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 *know* anything about cartesian closed categories; rather, one *learns* it, at least implicitly, 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, it must 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

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

consists of 0 alone, and that would immediately contradict our

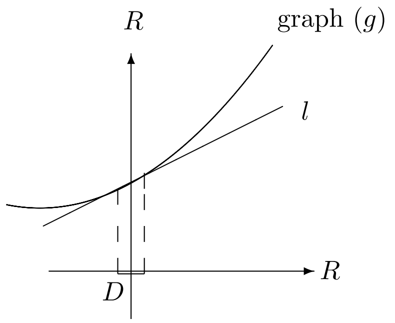
Axiom 1. *For any*‡ $g : D \to R$, *there exists a unique* $b \in R$ *such that*

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

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

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

‡ We really mean: "for any $g \in R^D \ldots$"; 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.



(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

"universally quantified ds 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

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

Then the axiom says

Axiom 1. *α is invertible (i.e. bijective).*

Let us further note:

Proposition 1.1. *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*

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

Proof. The pointwise product of the maps $D \to R$

$$d \mapsto a_1 + d \cdot b_1 \quad d \mapsto a_2 + d \cdot b_2$$

is

$$\begin{aligned} d &\mapsto (a_1 + d \cdot b_1) \cdot (a_2 + d \cdot b_2) \\ &= a_1 \cdot a_2 + d \cdot (a_1 \cdot b_2 + a_2 \cdot b_1) + d^2 \cdot b_1 \cdot b_2, \end{aligned}$$

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

If we let $R[\epsilon]$ 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*

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

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

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

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

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

If Axiom 1 holds, $D = \{0\}$ is impossible, hence, again by essentially using the law of excluded middle, we may assume $\exists d_0 \in D$ with $d_0 \neq 0$. By Axiom 1

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

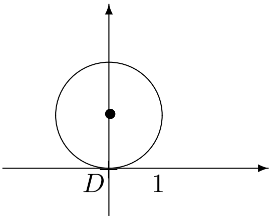
Substituting d_0 for d yields $1 = g(d_0) = 0 + d_0 \cdot b$, which, when squared, yields $1 = 0$.

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 usual, 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 1920s 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]

1.3. If $d \in D$ and $r \in R$, we have $d \cdot r \in D$. 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

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

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

I.2 Differential calculus

In this §, R is assumed to satisfy Axiom 1; and we assume that $2 \in R$ is invertible.

Let $f : R \to R$ be any function. For fixed $x \in R$, we consider the function $g : D \to R$ given by $g(d) = f(x + d)$. There exists, by Axiom 1, a unique $b \in R$ so that

$$g(d) = g(0) + d \cdot b \ \ \forall d \in D, \tag{2.1}$$

or in terms of f

$$f(x + d) = f(x) + d \cdot b \ \ \forall d \in D.$$

The b here depends on the x considered. We denote it $f'(x)$, so we have

Theorem 2.1 (Taylor's formula). *For any $f : R \to R$ and any $x \in R$,*

$$f(x + d) = f(x) + d \cdot f'(x) \ \ \forall d \in D. \tag{2.2}$$

Formula (2.2) characterizes $f'(x)$. Since we have $f'(x)$ for each $x \in R$, we have in fact defined a new function $f' : R \to R$, the derivative of f. The process may be iterated, to define $f'' : R \to R$, etc.

If f is not defined on the whole of R, but only on a subset $U \subseteq R$, then we can, by the same procedure, define f' as a function on the set $U' \subseteq U$ given by $U' = [[x \in U \mid x + d \in U \ \forall d \in D]]$. In particular, for $g : D \to R$, we may define $g'(0)$; it is the b occurring in (2.1). Also, there will in general exist many subsets $U \subseteq R$ with the property that $U' = U$, equivalently, such that

$$x \in U \wedge d \in D \ \Rightarrow \ x + d \in U. \tag{2.3}$$

For f defined on such a set U, we get $f' : U \to R$, $f'' : U \to R$, etc. In the following Theorem, U and V are subsets of R having the property (2.3).

Theorem 2.2. *For any $f, g : U \to R$ and any $r \in R$, we have*

$$(f + g)' = f' + g' \tag{i}$$

$$(r \cdot f)' = r \cdot f' \tag{ii}$$

$$(f \cdot g)' = f' \cdot g + f \cdot g' \tag{iii}$$

For any $g : V \to U$ *and* $f : U \to R$

$$(f \circ g)' = (f' \circ g) \cdot g' \quad \text{(iv)}$$

$$\text{id}' = 1 \quad \text{(v)}$$

$$r' \equiv 0 \quad \text{(vi)}$$

(where id : $R \to R$ *is the identity map and* r *denotes the constant function with value* r*).*

Proof. All of these are immediate arithmetic calculations based on Taylor's formula. As a sample, we prove the Leibniz rule (iii). For any $x \in U \subseteq R$, we have

$$(f \cdot g)(x + d) = (f \cdot g)(x) + d \cdot (f \cdot g)'(x) \;\; \forall d \in D,$$

by Taylor's formula for $f \cdot g$. On the other hand

$$\begin{aligned}(f \cdot g)(x + d) &= f(x + d) \cdot g(x + d) \\ &= (f(x) + d \cdot f'(x)) \cdot (g(x) + d \cdot g'(x)) \\ &= f(x) \cdot g(x) + d \cdot f'(x) \cdot g(x) + d \cdot f(x) \cdot g'(x);\end{aligned}$$

the fourth term $d^2 \cdot f'(x) \cdot g'(x)$ vanishes because $d^2 = 0$. Comparing the two derived expressions, we see

$$d \cdot (f \cdot g)'(x) = d \cdot (f'(x) \cdot g(x) + f(x) \cdot g'(x)) \; \forall d \in D.$$

Cancelling the universally quantified d yields the desired

$$(f \cdot g)'(x) = f'(x) \cdot g(x) + f(x) \cdot g'(x).$$

It is not true on basis of Axiom 1 alone that $f' \equiv 0$ implies that f is a constant, or that every f has a primitive g (i.e. $g' \equiv f$ for some g), cf. Part III.

What about Taylor formulae longer than (2.2)? The following is a partial answer for "series" going up to degree-2 terms. It generalizes in an evident way to series going up to degree-n terms. Again, f is a map $U \to R$ with U satisfying (2.3).

Proposition 2.3. *For any* δ *of form* $d_1 + d_2$ *with* d_1 *and* $d_2 \in D$ *we have*

$$f(x + \delta) = f(x) + \delta \cdot f'(x) + \frac{\delta^2}{2!} f''(x).$$

Proof.

$$f(x+\delta) = f(x+d_1+d_2)$$
$$= f(x+d_1) + d_2 \cdot f'(x+d_1)$$

(by (2.2))

$$= f(x) + d_1 \cdot f'(x) + d_2 \cdot (f'(x) + d_1 \cdot f''(x))$$

(by (2.2) twice)

$$= f(x) + (d_1+d_2) \cdot f'(x) + d_1 \cdot d_2 \cdot f''(x).$$

But since $d_1^2 = d_2^2 = 0$, we have $(d_1+d_2)^2 = 2 \cdot d_1 \cdot d_2$. Substituting this, and $\delta = d_1 + d_2$ gives the result.

The reason why this Proposition is to be considered a *partial* result only, is that we would like to state it for any δ with $\delta^3 = 0$, not just for those of form $d_1 + d_2$ as above. In the models (Part III), $\delta^3 = 0$ does not[2] imply existence of $d_1, d_2 \in D$ with $\delta = d_1 + d_2$. In the next §, we strengthen Axiom 1, and after that, the result of Proposition 2.3 will be true for all δ with $\delta^3 = 0$; similarly for still longer Taylor formulae.

EXERCISES

2.1. Assume R is a ring that satisfies the following axiom ("Fermat's Axiom") :

$$\forall f : R \to R \; \exists! g : R \times R \to R : \quad \forall x, y \in R : f(x) - f(y) = (x-y) \cdot g(x,y) \tag{2.4}$$

Define $f' : R \to R$ by $f'(x) := g(x,x)$, and prove (assuming $U = R$) the results of Theorem 2.2 (this requires a little skill). – The axiom and its investigation is mainly due to Reyes.

Use the idea of Exercise 1.1 to prove that the law of excluded middle is incompatible with Fermat's Axiom.

Moral. Fermat's Axiom is an alternative synthetic foundation for calculus, which does not use nilpotent elements.[3] The relationship between Axiom 1 and (2.4) is further investigated in §13 (exercises), and models for (2.4) are studied in III §8 and III §9.

I.3 Higher Taylor formulae (one variable)

In this §, we assume that $2, 3, \ldots$ are invertible in R (i.e. that R is a $\mathbb{Q}$-algebra).

We let $D_k \subseteq R$ denote the set

$$D_k := [[x \in R \mid x^{k+1} = 0]],$$

in particular, D_1 is the D considered in §§1 and 2. The following is clearly a strengthening of Axiom 1.

Axiom 1′. *For any $k = 1, 2, \ldots$ and any $g : D_k \to R$, there exist unique $b_1, \ldots, b_k \in R$ such that*

$$\forall d \in D_k : g(d) = g(0) + \sum_{i=1}^{k} d^i \cdot b_i.$$

Assuming this, we can prove

Theorem 3.1 (Taylor's formula). *For any $f : R \to R$ and any $x \in R$*

$$f(x+\delta) = f(x) + \delta \cdot f'(x) + \ldots + \frac{\delta^k}{k!} f^{(k)}(x) \quad \forall \delta \in D_k$$

(again it would suffice for f to be defined on a suitable subset U around x).

Proof. We give the proof only for $k = 2$, (cf. the exercises below, or [32], for larger k). We have, by Axiom 1′, b_1 and b_2 such that, for any $\delta \in D_2$

$$f(x+\delta) = f(x) + \delta \cdot b_1 + \delta^2 \cdot b_2; \tag{3.1}$$

specializing to δs in D_1, we see that $b_1 = f'(x)$. We have, by Proposition 2.3 for any $(d_1, d_2) \in D \times D$

$$f(x + (d_1 + d_2)) = f(x) + (d_1 + d_2) \cdot f'(x) + (d_1 + d_2)^2 \cdot \frac{f''(x)}{2!}. \tag{3.2}$$

For $\delta = d_1 + d_2$, we therefore have, by comparing (3.1) and (3.2) and using $b_1 = f'(x)$

$$\forall (d_1, d_2) \in D \times D : (d_1 + d_2)^2 \cdot b_2 = (d_1 + d_2)^2 \cdot \frac{f''(x)}{2!}$$

or

$$\forall (d_1, d_2) \in D \times D : 2 \cdot d_1 \cdot d_2 \cdot b_2 = 2 \cdot d_1 \cdot d_2 \cdot \frac{f''(x)}{2!}.$$

Cancelling the universally quantified d_1, and then the universally quantified d_2 (and the number 2), we derive

$$b_2 = \frac{f''(x)}{2!},$$