London Mathematical Society
Lecture Note Series 51

Synthetic Differential Geometry

ANDERS KOCK

CAMBRIDGE UNIVERSITY PRESS

London Mathematical Society Lecture Note Series. 51

Synthetic Differential Geometry

ANDERS KOCK
Lecturer in Mathematics
Aarhus University

CAMBRIDGE UNIVERSI
CAMBRIDGE
LONDON NEW YORK NEW ROCHELLE
MELBOURNE SYDNEY

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
32 East 57th Street, New York, NY 10022, USA
296 Beaconsfield Parade, Middle Park, Melbourne 3206, Australia

First published in 1981

Printed in Great Britain at the University Press, Cambridge

Library of Congress catalogue card number 81-6099

British Library cataloguing in publication data

Kock Anders
Synthetic Differential Geometry.-(London Mathematical Society
Lecture Note Series 51 ISSN 0076-0552.)
1.Geometry, Differential
I Title.   II Series
516'.3'6   QA641

ISBN 0 521 24138 3

PREFACE

The aim of the present book is to describe a foundation for synthetic reasoning in differential geometry. We hope that such a foundational treatise will put the reader in a position where he, in his study of differential geometry, can utilize the synthetic method freely and rigorously, and that it will give him notions and language by which such study can be communicated.

That such notions and language is something that till recently seems to have existed only in an inadequate way is borne out by the following statement of Sophus Lie, in the preface to one of his fundamental articles:

> "The reason why I have postponed for so long these investigations, which are basic to my other work in this field, is essentially the following. I found these theories originally by synthetic considerations. But I soon realized that, as expedient [zweckmässig] the synthetic method is for discovery, as difficult it is to give a clear exposition on synthetic investigations, which deal with objects that till now have almost exclusively been considered analytically. After long vacillations, I have decided to use a half synthetic, half analytic form. I hope my work will serve to bring justification to the synthetic method besides the analytical one."
>
> (Allgemeine Theorie der partiellen Differentialgleichungen erster Ordnung, Math. Ann. 9 (1876).)

What is meant by "synthetic" reasoning? Of course, we do not know exactly what Lie meant, but the following is the way we would describe it: It deals with space forms in terms of their structure, i.e. the basic geometric and conceptual constructions that can be performed on them. Roughly these constructions are the morphisms which constitute the base category in terms of which we work; the space forms themselves being objects of it.

This category is cartesian closed, since, whenever we have formed ideas of "spaces" A and B, we can form the idea of B^A, the "space" of all functions from A to B.

The category theoretic viewpoint prevents the identification of A and B with point sets (and hence also prevents the formation of "random" maps from A to B). This is an old tradition in synthetic geometry, where one, for instance, distinguishes between a "line" and the "range of points on it" (cf. e.g. Coxeter [8] p.20).

What categories in the "Bourbakian" universe of mathematics are mathematical models of this intuitively conceived geometric category? The answer is: many of the "gros toposes" considered since the early 60's by Grothendieck and others, - the simplest example being the category of functors from commutative rings to sets. We deal with these topos theoretic examples in Part III of the book. We do not begin with them, but rather with the axiomatic development of differential geometry on a synthetic basis (Part I), as well as a method of interpreting such development in cartesian closed categories (Part II). We chose this ordering because we want to stress that the axioms are intended to reflect some true properties of the geometric and physical reality; the models in Part III are only servants providing consistency proofs and inspiration for new true axioms or theorems. We present in particular some models $\mathcal{E}$ which contain the category of smooth manifolds as a full subcategory in such a way that "analytic" differential geometry for these corresponds exactly to "synthetic" differential geometry in $\mathcal{E}$.

Most of Part I, as well as several of the papers in the bibliography which go deeper into actual geometric matters with synthetic methods, are written in the "naive" style. By this, we mean that all notions, constructions, and proofs involved are presented as if the base category were the category of sets; in particular all constructions on the objects involved are described in terms of "elements" of them. However, it is necessary and possible to be able to understand this naive writing as referring to cartesian closed categories. It is necessary because the basic axioms of synthetic differential geometry have no models in the category of sets (cf. I §1); and it is possible: this is what Part II is about. The method is that we have to understand by an element b of an object B a generalized element, that is, a map $b: X \to B$, where X is an arbitrary object, called the stage of definition, or the domain of variation of the element b.

Elements "defined as different stages" have a long tradition in geometry. In fact, a special case of it is when the geometers say: A circle has no real points at infinity, but there are two imaginary points at infinity such that every circle passes through them. Here $\mathbb{R}$ and $\mathbb{C}$ are two different stages of mathematical knowledge, and something that does not yet exist at stage $\mathbb{R}$ may come into existence at the "later" or "deeper" stage $\mathbb{C}$. - More important for the developments here is passage from stage $\mathbb{R}$ to stage $\mathbb{R}[\varepsilon]$, the "ring of dual numbers over $\mathbb{R}$":

$$\mathbb{R}[\varepsilon] = \mathbb{R}[x]/(x^2).$$

It is true, and will be apparent in Part III, that the notion of elements defined at different stages does correspond to this classical notion of elements defined relative to different commutative rings, like $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{R}[\varepsilon]$, cf. the remarks at the end of III §1.

When thinking in terms of physics (of which geometry of space forms is a special case), the reason for the name "domain of variation" (instead of "stage of definition") becomes clear: for a non-atomistic point of view, a body B is not described just in terms of its "atoms" $b \in B$, that is, maps $\mathbf{1} \to B$, but in terms of "particles" of varying size X, or in terms of motions that take place in B and are parametrized by a temporal extent X; both of these situations being described by maps $X \to B$ for suitable domain of variation X.

The exercises at the end of each paragraph are intended to serve as a further source of information, and if one does not want to solve them, one might read them.

Historical remarks and credits concerning the main text are collected at the end of the book. If a specific result is not credited to anybody, it does not necessarily mean that I claim credit for it. Many things developed during discussions between Lawvere, Wraith, myself, Reyes, Joyal, Dubuc, Coste, Coste-Roy, Bkouche, Veit, Penon, and others. Personally, I want to acknowledge also stimulating questions, comments, and encouragement from Dana Scott, J. Bénabou, P. Johnstone, and from my audiences in Milano, Montréal, Paris, Zaragoza, Buffalo, Oxford, and in particular Aarhus. I want also to thank Henry Thomsen for valuable comments to the early drafts of the book.

The Danish Natural Science Research Council has on several occasions made it possible to gather some of the above-mentioned mathematicians for work sessions in Aarhus. This has been vital to the progress of the subject treated here, and I want to express my thanks.

Warm thanks also to the secretaries at Matematisk Institut, Aarhus, for their friendly help, and in particular, to Else Yndgaard for her expert typing of this book.

Finally, I want to thank my family for all their support, and for their patience with me and the above-mentioned friends and colleagues.

CONTENTS

PART II: Categorical Logic

PART III: Models

PART I

THE SYNTHETIC THEORY

INTRODUCTION

Lawvere has pointed out that "In order to treat mathematically the decisive abstract general relations of physics, it is necessary that the mathematical world picture involve a cartesian closed category $\mathcal{E}$ of smooth morphisms between smooth spaces".

This is also true for differential geometry, which is a science that underlies physics. So everything in the present Part I takes place in such cartesian closed category $\mathcal{E}$. The reader may think of $\mathcal{E}$ as "the" category of sets, because most constructions and notions which exist in the category of sets exist in such $\mathcal{E}$; there are some exceptions, like use of the "law of excluded middle", cf. Exercise 1.1 below. The text is written as if $\mathcal{E}$ were "the" category of sets. This means that to understand this part, one does not have to <u>know</u> anything about cartesian closed categories; rather, one <u>learns</u> it, at least implicitely, because the synthetic method utilizes the cartesian closed structure all the time, even if it is presented in set theoretic disguise (which, as Part II hopefully will bring out, is really no disguise at all).

Generally, investigating geometric and quantitative relationships brings along with it understanding of the logic appropriate for it. So, it also forces $\mathcal{E}$ (which represents our understanding of smoothness) to have certain properties, and not to have certain others. In particular, $\mathcal{E}$ must have finite inverse limits, and for some of the more refined investigations, to be a topos.

I.1: BASIC STRUCTURE ON THE GEOMETRIC LINE

The geometric line can, as soon as one chooses two distinct points on it, be made into a commutative ring, with the two points as respectively 0 and 1. This is a decisive structure on it, already known and considered by Euclid, who assumes that his reader is able to move line segments around in the plane (which gives addition), and who teaches his reader how he, with ruler and compass, can construct the fourth proportional of three line segments; taking one of these to be $[0,1]$, this defines the product of the two others, and thus the multiplication on the line. We denote the line, with its commutative ring structure* (relative to some fixed choice of 0 and 1) by the letter R.

Also, the geometric plane can, by some of the basic structure (ruler-and-compass-constructions again), be identified with $R \times R = R^2$ (choose a fixed pair of mutually orthogonal copies of the line R in it), and similarly, space with R^3.

Of course, this basic structure does not depend on having the (arithmetically constructed) real numbers $\mathbb{R}$ as a mathematical model for R.

Euclid maintained further that R was not just a commutative ring, but actually a field. This follows because of his assumption: for any two points in the plane, either they are equal, or they determine a unique line.

We cannot agree with Euclid on this point. For that would imply that the set D defined by

$$D := [\![x \in R \mid x^2 = 0]\!] \subseteq R$$

---

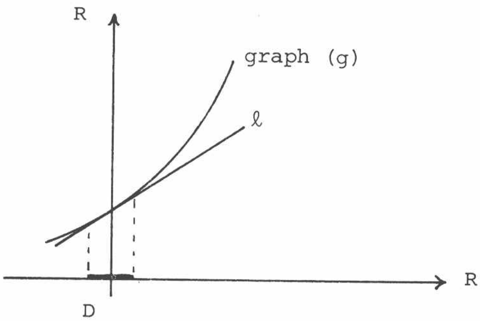
* Actually, it is an algebra over the rationals, since the elements 2 = 1+1, 3 = 1+1+1, etc., are multiplicatively invertible in R.

consists of 0 alone, and that would immediately contradict our

Axiom 1. For any* $g: D \to R$, there exists a unique $b \in R$ such that

$$\forall d \in D: g(d) = g(0) + d \cdot b.$$

Geometrically, the axiom expresses that the graph of g is a piece of a unique straight line ℓ, namely the one through $(0,g(0))$ and with slope b



(in the picture, g is defined not just on D, but on some larger set).

Clearly, the notion of slope, which thus is built in, is a decisive abstract general relation for differential calculus. Before we turn to that, let us note the following consequence of the uniqueness assertion in Axiom 1:

$$(\forall d \in D: d \cdot b_1 = d \cdot b_2) \Rightarrow (b_1 = b_2),$$

which we verbalize into the slogan

---

* We really mean: "for any $g \in R^D$..."; this will make a certain difference in the category theoretic interpretation with generalized elements. Similarly for the f in Theorem 2.1 below and several other places.

"universally quantified d's may be cancelled"

("cancelled" here meant in the multiplicative sense).

The axiom may be stated in succinct diagrammatic form in terms of Cartesian Closed Categories. Consider the map α:

$$R \times R \xrightarrow{\alpha} R^D \tag{1.1}$$

given by

$$(a,b) \longmapsto [d \mapsto a + d\cdot b].$$

Then the axiom says

<u>Axiom 1</u>. α is invertible (i.e. bijective).

Let us further note:

<u>Proposition 1.1</u>. The map α is an R-algebra homomorphism if we make $R \times R$ into an R-algebra by the "ring of dual numbers" multiplication

$$(a_1,b_1)\cdot(a_2,b_2) := (a_1\cdot a_2, a_1\cdot b_2 + a_2\cdot b_1) \tag{1.2}$$

<u>Proof</u>. The pointwise product of the maps $D \to R$

$$d \longmapsto a_1 + d\cdot b_1 \qquad\qquad d \longmapsto a_2 + d\cdot b_2$$

is

$$d \longmapsto (a_1 + d\cdot b_1)\cdot(a_2 + d\cdot b_2)$$

$$= a_1\cdot a_2 + d\cdot(a_1\cdot b_2 + a_2\cdot b_1) + d^2\cdot b_1\cdot b_2,$$

but the last term vanishes because $d^2 = 0 \quad \forall d \in D$.

If we let $R[\varepsilon]$ denote $R \times R$, with the ring-of-dual-numbers multiplication, we thus have

Corollary 1.2. Axiom 1 can be expressed: The map α in (1.1) gives an R-algebra isomorphism

$$R[\varepsilon] \xrightarrow[\cong]{} R^D.$$

Assuming Axiom 1, we denote by β and γ, respectively, the two composites

$$\begin{aligned} \beta &= R^D \xrightarrow{\alpha^{-1}} R \times R \xrightarrow{\text{proj}_1} R \\ \gamma &= R^D \xrightarrow{\alpha^{-1}} R \times R \xrightarrow{\text{proj}_2} R \end{aligned} \qquad (1.3)$$

Both are R-linear, by Proposition 1.1; β is just 'evaluation at $0 \in D$' and appears later as the structural map of the tangent bundle of R; γ is more interesting, being the concept of slope itself. It appears later as "principal part formation", (§7), or as the "universal 1-form", or "Maurer-Cartan form" (§18), on $(R,+)$.

EXERCISES AND REMARKS

1.1 (Schanuel). The following construction * is an example of a use of "the law of excluded middle". Define a function $g: D \to R$ by putting

$$* \begin{cases} g(d) = 1 & \text{if } d \neq 0 \\ g(d) = 0 & \text{if } d = 0 \end{cases}$$

If Axiom 1 holds, $d = \{0\}$ is impossible, hence, again by essentially using the law of excluded middle, we may assume $\exists d_o \in D$ with $d_o \neq 0$. By Axiom 1

$$\forall d \in D: \ g(d) = g(0) + d \cdot b \ .$$

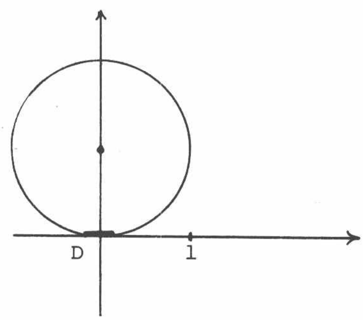
Substituting d_o for d yields $1 = g(d_o) = 0 + d_o \cdot b$, which, when squared, yields $1 = 0$.

Moral. Axiom 1 is incompatible with the law of excluded middle. Either the one or the other has to leave the scene. In Part I of this book, the law of excluded middle has to leave, being incompatible with the natural synthetic reasoning on smooth geometry to be presented here. In the terms which the logicians use, this means that the logic employed is 'constructive' or 'intuitionistic'. We prefer to think of it just as 'that reasoning which can be carried out in all sufficiently good cartesian closed categories'.

1.2 (Joyal). Assuming Pythagoras' Theorem, it is correct to define the circle around (a,b) with radius c to be

$$[\![(x,y) \in R^2 \mid (x-a)^2 + (y-b)^2 = c^2]\!] .$$

Prove that D is exactly the intersection of the unit circle around $(0,1)$ and the x-axis



(identifying, as usal, R with the x-axis in R^2).

Remark. This picture of D was proposed by Joyal in 1977. But earlier than that: Hjelmslev [26] experimented in the 1920's with a geometry where, given two points in the plane, there exists at least one line connecting them, but there may exist more than one without the points being identical; this is the case when the points are 'very near' each other. For such geometry, R is not a field, either, and the intersection in the figure above is, like here, not just $\{0\}$. But even earlier than that: Hjelmslev quotes the Old

Greek philosopher, Protagoras, who wanted to refute Euclid by the argument that it is evident that the intersection in the figure contains more than one point.

1.3. If $d \in D$ and $r \in R$, we have $d \cdot r \in R$. If $d_1 \in D$ and $d_2 \in D$, then $d_1 + d_2 \in D$ iff $d_1 \cdot d_2 = 0$ (for the implication '$\Rightarrow$', one must use that 2 is invertible in R).

(In the geometries that have been built based on Hjelmslev's ideas, $d_1^2 = 0 \wedge d_2^2 = 0 \Rightarrow d_1 \cdot d_2 = 0$, but this assumption is incompatible with Axiom 1, see Exercise 4.6 below.)

1.4 (Galuzzi and Meloni; cf.[50] p. 6). Assume $E \subseteq R$ contains 0 and is stable under multiplication by -1. If 2 is invertible in R, and if Axiom 1 holds for E (i.e., when D in Axiom 1 is replaced by E), then $E \subseteq D$.

1.5. If R is any commutative ring, and g is any polynomial (with integral coefficients) in n variables, g gives rise to a polynomial function $R^n \to R$, which may be denoted g_R or just g. For the ring R^X (X an arbitrary object), g_{R^X} gets identified with $(g_R)^X$. To say that a map $\beta: R \to S$ is a ring homomorphism is equivalent to saying that for any polynomial g (in n variables, say)

$$g_S \circ \beta^n = \beta \circ g_R .$$

This is the viewpoint that the algebraic theory consisting of polynomials is the algebraic theory of commutative rings, cf. Appendix A.

In particular, Proposition 1.1 can be expressed: for any polynomial g (in n variables, say), the diagram

$$
\begin{array}{ccc}
(R[\varepsilon])^n & \xrightarrow{\alpha^n} & (R^D)^n \cong (R^n)^D \\
\Big\downarrow {\scriptstyle g_{R[\varepsilon]}} & & \Big\downarrow {\scriptstyle g_{R^D}} \\
R[\varepsilon] & \xrightarrow[\alpha]{} & R^D
\end{array}
\qquad (1.4)
$$

commutes. In III § 4 ff. we shall meet a similar statement, but for arbitrary smooth functions $g\colon \mathbb{R}^n \longrightarrow \mathbb{R}$, not just polynomials.