Changing Spaces

Can synthetic differential geometry offer an autonomous foundation for mathematics?

Daniel Wood

Supervisor: Dr Richard Pettigrew
Word count: approx. 13,800

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Master of Arts by advanced study in the Philosophy and History of Science in the Faculty of Arts

Department of Philosophy
Submitted 15th September 2010

Abstract

We will assess whether synthetic differential geometry can offer a foundation for mathematics, and furthermore whether such a foundation can be autonomous with respect to the current orthodox set-theoretic foundation. We will answer both questions in the affirmative, although we will see that the autonomy of the two foundations is not straightforward and that our conclusions lend credence to a form of mathematical pluralism.

To my parents, without whose love and support this dissertation could never have been written

Acknowledgements

I owe a great debt of gratitude to my supervisor, Richard Pettigrew, for his help and encouragement whilst writing this essay and for his suggestion that I undertake the MA for which it is written. I would also like to thank Josephine Salverda for her brilliant proofreading and invaluable comments. Of course, all errors are entirely my own.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Taught Postgraduate Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, this work is my own work. Work done in collaboration with, or with the assistance of others, is indicated as such. I have identified all material in this dissertation which is not my own work through appropriate referencing and acknowledgement. Where I have quoted from the work of others, I have included the source in the references/bibliography. Any views expressed in the dissertation are those of the author.

SIGNED: .... *DATE:* ....
(Signature of student/candidate)

Contents

List of Figures

Chapter 1

Introduction

The need for secure foundations for mathematics became a pressing issue at the turn of the 20^{th} century when various mathematical paradoxes where discovered, the most famous being Russell's paradox in 1901, although the Burali-Forti paradox was discovered earlier in 1897. Previously, mathematics had been considered by its very nature to be consistent, and any concerns one might have had with it were thought to be purely philosophical, not mathematical. Investigations into the foundations of mathematics had been largely undertaken by philosophers, or mathematicians wearing philosophical hats, who wished to explain the apparent absoluteness and a priori nature of mathematics, as well as other phenomena, such as the enormous power of its application.[1] But now that inconsistencies had been found within mathematics itself, mathematicians, qua *mathematics*, were compelled to enter this debate to ensure that their subject did not collapse due to logically unstable foundations. This influx of mathematical input took the discussion in a new direction. Whereas the questions previously being asked were of the sort 'What are mathematical objects?' and 'How can we have knowledge of mathematics?', mathematicians asked new questions: 'Which premises contradict each other?' and 'How can we couch mathematical propositions in an entirely rigorous setting?'[2]

The solution proposed to this foundational problem was axiomatic set theory, which has proven to be both a successful foundation for mathematics as well as a fruitful area of

[1] These philosophical enquiries into the metaphysics and epistemology of mathematics go back to at least Plato (see Chapters VII & VIII of the *Republic*, for example) and carry on today.

[2] This last question had in fact been asked earlier with regard to infinitesimals, which we shall discuss in §4.1.2.

mathematical research in itself. The details of this theory will be explained in Chapter 3, but the idea is that one takes the fundamental objects of mathematics to be sets and specifies a rigorous theory of these sets to ensure consistency and mathematical power. In this essay we shall examine a proposed alternative foundational theory, synthetic differential geometry (henceforth 'SDG').

Now, one might wonder why we would want to go to the effort of examining SDG as a foundation for mathematics, as we have one already in the form of set theory. Well, the value comes from the different concepts involved. Set theory is by its nature discrete. For example, in set theory, one defines the real numbers $\mathbb{R}$ as a set of discrete points.[3] But what if we wish to start from a different concept, say that of the continuous? Indeed, it has only been for the past hundred years or so that the idea of a continuum being made up entirely from points has been the orthodoxy in mathematics. Previously most mathematicians considered a continuum to be a cohesive whole which could not be made up (entirely) of points, since how can a continuum, an entity with extension, be made up from points, entities without extension? This cohesive view of the continuum is one of the key concepts upon which SDG is based. By investigating SDG as a foundation for mathematics, we will learn more about the role of conceptual bases in mathematics and thus more about mathematics generally. We will see that set theory is just one possible starting point in mathematics and that different approaches to the foundations of mathematics are fruitful both philosophically and mathematically.

I have two theses. My primary thesis is that SDG can provide a foundation for mathematics autonomous from that provided by set theory. We shall start by outlining criteria a theory must meet in order to provide an autonomous foundation for mathematics in Chapter 2, which will facilitate our subsequent discussion. We shall then describe and investigate set theory and SDG as foundations for mathematics in Chapters 3 and 4 respectively, and in Chapter 5 we shall look at the autonomy of SDG from set theory.

My secondary thesis is that my primary thesis lends credence to a form of mathematical pluralism; this will be the topic of Chapter 6.

Before we move on to the next chapter, we need to make two remarks. The first regards ontology. As we mentioned earlier, the question of the metaphysics of mathematics is an

[3] See chapter 9 of [11] for a detailed demonstration of the set-theoretic construction of the reals.

age-old one that dates back to at least the ancient Greeks. However, in this essay we shall deal with this question only in passing, not because it is uninteresting – far from it! – but simply because we do not have room to address it fully. So, when discussing a foundational theory, we shall consider things such as its mathematical power and the concepts behind it,[4] but we will not consider the metaphysics of the theory. So, for example, when we discuss set theory, we will consider the concepts underlying it and examine different notions of sets, but we shall avoid the debate between platonists, intuitionists, formalists, and other philosophical schools regarding the ontological nature of sets: whether sets exist in some mathematical "heaven" or are merely creations of humanity's intellect will not concern us. As such, we shall talk of *a* foundation *for* mathematics, rather than *the* foundation *of* mathematics, so to remain ontologically neutral (if such a thing is possible). I do think the metaphysical consequences of our conclusions, especially those of Chapter 6, would make a good topic of investigation, but we will have to leave that for another day.

Our second remark concerns prerequisites. The nature of the subject matter of this essay necessitates technical exposition. As such, knowledge of elementary mathematical logic and undergraduate algebra on the part of the reader is essential. Further knowledge of set theory, category theory, topology, and real analysis will very much benefit the reader's understanding of our discussion, but I shall endeavour to explain the concepts involved in such a way that the philosophically salient notions can (at least partially) be understood by those unfamiliar with these areas. I will also provide references to texts that cover the technical material in more detail and depth.

[4] We will completely bracket the question of the metaphysical nature of a concept.

Chapter 2

Foundations for mathematics

In this chapter we shall outline two sets of criteria. The first will consist of conditions for a theory to provide a foundation for mathematics, which we shall present in §2.1, and the second will consist of conditions for a theory to provide an *autonomous* foundation for mathematics, which we shall present in §2.2. This second set of criteria is taken from [20]. These sets of criteria will provide a framework for our later discussion of set theory and SDG as autonomous foundations for mathematics.

Before we embark on the main topic of this chapter, a couple of remarks are in order. The first regards the nature of a theory. In the following two sections, we shall make generalised references to theories. But what do we mean by a *theory*? Since we are discussing foundations, we cannot simply restrict our attention to formal theories, such as Peano Arithmetic or the theory of complete ordered fields, since the definition of a formula in a formal theory requires a recursive definition, which in turn requires a foundation – you can't build a house from the first floor up. Thus we must deal with informal theories, although such an informal theory may have a formal component. We we will also allow a theory to have a philosophical component, say in the form of a conceptual basis. We also of course want a theory to be consistent.[1] So, a rough characterisation is this: by a *theory*, we shall mean a collection of consistent mathematical and philosophical propositions, some possibly couched in a formal way. We will not attempt to make this definition any more precise, which brings us to our second remark.

[1] In light of Gödel's second incompleteness theorem, (knowledge of) the consistency of (the formal component of) a theory is a non-trivial matter, but for the sake of brevity we shall bracket this issue.

The investigation of what it takes for a theory to provide a foundation for mathematics is a difficult one in which subtleties abound. Indeed, I am sure that an entire essay, if not a whole book, could fill its pages on this task alone. For this very reason, we shall not attempt to formulate such a general theory. Instead, we shall outline criteria in a level of detail that is sufficiently precise for our purposes, namely to analyse set theory and SDG as foundations for mathematics. Thus, while some of the notions that we employ would be quite inadequate for general application, such as our definition of *theory* in the previous paragraph, they will suffice for the two examples that we shall consider.

2.1 Criteria for a foundation

Perhaps the most obvious requirement for a theory to provide a foundation for mathematics is that of *technical strength*: in a proposed foundational theory we must be able to do most (if not all) of mathematics. For what good would a theory be to the differential geometer if it could carry out only elementary arithmetic?

The next criterion is that the theory have an *ontological or conceptual justification*.[2,3] As we mentioned in the previous chapter, we wish to avoid ontological questions, but a theory should make existential assertions.[4] The informal nature of this criterion makes it difficult to pin down, but the idea is that the theory should have some sort of reasonable and coherent ontological or conceptual basis – it must be talking about (possibly very abstract) "stuff".[5] A theory that only classifies entities, such as group theory or category theory,[6,7] cannot provide a foundation for mathematics.[8]

We will make two remarks at this point. The first regards uniqueness. It may be the case that an ontological or conceptual justification is not unique to a theory; that is, two

[2] I have chosen the term *ontological or conceptual justification*, rather than just *ontological justification*, in order to adhere to our ongoing abstinence from matters metaphysical.

[3] A small note of pedantry: Unless otherwise specified, throughout this essay the word 'or' is used in the inclusive sense.

[4] See Part I of [9], p. 2 of [20], p. 24 of [21], and §4 of [24].

[5] I use the word 'reasonable' to exclude dubious ontological or conceptual bases. For example, I would not deem a theory based on unicorns, lovely as they are, to provide a reasonable basis for a foundation of mathematics.

[6] We shall describe category theory in Chapter 4.

[7] I'm treading on thin ice here. By category theory I mean simply that, and not some specific category, such as ETCS or CCAF; see the references to [20] and [24] in footnote 4 above.

[8] c.f. Mayberry's distinction between *classificatory* and *eliminative* theories in [21].

different theories T_1 and T_2 may have the same conceptual or ontological basis.[9] In such a case, T_1 and T_2 are still justified. Such underdetermination of an ontological or conceptual basis may be the result of different ways of proceeding from the same basis.

Our second remark regards attempts to justify a theory. There is of course a difference between an attempt to justify a theory and a *completed* attempt to justify a theory. Obviously the latter is to be aspired to, but we should bear this distinction in mind, since our attempt to justify SDG will not be entirely complete. The degree to which an attempted justification fails to be complete is of course the crucial factor in assessing its justificatory merit, but the lack of a completed justification does not necessarily imply that the theory is not conceptually justified.

Our last criterion is that of *naturality*, which breaks down into two subcriteria: *parsimony* and *pathology*. A proposed foundational theory T should be parsimonious in that it should not contain unnecessary premises. For example, if T is a theory of the mathematical-logical kind, then parsimony would mean that its axioms are logically independent of each other, i.e. in T one cannot prove any axiom of T from any of the other axioms of T. This comes with a caveat: some axioms might not be independent of the other axioms but are still permitted for reasons of presentation or necessity of definition. The subcriterion of pathology is a negative one: a proposed foundational theory T should *not* be pathological. A theory should be talking about the "stuff" of its conceptual or ontological basis, rather than hidden higher-level concepts or entities, such as proofs or the theory itself, *à la* Isaacson ([14]). This is most easily demonstrated through an example: Zermelo–Fraenkel set theory is *not* pathological, while the theory consisting of ZF plus the axiom that (in the language of ZF) says 'ZF is inconsistent' *is* pathological.[10]

Let us summarise the criteria that we have discussed. For a theory T to provide a foundation for mathematics, it must satisfy the following criteria:

(i) *Technical strength*: We must be able to do mathematics in T.

(ii) *Ontological or conceptual justification*: T must have a conceptual or ontological basis.

[9] It may also be possible for the opposite to happen; that is, a single theory may have two different justifications.

[10] By Gödel's second incompleteness theorem, ZF cannot prove its own consistency, and so this theory is consistent (if ZF is consistent).

(iii) *Naturality*: T must be parsimonious and not pathological.

Now, these criteria are perhaps not the only criteria that we should consider. For example, one might wish to include aesthetic notions, such as the elegance of the theory, or computational ones, such as the efficiency of the theory. However, while these are indeed reasonable points to consider, especially if one wishes to develop a general theory of foundations of mathematics, the criteria listed above will be sufficient for our thesis, for they capture the issues germane to our discussion.

2.2 Criteria for autonomy

Now that we have set out criteria for a theory to provide a foundation for mathematics, we shall outline criteria for a theory to provide a foundation for mathematics that is *autonomous* from another foundational theory, i.e. for two theories to provide independent foundations. We take these criteria from [20].

Our first criterion is that of *logical autonomy.* A theory T_1 is logically autonomous with respect to a theory T_2 iff it is possible to formulate T_1 without appealing to notions belonging to T_2.[11] Exactly which notions belong to a theory is somewhat vague, but the idea is that T_1 should not use anything specific to T_2. For example, the theory of groups is logically autonomous from that of Boolean algebras, since one can formulate the former without an appeal to the latter, while linear transformations are not logically autonomous with respect to vector spaces, since one needs the latter to define the former.[12] Notice the asymmetry here: vector spaces are logically autonomous from linear transformations.[13]

The next criterion is that of *epistemic autonomy.*[14] A theory T_1 is epistemically autonomous from a theory T_2 iff it is possible to understand T_1 without first understanding the notions belonging to T_2. For example, arithmetic is epistemically autonomous from real

[11] We use the phrases 'autonomous from' and 'autonomous with respect to' interchangeably.

[12] This example is taken from [9] (p. 152). Note that in [25], McLarty points out that historically the notion of a linear transformation came before that of a vector space. Thus, in order for our example to hold, we must restrict our attention to the modern formulations of vector spaces and linear transformations.

[13] A terminological point: When we say that 'T_1 and T_2 are logically autonomous', we mean both that T_1 is logically autonomous from T_2 and that T_2 is logically autonomous from T_1. The same goes for the other criteria of autonomy.

[14] What we refer to as *epistemic autonomy* is what Linnebo & Pettigrew refer to in [20] as *conceptual autonomy.* We have changed the name in order to avoid confusion with the notion of *conceptual or ontological justification* outlined in the previous section.

analysis – just ask any schoolchild. The converse, however, does not hold: attempting to learn real analysis without first understanding arithmetic would be somewhat foolhardy.[15]

Our final criterion is that of *justificatory autonomy.* A theory T_1 has justificatory autonomy with respect to a theory T_2 iff it is possible to motivate and justify the claims of T_1 without appealing to those of T_2. This is the hardest criterion to pin down, and will be the one that requires the most work when we discuss the autonomy of SDG. The idea is that the conceptual or ontological bases of T_1 and T_2 should be independent. For example, arithmetic and topology have justificatory autonomy from each other, since they have different conceptual bases, namely the natural numbers and abstract space respectively.

A few comments on this criterion of justificatory autonomy are in order.

There can be common ground between the two theories without their justificatory autonomy being affected. For example, we shall argue that axiomatic set theory and SDG can both use naïve set theory in their justifications and still be autonomous.

In the previous section we highlighted the distinction between an attempted and a completed ontological or conceptual justification. The same applies for justificatory autonomy: one might attempt to give a justification of T_1 that is independent from T_2 without providing a complete justification.

One final point to bear in mind is that while justificatory autonomy and epistemic autonomy are related, they are distinct. For example, classical differential geometry and the study of Riemann surfaces are (to a degree) conceptually autononous, since the former is based on smooth structures and smooth maps between them, while the latter arises from the study of conformal maps on the (extended) complex plane.[16] However, they are not epistemically autonomous, since one needs to understand notions from differential geometry, such as those of a manifold and a tangent space, before one can understand Riemann surfaces.

Let us summarise our criteria for a theory T_1 to provide a foundation for mathematics autonomous from that provided by a theory T_2:

(i) *Logical autonomy*: T_1 can be formulated without appealing to T_2.

[15] I wonder if New Maths was an example of this?

[16] While conformal maps are smooth, they have a different conceptual starting point from the definition of a smooth map, namely preservation of angles.

(ii) *Epistemic autonomy*: T_1 can be understood without first having to understand T_2.

(iii) *Justificatory autonomy*: T_1 can be justified without appealing to T_2.

The above list is by no means exhaustive. There may be other criteria of autonomy that one would wish to consider, such as historical autonomy or computational autonomy. However, as we mentioned at the end of the previous section, while these criteria are not complete, they will be sufficient for our purposes and will enable us to address the key aspects of our discussion.

Chapter 3

Set theory

In this chapter we shall discuss set theory as a foundation for mathematics. Although our main topic of investigation in this essay is SDG, we need to study set theory in depth in order to demonstrate that the foundation offered by SDG is autonomous from that offered by set theory. This hard work will also help us in Chapter 6.

Our plan is as follows. We will first outline the current orthodox formulation of set theory in §3.1 and then in §3.2 we shall apply the criteria from §2.1 to set theory. In §3.3 we will briefly consider conceptions of sets that differ from the current orthodoxy.

3.1 An outline of set theory

In this section we shall outline the current orthodoxy in set theory, namely Zermelo–Fraenkel set theory and the cumulative hierarchy. We will go into some technical depth, although our exposition will not be comprehensive. Such an account can be found in [15] (my account is based upon this text).

The *language of set theory* is a first-order logical language with a non-logical binary predicate symbol $\in$, which is called the *membership symbol*.[1] Zermelo–Fraenkel set theory (ZFC)[2] is a theory in the language of set theory, and the cumulative hierarchy is the intended interpretation of ZFC, where $\in$ is interpreted as genuine set-membership. So ZFC is syntactic and the cumulative hierarchy is semantic.[3] But before we embark on a

[1] Throughout this essay we shall take $=$ to be a logical symbol, i.e. we shall only consider languages with identity.

[2] We'll explain what the 'C' in 'ZFC' stands for later on.

[3] The cumulative hierarchy is by no means the only model of set theory. Indeed, quite the contrary:

description of either ZFC or the cumulative hierarchy, we need to discuss ordinals.

Ordinal numbers, or just *ordinals*, can be thought of as generalisations of natural numbers. Heuristically, the idea is that you count $0, 1, 2, 3, \ldots,$ get to infinity, and then carry on counting. Now, prima facie, this makes absolutely no sense – Buzz Lightyear might say 'To infinity and beyond!', but that's just rhetoric, right? Well, it turns out that one can frame this idea completely rigorously in beautiful mathematics.

Before we formally define an ordinal, let us state some elementary definitions. Let X be a set equipped with a binary relation $<$. We say that $<$ is an *ordering* iff the following hold:

(i) $(\forall x \in X)(x \nless x)$ (*irreflexivity*);[4]

(ii) $(\forall x, y \in X)(x < y \rightarrow y \nless x$ (*antisymmetry*);

(iii) $(\forall x, y, z \in X)(x < y \wedge y < z \rightarrow x < z)$ (*transitivity*).[5]

We say that $<$ is a *well-ordering* iff the above conditions hold and

$$(\forall S \subseteq X)(\exists x \in S)(\forall y \in S)(x < y \vee x = y).^6$$

We can now define an *ordinal* as a transitive set that is well-ordered by membership;[7] that is, an ordinal is a transitive set X such that $\in$ satisfies the conditions on $<$ given above. Ordinals are usually denoted using lower-case Greek letters, and $\in$ is usually denoted by $<$ when in reference to ordinals (in order to highlight the ordering).

But how does this formal definition reflect the heuristic one given earlier of counting past infinity? Well, firstly, we can use ordinals to do arithmetic, for the finite ordinals turn out to model the natural numbers. This is achieved by defining 0 as the empty set, which is an ordinal, and $n + 1$ as the set of the previous n ordinals, which is itself an ordinal;[8]

the study of models of ZFC is an active area of mathematical research.

[4] '$a \nless b$' is an abbreviation for $\neg(a < b)$.

[5] Notice that irreflexivity and transitivity together imply antisymmetry. We state antisymmetry separately for clarity.

[6] Notice the inherently second-order nature of this statement, since it requires quantification over subsets S of X. We shall discuss this later.

[7] A set X is said to be *transitive* iff every element of X is also a subset of X, i.e. $x \in X \Rightarrow x \subset X$. This should not be confused with the transitivity of an ordering.

[8] I am skirting over the proofs and technical subtleties behind this construction; see Exercises 1.2–1.9 in chapter 1 and chapter 2 of [15] for an upliftingly transparent presentation.

that is:

$$0 := \varnothing, \quad n+1 := \{0, 1, 2, \ldots, n\}.^9$$

Okay, so we now have a way of doing ordinary arithmetic in the ordinals, but how do we count to *infinity*? The trick is to consider the finite ordinals together as a *set*, which we denote ω. As the lower-case Greek lettering suggests, ω is an ordinal. (This is not difficult to show: the reader might like to try it for themselves.) One then defines $\omega + 1$ as the set $\omega \cup \{\omega\}$ and $\omega + (n+1)$ as the set $\omega + n \cup \{\omega + n\}$. One then considers the set of all these ordinals to get to $\omega \cdot 2$, and so forth.

Now that we have a (rough) idea of what the ordinals are, we can define the cumulative hierarchy, which is denoted V. The first stage in the cumulative hierarchy, V_0, is the empty set $\varnothing$. The next stage, V_1, is the power set (the set of all subsets) of the empty set, $\mathcal{P}(\varnothing)$. V_2 is then the power set of this set, $\mathcal{P}(\mathcal{P}(\varnothing))$, and so on. We can write this recursively as

$$V_0 := \varnothing, \quad V_{\alpha+1} := \mathcal{P}(V_\alpha).$$

This covers the finite stages of the cumulative hierarchy. Before we move on to the infinite stages, we need two definitions. A *successor* ordinal is an ordinal of the form $\alpha + 1$ (where α is an ordinal). So $1, 2, 3, \ldots$ and $\omega + 1, \omega + 2, \omega + 3, \ldots$ are examples of successor ordinals. A *limit ordinal* is an ordinal that is neither $\varnothing$ nor a successor ordinal. So $\omega, \omega \cdot 2, \omega \cdot 3, \ldots$ are examples of limit ordinals. For a limit ordinal α, we define V_α as the union of all the previous V_γ's:

$$V_\alpha := \bigcup_{\gamma < \alpha} V_\gamma.$$

We can now define the cumulative hierarchy as a whole. Let On denote the class[10] of all ordinals. Then we define the cumulative hierarchy, to be

$$V := \bigcup_{\alpha \in \mathrm{On}} V_\alpha.$$

We can sketch the cumulative hierarchy pictographically: see Figure 3.1. The ordinals,

[9] Notice that this is equivalent to defining $n + 1$ as the set $n \cup \{n\}$.
[10] We shall explain the distinction between sets and classes shortly.

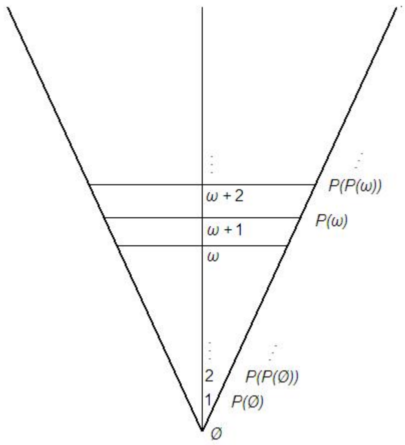


Figure 3.1: A sketch of the cumulative hierarchy

indicated by the vertical line, form the "backbone" of V,[11] and as we go up the ordinals the corresponding V_α's get larger and so the diagram gets wider.

The cumulative hierarchy is the picture behind ZFC. In ZFC, under this intended interpretation,[12] to exist is to be a set in the cumulative hierarchy, i.e. $\forall$ and $\exists$ quantify over V. The axioms of ZFC are ones that reflect this conception of sets.

The axioms can be grouped into two broad camps: those concerning the iterative conception of sets and those concerning the size of sets. There is also one straggler, the Axiom of Choice, which we shall deal with last.

The axioms concerning the iterative conception of sets are the Axiom of Power Set, the Axiom of Union, the Axiom of Pairing, the Axiom of Extensionality, the Axiom of Regularity, and the Axiom Schema of Separation. We shall cover them in that order.

The Axiom of Power Set says that if a set exists, then its power set exists; formally:

$$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (\forall u(u \in z \rightarrow u \in x))).$$

This very much reflects the cumulative hierarchy, since the operation of power set is a key tool in constructing it.[13]

[11] Notice that $\alpha \in V_{\alpha+1}$ for every ordinal α.

[12] There are of course models of ZFC that are not isomorphic to V (c.f. footnote 3 above).

[13] A general note is appropriate at this point. My use of verbs such as 'to construct' is not meant to have any philosophical implications; I use them simply for grammatical ease. I am not advocating a constructivist ontology, nor a platonic one, nor any other sort for that matter. As we said in Chapter 1, we will avoid metaphysical questions.

Notice that the *Axiom of Power Set* builds higher-order notions into ZFC (c.f. footnote 6), since it enables one to quantify over collections of sets.[14] We shall discuss this in the next section when we look at pathology.

The *Axiom of Union* asserts the existence of the union of any set; formally:

$$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow ((\exists u)(u \in x \wedge z \in u)).$$

This again reflects V, since we used unions to define V_α for limit ordinals.

The *Axiom of Pairing* says that for any sets x and y, the set $\{x, y\}$ exists; formally:

$$(\forall x)(\forall y)(\exists z)((\forall u)(u \in z \leftrightarrow u = x \vee u = y).$$

This axiom is crucial technically, since (as we shall see) it allows one to construct ordered pairs. It also allows one to define the operation $x \cup y := \bigcup\{x, y\}$, which is crucial in defining the ordinals, and it seems quite reasonable intuitively: surely given two sets we can consider them as a pair? But how is it justified *by the cumulative hierarchy*? The argument is as follows. Let $a, b \in V$. Then $a \in V_\alpha$ for some α and $b \in V_\beta$ for some β. Without loss of generality, assume that $\alpha \leqslant \beta$.[15] Then $V_\alpha \subset V_\beta$ (by the iterative definition of the cumulative hierarchy). Thus $a \in V_\beta$ and hence $\{a, b\} \subset V_\beta$. Therefore $\{a, b\} \in \mathcal{P}(V_\beta) = V_{\beta+1}$. So we are done.

The *Axiom of Extensionality* states that a set is completely defined by its members; formally:

$$(\forall x)(\forall y)(((\forall z)(z \in x \leftrightarrow z \in y)) \rightarrow x = y).^{16}$$

The Axiom of Extensionality certainly fits the iterative conception of sets, where sets are built from the bottom up and thus are entirely specified by their members. Indeed, such an axiom might initially seem obvious – how else could sets be built? – but in §3.3 we shall consider a conception of sets that is not extensional.

The *Axiom of Regularity* (or the *Axiom of Foundation*) states that every set has a

[14] In first-order Peano Arithmetic, for example, one can only quantifier over numbers; one cannot quantify over sets of numbers.

[15] A technical note: We're using the fact that the class of ordinals is (well-)ordered; see Lemma 2.11 in [15].

[16] The last implication is only left-to-right because the right-to-left implication is a theorem of logic: by equality-substitution, for any formula φ one can prove that $x = y$ implies $\varphi(x) \leftrightarrow \varphi(y)$.

$\in$-minimal element; formally:

$$(\forall x)(x \neq \varnothing \to (\exists y)(y \in x \land x \cap y = \varnothing)).^{17}$$

The Axiom of Regularity is justified by the bottom-up nature of V: in the cumulative hierarchy, a membership chain $(x_1 \ni x_2 \ni x_3 \ni \ldots)$ will eventually come to a end. Regularity captures this formally. It also allows one to define the *rank* of a set, which is (roughly speaking) the height of the set in V, e.g. $\varnothing$ has rank 0 and V_α has rank α. This is a very useful tool in ZFC.

The last axiom in this first group is the *Axiom Schema of Separation*. This allows us, given a formula φ and set x, to define the subset $\{z \in x : \varphi(z)\}$; formally: given a formula $\varphi(v)$ with one free variable v,[18]

$$(\forall x)(\exists y)((\forall z)(z \in y \leftrightarrow z \in x \land \varphi(z))).^{19}$$

Separation is justified by the fact that all subsets of a set in V are in V (recall that $V_{\alpha+1}$ is defined as the power set, the set of *all subsets*, of V_α). The Axiom Schema is a way of capturing this formally in ZFC (although it doesn't quite capture it perfectly – see footnote 31 below).

At this point, let us explain the set/class distinction. A *set* in ZFC is simply a variable which, under the intended interpretation, refers to a set in V. A *class* is a collection of the form

$$C = \{x : \varphi(x)\}$$

for some formula φ. We write $x \in C$ as an abbreviation for $\varphi(x)$; we do this because it is

[17] Note that the existence of the intersection of two sets follows from the Axiom Schema of Separation, which we shall discuss shortly. Please also forgive my cheating: $\varnothing$, $\neq$, and $\cap$ are, strictly speaking, not symbols in the language of set theory. I employed them for the sake of clarity, and I will continue to employ such abbreviations for this reason.

[18] It is an axiom *schema*, not just an axiom, because it is inherently second-order: for each formula $\varphi(v)$ there is an instance of the axiom schema.

[19] Note for the *cognoscenti*: For the sake of clarity, I have left out parameters from the statement of this axiom schema; for a full-fat version, see p. 7 of [15].

easier to work with classes than formulas.[20] We can consider any set S as the class

$$\{x : x \in S\}.$$

We cannot, however, treat all classes as sets. Consider the class $U := \{x : x = x\}$. This is the class of all sets, since for every set x we have $x = x$ (as a theorem of logic). Suppose that U was indeed a set; then we could apply Separation to it with the formula $x \notin x$ to get the set

$$S := \{x : x \notin x\}.$$

This leads to Russell's famous paradox, since both $S \in S$ and $S \notin S$ lead to contradiction. A class that is not a set is called a *proper* class.[21]

One of the key motivations behind the iterative notion of the cumulative hierarchy was to deal with the contradictory nature of proper classes. The Separation Schema allows one to specify only subsets of existing sets, and thus one cannot define a proper class, avoiding contradiction. This leads us to the next group of axioms of ZFC, those which deal with the size of the cumulative hierarchy.

The *Axiom of Infinity* asserts the existence of an infinite set; formally:

$$(\exists x)(\varnothing \in x \wedge ((\forall y)(y \in x \rightarrow y \cup \{y\} \in x)).$$ [22]

This axiom is easily justified by the cumulative hierarchy, since it asserts the existence of ω, the set that allowed us to proceed to the infinite in V.

Notice that the Axiom Schema of Separation and the Axiom of Infinity together imply the existence of the empty set, since we can apply Separation to the infinite set and the formula $x \neq x$.[23]

The last axiom of this group concerning the size of sets is the *Axiom Schema of Replacement*. This states that for any functional formula F and any set x, the image $F(x)$

[20] It is important to bear in mind that this is only an abbreviation: the use of the symbol $\in$ here is different from both its use in ZFC and the cumulative hierarchy.

[21] We mentioned earlier the class On of all ordinals. This is in fact a proper class. This fact is the content of the Burali-Forti paradox; we won't go into the details (see p. 20 of [15] for a proper account), but the rough idea is to consider On + 1.

[22] Note for the *cognoscenti*: The axiom actually asserts the existence of what is called an *inductive* set. The details are quite subtle, but this is equivalent to ω being a set (see Exercise 2.6 in [15]).

[23] Some presentations of ZFC list a separate axiom asserting the existence of $\varnothing$.

is a set; formally, given a functional formula $F(v, w)$,[24]

$$(\forall x)(\exists y)((\forall z)(z \in y \leftrightarrow (\exists u)(u \in x \wedge F(u, z))).^{25}$$

We won't get into the mathematics, but without Replacement the ordinals cannot get past $\omega \cdot 2 = \{0, 1, 2, 3, \ldots, \omega, \omega + 1, \omega + 2, \omega + 3, \ldots\}$. Thus, this Axiom Schema is justified by the idea that V should be bigger than $V_{\omega \cdot 2}$.

Note that the Axiom Schema of Replacement implies the Axiom Schema of Separation.[26] Separation is listed separately because a restricted version of ZFC without Replacement, but with Separation, is widely studied. Such a system is denoted ZC (or Z without the Axiom of Choice – see below), since it was Fraenkel who introduced Replacement.

There are in fact other axioms regarding the size of V that are not included in the usual presentation of ZFC, so-called *large cardinal* axioms. These axioms assert the existence of sets of different very large sets and their study is an active area of research. One has to be careful though: some large cardinal axioms turn out to be inconsistent with the rest of ZFC, e.g. Reinhardt cardinals ([17]).

The last axiom of ZFC, the *Axiom of Choice*, is a straggler, although it is perhaps affiliated with the group of axioms concerning the iterative conception of sets. Before we state it, we need to define functions in ZFC. One defines an *ordered pair* (a, b) in ZFC as

$$\{a, \{a, b\}\}.^{27}$$

We know this set exists by the Axiom of Pairing. (Note that we can apply Pairing to a to get $\{a, a\}$, which equals the singleton $\{a\}$ by Extensionality.) We then define a *function* as a set of ordered pairs f such that if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$. The *domain* of a function f is the set $\{x : (x, y) \in f\}$; the existence of this set is guaranteed by the Power Set Axiom and the Separation Schema. If $(x, y) \in f$, then we write $y = f(x)$.

[24] A *functional formula*, or just a *functional* in this context, is a formula $F(v, w)$ with two free variables v and w such that if $F(x, y)$ and $F(x, z)$, then $y = z$. The image of a set z under a functional is the class $F(z) := \{y : (\exists x)(x \in z \wedge \varphi(x, y)\}$.

[25] Note for the *cognoscenti*: I have again omitted parameters; see p. 13 of [15] for the unabridged version.

[26] Sketch proof: Given a set x and a formula $\varphi(v)$, define a functional F by $F(v, w) \leftrightarrow v = w \wedge \varphi(w)$. Then $F(x) = \{z \in x : \varphi(z)\}$.

[27] There are other ways of defining ordered pairs in ZFC. This is the conventional way, and is often referred to as the *Kuratowski ordered pair*.

We will not attempt to state the Axiom of Choice completely formally, since it would be somewhat long-winded and quite unclear. We shall state it semi-formally instead. The Axiom of Choice says that every family[28] of nonempty sets has a *choice function*. What this means is that if we have a family of sets $S = \{x_i : i \in I\}$ such that $x_i \neq \varnothing$ for every $i \in I$, then there exists a function f with domain S such that

$$f(x_i) \in x_i$$

for every $i \in I$. Put heuristically, the function f "chooses" an element from each x_i in S.

The Axiom of Choice is essential for various branches of mathematics, such as algebra and analysis. For example, without Choice one can have vector spaces with two bases of different cardinalities, i.e. the dimension of a vector space is no longer well-defined (p. 66 of [13]). It is, however, somewhat controversial. For example, the (in)famous *Banach–Tarski paradox* (1924) is a consequence of Choice. This paradox states that a solid three-dimensional sphere can be cut into finitely many pieces – five, in fact ([28]) – and, using only rotations and translation, can be reassembled into two spheres, each with the volume of the original sphere (see [30] for a detailed exposition). Some dispute whether this is in fact the fault of Choice, arguing that the definition of volume is to blame. We will not enter into this debate; I simply wish to highlight the controversy surrounding the Axiom of Choice.[29] Because of its controversial nature, the Axiom of Choice is usually listed separately from the other axioms: the 'C' in 'ZFC' stands for the axiom of *c*hoice, and the theory of Zermelo–Fraenkel set theory *without* the Axiom of Choice is denoted ZF.[30]

But all this controversy aside, how is Choice justified by the cumulative hierarchy? The argument is as follows. Consider a family of sets:

$$S = \{x_i : i \in I\}.$$

[28] The word *family* is merely a useful term to refer to a set of sets.

[29] For an excellent compilation of the good, the bad, and the downright odd aspects of Choice, see [13].

[30] When we wish to refer to either ZF or ZFC in a non-committal way, we shall write 'ZF(C)'. (We shall apply this notation more generally in §3.3.)

Take the union of S:

$$\bigcup S = \{x : (\exists i \in I)(x \in X_i)\}.$$

Let $f(S)$ denote the set of elements chosen from the x_i using the Axiom of Choice. By the Axiom of Union, $\bigcup S$ is in V. Now, each element chosen from the x_i's is in $\bigcup S$, and thus $f(S)$ is a subset of $\bigcup$. So, by the the iterative conception of sets, $f(S)$ should be in V.

Now, it turns out that this argument cannot be formalised in ZF,[31] hence my use of the word 'should' in the previous sentence (an unusual word to see in mathematical reasoning). However, it does provide a good justification for Choice.

In this section we outlined the cumulative hierarchy and showed how the axioms of ZFC are justified by this conception of sets. In the next section we shall examine set theory as a foundation for mathematics. But before we move on, we need to make an important terminological distinction (which we have been employing implicitly already). We shall call the theory that we have been describing in this section, that of ZFC and its concomitant justification from the cumulative hierarchy, *orthodox set theory*.[32] This is the theory which we shall show SDG to be autonomous from. We will also refer to *naïve set theory*, by which we shall mean the informal and heterogeneous collection of intuitions and elementary notions of membership, such as Venn diagrams and '1 is a member of $\{1, 2, 3\}$'. The boundary between the two is of course not well-defined, but is still an important distinction to maintain, since later we shall argue that SDG can take concepts from naïve set theory and still legitimately claim autonomy from orthodox set theory.

3.2 Set theory as a foundation for mathematics

In this section we shall apply the criteria outlined in §2.1. We shall conclude that orthodox set theory can provide a foundation for mathematics.

The technical strength of set theory is well attested. One can construct and carry

[31] One might think that one could use the Axiom Schema of Separation to prove it, but recall that to apply Separation one must have a formula to define the subset, i.e we must in some way be able to specify the elements of the subset. The Axiom of Choice is non-constructive – it does not tell you what the choice set is, only that it exists – and thus one cannot use Separation to define the choice set. Now, in the so-called *constructible hierarchy*, in which every set is defined by a formula, one can in fact prove that the Axiom of Choice holds (see chapter 13 of [15]).

[32] Unless otherwise specified, by *set theory* we shall always mean *orthodox set theory*.

out all "everyday" mathematics in set theory: real, complex, and functional analysis; algebraic and differential geometry; combinatorics – the list goes on. For example, one uses the ordinal ω to perform arithmetic, and one constructs the real line $\mathbb{R}$ from ω by taking various equivalence classes (see chapter 9 of [11]). However, there are some areas of mathematics that cannot be formalised in set theory. For example, in category theory (which we shall describe in the next chapter), one defines entities such as the category of *all* topological spaces and the category of *all* groups. Such objects cannot be built in ZFC, since they form proper classes. There are ways of partially dealing with such objects in ZFC though; for example, one can employ a large cardinal axiom and then restrict one's attention to *small* topological spaces and groups, ones that are smaller than this large cardinal. But all this aside, set theory is still a remarkably powerful theory that can carry out most of current mathematics, and thus it meets the criterion of technical strength.

Set theory has a strong conceptual basis, namely the cumulative hierarchy. It is based on the empty set and well-defined, easily understood operations. While one might question the ontology of these sets – do they exist in some sort of platonic sense, independently of humanity's study of them, or are they purely the creation of our intellect? – they certainly provide a sound justification for set theory.

Now, one might argue that ZFC is not the only theory that can be justified by the cumulative hierarchy; for example, one might think that certain large cardinal axioms are justified by the cumulative hierarchy. However, this is not important for orthodox set theory as a foundation for mathematics. As we pointed out in Chapter 2, two different theories might share a conceptual basis, but each theory still has a conceptual basis.

The naturality of set theory is straightforward. The axioms are parsimonious. While some axioms are not independent of each other, as we saw in the previous section, none is entirely redundant. For example, while Separation is a consequence of Replacement, its being stated separately is useful both for presentation and for distinguishing between the theories Z and ZF. ZFC is also not pathological: all of the axioms regard properties of sets. Now, one might worry that the Power Set Axiom is smuggling in higher-order concepts (remember that ZFC is a first-order theory), but one would be anxious unnecessarily. For set theory is a theory about sets and so it is quite reasonable that in ZFC one should quantify over sets of sets (since that is what it is, a theory of *sets*). It is not like Peano

Arithmetic, where quantifying over sets of numbers would be higher-order, since sets of numbers are not themselves numbers. Indeed, one needs to be careful to state ZFC in the correct first-order way: Separation and Replacement are axiom schemas, not axioms.

We have shown that set theory meets all the criteria from the previous chapter, and thus set theory provides a foundation for mathematics. Before we move on to the next theory of this essay, SDG, let us briefly consider other notions of sets.

3.3 Other conceptions of sets

So far we have only looked at orthodox set theory, but there are other conceptions of sets. In a slight digression from our main discussion, we shall briefly examine some examples of such different conceptions, namely variations of ZFC, the elementary theory of the category of sets, intuitionistic set theory, and non-well-founded set theory. While our discussion of these theories is not necessary for our primary thesis regarding SDG as an autonomous foundation for mathematics, we shall draw on them in Chapter 6 when we discuss our secondary thesis of mathematical pluralism.

There are many variations of orthodox set theory. For example, as we noted in §3.1, one can work in Z or ZF, rather than in full ZFC, or with constructible sets (see footnote 31 above), or one can add large cardinal axioms to ZFC. The systems ZF^- (ZF without Power Set) and ZF−inf (ZF with Infinity negated) are also studied. One can also introduce atoms, sets that have no members but are not equal to the empty set. In this formulation, denoted ZFA (or ZFAC with the axiom of choice),[33] one builds the the cumulative hierarchy from the empty set and the atoms, and so it starts with a line, rather than a point, as demonstrated in Figure 3.2.

A more major variation on ZF is that of *Bernays–Gödel set theory*, denoted BG (or BGC with the axiom of choice).[34] In this theory, there are two types of objects, sets and classes. The rough idea behind BG(C) is that classes are collections of sets that are "too big" to be sets themselves. BG(C) and ZF(C) are conceptually quite similar, and they have equal provability power over statements involving only sets (p. 70 of [15]).

[33] See p. 250 of [15] for a rigorous outline of ZFA(C). Note that ZFA is often referred to as 'ZFU', which stands for 'ZF with Urelemente' ('ZF with pure elements').

[34] See p. 70 of [15] for a rigorous outline of BG(C).

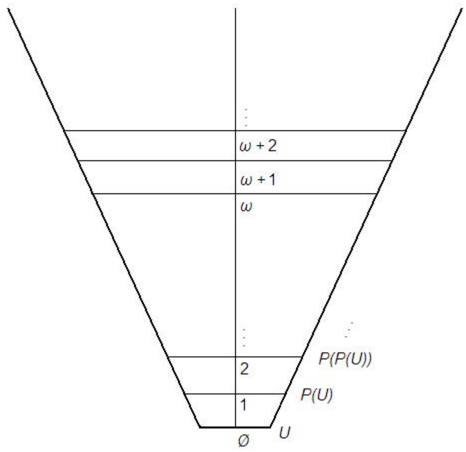


Figure 3.2: A sketch of the cumulative hierarchy with atoms (c.f. Figure 3.1)

The *elementary theory of the category of sets* (ETCS) was first formulated by William Lawvere in 1964 (see [18] and [19]). It is written in the language of category theory, which we shall describe in the next chapter. In ETCS, sets are not extensional: roughly speaking, a set is not defined by its members, but rather by the functions between it and other sets. This may sound contradictory, since one might think that one needs elements to define functions, but this turns out not to be the case. ETCS very much goes against the picture of sets painted by the cumulative hierarchy.

Intuitionistic set theory (IST) is a version of set theory built on intuitionism. We shall not go into the details,[35] but the main idea behind intuitionism is that we should only accept mathematics that can, at least in principle, be constructed or demonstrated. For example, the intuitionist takes a statement of the form '$\exists n\, \varphi(n)$' to mean that we actually have a number n for which φ holds. Thus, the Axiom of Choice is completely unacceptable to the intuitionist, since it does not specify which elements are chosen. The intuitionist also rejects completed infinities, since it is impossible for such a thing to actually be constructed (even in principle).[36] Thus, quite unlike ZF(C), all sets in IST are finite.

The final variant of set theory that we shall discuss is that of *non-well-founded set*

[35] A good introduction to intuitionism is to be found in chapters 4 and 5 of [10]. The comprehensive guide is of course [7].

[36] The intuitionist does however accept the existence of potential infinities. So, for example, it is intuitionistically valid to say that there are infinitely many numbers (since given any number one can always add 1), but it not intuitionistically valid to start talking about the *set* of natural numbers. Thus the Axiom of Infinity does not hold in IST.

theory. There are many different formulations (see [1]), but the common theme is a rejection of the Axiom of Regularity. The idea is that sets can have membership loops; such sets are often called *hypersets*. This goes completely against the well-founded nature of sets in the cumulative hierarchy.

Chapter 4

Synthetic differential geometry

We now come to the main topic of this essay, synthetic differential geometry. In §4.1 we shall outline SDG and then in §4.2 we shall examine SDG as a foundation for mathematics, applying the criteria from §2.1.

4.1 An outline of SDG

In this section we shall describe SDG and the concepts behind it: smoothness and isomorphism. Because of the length of the discussion, I have split this section up into subsections.

As with our description of set theory in the previous chapter, our outline of SDG will not cover all the technical details. The standard mathematical reference is Anders Kock's text, which is now in its second edition: [16]. Briefer introductions can be found in [22] and in [23], the latter also being an introduction to category theory and topos theory.

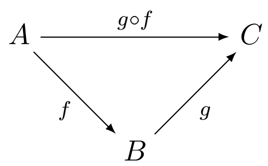
Before we can discuss SDG or its justification, we first need to describe two things: categories and toposes.

4.1.1 Categories and toposes

The idea behind category theory is to look at how things behave, rather than how they are made. So, for example, given two groups, a category theorist would be much more interested in the homomorphisms between them than in their individual elements. Pettigrew sums it up nicely: 'Ask not what a thing is; ask what it does' (p. 1 of [26]). Broadly speaking, the way this is achieved is to take functions as primitive, rather than as correspondences between elements of two sets.

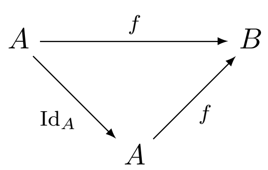
A *category* is defined in first-order logic. A category has *objects* and *arrows* (or *morphisms*). Each arrow is associated with precisely two (possibly equal) objects, one called its *domain* and the other its *codomain*. If f is an arrow with domain A and codomain B, then we write $\mathrm{Dom}(f) = A$ and $\mathrm{Cod}(f) = B$, which we denote by $f\colon A \to B$ or $A \xrightarrow{f} B$.

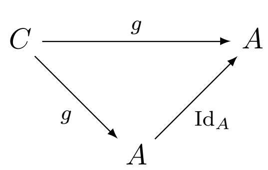
The first axiom of a category regards composition of arrows. If arrows f and g are such that $\mathrm{Cod}(f) = \mathrm{Dom}(g)$, then they have a *composite arrow*, denoted $g \circ f$, with $\mathrm{Dom}(g \circ f) = \mathrm{Dom}(f)$ and $\mathrm{Cod}(g \circ f) = \mathrm{Cod}(g)$. So, if $f\colon A \to B$ and $g\colon B \to C$, then $g \circ f\colon A \to C$. This can be expressed with a so-called *commutative diagram*:



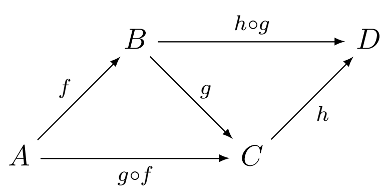
A diagram of objects and arrows, such as the one above, is said to *commute* iff any path from one object to another is the same under composition; such a diagram is called a *commutative diagram*. So, in the example above, we can go from A to C in one of two ways: along f and then along G, via B; or directly along $g \circ f$. By definition, these two paths are the same, since the direct path along $g \circ f$ is defined to be the composition of the two parts of the path via B.

The next axiom concerns *identity arrows*. Every object A has an identity arrow $\mathrm{Id}_A\colon A \to A$ such that for any objects B, C and arrows f, g with $f\colon A \to B$ and $g\colon C \to A$, we have $f \circ \mathrm{Id}_A = f$ and $\mathrm{Id}_A \circ g = g$; that is, the following diagrams commute:





The final axiom of a category is that composition is *associative*; that is, for any objects A, B, C, D and arrows f, g, h such that $A \xrightarrow{f} B \xrightarrow{g} C \xrightarrow{h} D$, we have $h \circ (g \circ f) = (h \circ g) \circ f$, i.e. the following diagram commutes:

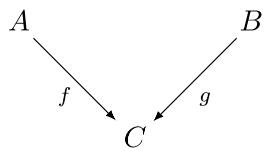


(Notice that we do not draw the composition arrows $h \circ (g \circ f)$ and $(h \circ g) \circ f$, since they would only clutter the diagram.)

A *topos* is a special kind of category that satisfies some additional axioms.[1] Our discussion in the next section necessitates that we describe these axioms in some detail, although the explanations given will not be completely rigorous, in particular that of exponential objects.

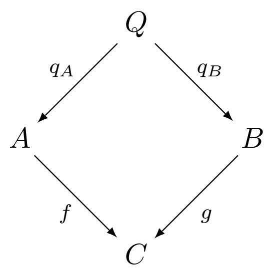
The first axiom of a topos states that there exists a *terminal object*, an object 1 such that for every object A there exists a unique arrow $1_A \colon 1 \to A$. Note that this terminal object need not be unique; when we use the symbol 1, we refer to an arbitrary terminal object.

The next axiom regards *pullbacks*. It comes in two parts. Consider the following diagram:



(4.1)

The first part of the axiom states that, for any such diagram, there exists an object Q along with two arrows $q_A \colon Q \to A$ and $q_B \colon Q \to B$ such that the following diagram commutes:

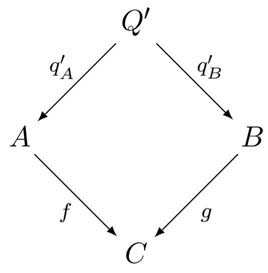


(4.2)

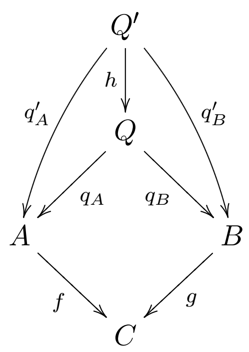
In other words, $f \circ q_A = g \circ q_B$. The second part of the axiom states that if there is another

[1] A technical note for topos theorists: There are many equivalent ways to define a topos. We shall follow that of [2], namely that a topos is a category which has a terminal object, pullbacks, exponentials, and a subobject classifier (p. 308). I have chosen this particular approach because I believe it is the easiest to explain and will be the most straightforward – or rather the least difficult! – to justify conceptually.

object Q' with arrows $q'_A\colon Q' \to A$ and $q'_B\colon Q' \to B$ such that the diagram



commutes, then there is a unique arrow $h\colon Q' \to Q$ such that the following diagram commutes:



If Q, $q_A\colon Q \to A$, and $q_B\colon Q \to B$ are such that they satisfy these conditions, then Q (with q_A and q_B) is called a *pullback* of (4.1), and (4.2) is called a *pullback square*.

At this point, a remark is in order. Notice that in each of the above axioms we do not define a unique object per se, but an object that is unique *only up to isomorphism*; that is, one states that an object with certain properties exists, but also that if there is another object with the same properties, then there is an unique arrow from one to the other.[2] The other topos axioms are of a similar vein. This notion of objects being unique only up to isomorphism is perhaps *the* crucial concept behind category theory; it is the formalisation of Pettigrew's 'Ask not what a thing is; ask what it does' motto that we mentioned earlier. This concept is important to bear in mind, and we shall come back to it later, but let us return to stating the axioms of a topos.

The third axiom asserts the existence of *exponential objects* (or just *exponentials* for short). The definition of an exponential is in a similar style to that of pullbacks given above,

[2] While we didn't quite state it in such terms, notice that this is true of the terminal object axiom: for if two terminal objects 1 and $1'$ exist then, by the definition of a terminal object, there is a unique arrow $1 \to 1'$.

but the details are more involved and thus I will not state their definition rigorously (a formal definition can be found on p. 57 of [23]). Instead, I shall explain them heuristically, which will suffice for our discussion. The exponential of two objects A and B is denoted by B^A. As the notation suggests, it behaves like the set of arrows from A to B. This axiom allows more advanced structures to be built from simpler ones and is crucial for doing mathematics in a topos, and thus in turn will be crucial for SDG to provide a foundation for mathematics (more on this later).

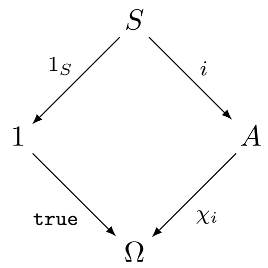
Before we explain the final axiom of a topos, we need to explain how one expresses membership in a topos. Given an object A, one defines a *member* of A to be an arrow $x\colon 1 \to A$ (recall that 1 denotes an arbitrary terminal object). This makes sense, since a terminal object is one to which each object has a unique map, and so the map $x\colon 1 \to A$ can be thought of as using this uniqueness to "pick out" a member of A. Notice here that an member is a *arrow*, *not* an object. This is quite different from membership set theory, where in fact quite the opposite is the case, since functions are made from sets. Notice also that one cannot have an member of a member, since a member is an arrow and so one would have to have an arrow to an arrow, which doesn't make sense in a category. This marks another difference from set theory.

One can generalise this notion of membership to define *subobjects*. Technically, a subobject of an object A is a monic[3] arrow from an object S to A, but one can think of this informally as a category-theoretic way of "picking out" several members of A.

We can now describe the last axiom of a topos, which states that a topos has a *subobject classifier*. The idea behind this axiom is to allow truth to be internalised into the topos. Formally, a subobject classifier consists of an object Ω and an arrow $1 \xrightarrow{\texttt{true}} \Omega$ such that for any object A and subobject $S \xrightarrow{i} A$, there is a unique arrow $A \xrightarrow{\chi_i} \Omega$ such that the

[3] An arrow $A \xrightarrow{f} B$ is *monic* iff for any $C \xrightarrow{g} A$ and $C \xrightarrow{h} A$, $f \circ g = f \circ h$ implies $g = h$. This can be thought of as the category-theoretic generalisation of an injection.

following diagram is a pullback square:



Let us try to explain this diagram heuristically. Notice that $1 \xrightarrow{\texttt{true}} \Omega$ is a member of Ω, and thus Ω can thought of as containing the truth-values of the topos.[4] The χ_i can be thought of as the characteristic function of i: it is the arrow that sends the subobject i to `true`.

Using the subobject classifier, one can build semantics within a topos, known as the *internal language*. The details of the construction are very involved,[5] but the idea is that using only the tools of category theory, namely objects and arrows, one can construct a notion of truth internal to the topos. This internal language is crucial to SDG, since it allows one to express concepts that classical logic is unable to; for example, in this internal language the law of the excluded middle does not necessarily hold, which, as we shall see, will be essential in expressing the axioms of SDG.

Now that we have described categories and toposes, let us discuss the justification behind SDG.

4.1.2 Smoothness and isomorphism

There are two key concepts behind SDG: smoothness and isomorphism. We shall discuss them in turn.

The concept of smoothness starts as an informal one. The idea is to view the continuum as a non-punctiform, cohesive whole that cannot be broken or bent perfectly at a point. A comparison with the set-theoretic continuum will be helpful. The set-theoretic continuum, $\mathbb{R}$, is made up of points. One can define non-smooth functions $\mathbb{R} \to \mathbb{R}$, such as the modulus function

$$f(x) = |x|,$$

[4] We shall not define it, but Ω also contains the truth-value `false` (see p. 29 of [26]).

[5] See chapter 14 of [23] for a detailed exposition.

and, moreover, discontinuous ones, such as the so-called *blip function*

$$f(x) = \begin{cases} -1 & \text{if } x < 0, \\ 1 & \text{if } x \geqslant 0. \end{cases}$$

We can do this because $\mathbb{R}$ is punctiform: we define functions by specifying their action at each point. Now, in the SDG-continuum, which we shall call R, one cannot define functions like this, since R is not punctiform. The informal idea is that R is like an indefinitely compressible and extendsble strip of elastic: one can manipulate it by pulling, stretching, and squashing it in various directions and magnitudes, but one cannot tear it or bend it perfectly at a point.[6]

This view of the continuum as being a non-punctiform, cohesive whole is an old one.[7] Mathematicians and philosophers as notable as Aristotle, Kant, Leibniz, and Poincaré all considered the continuum to be non-punctiform (pp. 1–2 of [4]). While the idea of the continuum (and matter more generally) being made up of atoms dates back to at least Democritus (p.15 of [3]), this view, until the late 19$^{\text{th}}$ century, was very much the minority opinion, for it seemed to be undermined by the following simple argument: how can the continuum, which has extension, be composed entirely from points, which have no extension?

The question is now this: how do we make this notion of smoothness more precise? This is where infinitesimals come in. The idea is to view the continuum not as consisting of a collection of points, as one does in set theory, but to view it as consisting of infinitesimals, infinitely short line segments. Now, the notion of an infinitesimal is a difficult one, since it immediately raises the following difficult problem. It seems as though the continuum should be infinitely divisible; that is, I should always be able to divide a line segment in two. The question is then: can one divide an infinitesimal in two? For if one can, then we seem to have a contradiction, since an infinitesimal has infinitely short length and yet we can find a shorter length, namely this "infinitely short" length divided by 2. But if we

[6] One could of course *approximate* the modulus function by bending the elastic around a nail or pin, but no matter how small, these objects would have width and thus the elastic would never be bent *at a point*.

[7] We do not have room to cover the history of the different views of the continuum in any serious depth. For such an account, see Part I of [3].

cannot, then we contradict the infinite divisibility of the continuum. Berkeley, a staunch critic of infinitesimals, famously described them as 'ghosts of departed quantities' (p. 22 of [32]).

This difficult problem didn't stop the use of infinitesimals, however. For example, they were fundamental in Newton's development of the calculus and the work of Riemann (pp. 82–86 and pp. 145–148 of [3] respectively).[8] Indeed, even after the work of Cauchy and Weierstrass, who eliminated infinitesimals from the calculus through the introduction of the now familar ε–δ notion of a limit, mathematicians of as high repute as Lie and Cartan continued to use them to help formulate concepts (p. 3 of [4]), and use amongst physicists and engineers is still widespread. However, as we alluded to earlier, the work of Cauchy and Weierstrass, along with that of Dedekind and Cantor, put a (temporary) end to the use of infinitesimals amongst mathematicians. Bertrand Russell's description of infinitesimals as 'unnecessary, erroneous, and self-contradictory' (p. 3 of [4]) sums up the opinion of most 20$^{\text{th}}$ century mathematicians quite accurately.

So where does this leave us? Without infinitesimals, how are we to make the concept of smoothness precise? Well, after a period of abeyance, infinitesimals began to reemerge. In the 1960s, Abraham Robinson outlined *non-standard analysis*, a rigorous system of analysis containing infinitesimals (as well as infinitely large numbers) (see [27]). While this approach is not the one that is employed for SDG, it marks a crucial stage in the development of infinitesimals, since it was the first time that they had been framed in an entirely rigorous way. We shall see that the internal language of a topos will enable us to frame infinitesimals in a rigorous way that captures the notion of smoothness.

We can now make the notion of smoothness more precise. We shall use Hellman's work in [12]. Hellman outlines two principles. The first is the *Principle of Microstraightness*:

> 'Given a smooth curve C and any point P on it, there is a nondegenerate microsegment about P which is straight.' (pp. 624 of [12])[9]

[8] I'm being a little blunt here. Infinitesimals have been posited in all sorts of different forms. For example, Newton talked of *fluxions*, which were conceptually different from Leibniz's notion of an infinitesimal, which in turn was different from the infinitesimals employed by the 17$^{\text{th}}$ century Dutch physicist Nieuwentijdt (see pp. 82–101 of [3]). Still, all the conceptions had one thing in common, the notion of (some sort) of infinitely short length, which is the notion germane to our discussion.

[9] Something is said to be *nondegenerate* iff it is not equal to zero.

The idea of Microstraightness is to view a curve as being made from infinitesimal line segments; so, for example, a circle is viewed an an infinitely-sided polygon. The microsegment at P is the derivative at that point, hence why we have smoothness, since we can differentiate at every point.

An important point to make here is that while the view of the continuum under Microstraightness is non-punctiform, it does not say that that curves do not contain points. That is, curves contain points, they just aren't solely composed from them.[10] To adapt an analogy from [5], one can view the continuum as points glued together by infinitesimals. The metaphysics of this does not concern us in this essay: whether or not continua are actually so composed is a topic for another day. What does matter for us is whether Microstraightness is coherent; if it is not, then how is it different from all the historical unrigorous approaches to infinitesimals? In the next section we shall see that with the internal language of a topos we can indeed make this concept rigorous. Indeed, it would seem that the lack of topos theory (or some other sufficiently powerful mathematical apparatus) was what held historical attempts back from rigour.

The second of Hellman's principles is the *Archimedian Method*, which is the principle that the area under a curve $y = f(x)$ is the sum of the areas of the infinitesimally thin rectangles between the curve and the x-axis. This is the two-dimensional version of the Principle of Microstraightness. Now, that this follows from our informal conception of smoothness is less clear than Microstraightness, and we shall discuss this in the next section, but for now we shall simply take the Archimedian Method as a principle behind SDG.

Before we move on to isomorphism, the other key concept behind SDG, a digression is in order. As we have been alluding to already, the notion of smoothness is related to that of cohesiveness, as articulated by Bell in [5]. Bell defines a space S to be *cohesive* (or *indecomposable*) iff for any parts U and V of S, if $U \cup V = S$ and $U \cap V = \varnothing$ then either $U = \varnothing$ or $V = \varnothing$ (p. 147).[11] We shall outline the relationship between smoothness and cohesiveness through two observations.

Our first observation regards the law of the excluded middle (LEM). If we have LEM,

[10] This seems to have been Poincaré's view of the continuum: '[T]he point is not prior to the line, but the line to the point.' (p. 2 of [4]).

[11] Bell goes on to define further notions of cohesiveness, which we shall not discuss.

then the only spaces that are cohesive are the empty space (trivially) and spaces with only one point. The proof is as follows. Consider a cohesive space S and a point $p \in S$. Define $U := \{p\}$ and $V := S \backslash \{p\}$. By LEM, for every $q \in S$, either $q = p$ or $q \neq p$. Thus $U \cup V = S$ and $U \cap V = \varnothing$, and so by cohesiveness either $U = \varnothing$ or $V = \varnothing$. But $p \in U$ and so $U \neq \varnothing$. Thus $V = \varnothing$ and so $S = \{p\}$. This brings us to our second observation.

Define a space to be *decomposable* iff it is not cohesive. Then the decomposability of a space S implies that we can define a discontinuous function on S. The argument is as follows. Let S be a decomposable space. Then there exist parts U and V of S such that $U \cup V = S$ and $U \cap V = \varnothing$ but $U \neq \varnothing$ and $V \neq \varnothing$. Define a function $f\colon S \to \{0, 1\}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \in U, \\ 1 & \text{if } x \in V. \end{cases}$$

This function is discontinuous. Now, by the contrapositive of our argument, continuity implies cohesiveness; that is, spaces upon which one can define only continuous functions are cohesive. Thus, since smoothness implies continuity, smoothness implies cohesiveness.[12] Therefore smoothness is a special case of cohesiveness and thus, by our first observation, if we wish to formalise the concept of smoothness then LEM cannot universally apply. This is why the internal language of a topos, in which LEM does not necessarily apply, is crucial for SDG. Let us now discuss isomorphism.

The idea behind isomorphism is to look at the structural properties of mathematical objects, rather than the objects themselves. For example, consider the sets $\{1, 2, 3, 4\}$ and $\{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$. We can turn these sets into groups by defining binary operations on them:

	1	2	3	4
1	1	2	3	4
2	2	1	4	3
3	3	4	1	2
4	4	3	2	1

	♣	♢	♡	♠
♣	♣	♢	♡	♠
♢	♢	♣	♠	♡
♡	♡	♠	♣	♢
♠	♠	♡	♢	♣

Now, these two groups are isomorphic (explicitly: $1 \mapsto \clubsuit$, $2 \mapsto \diamondsuit$, $3 \mapsto \heartsuit$, $4 \mapsto \spadesuit$) and thus, as groups, are the same – the elements themselves are unimportant. Compare this

[12] I am relying on intuitive notions of continuity and smoothness here, rather than on any particular formal definitions.

with the set-theoretic view of these groups, where the elements are very important, since they compose the underlying sets.[13]

The motivation behind isomorphism is to extend this outlook to all of mathematics. Such an idea has perhaps been implicit in mathematicians' minds since mathematics was first studied – counting bricks or counting sheep, what's the difference mathematically? – but I believe the origin of the modern notion can be traced back to the (independent) discovery of non-Euclidean geometry by Bolyai, Lobachevski, and Gauss in the early 19$^{\text{th}}$ century (see p. 166 of [31]). This discovery dislodged the old idea of the absoluteness of Euclidean geometry, as explicitly articulated by Kant, which lead to new ideas about the nature of geometry. One of these ideas was to no longer view geometry as about space per se, but rather to look at the transformations and the associated invariants between spaces, as articulated by Felix Klein in the *Erlangen programme* (named after the university Klein was working at at the time). Following on from this, more abstract notions were considered, moving further and further away from the study of the space itself. This lead to the publication in 1945 of Eilenberg's and Mac Lane's famous article ([8]) that marked the beginning of category theory, the branch of mathematics dealing with transformations in their most abstract form, which in turn lead to the development of topos theory in the 1960s, most notably by William Lawvere.

As with smoothness, we shall not consider the metaphysics of isomorphism. While questions regarding the nature of isomorphism and its relation to the ontology of structure are interesting to ask, we shall bracket them.[14] Let us now outline the axioms of SDG.

4.1.3 The axioms of SDG

SDG is a theory of three parts. Firstly, it consists of the axioms of a topos;[15] these axioms arise out of consideration of the concept of isomorphism. Secondly, using the internal language of the topos, one defines a set of axioms known as *smooth infinitesimal analysis*

[13] This is not to say that one cannot consider notions of isomorphism in set theory. My point is that in set theory one starts off with sets with given structures and *then* one studies the isomorphisms between them, while the idea behind the concept of isomorphism is to only consider the isomorphisms, as in category or topos theory, for example.

[14] Note that such questions are very much related to the current debate in the philosophy of mathematics over structuralism (see chapter 10 of [29]).

[15] One also specifies that the topos be *non-degenerate*, which, heuristically, says that the topos is not trivial.

(SIA); as the name suggests, these axioms reflect the concept of smoothness. The last part concerns discrete subspaces and the natural numbers. We shall start by outlining the axioms of SIA, which are stated in the internal language of the topos.[16] We will not go into the finer details of SIA; for such an account, see [4].

The first axiom of SIA asserts the existence of a nontrivial commutative ring $(R, 0, 1, +, \cdot)$ with various properties.

The crucial property is the existence of a neighbourhood $\triangle := \{\varepsilon \in R : \varepsilon^2 = 0\}$ of 0 in R. These are the infinitesimals which will allow us to formalise the Principle of Microstraightness. (The set of infinitesimals around an arbitrary point x is the translate $x + \triangle$.) We need another axiom to explain why these infinitesimals are defined to be *nilsquare* ($\varepsilon^2 = 0$). This axiom is the *Principle of Microaffineness*, which states that maps $\triangle \to R$ are affine; formally:

$$(\forall f \in R^{\triangle})(\exists! a \in R)((\forall \varepsilon \in \triangle)(f(\varepsilon) = f(0) + a \cdot \varepsilon)).^{17}$$

This is the formalisation of the Principle of Microstraightness: the point P, in this case 0, is sent to another point, $f(0)$, and the infinitesimal microsegments around 0 are mapped to microsegements around $f(0)$. The constant a is then the derivative at $f(0)$, since it is the gradient of the microsegments at $f(0)$. Using Microaffineness we can formally prove that $\triangle \neq \{0\}$: simply apply it to the curve $y = x^2$.

Let us now justify why the infinitesimals in $\triangle$ are nilsquare.[18] Consider Figure 4.1. Microstraightness states that about $f(x)$ there is a nondegenerate infinitesimal microsegment about which the curve is straight, and hence we must have $\varepsilon \neq 0$. Now, let us calculate the areas $\square$ and ∇:

$$\square = \varepsilon \cdot f(x)$$

[16] This is important to bear in mind, since LEM does not hold in the internal language of SDG. If it did then the axioms would be blatantly contradictory.

[17] Note that we are relying on the existence of exponentials in the topos in order to talk about $R^{\triangle}$.

[18] We take this justification from [12] (pp. 624–625).

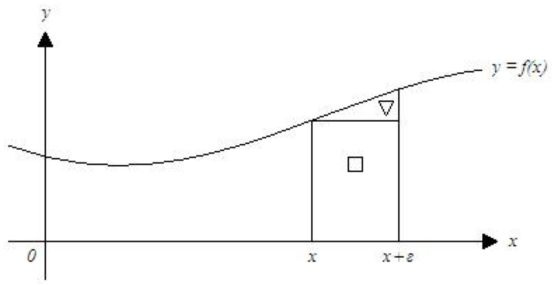


Figure 4.1: The justification for $\varepsilon^2 = 0$

and

$$\begin{aligned}\nabla &= \frac{1}{2}\varepsilon \cdot (f(x+\varepsilon) - f(x)) \\ &= \frac{1}{2}\varepsilon f(\varepsilon) \quad \text{(by Microaffineness).}\end{aligned}$$

So ∇ is proportional to ε^2. Now, by the Archimedian Method, the area under the curve is equal to the sum of the $\square$'s, and thus $\nabla = 0$. Thus we require that $\varepsilon^2 = 0$.

Let us now state the ordering properties of R. Trichotomy does not hold in R; that is, one does *not* have

$$(\forall x \in R)(\forall y \in R)(x < y \vee x = y \vee y < x).$$

Instead, one has

$$(\forall x \in R)(0 < x \vee x < 1)$$

and

$$(\forall x \in R)(\forall y \in R)(x \neq y \rightarrow x < y \vee y < x).$$

So, one can distinguish elements of R to a certain degree, and if one can distinguish two elements then one can order them precisely. This calls for a remark.

As we mentioned in the previous subsection, a necessary condition for continuity (and hence also smoothness) is that LEM does not hold. This is why in SIA not all elements of R can be distinguished; that is, trichotomy fails because it is a consequence of LEM: if we could distinguish all points from 0, say, then we could define the blip function, going against the concept of smoothness. Let us return to the axioms.

In R, one has has transitivity

$$(\forall x \in R)(\forall y \in R)(\forall z \in R)(x < y \wedge y < z \rightarrow x < z$$

and irreflexivity

$$(\forall x \in R)(x \nless x).$$

The final axiom is that R is a field in a partial sense:

$$(\forall x \in R)(x \neq 0 \rightarrow ((\exists y)(x \cdot y = 1))).$$[19]

The idea here is that elements that can be distinguished from 0 have inverses, which calls for a remark.

We shall now discuss the topos axioms. Since we already defined them in §4.1.1 we shall now only justify them. Before we examine each axiom individually, let us make a remark regarding isomorphism. As we highlighted when we outlined the axioms of a topos, objects are defined uniquely only up to isomorphism. That this holds in SDG is clearly justified by the concept of isomorphism that we outlined in the previous subsection: we are concerned with how the smooth objects behave, rather than the smooth objects themselves.

The existence of terminal objects is justified by the Microstraightness Principle. Given that we have a points-and-glue concept of the continuum, the terminal objects are the 0-dimensional spaces containing just one point of R.

The existence of pullbacks and exponentials is a consequence of smoothness. We will not be entirely precise, since the reasoning is mathematically quite involved. The idea with exponentials is that smooth maps between smooth spaces vary smoothly, and thus exponentials themselves are smooth objects. The pulback construction is more complicated, but the idea is that it acts like the intersection of the two smooth spaces and hence is itself smooth. Note that these axioms allow us to explicitly construct higher-dimensional smooth spaces from R, uch as R^2 and R^R, as well as subspaces defined by equations (see

[19] In this regard Bell calls R an *intuitionistic field* (p. 298 of [3]).

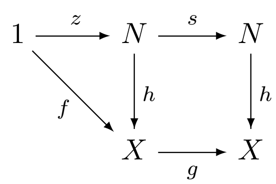
p. 81 of [22]).

We now come to the subobject classifier axiom. While its technical role is crucial to SDG, since without it we could not construct the internal language of a topos and thus we could not consistently state the SIA axioms, it is not clear how it is justifed by the concept of smoothness.[20] Why should there be a smooth object that can express truth values within SDG? At this point I have to admit defeat: I cannot think of a precise argument as to why smoothness justifies the existence of the subobject classifier. I suspect that it can be justified, perhaps along the following lines: since smoothness implies that LEM cannot hold, the very process of formalising the concept of smoothness in some way justifies the existence of a smooth method for building the internal language of a topos.

We now come to the the discrete subspace axioms of SDG. These were first put forward by McLarty in [22] and are not usually included in presentations of SDG.[21] However, they are crucial for SDG to provide a foundation for mathematics.

The two discrete subspace axioms state that every space is either empty or contains a point; and that every space M has a unique discrete subspace, denoted ΓM. These are justified for R by the points-and-glue picture of the continuum: the discrete subspace is the discrete space of points (without the infinitesimal glue). That all spaces have such a discrete subspace is justified by our remark earlier that in SDG one builds higher-dimensional spaces from R, and thus the discrete subspace of R is transferred to these higher-dimensional objects.

The last axiom of SDG asserts the existence of a *natural number object*, which is an object N together with arrows $1 \xrightarrow{z} N \xrightarrow{s} N$ such that for any object X with arrows $1 \xrightarrow{f} X \xrightarrow{g} X$, then there is a unique arrow $N \xrightarrow{h} X$ such that the following diagram commutes:



The way to think of this heuristically is to view $1 \xrightarrow{z} N$ as the number 0 and $N \xrightarrow{s} N$ as

[20] If we can justify the existence of a subobject classifier by smoothness, then we can justify its uniqueness only up to isomorphism by the concept of isomorphism underpinning SDG.

[21] Note that Bell includes them though in [3].

the successor operation. This is justified by the discrete subspace of the continuum, since one can embed a model of the natural numbers in R using these discrete points. Indeed, we defined R as a commutative ring with a distinguished element 1, and so the numbers $1, 1+1, 1+1+1, \ldots$ exist explicitly in R. That we have a separate axiom asserting the existence of a natural number object unique up to isomorphism is justified by the concept of isomorphism.

4.2 SDG as a foundation for mathematics

We will now apply the criteria from §2.1 to SDG. Let us start with technical strength.

SDG is a powerful theory. Its geometrical power is clear (see [16]) and, using the natural number object axiom, one can do arithmetic and combinatorics. But how does SDG's strength fare with regard to other branches of mathematics, for example analysis and set theory?[22]

SIA is a powerful system of analysis (see [4] and [16]), but the analysis is, as its name suggests, all smooth and thus is incomplete from a classical standpoint. One cannot simply stipulate that all analysis should be smooth – just ask someone studying catastrophe theory or singular algebraic curves! This is where the discrete subspace axioms come in.[23] One can construct the classical reals by taking the discrete subspace of R; that is, $\Gamma R = \mathbb{R}$ (p. 319 of [3]). More complex classical objects are constructed similarly, e.g. $(\Gamma R)^{(\Gamma R)} = \mathbb{R}^{\mathbb{R}}$ (p. 320 of [3]). Thus SDG does indeed have enough technical strength to carry out classical analysis, not just smooth analysis.

The discrete subspace axioms also allow one to carry out at least elementary set theory (see [23]), but how does SDG compare to ZFC with regard to higher cardinals? Unfortunately I have not been able to ascertain the answer,[24] but I suspect that it is equivalent in power to Z (ZF without the Axiom Schema of Replacement). My (very rough) reasoning is this. To generate higher cardinalities, we need to be able to iterate the power set operation

[22] While we are generally viewing set theory as a foundational theory, it is also an important branch of mathematics in its own right, and thus SDG must be powerful enough to carry out set theory if it is to provide a foundation for mathematics.

[23] In [20], Linnebo & Pettigrew claim that SDG can only act as a foundation for a small part of mathematics (p. 6). They were not wrong to write this, since they did not include the discrete subspace axioms as part of SDG.

[24] As far as I can see, no one has actually published on this.

along the ordinals, taking unions at limits, as we did with the cumulative hierarchy. In SDG we can use exponentials to mimic the construction of power sets and we have a natural number object, and thus we can get to at least (the equivalent of) V_ω, and thus SDG is at least as powerful as Z. However, I cannot see how we could construct $\omega \cdot 2$, and hence SDG must be weaker than ZF.[25] My reasoning is of course woefully unrigorous; I think the question of the set-theoretic power of SDG would make a fine area of mathematical research.

In conclusion, it would seem that while SDG cannot compare with ZFC in terms of set-theoretic strength, it is capable of carrying out most "everyday" mathematics, such as analysis and arithmetic, and thus it meets the requirement of technical strength to a reasonable degree. Let us move on to the conceptual justification for SDG.

SDG has a conceptual basis, namely that of smoothness and isomorphism. The informal concept of smoothness was made more precise in the form of the Principle of Microstraightness and the Archimedian Method. As we suggested earlier, these are not the only ways to proceed from the informal notion of smoothness. For example, one can consider variants of SDG in which infinitesimals are not nilsqaure, but nilcube, or indeed $\varepsilon^n = 0$ for any n (see p. 224 of [23]). However, a theory's conceptual basis does not have to be unique, and Microstraightness and the Archimedian Method are coherent ways of proceeding from smoothness. Thus SDG does indeed has a reasonable conceptual basis.[26]

Let us lastly consider naturality. The axioms of SDG are clearly parsimonious. SDG is also not pathological, since the axioms assert the existence of smooth objects, the stuff of SDG's coceptual basis.[27]

In conclusion, then, we can see that SDG meets all the criteria laid out in §2.1 and thus SDG can provide a foundation for mathematics.

[25] If we had a notion of Choice in SDG then we could well-order ΓR and thus get to at least V_{ω_1}, but I cannot see how we could go about constructing a notion of Choice in SDG.

[26] At this point I should be honest and remind the reader that I not fully justify the subobject classifier axiom. While this is of course a hole in my thesis, I do not believe it to be a fatal one.

[27] The one problem with this might be the subobject classifier axiom, since one might argue that building a notion of truth into SDG is akin to pathologically smuggling in a higher-order concept. This may be the case, but since I have been unable to justify this axiom conceptually this question of pathology too remains unanswered. I will appeal to the reader's charitable side and suggest that a successful filling of the first hole in my thesis will also fill this second hole (that is, the second hole is a subhole of the first!).

Chapter 5

The autonomy of set theory and synthetic differential geometry

We now come to the second part of my primary thesis, that SDG can provide an *autonomous* foundation for mathematics. Before we apply the criteria from §2.2 to orthodox set theory and SDG, we need to address the distinction between orthodox and naïve set theory. Naïve notions of membership are based on our intuitions and to not appeal to notions specific to *orthodox* set theory. For example, when we say 'SDG contains an object R', we are in no way appealing to ZFC or the cumulative hierarchy. As such, when we consider the different criteria of autonomy, SDG can appeal to naïve notions without being dependent on orthodox set theory.

The logical autonomy of SDG and set theory is the easiest criterion of autonomy to demonstrate. Set theory is formulated in the language of set theory, while SDG is formulated in the language of category theory. While both languages are built from standard first-order logic, the non-logical symbols are different: set theory has $\in$ and only one type of variable (sets), while category theory has $\circ$ (composition) and two types of variable (objects and arrows). Thus, since the very languages in which SDG and set theory are couched are different, the criterion of logical autonomy of the two theories is satisfied.

We turn to epistemic autonomy. The epistemic autonomy of set theory from SDG is readily demonstrated: there are countless mathematicians and philosophers – myself included! – who have learnt set theory without even knowing what SDG stands for. The converse is more involved, however. The issue of the epistemic autonomy of SDG from set theory is complicated by historical developments and pedagogy. Set theory came about over fifty years before SDG did, so by the time SDG was on the scene, all the textbooks were (implicitly or explicitly) written in terms of set theory. Thus, when mathematicians come to learn SDG, they have already learnt set theory and thus it is difficult to determine exactly what role set theory plays in the learning of SDG. However, we can still see that SDG is epistemically autonomous with respect to set theory. For example, if one studies [23], one notices the great care that McLarty has taken great care to develop the whole of topos theory, including SDG, without appealing to *orthodox* set-theoretic notions; indeed, if one examines my exposition one will find the same. Now, in order to understand SDG, one certainly requires an understanding of *naïve* set-theoretic notions, but this does not make SDG epistemically dependent upon orthodox set theory. Thus SDG does indeed have epistemic autonomy from orthodox set theory.

We now come to justificatory autonomy. The conceptual bases for set theory and SDG are quite distinct. Set theory is based upon the cumulative hierarchy, a discrete, (mathematically) concrete structure, while SDG is based upon the notions of smoothness and isomorphism. Figuratively speaking, the two conceptual bases are polar opposites. As such, SDG has justificatory autonomy with respect to set theory (and vice versa).

In conclusion, SDG meets all the criteria of autonomy with respect to set theory and thus provides an autonomous foundation for mathematics. Thus our primary thesis is shown. Let us now consider our secondary thesis, that of mathematical pluralism.

Chapter 6

Conceptual bases and mathematical pluralism

In this last chapter we show that my primary thesis lends credence to a form of mathematical pluralism.

We saw in the previous chapter that set theory and SDG have distinct conceptual bases. Set theory is based upon the discrete and the concrete.[1] Sets live in the cumulative hierarchy, a rigid structure, and the set-theoretic real line is punctiform. SDG, on the other hand, is based upon smoothness, a special case of cohesiveness, and isomorphism. Objects in SDG exist uniquely only up to isomorphism, and in SIA the continuum is a cohesive, non-punctiform whole. These two theories point to two general conceptual dichotomies: discreteness and cohesiveness; and concreteness and isomorphism. These two dichotomies are orthogonal; that is, any one of the four combinations can be embodied in a theory, for example:

	discreteness	**cohesiveness**
concreteness	set theory	intuitionism[2]
isomorphism	ETCS	SDG

This lends evidence to a form of conceptual mathematical pluralism, since different theories capture different conceptual bases. However, one might be able to capture a

[1] By describing a mathematical entity to be *concrete*, we mean that it is determined absolutely, and not just up to isomorphism.

[2] The intuitionistic continuum is *very* cohesive: one can remove countably many points and retain cohesiveness ([6]). Van Dalen describes the intuitionistic continuum as being "syrupy" in nature. This is a nice analogy: no matter how hard you try, you can't cut treacle in two.

concept X in a theory based on concept Y. For example, in set-theory one builds $\mathbb{R}$ out of points, but to reflect conceptions of cohesiveness one imposes the Euclidean topology on $\mathbb{R}$, which is cohesive in this topology sense: one cannot decompose $\mathbb{R}$ into two disjoint *open* sets. Likewise, in SDG one can perform set theory by using discrete subspaces. Thus one might think the choice of conceptual basis is arbitrary. Indeed, why not stick with set theory? Its axioms are straightforward, stated in simple classical logic, not the recondite internal language of a topos. I have two arguments against this, one philosophical and one mathematical.

There are various philosophical reasons why one might wish to explore theories based on different conceptual bases. For example, although we have been avoiding metaphysical issues, one might want to explore the mathematics of a certain ontological position. For example, an atomist might like to explore the mathematics of discrete theories, while a synechist[3] may wish to see how much mathematics one can develop starting from the concept of cohesiveness. Another philosophical reason for exploring different conceptual bases is for the very reason that we can express one concept using another: is it not interesting that one can formulate notions of continuity in set theory, a discrete theory? This brings us to our mathematical argument.

Studying different conceptual bases is, quite simply, mathematically interesting. For example (as I suggested earlier) investigating the set-theoretic power of SDG would make an excellent topic of research, and the study of models of SDG in set theory is well developed (see pp. 224–226 of [23]). Furthermore, recall the many notions of sets discussed in §3.3. It is through the exploration of new concepts that new and exciting discoveries are made; category and topos theory are excellent examples of this.

So there are both philosophical and mathematical reasons for investigating theories based on different concepts. But what about exploring different theories based on the same concept? As we saw with set theory and SDG there are different ways of formalising the same conceptual basis. For example, the Principle of Microstraightness and the Archimedian Method are just one way of proceeding from the concept of smoothness; there are many different axiomatisations of SDG (see p. 224 of [23]). Exploring the different

[3] This is a useful term introduced by Bell in [3]: he defines *synechism* to be the doctrine that nature is continuous, i.e. the opposite of atomism.

ways in which one can turn a concept into rigorous mathematics is of course interesting mathematically, and can also be so philosophically. For example, Hilbert's programme, which was an attempt to base mathematics on a sophisticated synthesis of finitism and formalism, was knocked down in a stroke by a purely technical result, Gödel's incompleteness theorems. Gödel only discovered these beautiful results because of the move in (the philosophy of) mathematics at that time towards formalism, a philosophical concept.

We can see then that there are many reasons, both philosophical and mathematical, for adopting a form of conceptual mathematical pluralism. Sticking to just one theory, be it set theory or SDG, is like viewing a landscape from just one place: to get the full picture, you need to move around.

Chapter 7

Conclusion

We developed various criteria that a theory must meet in order to provide a foundation for mathematics and, furthermore, to provide an autonomous foundation. We applied these criteria to set theory and synthetic differential geometry, demonstrating our primary thesis that synthetic differential geometry can provide a foundation for mathematics autonomous from that provided by ZFC and the cumulative hierarchy. We then drew conclusions from this, arguing for a form of conceptual mathematical pluralism.

Bibliography

[1] Barwise, J. and Moss, L., *Vicious Circles*, Stanford: CSLI Publications, 1996.

[2] Bell, J.L., 'Categories, Toposes and Sets', *Synthese*, vol. 51: pp. 293–337, 1982.

[3] Bell, J.L., *The Continuous and the Infinitesimal in Mathematics and Philosophy*, Milan: Polimetrica, 2006.

[4] Bell, J.L., *A Primer of Infinitesimal Analysis*, New York: Cambridge University Press, 2nd ed., 2008.

[5] Bell, J.L., 'Cohesiveness', *Intellectica*, vol. 51, no. 1: pp. 145–168, 2009.

[6] van Dalen, D., 'How Connected is the Intuitionistic Continuum?', *The Journal of Symbolic Logic*, vol. 62, no. 4: pp. 1147–1150, 1997.

[7] Dummett, M., *Elements of Intuitionism*, New York: Oxford University Press, 2nd ed., 2000.

[8] Eilenberg, S. and Mac Lane, S., 'General theory of natural equivalences', *Transactions of the American Mathematical Society*, vol. 58: pp. 231–294, 1945.

[9] Feferman, S., 'Categorical Foundations and Foundations of Category Theory', *Logic, Foundations of Mathematics, and Computability Theory*, edited by R.E. Butts and J. Hintikka, pp. 149–169, Dordrecht: D. Reidel Publishing, 1977.

[10] George, A. and Velleman, D.J., *Philosophies of Mathematics*, Malden: Blackwell, 2002.

[11] Gleason, A.M., *Fundamentals of Abstract Analysis*, Boston: Jones and Bartlett, 1991.

[12] Hellman, G., 'Mathematical Pluralism: The Case of Smooth Infinitesimal Analysis', *Journal of Philosophical Logic*, vol. 35: pp. 621–651, 2006.

[13] Herrlich, H., *Axiom of Choice*, Berlin: Springer, 2006.

[14] Isaacson, D., 'Arithmetical Truth and Hidden Higher-Order Concepts', *Logic Colloquium '85*, edited by The Paris Logic Group, pp. 147–169, Amsterdam: North Holland, 1987. Reprinted in Hart, W.D. (ed.), *The Philosophy of Mathematics*, New York: Oxford University Press, 1996, pp. 203–224.

[15] Jech, T., *Set Theory*, Berlin: Springer, 3rd ed., 2002.

[16] Kock, A., *Synthetic Differential Geometry*, New York: Cambridge University Press, 2nd ed., 2006.

[17] Kunen, K., 'Elementary Embeddings and Infinitary Combinatorics', *Journal of Symbolic Logic*, vol. 36: pp. 407–413, 1971.

[18] Lawvere, F.W., 'An elementary theory of the category of sets', *Proceedings of the National Academy of Sciences of the USA*, vol. 52: pp. 1506–1511, 1964.

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

[20] Linnebo, Ø. and Pettigrew, R., 'Only up to Isomorphism? Category Theory and the Foundations of Mathematics', unpublished, 2010. Available at <http://seis.bris.ac.uk/~rp3959/papers/onlyuptoiso.pdf> (accessed 4^{th} August 2010).

[21] Mayberry, J., 'What is Required of a Foundation For Mathematics?', *Philosophia Mathematica*, vol. 2: pp. 16–35, 1994.

[22] McLarty, C., 'Defining Sets as Sets of Points', *Journal of Philosophical Logic*, vol. 17: pp. 75–90, 1988.

[23] McLarty, C., *Elementary Categories, Elementary Toposes*, New York: Oxford University Press, 1992.

[24] McLarty, C., 'Exploring Categorical Foundations', *Philosophia Mathematica*, vol. 12: pp. 37–53, 2004.

[25] McLarty, C., 'Learning from Questions on Categorical Foundations', *Philosophia Mathematica*, vol. 13: pp. 44–60, 2005.

[26] Pettigrew, R., 'An introduction to toposes', unpublished, 2008. Available at <http://bristol.academia.edu/documents/0045/9469/toposes.pdf> (accessed 5^{th} September 2010).

[27] Robinson, A., *Non-Standard Analysis*, Princeton: Princeton University Press, 1996. Reprint of the seocond edition (1974). First published 1966.

[28] Robinson, R.M., 'On the decomposition of spheres', *Fundamenta Mathematicae*, vol. 34: pp. 246–260, 1947.

[29] Shapiro, S., *Thinking about Mathematics: the Philosophy of Mathematics*, New York: Oxford University Press, 2000.

[30] Wagon, S., *The Banach–Tarski Paradox*, Cambridge: Cambridge University Press, 1986.

[31] Wilder, R.L., *Evolution of Mathematical Concepts*, New York: John Wiley & Sons, 1968.

[32] Wisdom, J., 'Berkeley's Criticism of the Infinitesimal', *The British Journal for the Philosophy of Science*, vol. 4: pp. 22–25, 1953.