Synthetic Differential Geometry

An application to Einstein's Equivalence Principle

Tim de Laat

Bachelor's thesis for Mathematics and Physics & Astronomy

Supervisor: Prof. Dr. N.P. Landsman
Second Reader: Dr. M.H.A.H. Müger

Radboud University Nijmegen

Institute for Mathematics, Astrophysics and Particle Physics

Preface

This thesis is the result of my bachelor project in both Mathematics and Physics & Astronomy. The aim of this project was to give a satisfactory and rigorous formulation of the equivalence principle of the general theory of relativity (GR) in terms of synthetic differential geometry (SDG). SDG is a "natural" formulation of differential geometry in which the notion of "infinitesimals" is very important. Smooth infinitesimal analysis (SIA) is the mathematical analysis corresponding to these infinitesimals and it forms an entrance to SDG. Both SIA and SDG are formulated in terms of categories and topoi. As I was quite new to these subjects, I first needed to study them thoroughly before I could start studying SDG.

Besides using synthetic differential geometry to reformulate Einstein's equivalence principle, I intend to give an introduction to SIA and SDG. I will also explain the special aspects of these theories and point out the contrasts with the usual theories and structures. I assume that the reader has some background in mathematical reasoning, logic, abstract algebra and classical analysis. Background in category theory and classical differential geometry is not assumed, but may make things easier. I wrote an appendix covering basic category theory in a concise way. However, this should not be regarded as an introductory text to category theory.

My project was supervised by Prof. Dr. Klaas Landsman. I want to thank him for the original idea and the enthusiastic supervision. I want to thank Dr. Michael Müger for being the second reader of this thesis. I also want to thank Prof. Dr. A. Kock from Aarhus University, Prof. Dr. I. Moerdijk from the University of Utrecht and Prof. Dr. G.E. Reyes from the Université de Montréal for kindly answering the questions Klaas and I asked them.

Tim de Laat,
Nijmegen, July 2008.

Contents

1 Introduction

Natural scientists have made models to describe the world around us for a very long time. Physics, in the way it is still being taught, came up in the seventeenth century. One of the most important events was the publication of Isaac Newton's book, *Philosophiae Naturalis Principia Mathematica*, in 1687. One can still buy modern printings of this book [16]. In this book, Newton invented what is nowadays appropriately called Newtonian mechanics. For the formulation of this theory, he also invented basic calculus, but this nice piece of mathematics was, curiously, left out of his book. In his other work, the most important calculus statement was the fundamental theorem of calculus, which in its modern form reads as follows:

Theorem 1.1. *Let $f : [a, b] \longrightarrow \mathbb{R}$ be a continuous function. Let $F : [a, b] \longrightarrow \mathbb{R}$ be differentiable such that $\forall x \in [a, b]\ f(x) = F'(x)$. Then*

$$\int\limits_a^b f(t)dt = F(b) - F(a).$$

The idea to build a mathematical framework to solve physical problems and to formulate models of physical systems in a mathematical way, was completely new at that time. The way in which Newton developed this idea connected closely to his scientific intuition. In the centuries that followed, mathematics and (theoretical) physics were practically inseparable.

Some years before Newton published his calculus, another brilliant philosopher and scientist, Gottfried Wilhelm Leibniz, also published the basic calculus. Also his main result was the fundamental theorem of calculus. The notation he invented is the one we still use. However, Leibniz did not invent the calculus for the description of physical systems.

The question who deserves the credit for inventing the calculus is a topic of ongoing debate [6]. It is often said that Newton derived his results first, but published them after Leibniz, and we also know that Leibniz knew Newton's work to some extent. For us, it is not at all necessary to settle this debate, since both Newton and Leibniz have made contributions to mathematics which are still of great importance to us: Newton described physical systems in a mathematical way and Leibniz used so-called infinitesimals in his formulation of calculus. Newton also used them, but changed to another formulation during his later work. Infinitesimals are "infinitely small quantities"[1]. Leibniz considered the derivative of a function by calculating its increase on a certain interval. If one takes the interval infinitely small, say $[x, x + dx]$, then the slope of the function on that interval is the derivative of the function at the point x.

Intuitively, mathematicians and in particular physicists and other scientists use infinitesimals to derive results which are often correct. There are also methods to solve equations which make use of infinitesimals, which have been proved to be correct by different means. An example is given by a method which is widely used to solve ordinary differential equations: separation of variables.

Any ordinary differential equation (in Leibniz's notation) of the form

$$\frac{dy}{dx} = g(x)h(y),$$

---

[1] We will give a precise definition later.

where y is a function depending on x and g and h are functions of one variable such that $h(y) \neq 0$, can be solved using the following method. We separate the variables, i.e. we rewrite the equation such that the left-hand side only depends on y and the right-hand side only depends on x:

$$\frac{dy}{h(y)} = g(x)dx.$$

(Formally) integrating this equation then yields a class of solutions to the differential equation. In general, for the uniqueness of a solution, an initial value is needed.

Note that the method above is a formal method in the way it is given, since we did not formulate it in a rigorous way. It shows us that the ideas of Leibniz are still used in practice, at least as a heuristic principle. Note, however, that in standard analysis, infinitesimals are not defined. The most important objects defined in standard analysis are numbers, functions, various kinds of spaces and, what is very important, limits, but infinitesimals are avoided. One could say that in the standard formulation of analysis, limits play the role of infinitesimals. The reason that the concept of infinitely small quantities is dropped, is that you do not really need them in modern analysis. But are infinitesimals not just easier to understand than limits?

In modern physics, which may be considered to have started with the publication of the special theory of relativity by Einstein in 1905, there are many physicists who do not use mathematics in a rigorous way. Results need not be proved mathematically to be physically correct. There are still mathematicians who try to formulate physical theories in a rigorous mathematical framework, but that is not at all a goal of modern physics itself.

Roughly speaking, in physics, a theory is considered correct if its predictions are confirmed by experiments; in mathematics, a theory is considered correct if all symbols have been defined properly and if all statements have been proved. A way to combine mathematics and physics is done in what is called mathematical physics. Theories in mathematical physics have to make verifiable predictions as well, but the symbols have to be defined properly and all statements have to be proved mathematically (and not by experiments). So you could think of it as adopting the truth interpretation of both mathematics and physics.

In what follows, we will use the ideas of both Newton and Leibniz. We will consider a physical theory and describe it mathematically (Newton) and in our description we will make use of infinitesimals in a rigorous way (Leibniz). The great advantage of this method is that it is both natural, intuitive and rigorous. We will use a framework that came up in mathematics in the 1940s: category theory. Category theory is a theory that describes mathematical structures starting from the notion that many of these structures "behave in the same way" [13]. Some of these structures are very similar to sets and the functions between them. Such a structure is called a topos [15]. This statement is not very precise, but it will be made so in what follows. Topos theory provides a framework in which it is possible to define a category which roughly behaves like sets and functions and which consists of "smooth" objects and morphisms. This gives us a structure to describe physics mathematically, as was already anticipated by Lawvere.

We will define a topos in which the objects and arrows are smooth, whatever that may be [9]. This topos is called the smooth world. Our main goal is, after giving a survey of the most important results of relativity theory and the foundations of general relativity, to formulate Einstein's equivalence principles, one of the foundations of general relativity, in the smooth

world. It turns out that in this topos we can even say more than the classical formulation of the principle.

Einstein's general relativity theory was published in 1915 and 1916. It is a geometric theory of spacetime, in which gravitation is not an action-at-distance interaction, but a property of the geometry of spacetime [19], [23]. The theory combines the special theory of relativity and Newton's universal law of gravity and can be considered as one of the most brilliant contributions to modern science.

The foundations of general relativity are still being discussed. Einstein himself changed his opinion about them during his life, and still, scientists do not agree what the foundations are [17], [18]. One of them, given by Einstein, is the equivalence principle. It says that, if one is isolated, i.e. if one cannot look around, one cannot distinguish between gravitational forces and acceleration. This heavily relies on Galilei's observation that inertial and gravitational mass is physically the same.

Later, it was pointed out that the equivalence principle is wrong in the way it was formulated by Einstein. Because of so-called tidal forces, one would be able to decide if a force is gravitational or caused by acceleration, since gravitational forces are in general not uniform. Only if they were uniform, they would be indistinguishable from acceleration.

There are other, so-called infinitesimal formulations of the principle. An important one was given by Pauli [17], [18]. Mathematically, probably the best formulation says that the metric, which describes the local curvature of spacetime, is locally Lorentz [19], [23]. The Lorentz metric is the metric of special relativity. So the infinitesimal principle says that spacetime in general relativity locally behaves like special relativity. The best you can get in classical geometry, is that the metric and its derivatives are Lorentz and zero respectively at an arbitrary point in spacetime. This is not a very strong statement.

Our observation is that synthetic differential geometry may give a satisfactory and rigorous way to formulate general relativity. We will reformulate the principle and say even more than the classical formulation. We have chosen the equivalence principle as our object of study, because it may be the most important foundation of the general theory of relativity, despite of which it was never formulated well mathematically.

2 Topoi

In this chapter, we define the concept of a topos. A brief overview of the basic category theory needed for the rest of this thesis, is given in appendix A.

Definition 2.1. *An **elementary topos** (plural: topoi) $\mathcal{E}$ is a category such that:*

1. *$\mathcal{E}$ has all finite limits and colimits,*

2. *$\mathcal{E}$ has exponentiation,*

3. *$\mathcal{E}$ has a subobject classifier.*

In what follows, we will drop the adjective "elementary" and just talk about "topoi". There are various other (equivalent) definitions of a topos, but we will use the one stated above. Other definitions turn out to be equivalent, as follows from basic category theory.

The above definition of a topos is quite complex. It requires the existence of a number of categorical constructions. However, there is an intuitive way to think of topoi. They are generalizations of the category of sets and functions. Roughly speaking, topoi have the same properties, namely the three given above in the definition of a topos. Like for any generalization in mathematics, one can prove that the thing generalized is a special case of the generalization.

Definition 2.2. *$\textbf{Set}$ is the category with as objects all small sets and as arrows all functions between them.*

Proposition 2.1. *$\textbf{Set}$ is a topos.*

To prove this proposition, one needs to check the three conditions of a topos given above. The finite limits and colimits are technical categorical properties of the category **Set**. We do not give them explicitly.

Any two-element set is a subobject classifier in **Set**. One can identify such a set with the collection of classical truth values (e.g. $\{0, 1\}$ or $\{true, false\}$). The characteristic function from an object A to the subobject classifier Ω then classifies subobjects.

The exponentiation of two sets A and B is given by $A^B = \{f : B \longrightarrow A\}$. Such exponents are well-defined for any two sets A and B.

For a complete proof of this proposition, cf. [12].

When working with sets, we use symbols like $\{\}$, $\in$, $\subset$, $\bigcap$ and $\bigcup$ as defined in axiomatic set theory. Together with the logical operators ($\wedge, \vee, \rightarrow, \neg$) and the universal and existentional quantifiers ($\forall$ and $\exists$), this language is called the **internal language** of sets. In a topos we can do more or less the same, since a topos defines its own internal language. We will not explain how this works in detail, but we will say something about it when defining our topos of interest. For more information, cf. [3], [4] or [15]. We will simply use the internal language in a somewhat intuitive way.

In the category **Set**, the objects are sets, which contain elements. In the categorical description we are not bothered with determining what elements a set has, but in determining if and how a set can be mapped to another one (injective, surjective, bijective). Actually, we are in particular interested in how many elements a set has, since any two sets with the same cardinality are isomorphic objects in **Set**, which is very straightforward to prove by defining a bijection between them.

There is a very nice way of denoting elements of sets in terms of arrows from the initial object to another object. This way of "denoting" elements works well for arbitrary topoi. An element of an object A is then considered to be an arrow from the terminal object[2] to A: $a : 1 \longrightarrow A$. This corresponds to the notion of **point** or **location** which we adopt in the topos $\mathbb{S}$ that we will define later on. For more information about this way of denoting elements, cf. [12].

2.1 Topoi in physics

As already mentioned in the introduction, a reason to use topos theory in physics was given by Lawvere. He said [9]:

> "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."

Another reason, which came up later, was the mathematical formulation of quantum theory. The logic in quantum theory is not classical, and if the logic is already inappropriate, why try to give an all-encompassing formulation of quantum theory in terms of sets.

In topoi one has a minimal, more primitive logic than the classical one, namely **intuitionistic** logic[3]. This is currently also being considered to be the logic of quantum theory by some, so a formulation of quantum theory in terms of topoi may be very successful [7].

When specifying the objects and arrows of a topos, more logical rules may be adopted. For example, if one defines the topos **Set**, then the law of excluded middle is adopted and we end up with classical logic in this topos. We will say more about that later.

There is, however, a problem in combining topos theory and physics. Topos theory is a very hard mathematical theory. It requires the notion of category theory and many abstract structures. Most physicists do not know anything about categories and topoi. Even Lee Smolin, who is considered to be one of the leading theoretical physicists of this time by some, says the following in his book [21]: "As a mathematical formalism, topos theory is not easy. It is perhaps the hardest mathematical subject I've yet encountered." So purely sociological it will take some time before topos theory and physics can be combined.

---

[2]In an arbitrary topos there exists an initial object, since an initial object can be given as a finite colimit, and all topoi have all finite limits and colimits.

[3]This kind of logic will be described in the next chapter.

3 Axiomatic Smooth Infinitesimal Analysis

3.1 Motivation

Smooth infinitesimal analysis (SIA) is a formulation of mathematical analysis in terms of infinitesimals, which gives us an entrance to synthetic differential geometry (SDG). This formulation is an attempt to rigorously formulate analysis in a way that is closer to the ideas of Leibniz than ordinary analysis. The axiomatisation defines analysis in a so-called smooth world $\mathbb{S}$, which is a topos. This $\mathbb{S}$ is a model of SIA. It first came up in Lawvere's work in the 1970s. Lawvere is in fact, together with Grothendieck and Tierney, one of the founders of topos theory.

It is important not to confuse SIA with non-standard analysis, which was axiomatised in the 1960s by Robinson. There, infinitesimally small quantities come up as reciprocals of infinitesimally large quantities, which is not the case in SIA. For more information about non-standard analysis, cf. [20].

Two important concepts in SIA are *continuum* and *infinitesimal*. A continuum is the domain over which a continuously varying magnitude varies. A connected continuum coheres and is indefinitely divisible. It is not composed of discrete points. It has a so-called *non-punctuate nature*. Leibniz already said: "A point may not be considered a part of a line," which clearly illustrates this point of view.

Some prominent mathematicians think that set theory is not capable of describing continua [1], because set theory is based on sets which contain elements, which have a punctuate nature. As pointed out earlier, the category of sets is generalized to topoi, categories which are very similar to sets. However, topoi need not have a punctuate nature and hence turn out to be very useful structures for giving models of SIA.

An infinitesimal is often thought of as a quantity that is not necessarily equal to zero and smaller than any finite quantity. For our purposes, we assume that an infinitesimal quantity x is *nilsquare*, i.e. $x^2 = 0$, which clearly implies that $x^n = 0$ for $n \geq 2$. We want to formulate mathematical analysis such that the infinitesimals are parts of continua. They permit a non-punctuate nature, since e.g. a line will not be thought as composed of points, but of infinitesimal line segments, called *linelets*. This gives us the reason that we need topoi for our description of infinitesimals. They simply do not exist in set theory, but can in fact be constructed in topoi.

3.2 Logic

Many definitions in mathematics assume that the logic of the system has already been defined. Roughly speaking, the logic of a mathematical system or framework is the collection of rules that are allowed in definitions, statements and proofs. There are several ways to define these rules in a rigorous way. In topos theory, logic occupies a prominent position. In general, there are two ways of defining a topos:

1. Define the objects and morphisms of the topos and derive the logical rules.

2. Define the logical rules and construct a topos model corresponding to these rules.

One could say that the first way is very natural and intuitive, since one first defines the most important properties of the structure, and then derives what one is allowed to do.

In any topos, we have a "minimal" logic in the sense that the logic of every topos contains the logical rules and rules of inference of this logic. This minimal logic turns out to be intuitionistic logic.

We will use the logical operators $(\wedge, \vee, \rightarrow, \neg)$ and the universal and existential quantifiers ($\forall$ and $\exists$), which form the essential part of the internal language of toposes. We will not rigorously define these notions, although they are by no means trivial. For more information, cf. [3], [4] or [14].

The axioms of (first-order) intuitionistic logic as given in [1] are as follows:

1. $A \rightarrow (B \rightarrow A)$;

2. $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$;

3. $A \rightarrow (B \rightarrow (A \wedge B))$;

4. $(A \wedge B) \rightarrow A$ and $(A \wedge B) \rightarrow B$;

5. $A \rightarrow (A \vee B)$ and $B \rightarrow (A \vee B)$;

6. $(A \rightarrow C) \rightarrow ((B \rightarrow C) \rightarrow ((A \vee B) \rightarrow C))$;

7. $(A \rightarrow B) \rightarrow ((A \rightarrow \neg B) \rightarrow \neg A)$;

8. $\neg A \rightarrow (A \rightarrow B)$;

9. $D(y) \rightarrow \exists x D(x)$ and $\forall x D(x) \rightarrow D(y)$;

10. $x = x$ and $(D(x) \wedge (x = y)) \rightarrow D(y)$.

Here A, B and C are propositions and D is a predicate on the variable x. We also assume two rules of inference:

1. From A and $A \rightarrow B$ we can conclude B (Modus ponens).

2. $(B \rightarrow A(x))$ implies $B \rightarrow \forall x A(x)$ and $A(x) \rightarrow B$ implies $\exists x A(x) \rightarrow B$.

Note that if we add the law of excluded middle — $A \vee \neg A$ — as an axiom, we obtain (first-order) classical logic. The law of double negation, $(\neg\neg A) \rightarrow A$, is logically equivalent to the law of excluded middle, so we cannot use this law either. This is a serious restriction in giving proofs.

3.3 Axiomatic construction of $\mathbb{S}$

We will use the second way described above to define SIA, i.e. we will define the logical rules, give some axioms and give a model for it. If you want to know more about model theory, cf. [8]. If one wants to use a more natural and intuitive way of defining $\mathbb{S}$ —and that was actually our reason to use infinitesimals—, the first way of defining a topos perfectly makes sense as well. Cf. [1] for more information.

We start with saying that in any case in SIA we can use intuitionistic logic[4] (as defined above) and that the smooth world $\mathbb{S}$, being a topos, contains an object Ω that plays the role of the

---

[4]This includes the logical operators $(\wedge, \vee, \rightarrow, \neg)$ and the universal and existential quantifier ($\forall$ and $\exists$).

collection of **truth values**. We assume that Ω contains at least two elements, and call these "true", denoted $\top$ and "false", denoted $\bot$.[5] Maps from an object X to Ω correspond to parts of X. We say that the map from X to Ω that takes the constant value "true" corresponds to X itself and that the map that takes the constant value "false" corresponds to the empty part of X.

There is an object R in our topos $\mathbb{S}$, for which the following three axioms hold:

R_1 : $(R, +, \cdot, 0, 1)$ is a commutative ring with identity. There is a notion of location[6] on R and we assume we can say when two points in R are equal.[7] We assume that $0 \neq 1$.

R_2 : There is an order relation $<$ on R which makes into an ordered ring, in which from every positive element we can extract a square root.

Kock-Lawvere axiom : For all f in R^Δ there exists a unique $b \in R$ such that $\forall \varepsilon \in \Delta$ we have $f(\varepsilon) = f(0) + \varepsilon b$.

Here, $\Delta := \{x \in R | x^2 = 0\}$. The notion of a function f will be given in the construction of a model of SIA. However, note that we can state this axiom without defining functions as long as the definition will not lead to inconsistencies.

We will see that the Kock-Lawvere axiom together with the intuitionistic logic is what makes smooth infinitesimal analysis differ from classical analysis. The consequences of these axioms are given in the next chapter.

These axioms give us the notion of points in R. We will see that there are also other (generalized) elements in R: the infinitesimals. These two kinds of elements take into account the undecidability of the relation $(\varepsilon = 0) \vee (\varepsilon \neq 0)$. For more information about generalized elements, cf. [12].

The existence of SIA is proved by constructing models of it. A model can be regarded as an explicit framework in which the given logic and the given axioms hold. We will briefly describe a way to construct a model for SIA, without proofs. We will only heuristically explain what we need to do, and describe the first step in some detail, but we prefer not to enter into a discussion about sheaves and presheaves. For a thorough explanation, cf. [14].

Definition 3.1. *A topos $\mathcal{E}$ is a model for* SIA *if in $\mathcal{E}$ we have (at least) intuitionistic logic and if $\mathcal{E}$ contains $(R, +, \cdot, 0, 1)$ for which the three axioms given above hold.*

It can be proved that any topos is a model of intuitionistic logic.

For the construction of a model of SIA, we need quite advanced category theory, but the construction can be skipped if you are primarily interested in the consequences of SIA and results which are proved in the model itself. We will not use much of the following construction after this chapter.

We will start our construction vy repeating a definition and giving two others.

Definition 3.2. $\Delta := \{x \in R \mid x^2 = 0\}$.

---

[5]Note that the fact that Ω contains at least two elements is very natural, because the logic would be trivial if there were only one truth value.

[6]What is meant by this, is explained in the next chapter.

[7]A point in R can categorically be regarded as the map $1 \longrightarrow R$. Thus we really have a notion of points on R. For more information about this, cf. [12].

Definition 3.3. *The category **Man** has as objects all small smooth manifolds and as arrows all smooth, i.e. infinitely differentiable, mappings between them.*

Definition 3.4. *The category **CRng** has as objects all small commutative rings with identity and as arrows all unital ring homomorphisms between them.*

Man is neither a topos, nor does it contain an object like Δ. Our aim is to embed **Man** into a topos $\mathcal{E}$ that contains our R such that this embedding preserves the most important properties of **Man**.

We will first embed **Man** in a category $\mathcal{A}$ in a certain way. Note that **Man** is the category of manifolds in the formulation of standard differential geometry, so the field we work over is $\mathbb{R}$, the field of real numbers, in which infinitesimals other than zero do not occur.

The set of smooth mappings from a manifold M to $\mathbb{R}$, together with the binary operations pointwise addition and pointwise multiplication, forms a ring, the coordinate ring of M, denoted CM.

The embedding of **Man** in $\mathcal{A}$ is given in the following way. Each manifold M has a coordinate ring in **CRng** as described above. We identify each manifold M, i.e. each object in **Man**, with its coordinate ring CM and add this ring to a collection $\mathcal{A}$, which was empty before we started. Then we will add some other rings, which we identify with the micro objects —roughly speaking, these are objects like Δ.

We define a ring $\mathbb{R}^* := (\mathbb{R} \times \mathbb{R}, \oplus, \otimes, (0,0), (1,0))$, where the binary operations are defined as follows:

$$
\begin{aligned}
(a,b) \oplus (c,d) &= (a+c, b+d), \\
(a,b) \otimes (c,d) &= (ac, ad+bc).
\end{aligned}
\tag{1}
$$

We define this ring to be the coordinate ring of Δ, denoted $\mathbb{R}^{\Delta}$, and add it to $\mathcal{A}$.

This means that we assign to each smooth map between manifolds $g : M \longrightarrow N$ a ring homomorphism $Cg : CM \longrightarrow CN$. This C is a contravariant functor from **Man** to $\mathcal{A}$, which can be regarded as functorial embedding of **Man** in $\mathcal{A}^{op}$, the dual category of $\mathcal{A}$.

$\mathcal{A}^{op}$ can be identified with all smooth manifolds, micro objects and smooth mappings between them, but is alas not a topos. To achieve this, there is a natural embedding from $\mathcal{A}^{op}$ into the (presheaf[8]) topos $\mathbf{Set}^{\mathcal{A}}$, called the Yoneda embedding[9].

Then we need another embedding $L : \mathbf{Set}^{\mathcal{A}} \longrightarrow E$, because $\mathbf{Set}^{\mathcal{A}}$ still does not have the desired properties, e.g. the relation $\forall x \in R(x < 1 \vee x > 0)$, which is an assumption on $<$, is not true. The composite $s = LYC$ eventually gives us the embedding of **Man** into a topos $\mathcal{E}$ that has all desired properties:

$$
\begin{aligned}
s(\mathbb{R}) &= R, \\
s(\mathbb{R} - \{0\}) &= \{\text{invertible elements of } R\}, \\
s(f') &= s(f)', \\
s(TM) &= (sM)^{\Delta}.
\end{aligned}
\tag{2}
$$

This shows us not only that $\mathcal{E}$ closely resembles **Man**, but that $\mathcal{E}$ is a topos model for SIA as well. We will call this topos $\mathbb{S}$ from now on. From equation (2) we conclude that the object R

---

[8]For more information about presheaves and sheaves, cf. [15].
[9]Cf. [13]

of $\mathbb{S}$ is the image of the real numbers $\mathbb{R}$ under the functor s. This is how our real line R is given by the model, starting from standard differential geometry. The notion of functions in $\mathbb{S}$ is given by the images of functions between the coordinate rings under the functor LY.

4   Smooth infinitesimal analysis

Given the topos construction of the previous chapter, in this chapter we will derive some important results in SIA. We will do so by using the axioms, giving definitions and proving theorems with the rules of intuitionistic logic. The mathematical reasoning takes place within our topos $\mathbb{S}$, so we will not be bothered by the construction of $\mathbb{S}$ any more, but we will simply use anything we know about it. This chapter is based on [1].

4.1   Basics

We will now briefly describe the meaning of the axioms and some other concepts.

A very important object in any smooth world is the indefinitely extensible homogeneous straight line R, called the *smooth*, *real* or *affine line*. We have assumed that we are given the notion of a *location* or *point* in R, together with the relation "$=$" of *identity*, i.e. coincidence of locations. The identity relation is an equivalence relation. The notation $a \neq b$ now means that a and b are distinguishable. In smooth worlds, the relation "$=$" is not decidable, i.e. we can not always say if the relation $a = b$ is true or false. This is typical for intuitionistic logic.

We assume that $(R, +, \cdot, 0, 1)$ is a commutative ring with identity, where $0 \neq 1$. In what follows, we will, when writing down a product, drop the $\cdot$ and just write ab instead of $a \cdot b$.

Proposition 4.1. $0 \cdot a = 0 \; \forall a \in R$.

Proof. $0a = (0 + 0)a = 0a + 0a$ directly implies $0a = 0$. $\qquad\qquad\qquad\qquad\qquad\qquad$ □

We assume that for any two points a, b there is an entity $a \wedge b$, called the *oriented* $(a, b)-segment$ of R. We say that $a \wedge b$ and $c \wedge d$ are identical if and only if $a = c$ and $b = d$. Furthermore, $a* := 0 \wedge a$ is called the segment of length a.

We furthermore assume that for any two points a and b we can form a segment $a* : b*$, which is obtained by "connecting" the oriented segment of length b to the oriented segment of length a, in the given order and preserving orientation. As expected, we assume that $a* : b*$ is of the form $c*$ for some c on the real line. We call c the sum of a and b and write $c = a + b$. We also say $a - b = a + (-b)$.

The reason to give the line segment construction is that many of the results we will give and prove perfectly make sense in a geometrical way of thinking.

We can extend the theory to R^n, the n-fold Cartesian power of R. This is possible, since $\mathbb{S}$ is a topos, and hence Cartesian closed. A point x in R^n is denoted by the n-tuple $(x_1, ..., x_n)$ and $x = y$ if and only if $x_i = y_i$ for all $i = 1, ..., n$.

In R^n we can define so-called k-monads [9], which give us a notion of neighbours in R^n.

Definition 4.1. *The k-**monad** $D_k(n)$ is defined as*
$D_k(n) = \{(x_1, ..., x_n) \in R^n \mid \text{the product of any } k+1 \text{ of the } x_i\text{'s is zero.}\}$.

Definition 4.2. *We say that two points in R^n are k-**neighbours**, and write $x \sim_k y$, if $x - y \in D_k(n)$.*

The second axiom says that there is an ordering $<$ on R, which can be interpreted as follows: $a < b$ means that a is strictly to the left of b on R. We assume that:

1. $\neg(a < a)$;

2. $a < b$ and $b < c$ implies $a < c$ (transitivity);

3. $a < b$ implies $a + c < b + c$ for any c;

4. $a < b$ and $0 < c$ implies $ac < bc$;

5. either $0 < a$ or $a < 1$;

6. $a \neq b$ implies $a < b$ or $b < a$.

It is, in general, not true that for any two a and b in R the relation $(a < b \lor a = b \lor a > b)$ holds.

We say $a \leq b$ if and only if $\neg(b < a)$. The relation $a \geq b$ is defined in the same manner. We define the open interval (a, b) as the collection of points in R satisfying both $a < x$ and $x < b$ and we define the closed interval $[a, b]$ as the collection of points in R satisfying both $a \leq x$ and $x \leq b$.

It is straightforward to prove some results about the R and the order relation. Let us give a typical example.

Proposition 4.2. $((a < 0)$ or $(0 < a))$ implies $(0 < a^2)$.

Proof. Assume $0 < a$. Then $0 \cdot a < a \cdot a$. Since $0 \cdot a = 0$, we have $0 < a^2$. Assume $a < 0$. Then $a \cdot (-a) < 0 \cdot (-a)$. Then $-a^2 < 0$. So $0 < a^2$. Since both $0 < a$ and $a < 0$ imply $0 < a^2$, it is obvious that $((a < 0)$ or $(0 < a))$ implies $0 < a^2$. □

We define a part X of R to be *microstable* if for any x in X and any ε in Δ we have that $x + \varepsilon$ is in X. Microstable parts are objects in the topos $\mathbb{S}$.
We will soon see that all maps are continuous by the Kock-Lawvere axiom, since varying x infinitesimally also changes y infinitesimally.

The *graph* of a function is the collection of points in $R \times R$ of the form $(x, f(x))$. There is also a categorical definition of a graph [12]. In what follows, we will, however, do mathematics in $\mathbb{S}$ without using too much category theory and therefore we use the first definition of a graph.

For any two functions $f, g : J \longrightarrow R$ from a part J of R to R, we can define a new function $(f + g)(x)$ by $x \mapsto (f(x) + g(x))$ and another new function $(f \cdot g)(x)$ by $x \mapsto (f(x)g(x))$.

The axiom that makes smooth infinitesimal analysis different from standard analysis is the Kock-Lawvere axiom, which we will state shortly. First, however, we define what we mean by infinitesimals.

Definition 4.3. *An **infinitesimal** on R is any nilsquare element of R, i.e. $x^2 = 0$. We denote the collection of infinitesimals on R by $\Delta := \{x \in R \mid x^2 = 0\}$.*

Axiom 4.1. *(Kock-Lawvere) For any mapping $g : \Delta \longrightarrow R$ there exists a unique b in R such that for all ε in Δ we have $g(\varepsilon) = g(0) + b\varepsilon$.*

From this, it follows that all curves that come from a mapping in $\mathbb{S}$ satisfy the principle of *microstraightness*, i.e. for any smooth curve and for any point P on it, there is a nondegenerate segment in C, called a microsegment, around P which is straight (microstraight) around P, where nondegenerate means that the segment is not a point.

We are now in a position to state a substantial theorem.

Theorem 4.1. *All functions in $\mathbb{S}$ are continuous in the following sense: if a and b in R are such that $a - b$ is in Δ, then $f(a) - f(b) \in \Delta$.*

Later on, this theorem will follow from the fact that every function is differentiable. For that, we do not need a function to be continuous.

The following theorem guarantees the existence of non-zero infinitesimals.

Theorem 4.2. $\Delta \neq \{0\}$.

Proof. Suppose $\Delta = \{0\}$. Consider the mapping $f : \Delta \longrightarrow R$ defined by $x \mapsto x$. That this mapping is in $\mathbb{S}$ follows from the fact that for $J \subset M$ the identity injection $f : J \longrightarrow M$ is in **Man**. Then $g(\varepsilon) = g(0) + \varepsilon b$ for $b = 0$, since $g(\varepsilon) = g(0) = 0 = 0 + \varepsilon 0$, but also for $b = 1$, since $g(\varepsilon) = g(0) + \varepsilon = \varepsilon$. So this b is not unique (because $0 \neq 1$), which contradicts the Kock-Lawvere axiom. So Δ cannot coincide with $\{0\}$. $\square$

It is important to notice that this proof does not make use of the rule of double negation, $(\neg\neg A) \to A$, since we want to prove that something is not true and suppose it is true. If we wanted to prove that something is true and would suppose that it is not, then we would make use of the rule of double negation.

The following property is called the Principle of Microcancellation.

Proposition 4.3. $\forall a, b \in R$, *if* $\varepsilon a = \varepsilon b \; \forall \varepsilon \in \Delta$, *then* $a = b$.

Proof. Suppose that $\forall a, b \in R$, if $\varepsilon a = \varepsilon b \; \forall \varepsilon \in \Delta$ and consider the mapping $f : \Delta \longrightarrow R$ defined by $x \mapsto xa$. Hence $g(\varepsilon) = g(0) + \varepsilon a = \varepsilon a$, which is by assumption equal to εb. By the Kock-Lawvere axiom, $a = b$, which directly proves the claim. $\square$

Corollary 4.1. *If* $\varepsilon a = 0 \; \forall \varepsilon \in \Delta$, *then* $a = 0$.

So this Δ contains more points than only 0. There is a very elegant geometric illustration of Δ given by Joyal. It says that if you consider the unit circle in R^2 centered at the point $(0, 1)$, then the collection of points where it is tangent to the x-axis is precisely Δ. To see this, we remember that the given circle is parametrized by $x^2 + (y - 1)^2 = 1$. On the x-axis, we have $y = 0$. So the points at which the given circle is tangent to the x-axis satisfy $x^2 = 0$, which are precisely the infinitesimals on R. This illustrates that tangency is geometrically not something that happens only at a point. This example is another indication that our way of thinking may be useful.

4.2 Calculus

4.2.1 Differential calculus

For the definition of differentiability, we need a mapping to be defined on a microstable domain, so we will always assume that this is the case, unless stated otherwise.

With the concept of a function and the Kock-Lawvere axiom in mind, we are now ready to define what it means for a function to be differentiable. Let $f : J \longrightarrow R$ be any function from a microstable part J of R to R. If we fix x, we can define a function $g_x : \Delta \longrightarrow R$ by $g_x(\varepsilon) = f(x + \varepsilon)$. This is a well-defined function, since it is defined for all ε in Δ, because J is microstable in R.

As before, let R be the real line in $\mathbb{S}$, $\Delta = \{x \in R \mid x^2 = 0\}$ and let $f : R \longrightarrow R$ be a function with domain R and codomain R. Let $g_x : \Delta \longrightarrow R$ denote the function $g_x(\varepsilon) = f(x + \varepsilon)$.

18

From the Kock-Lawvere axiom, we conclude that there exists a unique b in R —denoted b_x—, such that for all ε in Δ:

$$f(x + \varepsilon) = g_x(\varepsilon) = g_x(0) + b_x \varepsilon = f(x) + b_x \varepsilon. \tag{3}$$

The rule defined by $x \mapsto b_x$ is a well-defined function, denoted f' and called the derivative of f. This notion can be defined for any microstable part of R. With this definition in mind, equation (3) becomes:

$$f(x + \varepsilon) = f(x) + \varepsilon f'(x). \tag{4}$$

This is the fundamental equation of differential calculus in $\mathbb{S}$. In $\mathbb{S}$ every function from R to R has derivatives up to any order, which means that the process of taking derivatives can be repeated arbitrarily often.

Corollary 4.2. *All functions in $\mathbb{S}$ are continuous, i.e. if a and b in R are such that $a - b$ is in Δ, then $f(a) - f(b) \in \Delta$.*

Proof. Suppose we have a and b in R are such that $a - b$ is in Δ. Then there exists an $\varepsilon \in \Delta$ such that $b = a + \varepsilon$. By the differentiability of an arbitrary function f from R to R we now have $f(b) - f(a) = f(a + \varepsilon) - f(a) = \varepsilon f'(a)$, which clearly is an infinitesimal. $\square$

The derivative has some familiar arithmetic properties:

Proposition 4.4. *Let J be a microstable part of R. For any two functions $f, g : J \longrightarrow R$ and for any c,d in R we have:*

$$(cf + dg)' = cf' + dg'$$

Proof. By definition we have

$$(cf + dg)(x + \varepsilon) = (cf + dg)(x) + \varepsilon(cf + dg)'(x) = cf(x) + dg(x) + \varepsilon(cf + dg)'(x),$$

and

$$cf(x + \varepsilon) + dg(x + \varepsilon) = cf(x) + \varepsilon cf'(x) + dg(x) + \varepsilon dg'(x).$$

Since $(cf + dg)(x + \varepsilon) = cf(x + \varepsilon) + dg(x + \varepsilon)$, we have:

$$(cf + dg)' = cf' + dg'.$$

$\square$

We will call this property the *linearity* of the derivative. In the same way, we can prove the so-called Leibniz or product rule:

Proposition 4.5. *Let J be a microstable part of R. For any two functions $f, g : J \longrightarrow R$ we have:*

$$(fg)' = f'g + fg'.$$

Corollary 4.3. *Let J be a microstable part of R. For any two functions $f, g : J \longrightarrow R$ such that $g(x) \neq 0$ for all x in R, we have:*

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}.$$

19

Another important rule for differentiation is the *chain* rule:

Proposition 4.6. *Let I and J be microstable parts of R, let $f : J \longrightarrow R$ and $g : I \longrightarrow J$. Then we have:*

$$(f \circ g)' = (f' \circ g)g'$$

For proofs of these propositions, cf. [1].

4.2.2 Integral calculus

To be able to integrate functions in $\mathbb{S}$, we adopt an integration axiom.

Axiom 4.2. *Let $f : [0,1] \longrightarrow R$ be a function. Then there exists an unique $g : [0,1] \longrightarrow R$ with $g' = f$ and $g(0) = 0$.*

This turns out to make integration rather easy. We will often write $\int_0^x f$ for $g(x)$, as we are used to from standard analysis. It is clear that if we integrate f', then we have $\int_0^x f' = f(x) - f(0)$.

It is straightforward to prove the following rules.

Proposition 4.7. $\int_0^x (cf + dg) = c \int_0^x f + d \int_0^x g$ *and* $\int_0^x f'g = fg|_0^x - \int_0^x fg'$.

By rescaling we can extend the theory of integration to arbitrary intervals in R.

This notion of integration is a bit surprising, since it is given by an axiom. This is also the way in which it is done in the texts we considered about it. Note that in ordinary analysis, first the definition of the Riemann or Lebesgue integral is given. Then, what we stated as an axiom above, is in fact a provable statement for continuous functions.

4.2.3 Minima and maxima

A very important application of differential calculus is determining *extrema*: *minima* and *maxima*.

Definition 4.4. *We say that a function $f : J \longrightarrow R$ has an **extremum** in a if $f(a + \varepsilon) = f(a)$ for all $\varepsilon \in \Delta$.*

Theorem 4.3. *A function f has an extremum in a if and only if $f'(a) = 0$.*

Proof. Suppose f has an extremum in a. Then for all $\varepsilon \in \Delta$, $f(a + \varepsilon) = f(a)$. Hence for all ε we have $f(a) = f(a + \varepsilon) = f(a) + \varepsilon f'(a)$, from which it follows that $f'(a) = 0$.

Suppose $f'(a) = 0$. Then for all ε we have the relation $f(a + \varepsilon) = f(a) + \varepsilon f'(a) = f(a)$. $\square$

This is a very important result, since the concepts of minima and maxima correspond to the notion of these concepts in standard analysis.

5 Synthetic Differential Geometry

With the results of SIA in mind, we are now ready to interpret differential geometry in this setting. The resulting theory is called synthetic differential geometry (SDG). The way of reasoning in SDG is very intuitive, so this may be a very useful mathematical framework for the formulation of physical theories.

5.1 Basic notions

If we want to study differential geometry, we first need a definition of the spaces we are working with, i.e. the analogues of manifolds in the classical theory. The explicit analogue of a (standard) manifold in SDG is given by the image of this manifold under the functor $s : \mathbf{Man} \longrightarrow \mathbb{S}$ given in chapter 3. Basically, a manifold M is replaced by its coordinate ring under this functor. So the construction of a model provides us with the notion of a manifold. This is all we need, since the objects we will consider in relativity theory are manifolds in standard differential geometry. We will just talk about manifolds in $\mathbb{S}$, but what we consider are actually images of ordinary manifolds in $\mathbf{Man}$ under the functor s. For a very thorough text about the identification of manifolds with objects in $\mathbb{S}$, cf. [9] or [14].

There are, however, also direct definitions of manifold(-like) objects in $\mathbb{S}$, but these are very special and technical, since they use a lot of category theory. For more information, cf. [11].

One of the main concepts of standard differential geometry is tangency.

Definition 5.1. *A **tangent** vector to M at a point m is a map $t : \Delta \longrightarrow M$ with $t(0) = x$.*

The image of a tangent vector lies in M. Note that tangent vectors always exist, since the objects of study are smooth.

There is a rather subtle, but important difference between standard and synthetic differential geometry concerning tangency. In the standard formulation, a tangent vector at a point $m \in M$ is given by an equivalence class of arbitrarily short short paths $t : (-\alpha, \alpha) \longrightarrow M$. In that case, the tangent vectors are paths with a certain length (2α), but in the synthetic case, the tangent vectors have "length" 0, in the sense of the notion of metrics defined below.

A related notion is the tangent bundle.

Definition 5.2. *There is a map $\pi : M^{\Delta} \longrightarrow M$ from the collection M^{Δ}, consisting of all tangent vectors to M, to M, defined by sending each tangent vector to its base point, $\pi(t) = t(0)$. The object (M^{Δ}, π) is called the **tangent bundle** of M.*

For our purposes however, the most important definition may be the one of a tangent space.

Definition 5.3. *The **fibre** over $m \in M$, i.e. the set of tangent vectors with m as its base point, is called the **tangent space** to M at m and is denoted $(M^{\Delta})_m$ or $T_m M$.*

According to Bell [2], the tangent space can be regarded as "locally lying" in M. This is a very important difference with standard differential geometry, in which this is clealy not the case. There, in general the intersection of a manifold and the tangent space is the point at which the tangent space is considered. We do not go into this, since we do not really need this, but we wanted to give this contrasting result.

5.2   Metrics

In standard differential geometry, a very important concept is the metric. As will be explained later, the metric is also very important in relativity theory. We will therefore define this concept in SDG.

To do so, we extend the theory of k-monads, which is given in chapter 4. Since any n-dimensional manifold in SDG is locally isomorphic to R^n, we can extend the relation $\sim_k$, which is given in chapter 4, to manifolds, by defining it on M via the atlas of charts $\{f_i : U_i \longrightarrow M\}$. This turns out to be independent on the choice of the chart. We will not give this construction explicitly, but for a thorough description cf. [10].

Definition 5.4. *For each $k \geq 1$, $M_{[k]}$ is the collection of pairs $(x, y) \in M$ such that $x \sim_k y$.*

So, $M_{[k]}$ is a (generalized) relation on M. An ordered pair is in the relation if the components of the ordered pair are k-neighbours on the manifold M.

We follow Kock [10] in defining the metric via a quadratic differential form.

Definition 5.5. *A **quadratic differential form** g on a manifold M is a mapping $g : M_{[2]} \longrightarrow R$ that vanishes on $M_{[1]}$.*

Such a mapping turns out to be symmetric if $x \sim_2 y$. We can now define the notion of a metric.

Definition 5.6. *A **(pseudo-Riemannian) metric** is a nondegenerate[10] quadratic form on M.*

Note that the vanishing condition on the quadratic differential form establishes that the distance between two points "which have an infinitesimal distance to each other" is zero, which is not very surprising.

A (global) metric g satisfies the following usual conditions:

1. g is nondegenerate (by definition).

2. g is symmetric.

3. g satisfies the triangle inequality.

A metric in SDG gives us a notion of distance, in just the same way as in standard differential geometry.

There are other ways to define metrics in SDG. For example, we can define inner products and norms and say how an inner product induces a norm and how a norm in turn defines a metric. We did not follow this way of reasoning here, because Kock's procedure is more general. We also do not need inner products and norms, and in general, they are not assumed to exist on manifolds.

With these concepts of synthetic differential geometry in mind, we can at last focus on the theory of general relativity.

---

[10]This is a condition that arises as a technical consequence of the generalization of k-neighbours to manifolds.

6 Mechanics

In this chapter, we give a brief overview of the physical theory of mechanics and its history. We start with Newtonian mechanics and end up with general relativity. We do not discuss quantum theory.

6.1 Classical Mechanics

Until 1905, the theory of mechanics that was used in all areas of physics, was classical mechanics, which is today considered as the mechanical theory of macroscopic objects moving at low velocities. The aim of classical mechanics is to calculate the trajectory of a moving object. The subject started in 1687 with the publication of Isaac Newton's book. In *Philosophiae Naturalis Principia Mathematica*, he describes two important theories: the motion of bodies and gravitation. He formulates his three laws of mechanics and the law of universal gravity, which in modern forms read as follows[5]:

1st law Free particles move at constant velocity.

2nd law The force on a particle is proportional to the acceleration of the particle with proportionality constant m, i.e. $\vec{F} = m\vec{a}$.

3rd law The forces of action and reaction are equal in magnitude and opposite in direction.

Law of Universal Gravity A point mass A attracts a pointmass B in A's direction. The force that A acts on B is given by

$$\vec{F}_{AB} = -G\frac{m_A m_B}{r^2}\hat{r}, \tag{5}$$

where F is the force, G is Newton's universal constant of gravity, m_A is the mass of A, m_B is the mass of B, r is the distance between A and B and $\hat{r}$ is the unit vector directed from A to B.

Cf. [16] for more information.

In 1788, there was a reformulation of Newtonian mechanics by Joseph Louis Lagrange. It relies on two important results of Newtonian mechanics, which both do not appear in Newton's own work: conservation of momentum and conservation of energy. The method relies on the fact that one can choose independent generalized coordinates such that the results found are valid for each well-defined coordinate system. The idea is to solve the so-called Euler-Lagrange equations,

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = Q', \tag{6}$$

where $\mathcal{L} := T - V$, the co-called Lagrangian, is defined as the kinetic energy minus the potential energy, q is a generalized coordinate and Q' is a term which corresponds to the non-conservative[11] interactions within the system. The system of equations needs to be solved for each of the generalized coordinates to get a set of equations which describes the particle trajectory.

A reformulation of Lagrangian mechanics was given in 1833 by William Rowan Hamilton. He used the fact that in the Lagrangian method we have, say, n second order differential equations

---

[11]An interaction is non-conservative if it cannot be written as the gradient of a scalar.

with constraints. According to Hamilton, this is equivalent to $2n$ first order differential equations with constraints. These equations are called the Hamilton equations:

$$\dot{q} = \frac{\partial \mathcal{H}}{\partial p},$$
$$\dot{p} = -\frac{\partial \mathcal{H}}{\partial q}, \tag{7}$$

where $\mathcal{H} := T + V$ is the so-called Hamiltonian and p is the generalized impuls conjugate to q.

Hamilton's theory can be formulated completely geometrically in terms of symplectic spaces, but this goes beyond the scope of this thesis.

A very important Newtonian concept in classical mechanics is absolute space and time, which says that space can be regarded as an inert scene on which physical events take place and that time is the same for all observers in space. This is by no means evident, and has been rejected by Einstein's work, as will be explained below. It has even been rejected by Leibniz before.

For a comprehensive text about classical mechanics, cf. [5].

6.2 Special Theory of Relativity

In 1905, the special theory of relativity (SR) was published by Albert Einstein, which forced physicists to not only use classical mechanics and classical field theory[12]. Classical mechanics was (and is) still a good theory for describing macroscopic objects at low velocities —compared to the speed of light—, but at higher velocities, one needs to use special relativity. The formulation of this theory by Einstein relies heavily on the work of Poincaré and Lorentz.

The two postulates of SR are:

1. All inertial frames are equivalent for the performance of all physical experiments.

2. Light travels rectilinearly at speed c in all directions in all inertial frames.

Both postulates clearly break with the tradition of absolute space and time, and this may be the most important assertion of SR. Space and time together form a continuum which we will call spacetime. From the postulates we can derive transformations that transform a coordinate vector $(x, y, z, t)^T$ in an inertial frame S to a coordinate vector $(x', y', z', t')^T$ in another inertial frame S' moving with a velocity $\vec{v}$ with respect to S. If S' moves with a velocity $\vec{v}$ in the x-direction and if at $t = t' = 0$ the origin of both inertial frames cross, then the coordinate transformation is given by:

$$x' = \gamma x - \beta \gamma ct,$$
$$y' = y,$$
$$z' = z, \tag{8}$$
$$t' = \gamma t - \beta \gamma \frac{x}{c}.$$

Such a transformation is called a Lorentz transformation. Here $\beta = \frac{v}{c}$ and $\gamma = \frac{1}{\sqrt{1-\beta^2}} = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$.

---

[12]Classical mechanics and classical field theory together formed physics before 1905.

From the Lorentz transformation we obtain the astonishing facts that moving objects "become shorter" in the direction of their velocity and that moving clocks "run more slowly". Together with the fact that an object cannot accelerate to a velocity which is greater than the speed of light and that mass and energy are equivalent as in the following equation:

$$E = \gamma m c^2, \tag{9}$$

where E is the energy, m the mass, c the velocity of and γ is defined as above, these are the most important results of SR.

6.3 General Theory of Relativity

In our description of SR we only discussed inertial frames of reference, but what if the frames accelerate and what if there is gravity, both of which make frames non-inertial?

Einstein himself generalized his special theory of relativity to the general theory of relativity (GR), which deals with accelerating systems and systems in which gravity plays a role, and published it in 1915 and 1916. GR is a geometrical theory of spacetime, in which gravitation is not an action-at-distance interaction, but a property of the geometry of spacetime. The mathematical framework in which this theory is formulated is differential geometry. This is the main reason why the mathematical concepts of the theory are not familiar to all physicists, since this framework is quite difficult and technical.

The local geometry of spacetime is described by the **metric tensor** $g_{\mu\nu}$. The metric tensor on a manifold M is a covariant, symmetric and nondegenerate tensor of rank two. This is equivalent to saying that the **metric** $g = g_{\mu\nu} dx^\mu \otimes dx^\nu$ defines an inner product (not necessarily positive definite) on the tangent space of each point. The metric tensor is transformed to its contravariant form by the following transformation:

$$g^{\alpha\nu} g_{\nu\beta} = \delta^\alpha_\beta. \tag{10}$$

The metric can be regarded as the description of the (local) geometric and causal structure of spacetime. It gives a complete local general relativistic description of it.

In GR we can of course also calculate trajectories of particles, but we will not describe how this is done, since this is rather technical and we do not use it for what follows. A good book about relativity (for physicists) is [19]; for mathematicians we would recommend [23], but the latter does not deal with SR.

In the next chapter, we will take a closer look at the foundations of general relativity.

7 Einstein's Equivalence Principle

7.1 Foundations of General Relativity

In 1918, after the publication of the general theory of relativity (GR) in 1916, Einstein again discussed the foundation of his own theory. He listed the following three principles [17]:

1. **Principle of relativity**: All systems of reference are equivalent with respect to the formulation of the fundamental laws of physics; therefore they find their unique, natural expression in generally covariant equations.

2. **Principle of equivalence**: Inertia and weight are identical in essence. From this and from the results of the special theory of relativity, it follows necessarily that the symmetric "fundamental tensor" $g_{\mu\nu}$ determines the metric properties of space, the inertial relations of bodies in it, as well as gravitational effects. We will call the condition of space, described by the fundamental tensor, the G-field.

3. **Mach's principle**: The G-field is determined entirely by the masses of bodies. Since mass and energy are equivalent according to the results of the special theory of relativity and since the energy is described formally by the symmetric energy tensor $T_{\mu\nu}$, this means that the G-field is conditioned and determined by the energy tensor.

Although the foundations of GR are still an interesting topic of discussion, we will not go into all three of them. For a very thorough view on these three principles, and the differences between them, cf. [17]. We will only discuss the principle of equivalence (EP), and of that only its mathematical part, rather than the physical and philosophical aspects. For a thorough discussion of the EP as a whole, cf. [18].

7.2 Equivalence Principle: standard formulation

The first problem that arises is that the equivalence principle formulated by Einstein is simply not true. Because of so-called tidal forces, one would be able to decide if a force is gravitational or caused by acceleration, since gravitational forces are in general not uniform [18]. Only if it is uniform, then it would be indistinguishable from acceleration.

There have been various attempts to formulate the principle infinitesimally, starting with Pauli's formulation [18]. They say:

Spacetime is locally flat.

It would be nice to have a mathematical formulation of the EP in terms of (standard) differential geometry[13], since GR has been formulated in this mathematical framework.

In terms of the metric, the mathematical content of the EP is as follows [18]: at a point p of

---

[13]Or better, semi-Riemannian geometry, because the metric is not positive definite. If you are not familiar with (standard) differential geometry, cf. e.g. [22].

spacetime[14] M we can transform to normal coordinates [23] such that in general we have:

$$
\begin{aligned}
g_{\mu\nu}(p) &= \eta_{\mu\nu}(p) = \eta_{\mu\nu}, \\
g_{\mu\nu,\sigma}(p) &= 0, \\
\Gamma^{\sigma}{}_{\mu\nu}(p) &= 0, \\
g_{\mu\nu,\sigma\tau}(p) &\neq 0,
\end{aligned}
\tag{11}
$$

where $\eta_{\mu\nu} = diag(1,-1,-1,-1)$ is the Lorentzian metric (the metric of the special theory of relativity) and an index after a comma denotes differentiation with respect to this index.

The fact that the Christoffel symbols vanish at p implies that the geodesics through p are locally linear.

Physicists often use the principle of equivalence as if the equations (11) holds in a neighbourhood of p, but that is simply not true. The maximal mathematical conclusion that can be drawn is that spacetime is locally Lorentzian any point p of spacetime, where "locally" is interpreted as "at the point p".

7.3   Equivalence Principle: topos formulation

As defined in chapter 5, there is a notion of metric in SDG. Such a metric g is interpreted in the same way as the metric in standard differential geometry: it is responsible for the local geometric and causal structure of spacetime. It can also be written in terms of the metric tensor as in standard differential geometry:

$$
g = g_{\mu\nu} dx^{\mu} \otimes dx^{\nu}.
\tag{12}
$$

In our formulation, it is now clear that for a given point p in spacetime, we also have normal coordinates such that at the point p in the spacetime manifold, one has:

$$
\begin{aligned}
g_{\mu\nu}(p) &= \eta_{\mu\nu}(p), \\
g_{\mu\nu,\sigma}(p) &= 0, \\
\Gamma^{\sigma}{}_{\mu\nu}(p) &= 0, \\
g_{\mu\nu,\sigma\tau}(p) &\neq 0.
\end{aligned}
\tag{13}
$$

However, and this is the whole point, the first equation holds in more points, namely throughout $\mathcal{O}(p) = \{x \in M \mid x \sim_1 p\}$, for we know that any $x \in \mathcal{O}(p)$ can be written as $x = p + \varepsilon$, where $\varepsilon \in M_{[1]}$. Then we have:

$$
g_{\mu\nu}(x) = g_{\mu\nu}(p + \varepsilon) = g_{\mu\nu}(p) + \varepsilon^{\sigma} g_{\mu\nu,\sigma}(p) = \eta_{\mu\nu}.
\tag{14}
$$

As a consequence of this, also the second —and therefore the third– equation holds at $\mathcal{O}(p)$, since $\forall x \in \mathcal{O}(p)$ we have $g_{\mu\nu}(x) = \eta_{\mu\nu}$ and equations (13) hold. Our conclusion is that $\forall x \in \mathcal{O}(p)$:

$$
\begin{aligned}
g_{\mu\nu}(x) &= \eta_{\mu\nu}, \\
g_{\mu\nu,\sigma}(x) &= 0, \\
\Gamma^{\sigma}{}_{\mu\nu}(x) &= 0, \\
g_{\mu\nu,\sigma\tau}(x) &\neq 0.
\end{aligned}
\tag{15}
$$

---

[14]This is a manifold.

This is a nice result, for we have both reformulated the equivalence principle in SDG and we conclude that in our formulation the infinitesimal equivalence principle holds at more points than just one. A problem of our formulation, however, is that we are not able to really construct the neighbourhood $\mathcal{O}(p)$.

We think that this formulation is succesful. This contributes to the work that has already been done on describing metrics in general relativity.

8   Conclusion

As we have seen, Einstein's principle of equivalence can be formulated in synthetic differential geometry in the following way: for all points x in $\mathcal{O}(p) = \{x \in M \mid x \sim_1 p\}$, we have:

$$\begin{aligned}
g_{\mu\nu}(x) &= \eta_{\mu\nu}, \\
g_{\mu\nu,\sigma}(x) &= 0, \\
\Gamma^{\sigma}{}_{\mu\nu}(x) &= 0, \\
g_{\mu\nu,\sigma\tau}(x) &\neq 0.
\end{aligned} \tag{16}$$

We conclude that it is possible to formulate Einstein's equivalence principle in terms of synthetic differential geometry. We have even said more than the classical formulation, as our formulation holds in more points than one. However, if we add the law of excluded middle to our logic, we recover the classical formulation of the equivalence principle.

Formulating the equivalence principle in synthetic differential geometry gives us an example of using topos theory in describing physics. SDG is a natural and intuitive formalism for differential geometry. Thinking this way is easier than in the standard theory, because the arguments are typically geometric and quite simple. We think that applying topos theory may give very interesting new formulations of physics. It may even be a very succesful framework for quantum gravity, which is supported by the phenomenon that there are already a few mathematical physicists working in that way.

A Category Theory

This appendix is a short overview of basic category theory. We will introduce most main concepts and state some results about categories without proof. This text is not meant as an introduction to category theory, but it presents basic category in such a way that one is able to understand the sections about topoi, smooth infinitesimal analysis and synthetic differential geometry after reading this appendix. It is based on [13]. For a thorough introduction in Category Theory, cf. [4], [12] or [13].

There is much to be said about the foundations and history of category theory, but we will not go into that. For further reading, cf. [4] or [12].

Roughly speaking, category theory is a theory which describes mathematical structures in a very abstract way. It starts with the notion that many properties of mathematical structures can be described with diagrams, consisting of "objects" and "arrows".

A.1 Basics

A.1.1 Categories and objects

Before giving the definition of a category, we will define some other important concepts. A **category** $\mathcal{C}$ contains a class of **objects** $\mathrm{ob}\,(\mathcal{C})$ and a class of **arrows** or **morphisms** $\hom(\mathcal{C})$ between objects.

Definition A.1. *To a morphism f in a category, we assign an object $A = \mathrm{dom}\,(f)$ in that category, which is called the **domain** of f, and an object $B = \mathrm{cod}\,(f)$, which is called the **codomain** of f.*

Usually, we will denote a morphism with domain A and codomain B by

$$f : A \longrightarrow B. \tag{17}$$

Definition A.2. *For any two arrows $f : A \longrightarrow B$ and $g : B \longrightarrow C$, such that $\mathrm{dom}\,(g) = \mathrm{cod}\,(f)$, the **composite** morphism $g \circ f : A \longrightarrow C$ is defined.*

The explicit "calculation" of this composite depends on the structure of the morphisms of the category.

Definition A.3. *An **identity** morphism for an object X is a morphism $1_X : X \longrightarrow X$ such that for every morphism $f : A \longrightarrow B$ we have $1_B \circ f = f = f \circ 1_A$.*

It is straightforward to prove that if there is an identity morphism for an object X, then it is unique.

Definition A.4. *A **category** $\mathcal{C}$ consists of a class of objects $\mathrm{ob}\,(\mathcal{C})$, a class of morphisms $\hom(\mathcal{C})$ between objects and a binary operation $\circ$ (composition) such that to every arrow in $\mathcal{C}$ we can assign a domain and a codomain, the composition is associative, i.e. $(h \circ g) \circ f = h \circ (g \circ f)$ (when defined) and for every object X of $\mathcal{C}$ there exists an identity morphism $1_X : X \longrightarrow X$.*

Because of the associativity of the composition, it is not necessary to place parentheses. If no confusion arises, we may drop the notation $\circ$ in a composition and just write gf for $g \circ f$. We will also use the words **map**, **mapping** and **function**[15] to denote morphisms.

There are many examples of categories. We will give two important ones. The first of them is of great importance for this text, since it is a topos[16].

Example A.1. *Set: The class of sets together with the usual functions defined between them form a category.*

Example A.2. *Grp: The class of groups together with the group homomorphisms between them form a category.*

We take a certain class of objects in these examples, because for example one cannot speak about the set of all sets. This is solved by the idea that there exist so-called small and large categories. We will not go into that. For more information, cf. [13].

There is a straightforward notion of a subcategory $\mathcal{D}$ of a category $\mathcal{C}$.

Definition A.5. *A **subcategory** $\mathcal{D}$ of $\mathcal{C}$ is a category such that $\mathrm{ob}\,(D) \subset \mathrm{ob}\,(C)$ and $\mathrm{hom}\,(D) \subset \mathrm{hom}\,(C)$.*

There is a useful way of denoting (parts of) categories as diagrams. These diagrams have vertices and arrows between the vertices. The vertices correspond to objects and the (diagram) arrows to morphisms. We will say that a diagram is commutative if every "arrow path" from an object to another yields the same mapping.

Example A.3. *The identity law, i.e. for every morphism $f : A \longrightarrow B$ and for every morphism $g : B \longrightarrow C$ we have $1_B f = f$ and $g 1_B = g$, can be pictured in a commutative diagram in the following way:*

$$
\begin{array}{ccc}
A & \xrightarrow{\ f\ } & B \\
 & \searrow^{f} \quad \downarrow^{1_B} \quad \searrow^{g} & \\
 & B \xrightarrow{\ g\ } & C
\end{array}
$$

A.1.2 Functors

Definition A.6. *For categories $\mathcal{B}$ and $\mathcal{C}$, a **(covariant) functor** $T : \mathcal{B} \longrightarrow \mathcal{C}$ consists of two functions: the **object function**, which assigns to every object B in $\mathcal{B}$ an object TB in $\mathcal{C}$, and the **arrow function**, which assigns to each arrow $f : B \longrightarrow B'$ in $\mathcal{B}$ an arrow $Tf : TB \longrightarrow TB'$ in $\mathcal{C}$ such that $T(1_B) = 1_{TB}$ and $T(g \circ f) = Tg \circ Tf$ for any two morphisms f and g such that the composition gf is defined.*

A functor is actually a morphism of categories. The term *covariant* suggests that there exist *contravariant functors* as well:

---

[15]The word function will be used in particular in **Set**. A definition of this category will be given very soon.
[16]Cf. chapter 2 for a definition.

Definition A.7. *For categories $\mathcal{B}$ and $\mathcal{C}$, a **contravariant functor** $T : \mathcal{B} \longrightarrow \mathcal{C}$ consists of two functions: the **object function**, which assigns to every object B in $\mathcal{B}$ an object TB in $\mathcal{C}$, and the **arrow function**, which assigns to each arrow $f : B \longrightarrow B'$ in $\mathcal{B}$ an arrow $Tf : TB' \longrightarrow TB$ in $\mathcal{C}$ such that $T(1_B) = 1_{TB}$ and $T(g \circ f) = Tf \circ Tg$ for any two morphisms f and g such that the composition gf is defined.*

A contravariant functor reverses the direction of the arrows in the categories and the composition of morphisms is done in reversed order.

A.1.3 Natural transformations

Definition A.8. *If S and T are functors from $\mathcal{B}$ to $\mathcal{C}$, then a **natural transformation** $\tau : S \longrightarrow T$ assigns to every object B of $\mathcal{B}$ an arrow $\tau_B : SB \longrightarrow TB$ in $\mathcal{B}$ in such a way that every arrow f in $\mathcal{B}$ yields the following commutative diagram:*

$$
\begin{array}{ccc}
B & \quad SB \xrightarrow{\ \tau_B\ } TB \\
{\scriptstyle f}\Big\downarrow & \quad {\scriptstyle Sf}\Big\downarrow \qquad \Big\downarrow {\scriptstyle Tf} \\
B' & \quad SB' \xrightarrow[\ \tau_{B'}\ ]{} TB'
\end{array}
$$

We also say that $\tau_B : SB \longrightarrow TB$ is **natural** in B.

A.1.4 Properties of morphisms

Definition A.9. *A morphism $f : A \longrightarrow B$ in a category $\mathcal{C}$ is called **invertible** if there exists a morphism $f' : B \longrightarrow A$ in $\mathcal{C}$ such that $ff' = 1_B$ and $f'f = 1_A$. The map f' is then called an inverse of f in $\mathcal{C}$.*

It is straightforward to prove that if there exists an inverse of a morphism, then it is unique, so we will talk about the inverse of f from now on and denote it by f^{-1}.

If in definition A.9 one only has $ff' = 1_B$, then f' is called a **right inverse**. If, on the other hand, only $f'f = 1_A$, then f' is called a **left inverse**.

Definition A.10. *Two objects A and B are called **isomorphic** (notation: $A \cong B$) if there exists an invertible morphism (**isomorphism**) $h : A \longrightarrow B$.*

Definition A.11. *A morphism $f : A \longrightarrow B$ is called **left cancellable**, **monic** or a **monomorphism** if for any two arrows $g_1, g_2 : C \longrightarrow A$ the equality $fg_1 = fg_2$ implies $g_1 = g_2$.*

Definition A.12. *A morphism $f : A \longrightarrow B$ is called **right cancellable**, **epi** or an **epimorphism** if for any two arrows $g_1, g_2 : B \longrightarrow C$ the equality $g_1 f = g_2 f$ implies $g_1 = g_2$.*

In **Set**, monomorphisms and epimorphisms correspond to injective and surjective functions respectively.

Definition A.13. *A morphism $f : A \longrightarrow A$ for which $f^2 \equiv ff = f$ is called **idempotent**.*

Idempotent morphisms are categorical generalizations of projections.

Definition A.14. *An object A is called **initial** in $\mathcal{C}$ if for each object B in $\mathcal{C}$ there exists exactly one morphism from A to B.*

Definition A.15. *An object A is called **terminal** in $\mathcal{C}$ if for each object B in $\mathcal{C}$ there exists exactly one morphism from B to A.*

An object which is both initial and terminal is called a **null** or **zero** object. In **Set**, the empty set is the (only) initial object and each one-element set is a terminal object.

A.2   Duality

When in a given category all arrows are reversed, a new category is obtained. This property is called **categorical duality**. For our purposes, it is sufficient to note that the dual category $\mathcal{C}^{op}$ of a category $\mathcal{C}$ is obtained by replacing a morphism $f : A \longrightarrow B$ by a morphism $f : B \longrightarrow A$, i.e. $A = \mathrm{dom}\,(f)$ is replaced by $A = \mathrm{cod}\,(f)$ and $B = \mathrm{cod}\,(f)$ by $B = \mathrm{dom}\,(f)$. In the dual category, the order of composition is reversed. Monic arrows become epi, left inverses become right inverses, initial objects become terminal and vice versa. It is important to mention that the dual of an axiom for categories is also an axiom. In proofs, it is therefore allowed to state dual statements without proofs if the original statements are true.

We can say more about duality of categories, but we cannot go into that now, because we have not yet introduced some very important concepts.

A.3   Universal properties

Definition A.16. *Let $S : \mathcal{D} \longrightarrow \mathcal{C}$ be a covariant functor and let C be an object of $\mathcal{C}$. A **universal arrow** from C to S is a pair (R, u) consisting of an object R of $\mathcal{D}$ and an arrow $u : C \longrightarrow SR$ in $\mathcal{C}$ such that to every pair (D, f) with D an object of $\mathcal{D}$ and $f : C \longrightarrow SD$ an arrow of $\mathcal{C}$, there is a unique arrow $f' : R \longrightarrow D$ of $\mathcal{D}$ such that $Sf' \circ u = f$, i.e. the following diagram commutes:*

$$
\begin{array}{ccccc}
C & \xrightarrow{\ u\ } & SR & \quad & R \\
 & & \big| & & \big| \\
 & {\scriptstyle f} \searrow & \big|\, {\scriptstyle Sf'} & & \big|\, {\scriptstyle f'} \\
 & & \downarrow & & \downarrow \\
 & & SD & & D
\end{array}
$$

We say that if this diagram is commutative, then the **universal property** is satisfied. Universal arrows are unique up to unique isomorphism. There is also a dual of a universal arrow. The definition is very straightforward by reversing all arrows. We do not give it explicitly, but we will give the definitions of the universal properties which bother us and also give their duals.

Important examples of universal arrows are limits and colimits. These are explained in the next section.

A.4   Limits and colimits

Definition A.17. *A **cone** for a diagram D consists of an object C in $\mathcal{C}$ together with an arrow $f_i : C \longrightarrow D_i$ for each object D_i in D such that, when g is an arrow in D, the following diagram commutes:*

$$D_i \xrightarrow{\quad g \quad} D_j$$
$$f_i \nwarrow \qquad \nearrow f_j$$
$$C$$
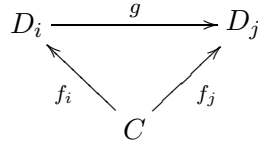
We will denote this cone by $\{f_i : C \longrightarrow D_i\}$.

Definition A.18. *A **limit** for a diagram D is a cone for D, $\{f_i : C \longrightarrow D_i\}$, such that for every other cone for D, $\{f_i' : C' \longrightarrow D_i\}$ there exists a unique arrow $f : C' \longrightarrow C$ such that for every object D_i the following diagram commutes:*

$$D_i$$
$$f_i' \nearrow \qquad \nwarrow f_i$$
$$C' \xrightarrow{\quad f \quad} C$$
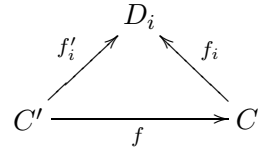
It is said that this cone has the **universal property**. Any other cone factors uniquely through it. We are now ready to use the duality property to define the dual of a limit.

Definition A.19. *A **cocone** for a diagram D consists of an object C in $\mathcal{C}$ together with an arrow $f_i : D_i \longrightarrow C$ for each object D_i in D such that, when g is an arrow in D, the following diagram commutes:*

$$D_i \xleftarrow{\quad g \quad} D_j \ .$$
$$f_i \searrow \qquad \swarrow f_j$$
$$C$$
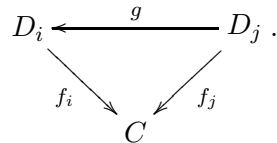
We will denote this cocone by $\{f_i : D_i \longrightarrow C\}$.

Definition A.20. *A **colimit** for a diagram D is a cocone for D, $\{f_i : D_i \longrightarrow C\}$ such that for every other cocone for D, $\{f_i' : D_i \longrightarrow C\}$ there exists a unique arrow $f : C \longrightarrow C'$ such that for every object D_i the following diagram commutes:*

$$D_i$$
$$f_i \swarrow \qquad \searrow f_j$$
$$C' \xrightarrow{\quad f \quad} C$$
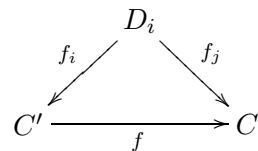
So the colimit is the dual of a limit. If we construct a dual category from a given category, then a limit turns into its corresponding colimit and vice versa.

In the following sections, we will give some examples of limits and colimits which are of great importance for the definition of a topos.

A.4.1 Products and coproducts

A very important example of a limit is the product.

Definition A.21. *In a category $\mathcal{C}$, a **product** of two objects A and B is a pair of mappings $A \xleftarrow{\pi_A} A \times B \xrightarrow{\pi_B} B$ such that $\forall X$, $f : X \longrightarrow A$, $g : X \longrightarrow B$, $\exists! h : X \longrightarrow A \times B$ such that $f = \pi_A h$ and $g = \pi_B h$.*
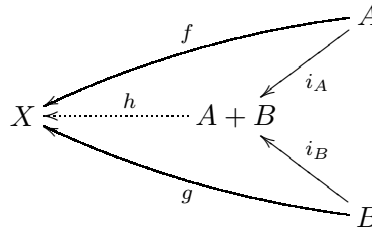
This definition is illustrated by the following commutative diagram:
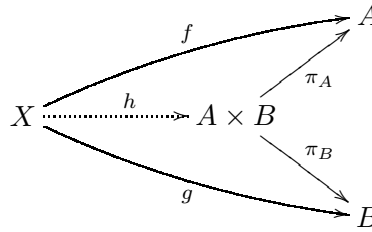
Dual to the product is the coproduct, which is a colimit.

Definition A.22. *In a category $\mathcal{C}$, a **coproduct** of two objects A and B is a pair of mappings $A \xrightarrow{i_A} A + B \xleftarrow{i_B} B$ such that $\forall X$, $f : A \longrightarrow X$, $g : B \longrightarrow X$, $\exists! h : A + B \longrightarrow X$ such that $f = h i_A$ and $g = h i_B$.*

This definition is illustrated by the following commutative diagram:

This construction is often regarded as the sum of the objects A and B.

Example A.4. *In the category **Grp**, the product of two groups G and H is the direct product of the groups and the coproduct is the free product of the groups.*

Example A.5. *In the category **Vec**, in which the objects are all vector spaces and the arrows are linear transformations of vector spaces, the product and the coproduct of two objects V and W is the same: $V \times W = V \oplus W$. Here the coproduct is denoted by $\oplus$. It is equal to the usual direct product or direct sum of vector spaces.*

A.4.2 Equalisers and coequalisers

Definition A.23. *For two arrows f and g from an object A to an object B, we define an **equaliser** for f and g as a map $i : E \longrightarrow A$ such that $fi = gi$ and $\forall x : T \longrightarrow A$ with $fx = gx$ we have $\exists! \bar{x}$ such that $x = i\bar{x}$.*

This is illustrated by the following diagram:

$$
\begin{array}{ccc}
 & T & \\
\bar{x} \swarrow & & \searrow x \\
E \xrightarrow{\quad i \quad} & A & \underset{g}{\overset{f}{\rightrightarrows}} B
\end{array}
$$

The arrow i is called an equaliser because composition with it equalises the morphisms f and g. The concept of a coequaliser is dual to the concept of an equaliser.

Definition A.24. *For two arrows f and g from an object A to an object B, we define a **co-equaliser** for f and g as a map $q : B \longrightarrow Q$ such that $qf = qg$ and $\forall y : B \longrightarrow T$ with $yf = yg$ we have $\exists! \bar{y}$ such that $y = \bar{y}q$.*

This is illustrated by the following diagram:

$$
\begin{array}{ccc}
A \underset{g}{\overset{f}{\rightrightarrows}} B & \xrightarrow{\quad q \quad} & Q \\
& \underset{y}{\searrow} \quad \swarrow \bar{y} & \\
& T &
\end{array}
$$

A.4.3 Pullbacks and pushouts

Definition A.25. *A **pullback** of the pair of morphisms $f : A \longrightarrow C$, $g : B \longrightarrow C$ consists of an object P and a pair of morphisms $p_1 : P \longrightarrow A$, $p_2 : P \longrightarrow B$, such that the diagram*

$$
\begin{array}{ccc}
P & \xrightarrow{\ p_2\ } & B \\
{\scriptstyle p_1}\downarrow & & \downarrow{\scriptstyle g} \\
A & \xrightarrow{\ f\ } & C
\end{array}
$$

is commutative. Moreover, for any other object T and pair of morphisms $t_1 : T \longrightarrow A$, $t_2 : T \longrightarrow B$ with this property, there is a unique morphism $u : T \longrightarrow P$ such that the following diagram is commutative:

$$
\begin{array}{ccc}
T & & \\
{\scriptstyle t_1}\Big\downarrow \quad \overset{u}{\searrow} & \overset{t_2}{\searrow} & \\
& P & \xrightarrow{\ p_2\ } B \\
& {\scriptstyle p_1}\Big\downarrow & \Big\downarrow{\scriptstyle g} \\
& A & \xrightarrow{\ f\ } C
\end{array}
$$

The dual of a pullback is called a pushout.

Definition A.26. *A **pushout** of the pair of morphisms $f : C \longrightarrow A$, $g : C \longrightarrow B$ consists of an object P and a pair of morphisms $i_1 : A \longrightarrow P$, $i_2 : B \longrightarrow P$ such that the diagram*

$$
\begin{array}{ccc}
C & \xrightarrow{\;g\;} & B \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle i_2} \\
A & \xrightarrow{\;i_1\;} & P
\end{array}
$$

is commutative. Moreover, for any other object T and pair of morphisms $j_1 : A \longrightarrow T$, $j_2 : B \longrightarrow T$ with this property, there is a unique morphism $u : P \longrightarrow T$ such that the following diagram is commutative:

$$
\begin{array}{ccc}
C & \xrightarrow{\;g\;} & B \\
{\scriptstyle f}\downarrow & & \downarrow{\scriptstyle i_2} \\
A & \xrightarrow{\;i_1\;} & P \quad\searrow{\scriptstyle j_2} \\
 & {\scriptstyle j_1}\searrow & \;\;\downarrow{\scriptstyle u} \\
 & & T
\end{array}
$$

Pullbacks and pushouts might seem quite abstract, but are fundamental concepts in category theory. In our context, they are particularly useful for showing that a category is a topos.

A.5 Exponentiation

In **Set** we denote by B^A the set of all morphisms from a set A to a set B. We will now describe a way to generalize this to arbitrary categories.

Definition A.27. *A category $\mathcal{C}$ has **exponentiation** if it has a product for each two objects and if for arbitrary objects A and B there is an object B^A and an arrow $ev : B^A \times A \longrightarrow B$ in $\mathcal{C}$ such that for any object C and each arrow $f : C \times A \longrightarrow B$ there is a unique arrow $\hat{f} : C \longrightarrow B^A$ such that the following diagram is commutative:*

$$
\begin{array}{ccc}
B^A \times A & & \\
\;\uparrow & \searrow{\scriptstyle ev} & \\
{\scriptstyle \hat{f}\times 1_A}\; & & P \\
\;\vdots & \nearrow{\scriptstyle f} & \\
C \times A & &
\end{array}
$$

The morphism ev is called **evaluation**.

A.6 Subobject classifiers

A subobject classifier is a categorical generalization of the classical set of truth values (e.g. $\{0, 1\}$ or $\{true, false\}$). Morphisms from an object to the subobject classifier correspond to parts of the object.

Definition A.28. *In a category $\mathcal{C}$ with a terminal object 1, an object Ω is called a* ***subobject classifier*** *if there exists a monomorphism $t : 1 \longrightarrow \Omega$ such that for each monomorphism $j : U \longrightarrow X$ there exists $\chi : X \longrightarrow \Omega$ such that the following diagram is a pullback:*

$$
\begin{array}{ccc}
U & \longrightarrow & 1 \\
{\scriptstyle j}\downarrow & & \downarrow{\scriptstyle t} \\
X & \xrightarrow{\;\chi\;} & \Omega
\end{array}
$$

Here, χ can be regarded as a generalized characteristic function.

We have now precisely defined all notions necessary to state the definition of an **elementary topos**. Cf. chapter 2.

References

[1] J.L. Bell, *A Primer of Infinitesimal Analysis*. Cambridge University Press, New York, 1998.

[2] J.L. Bell, *The Incredible Shrinking Manifold*. URL: http://publish.uwo.ca/∼ jbell.

[3] F. Borceux, *Handbook of Categorical Algebra*. Cambridge University Press, Cambridge, 1994.

[4] R. Goldblatt, *Topoi, The Categorial Analysis of Logic*. North-Holland Publishing Company, Amsterdam, 1979.

[5] H. Goldstein, C. Poole, J. Safko, *Classical Mechanics*. Addison-Wesley, San Fransisco, 2002.

[6] A.R. Hall, *Philosophers at War: The Quarrell between Newton and Leibniz*. Cambridge University Press, Cambridge, 2002.

[7] C. Heunen, N.P. Landsman, B. Spitters, *A Topos for Algebraic Quantum Theory*. arXiv:0709.4364.

[8] W. Hodges, *Model Theory*. Cambridge University Press, Cambridge, 1993.

[9] A. Kock, *Synthetic Differential Geometry*. Cambridge University Press, New York, 1981.

[10] A. Kock, *Geometric Construction of the Levi-Civita Parallelism*. Theory and Applications of Categories, **4**, 195-207 (1998).

[11] R. Lavendhomme, *Basic Concepts of Synthetic Differential Geometry*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[12] F.W. Lawvere, R. Rosebrugh, *Sets for Mathematics*. Cambridge University Press, Cambridge, 2003.

[13] S. Mac Lane, *Categories for the Working Mathematician*. Springer-Verlag, New York, 1969.

[14] I. Moerdijk and G.E. Reyes, *Models for Smooth Infinitesimal Analysis*. Springer-Verlag, New York, 1991.

[15] S. Mac Lane and I. Moerdijk, *Sheaves in Geometry and Logic*. Springer-Verlag, New York, 1992.

[16] Sir I. Newton, *The Principia: Mathematical Principles of Natural Philosophy*. University of California Press, Berkeley, 1999.

[17] J.D. Norton, *General Covariance and the Foundations of General Relativity: Eight Decades of Dispute*. Reports on Progress in Physics, **56**, 791-858 (1993).

[18] J.D. Norton, *What was Einstein's Principle of Equivalence?* Studies in History and Philosophy of Science, **16**, 203-246 (1987).

[19] W. Rindler, *Relativity*. Oxford University Press, Oxford, 2006.

[20] A. Robinson, *Non-standard Analysis*. North-Holland Publishing Company, Amsterdam, 1974.

[21] L. Smolin, *Three Roads to Quantum Gravity.* Basic Books, New York, 2001.

[22] M. Spivak, *A Comprehensive Introduction to Differential Geometry.* Publish or Perish, Houston, 1979.

[23] R.M. Wald, *General Relativity.* University of Chigaco Press, Chicago, 1984.