vector is just a number, a function $f:\mathbb{R}^n \to \mathbb{R}^1$ is a scalar-valued function of n variables.

Take as orthogonal reference frame in $\mathbb{R}^k$ the unit vectors $\vec{e}_1 = (1,0,\ldots,0)$, $\vec{e}_2 = (0,1,0,\ldots,0)$, ..., $\vec{e}_k = (0,\ldots,0,1)$. The lines on which they lie are then a system of axes for a Cartesian coordinate system in $\mathbb{R}^k$. Every point $\vec{x} \in \mathbb{R}^k$ has a unique expression $\vec{x} = \sum x_i \vec{e}_i$ in this reference frame, its Cartesian coordinates in this frame being $(x_1, x_2, \ldots, x_k)$. To each coordinate direction is associated a partial differential operator $\frac{\partial}{\partial x_i}$, which acts upon scalar-valued functions $f:\mathbb{R}^k \to \mathbb{R}^1$ of a vector argument thus:

$$[\frac{\partial}{\partial x_i} f](\vec{x}) = \lim_{\delta \to 0} \frac{f(\vec{x} + \delta\vec{e}_i) - f(\vec{x})}{\delta} . \tag{1}$$

Note well that the result of applying the operator $\frac{\partial}{\partial x_i}$ to a function $f:\mathbb{R}^k \to \mathbb{R}^1$ is another function $[\frac{\partial}{\partial x_i} f] : \mathbb{R}^k \to \mathbb{R}^1$ of the same type. We will denote this new function, whenever possible, by the short notation f_i. Thus

$$f_i(\vec{x}) = [\frac{\partial}{\partial x_i} f](x) ; \tag{2}$$

and f_i is called the i^{th} partial derivative of f. Its geometric significance is as follows. Given $f:\mathbb{R}^k \to \mathbb{R}^1$, we may interpret the equation

$$z = f(\vec{x}) \tag{3}$$

as defining a k-dimensional surface in $\mathbb{R}^{k+1}$. If we fix all but the i^{th} coordinate of a point $\vec{x} \in \mathbb{R}^k$, and let the i^{th} coordinate x_i vary freely, we generate a straight line in $\mathbb{R}^k$ passing through x and lying parallel to $\vec{e}_i$. Its coordinates are $(x_1, x_2, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_k)$, with $-\infty < t < +\infty$. Its image under f is a 1-dimensional curve (it has one degree of freedom; namely, the variation of t) lying on the surface (3). The slope of this curve as a function of t is $f_i(x_1, x_2, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_k)$.

Since the partial derivatives f_i of a function $f : \mathbb{R}^k \to \mathbb{R}^1$ are again functions of the same type, they too may be differentiated partially. The result of differentiating f_i by its j^{th} argument may (and will, whenever possible) be denoted f_{ij}. The concordance with the standard notation is

$$f_{ij} = \frac{\partial^2}{\partial x_j \partial x_i} f = \frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i} f\right). \quad (4)$$

One should note the reversal of order of the subscripts. Since the f_{ij} are again functions of the same type, they may be differentiated partially; denote these functions by f_{ijk}. In this hierarchy, the f_i are called first partials, the f_{ij} second partials, and so on. We are assuming for the purposes of this discussion that all limits involved (they are all of the general form (1)) exist. This being assumed, there are in principle k first partials $f_1, f_2, \ldots, f_k$; k^2 second partials $f_{11}, f_{12}, \ldots, f_{1k}, f_{21}, f_{22}, \ldots, \ldots, f_{kk}$; k^3 third partials; and so on.

On the other hand, it is not hard to show, using the mean value theorem, that when all partial derivatives involved are themselves continuous functions of their several variables, then the order in which the differentiations are done does not matter.(1) For example, $f_{12} = f_{21}$ if both are continuous functions; similarly $f_{112} = f_{121} = f_{211}$ if all are continuous. Thus the number of distinct higher partials is sharply reduced if they are continuous. The rule for the equality of mixed higher partials may be stated thus: two higher partials are equal (when they are continuous functions) if they involve the same indices with the same multiplicities.

1.2 Differentiability, Chain Rule

By definition a real function $f : \mathbb{R}^1 \to \mathbb{R}^1$ of a real argument is differentiable at x if

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \varepsilon(h) , \tag{1}$$

$$\varepsilon(h) \to 0 \quad \text{as} \quad h \to 0 ; \tag{2}$$

which is to say, the increment

$$(\Delta f)(x,h) = f(x+h) - f(x)$$

(1) See Courant [1], volume II, pp. 55-58.

is approximated by the differential

$$(df)(x,h) = f'(x) \cdot h \tag{3}$$

so well that the error $\eta(h) = h \cdot \varepsilon(h)$ vanishes faster than h, $\lim_{h \to 0} \frac{\eta(h)}{h} = 0$. Note that the differential is a linear function of the increment h.

Suppose, for a given function $f : \mathbb{R}^1 \to \mathbb{R}^1$ and a given $x \varepsilon \mathbb{R}^1$, that there exists a constant A and a function $\alpha : \mathbb{R}^1 \to \mathbb{R}^1$ such that

$$(\Delta f)(x,h) = Ah + \alpha(h),$$
$$\lim_{h \to 0} \frac{\alpha(h)}{h} = 0; \tag{4}$$

which is to say, the increment $(\Delta f)(x,h)$ can be approximated with the stated accuracy by a linear function of h. This would imply at once that f is differentiable at x and that $f'(x) = A$.

Putting together the two foregoing paragraphs, we see that, for functions of one variable, differentiability is the same as linear approximability. The generalization of differentiability to functions of several variables is made by generalizing to several variables the idea of linear approximability, as follows.

A function $f : \mathbb{R}^k \to \mathbb{R}^1$ is (by definition) differentiable at $x \varepsilon \mathbb{R}^k$ if there exist constants $A_1, A_2, \ldots, A_k$ and a function $\alpha : \mathbb{R}^k \to \mathbb{R}^1$ such that

$$(\Delta f)(\vec{x},\vec{h}) = \sum A_i h_i + \alpha(\vec{h}) , \tag{5}$$

$$\lim_{\vec{h}\to\vec{0}} \frac{\alpha(\vec{h})}{\|\vec{h}\|} = 0 , \tag{6}$$

where $\vec{h} = (h_1,h_2,\ldots,h_k)$ is the increment vector and $(\Delta f)(\vec{x},\vec{h}) = f(\vec{x}+\vec{h}) - f(\vec{x})$ is the corresponding increment of f.

We have replaced the linear function Ah of (4) by a linear function of the components $h_1,\ldots,h_k$ of the increment vector $\vec{h}$, and "$h \to 0$" becomes "$\vec{h} \to \vec{0}$". This is the exact analog of (4). Note that for any $\vec{x} \in \mathbb{R}^k$ one has

$$|x_j| \leq \|\vec{x}\| \leq \sqrt{k} \cdot \mathrm{Max}\{|x_1|,\ldots,|x_k|\}, \tag{7}$$

for $x_j^2 \leq \sum x_i^2 \leq k \cdot \mathrm{Max}\{x_1^2,\ldots,x_k^2\}$. Therefore all components of $\vec{x}$ are small independently if and only if $\|\vec{x}\|$ is small, and so $\vec{h} \to \vec{0}$ if and only if $\|\vec{h}\| \to 0$ if and only if $h_j \to 0$ for each j.

Assume f is differentiable in this sense, and put $\vec{h} = \delta\vec{e}_i$ in (5). This yields

$$\frac{f(\vec{x}+\delta\vec{e}_i) - f(\vec{x})}{\delta} = A_i + \frac{\alpha(\vec{h})}{\delta} ,$$

whence the limit as $\delta \to 0$ (i.e., $\vec{h} \to \vec{0}$) exists, and $A_i = f_i(\vec{x})$. That is to say, differentiability as defined above implies the existence of the first partial derivatives. Conversely, if f has continuous first partial derivatives near a point $\vec{x}$, then f is differentiable at $\vec{x}$, as we shall now show. The proof is perhaps forbidding in notation, but the idea is quite simple: one makes the change from $\vec{x}$ to $\vec{x}+\vec{h}$ in successive steps involving one variable at a time, so that the definition 1.1(1) and basic property (1), (2) of differentiation may be invoked. Here is the proof. To establish (5), (6) we put $\vec{h} = \sum_{i=1}^{k} h_i\vec{e}_i$ and decompose $(\Delta f)(\vec{x},\vec{h})$ as $(\Delta f)(\vec{x},\vec{h}) = f(\vec{x}+\sum h_i\vec{e}_i) - f(\vec{x}) = \sum_{j=1}^{k} \{f(\vec{x}+\sum_{i=j} h_i\vec{e}_i) - f(\vec{x}+\sum_{i=j+1} h_i\vec{e}_i)\}$, where the last term $(j=k)$ is to be interpreted as $\{f(\vec{x}+h_k\vec{e}_k) - f(\vec{x})\}$. Now $f(\vec{x}+\sum_{i=j} h_i\vec{e}_i) - f(\vec{x}+\sum_{i=j+1} h_i\vec{e}_i) = h_j f_j(\vec{x}+\sum_{i=j+1} h_i\vec{e}_i) + \alpha_j$, where $\alpha_j/h_j \to 0$ as $h \to 0$, by 1.1(1) and (1), (2). By the assumed continuity of the first partials, we have $h_j f_j(\vec{x}+\sum_{i=j+1} h_i\vec{e}_i) = h_j f_j(\vec{x}) + h_j\varepsilon_j$, where $\varepsilon_j \to 0$ as $h_j \to 0$. Therefore $(\Delta f)(\vec{x},\vec{h}) = \sum_{j=1}^{k} h_j f_j(x) + \sum_{j=1}^{k} \{\alpha_j + h_j\varepsilon_j\}$. By means of the first part of (7) one sees that the error term $\sum_{j=1}^{k} \{\alpha_j + h_j\varepsilon_j\}$ satisfies (6), and the proof is complete.

The following notation is suggestive. Put $d\vec{x} = (dx_1, dx_2, \ldots, dx_k)$ for the increment vector, formerly called $\vec{h}$, and write

$$(df)(\vec{x},d\vec{x}) = \sum f_i(\vec{x})dx_i \tag{8}$$

for the linear term in (5). This quantity, which we emphasize is a function of $\vec{x}$ and of $d\vec{x}$, is called the differential of the function $f:\mathbb{R}^k \to \mathbb{R}^1$, and one should note the similarity with the corresponding formula (3) for a function of one argument. The definition of differentiabili may now be given essentially the same formulation, for functions of one or of several arguments: f is differentiable at a point if the increment Δf near the point is approximated by the corresponding differential df within the prescribed accuracy (4) or (6) respectively. For functions of one argument the existence of the derivative is sufficient for this accuracy, and for functions of several arguments we have shown that the existence and continuity of the first partial derivatives is sufficient.

Suppose we have a function $\vec{g}:\mathbb{R}^1 \to \mathbb{R}^k$ with components $g^i:\mathbb{R}^1 \to \mathbb{R}^1$, $i = 1,2,\ldots,k$. That is, $\vec{g}(t) = (g^1(t), g^2(t), \ldots, g^k(t))$, $t \in \mathbb{R}^1$. If each g^i is differentiable, $(\Delta g^i)(t,h) = h \cdot \frac{dg^i}{dt} + \alpha_i(h)$, where $(\alpha_i(h)/h) \to 0$ as $h \to 0$, then $(\Delta\vec{g})(t,h) = h \cdot \frac{d\vec{g}}{dt} + \vec{\alpha}(h)$, where $\frac{d\vec{g}}{dt} = \left(\frac{dg^1}{dt}, \ldots, \frac{dg^k}{dt}\right)$ and $\vec{\alpha}(h) = (\alpha_1(h), \ldots, \alpha_k(h))$, and clearly $(\|\vec{\alpha}(h)\|/h) \to 0$ as $h \to 0$. It is therefore a natural extension of the terminology to say in this circumstance (namely, each g^i is differentiable) that $\vec{g}$ is differentiable.

If $\vec{g}:\mathbb{R}^1 \to \mathbb{R}^k$ is differentiable and $f:\mathbb{R}^k \to \mathbb{R}^1$ is differentiable, then the composed function $f \circ \vec{g} : \mathbb{R}^1 \to \mathbb{R}^1$ is differentiable, and

$$\frac{d}{dt}(f \circ \vec{g})(t) = \sum f_i(\vec{g}(t)) \cdot \frac{d}{dt} g^i(t) . \qquad (9)$$

This formula may be guessed from (8), as follows. Putting $\vec{x} = \vec{g}(t)$ there, we have $df(\vec{g}(t), d\vec{g}) = \sum f_i(\vec{g}(t))dg^i$, and dividing now by dt we get (9), more or less. Here is a proof of (9). By the differentiability of $\vec{g}$ we have $\vec{g}(t+h) = \vec{g}(t) + h \cdot \frac{d}{dt}\vec{g}(t) + \vec{\alpha}(h)$. Now $(f \circ \vec{g})(t+h) = f(\vec{g}(t+h))$, so by the differentiability of f we have $(f \circ \vec{g})(t+h) =$

$(f\circ\vec{g})(t) + \sum f_i(\vec{g}(t))\cdot\{h\frac{d}{dt}g^i(t) + \alpha_i(h)\} + \beta(\Delta\vec{g}(t)) =$

$(f\circ\vec{g})(t) + h\sum f_i(\vec{g}(t))\cdot\frac{d}{dt}g^i(t) + \sum f_i(\vec{g}(t))\cdot\alpha_i(h) + \beta(\Delta\vec{g}(t))$. Thus $\Delta(f\circ\vec{g})(t,h)$ is approximated by $h\{\sum f_i(\vec{g}(t))\cdot\frac{d}{dt}g^i(t)\}$ with error $\sum f_i(\vec{g}(t))\alpha_i(h) + \beta(\Delta\vec{g}(t))$. Now $\frac{1}{h}\sum f_i(\vec{g}(t))\cdot\alpha_i(h) \to 0$ as $h\to 0$ because $\frac{1}{h}\alpha_i(h)\to 0$ as $h\to 0$; and $\frac{1}{h}\beta(\Delta\vec{g}(t)) = \frac{\beta(\Delta\vec{g}(t))}{\|\Delta\vec{g}(t)\|}\,\frac{\|\Delta\vec{g}(t)\|}{h} \to 0$ as $h\to 0$ because the first factor vanishes by the differentiability of f, and the second factor equals $\|\frac{d}{dt}\vec{g}(t) + \frac{\vec{\alpha}(h)}{h}\|$, which approaches $\|\frac{d}{dt}\vec{g}(t)\|$ as $h\to 0$. Thus the total error vanishes with the required rate (4); which is to say $f\circ\vec{g}$ is differentiable and (9) holds.

This formula is an extension to several variables of the chain rule, and we shall refer to it by that name.

1.3 Taylor's Theorem

Given $f:\mathbb{R}^k\to\mathbb{R}^1$, let $F:\mathbb{R}^1\to\mathbb{R}^1$ be defined as

$$F(t) = f(\vec{x}+t\vec{h}) \tag{1}$$

for $\vec{x},\vec{h}\in\mathbb{R}^k$ arbitrary and fixed. Considerations of differentiability and convergence aside, the Taylor series for F is

$$F(t) = \sum\frac{t^n}{n!}F^{(n)}(0)\,. \tag{2}$$

By the chain rule 1.2(9) we have