COLIN McLARTY

DEFINING SETS AS SETS OF POINTS OF SPACES

By a "set" we understand any collection into a whole M of definite well distinguished objects m of our intuition or thought. (Cantor [4], 1895)

§0. INTRODUCTION

Cantor's definition of a set as "any collection" has come under intense investigation, occasioned by the paradoxes and leading to type theories and comprehension axioms. His requirement that the elements be "well distinguished," on the other hand, has seemed a near banality. Yet in the dominant nineteenth century conception, while the continuum contained points, the continuum was not made up of well distinguished points. Cauchy, Weierstrass, Dedekind, Cantor, and others made tremendous efforts to understand the continuum and it took most of the century for these efforts to culminate in modelling the continuum by a set.

This paper uses topos theory to formalize a development of sets out of spaces and compares it to nineteenth century developments. In the nineteenth century sets were first conceived as sets of points of spaces, and then various assumptions and discoveries were made relating spaces to their sets of points until eventually spaces could be defined as sets of points with some structure. Then set theoretic thinking displaced geometric intuition in the foundations of mathematics, although neither was rigorously formalized as of 1900. In the present formal development sets are defined precisely by the requirement that their elements be well distinguished: The sentence "$x = y \vee \neg(x = y)$" is true when "x" and "y" are variables ranging over a set and not when they range over any other space.

Two axiom systems are used. CS describes a category of sets, in fact a topos to be called *Set*. The axioms of synthetic differential

Journal of Philosophical Logic **17** (1988) 75–90.

geometry, SDG, describe a topos of smooth spaces and differentiable maps. This topos is called *Spaces*. This version of SDG makes *Set* definable over *Spaces*.

Section one describes toposes briefly. Section two defines *Set*, and section three introduces *Spaces*. Section four defines *Set* as a subcategory of *Spaces*. Section five makes some historical comparisons. Section six wraps up by describing another way to get *Set* from *Spaces*, and shows that in a certain sense the difference between *Set* and *Spaces* depends entirely on the failure of the law of excluded middle in *Spaces*.

§1. TOPOSES IN BRIEF

This paper does not assume familiarity with category theory or toposes. It only requires enough background in logic to make the following précis of topos logic reasonably comprehensible. For fuller treatment of topos logic see [11], or [1] and [24].

A topos can be thought of as a universe of 'sets' and 'functions.' Any model of the Zermelo Fraenkel axioms (or any reasonable set theory) gives a topos where the 'sets' actually are sets, and the 'functions' actually functions. But a topos does not have to be very much like a model of ZF. The standard terminology calls the 'sets' objects, and the 'functions' arrows, and we will use this terminology to avoid prejudging the relation between any given topos and classical set theory.

In a topos any two objects A and B have a product $A \times B$, comparable to the Cartesian product in a model of ZF. Given two arrows $f\colon A \to B$ and $g\colon A \to B$ there is a subobject of A called the equalizer of f and g, and often written $\{a \in A \mid f(a) = g(a)\}$. The notation is correctly suggestive. Think of the equalizer as the 'subset' of A where the 'functions' f and g agree. There is also an object B^A, intuitively the 'set' of all 'functions' from A to B.

In fact, each topos has a kind of set theory suited to it, called the internal set theory of the topos. It is a multi-sorted set theory, where each variable has one of the objects as sort. For an object B there are variables ranging over B and quantifiers $(\forall x \in B)$ and $(\exists x \in B)$ ranging over B. There are no unrestricted variables or quantifiers in the internal set theory of a topos. Sentences in the internal set theory of a topos will always be put in quote marks. The law of excluded middle is not valid in all toposes. That is, in the internal set theories of some toposes there are sentences "φ" such that "$\varphi \vee \neg\varphi$" is not true and neither is "$(\neg\neg\varphi) \rightarrow \varphi$." We will see an example in section three, together with a proof that excluded middle must fail for it. In other toposes, including all models of ZF, excluded middle is valid. The axiom of choice may also fail in a topos, but we will not be concerned with that.

In any topos, proof by contradiction is valid in proving a negation: If "$\varphi \wedge \psi$" implies a contradiction then "φ" implies "$\neg\psi$." So "$\varphi \rightarrow (\neg\neg\varphi)$" is always true.

The main conceptual difference between the topos approach and set theory is that in the topos approach arrows are primitive, not elementhood. For example, a singleton object in a topos, usually called a terminal object, is defined as an object such that every object B in the topos has exactly one arrow to it. Compare singletons in set theory, where every set has exactly one function to each singleton. Every topos has singletons. We can take one singleton, call it 1, and *define* an element of an object B to be an arrow $x: 1 \rightarrow B$ from 1 to B. We write $x \in B$. (These are often called "global elements," as compared with "generalized elements" which we will not use.)

Directly from the definition, 1 has exactly one arrow to each singleton, so each singleton has exactly one element. The terminology is justified. But the definition is much stronger than just saying a singleton has one element. Section three describes an object D in *Spaces* which has at least two arrows to itself. In fact it has infinitely many but we will not prove that. So D is not a singleton. Yet, by the axioms of section four, D has only one element. The situation is simply that the axioms imply there are many arrows from D to D and exactly one arrow from 1 to D. The axioms are consistent with themselves – they are only inconsistent with the idea that an arrow ought to be determined by its effect on elements, the way a function in classical set theory is. Section four gives some intuitive geometric motivation for the properties of D.

An object is empty, or initial, if it has exactly one arrow to every object in the topos. Compare the empty set in a model of ZF. It has exactly one function to each set. It follows in topos theory that if the topos is not trivial (if the objects are not all isomorphic to each other) then an empty object has no elements. But since an object is not defined by its elements, there can be many empty objects in a topos. In general there can also be objects with no elements, which are nevertheless not empty in this sense. But both the axiom systems used here include axioms saying an object is either empty or has an element.

The definition of the equalizer of f and g is also much stronger than saying every element $a \in A$ with $f(a) = g(a)$ is in $\{a \in A | f(a) = g(a)\}$.

Finally, we come to quantifiers. Saying "$(\forall x \in B)\varphi(x)$" is true in the internal set theory of a topos is equivalent to saying "$\varphi(x)$" is true for "x" a variable over B. It is stronger than saying $\varphi(b)$ is true for each element $b \in B$. The sentence "$(\exists x \in B)\varphi(x)$" is weaker than asserting there is some element $b \in B$ with "$\varphi(b)$" true. (See [12] or [25] for the correct interpretation of quantifiers, using generalized elements.) These distinctions collapse in the special case of *Set*.

The topos axioms are given in full in many of the references listed below. [19] is the clearest source, although Mac Lane specializes very quickly to a topos of sets with the axiom of choice.

§2. THE TOPOS OF SETS

The categorial definition of sets was first conceived by Lawvere, [13], and the version given here is based on [28]. The axioms CS consist of the topos axioms plus four more, intuitively true of sets:

(CS_1) An empty object has no elements.

(CS_2) Every object either is empty or has elements.

(CS_3) The law of excluded middle is internally valid.

(CS_4) There is a natural number object – intuitively a set of natural numbers.

Neither of CS_2 and CS_3 implies the other. Omitting CS_2 gives a theory satisfied by any Boolean valued model of ZF. In such a model with a nontrivial Boolean algebra there will be sets A such that $x \in A$ is not true for any x, but for some x its truth value is intermediate between 'true' and 'false.' Such a set A is not empty in our sense, and yet has no elements. All the axioms except CS_3 are satisfied by *Spaces*, and we will see the axioms of *Spaces* are inconsistent with CS_3.

From CS it follows that: For every pair $f: A \to B$ and $g: A \to B$, if for every $a \in A\, f(a) = g(a)$, then $f = g$. So an arrow is fully determined by its effect on elements. Consequently the structure of an object is fully determined by its elements, since in the topos approach all structure is determined by arrows.

So in *Set* we can interpret the internal quantifiers as simply quantifying over elements. A singleton is precisely an object with one element, and equalizers can be defined just in terms of their elements. Any model of ZF gives a model of CS. But a model of CS does not necessarily give a model of ZF. It will give a model for the axioms of extensionality, pair set, sum set, powerset, and infinity, plus a restricted axiom scheme of separation where only bounded quantifiers are allowed. The axiom of choice does not follow from CS and we will not take it as an axiom. Stronger axioms can be formulated. See [13], [23], and [28].

§3. SYNTHETIC DIFFERENTIAL GEOMETRY

Lawvere introduced the project of axiomatizing differential geometry in [14], and [12] gives an excellent introduction. [9] and [22] construct models for the axioms. The axioms SDG begin with the topos axioms. Objects may be called spaces, and arrows may be called maps, since that is how they should be thought of. Elements may be called points. So 1 is a one point space. The axioms call for a space R with specified points $0 \in R$ and $1 \in R$. (So "1" names an arrow 1: $1 \to R$, not to be confused with the singleton space 1.) And there are maps $+: R \times R \to R$ and $\cdot: R \times R \to R$. Think of R as a line with arithmetic structure something like the usual real numbers plus infinitesimals.

(SDG_1) R, 0, 1, $+$, $\cdot$ forms a commutative ring. That is, the usual axioms of commutativity, distributivity, $x + 0 = x$, and so on are true internally for variables over R.

(SDG_2) The ring is nontrivial. $0 \neq 1$.

(SDG_3) R is a field in this sense:
"$(\forall x \in R)[\neg(x = 0) \rightarrow (\exists y \in R) x \cdot y = 1]$" is internally true.

There is a constant zero map from R to R, and a squaring map defined by "$r^2 = r \cdot r$." The equalizer of these is named $D = \{r \in R \mid r^2 = 0\}$. Call D the space of infinitesimals of square zero. The central axiom of SDG is the axiom of line type:

(SDG_4) "$(\forall f \in R^D)(\exists! b \in R)(\forall d \in D) f(d) = f(0) + b \cdot d$"

Intuitively, SDG_4 says every map f from D to R is linear around $f(0)$, with a uniquely determined slope "b." This captures one classical idea of the derivative. In slightly anachronistic terms, the derivative of $g: R \rightarrow R$ at 0 was sometimes defined as the slope of the graph of g on an infinitesimal interval around 0, an interval so small g is linear on it and so has a well defined slope. We take D to be that interval. Then D can be translated along R to define the derivative of g anywhere on R. To find the derivative at $x \in R$ we look at the interval $\{x + d \mid d \in D\}$. In fact, it follows from the topos axioms plus SDG_1 – SDG_4 that every map between any two spaces has a unique derivative defined on the whole domain of the map, the chain rule holds, and the usual rules $(g + h)' = g' + h'$ and $(g \cdot h)' = g' \cdot h + g \cdot h'$ hold for maps $g: R \rightarrow R$ and $h: R \rightarrow R$ and their derivatives g' and h'.

We can prove "$D \neq \{0\}$" by contradiction. Intuitively, we will prove that a map from $\{0\}$ cannot have a uniquely determined slope. Suppose "$D = \{0\}$" is true, or more explicitly "$(\forall d \in D)\ d = 0$." Then trivially "$(\forall d \in D) f(d) = f(0) + 0 \cdot d = f(0) + 1 \cdot d$." By the uniqueness of "b" in SDG_4 we conclude $0 = 1$ and that contradicts SDG_2. So "$D \neq \{0\}$" and the equivalent "$\neg(\forall d \in D)\ d = 0$" are internally true. And D has at least two distinct maps to itself: the constant zero map and the identity map defined by "$I(d) = d$."

On the other hand, "$(\forall d \in D)\neg\neg(d = 0)$" is also true. By SDG_3, since D is a subspace of R, we have "$\neg(d = 0) \rightarrow (\exists y \in R)\ d \cdot y = 1$." By algebra this implies "$\neg(d = 0) \rightarrow (\exists y \in R)\ 0 = d^2 \cdot y^2 = 1^2 = 1$." From the contradiction to SDG_2 conclude "$\neg\neg(d = 0)$." Nothing in D is affirmatively different from 0.

As a consequence, "$d = 0 \vee \neg(d = 0)$" is not true for "d" a variable over D. More generally, since D is a subspace of R, "$x = y \vee \neg(x = y)$" fails for "x" and "y" variables over R. In any topos an object B is called decidable if and only if "$x = y \vee \neg(x = y)$" is true for variables over B. By the axioms of SDG most spaces are not decidable. We will call a space discrete if and only if it is decidable.

Starting with the line R one gets the plane $R \times R$, or R^2, and all the euclidean spaces R^n and subspaces of them defined by equations, plus infinite dimensional spaces such as the function space R^R or more generally B^A for any spaces B and A. The axioms of SDG can be extended to cover integrals, differential equations, continuity and compactness, and so on: [2], [20], [22]. Different extensions may suit different purposes. No such extension is important here except that the axioms of [20], for example, imply *Spaces* has a natural number subject. So we add an axiom:

(SDG_5) There is a natural number object N.

N is provably discrete, as it intuitively ought to be.

§4. DEFINING *SETS* OVER *SPACES*

We need two further axioms:

(SDG_6) Every space either is empty or has points.

This implies for every space B and points $b_1 \in B$ and $b_2 \in B$, "$b_1 = b_2 \vee \neg(b_1 = b_2)$" is true even if B is not discrete and the corresponding sentence with variables over B is not true. (The topos theoretic proof is simple, but technical. Intuitively, SDG_6 says points are the smallest spaces, so two points in any space are either wholly coincident or wholly disjoint.) It follows that $0 \in D$ is the only point of D, since any point $p \in D$ has "$\neg\neg(p = 0)$."

You could picture D as a point and a rate of motion along the line. Then a map f; $D \rightarrow R$ is specified by giving its base point $f(0)$, and a scalar "b" saying how the rate of motion is scaled up or down by f. So if the scalar is negative, the motion is reversed, and so on. Or, if you want to keep actual motion out of it, think of D as just enough of the line to allow the possibility of a rate of motion, but not enough to allow even the possibility of a rate of change in the rate of motion. (There is a space just enough bigger to also make a rate of change in a rate of motion possible; Kock calls it D_2 and uses it to define second derivatives in [12].) This space cannot include two distinct points, since then it would be a finite interval and all sorts of complex motions would be possible in it. See the discussion with intuitive diagrams of D and other infinitesimal structures in [26].

SDG_6 has been studied before, especially in [10], but not for the present purpose. The next axiom is new in SDG:

(SDG_7) Every space M has a unique discrete subspace ΓM such that every point of M is in ΓM.

Call ΓM the set of points of M. The word "set" is justified because the discrete spaces collectively, plus all arrows between them, form a model of CS. (See [21], which also shows which of the usual models of SDG verify SDG_6 and SDG_7.) They form a category of sets, a subcategory of *Spaces*, and we will call this category *Set*.

The law of excluded middle is valid in *Set*, since it is valid in any model of CS. But more than that, in the internal set theory of *Spaces* any sentence of the form "$\varphi \vee \neg\varphi$" or "$(\neg\neg\varphi) \rightarrow \varphi$" is true if all free variables of "φ" range over discrete spaces. "φ" may have quantifiers over any space and constants referring to points of any space. Conversely, if "$\varphi \vee \neg\varphi$" is true for every sentence "φ" whose free variables range over A then let "φ" be "$x = y$" for "x" and "y" variables over A to see A is discrete. In this sense the law of excluded middle is 'valid over discrete spaces' in *Spaces* itself.

Lawvere has maintained that classical logic is associated with constancy – that is, with objects which admit only trivial variation [16]. This applies to discrete spaces, in that a map from any space to a discrete space is locally constant. (This should be intuitively plausible. Compare continuous maps of classical topological spaces. It has technical explications either in terms of sheaf semantics generally [25] or in terms of the usual sites for models of SDG [12], [22].) And here we find classical logic is valid precisely over the discrete spaces of *Spaces*.

§5. GEOMETRY AND THE BEGINNINGS OF SET THEORY

Listen to one of the mathematical pioneers of set theory:

The development of set theory had its source in the effort to produce clear analyses of two fundamental mathematical concepts, namely the concepts of argument and of function. Both concepts have undergone quite essential changes through the course of years. The concept of argument, specifically of independent variable, originally coincided with no further defined, naive concept of the geometric continuum; today it is common everywhere, to allow as domain of arguments any chosen value-set or point-set, which one can make up out of the continuum by rules defined in any way at all. Even more decisive is the change which has befallen the concept of function. This change may be tied internally to Fourier's theorem, that a so-called arbitrary function can be represented by a trigonometric series; externally it finds expression in the definition which goes back to Dirichlet, which treats the general concept of function, to put it briefly, as equivalent to an arbitrary table . . . . It was left for Cantor to find the concepts which proved proper for a methodical investigation, and which made it possible to force infinite sets under the dominion of mathematical formulas and laws . . . .

Sets of infinitely many elements had already been objects of mathematical operations, especially in geometry, where one was long accustomed to comparing sets in respect of their power [*Mächtigkeit*, cardinality]. Yet this practice failed to find more than an external analogy. Even if Cantor, as he incidentally stated, borrowed the notation or the concept of power from Steiner, still the corresponding geometric formulations have little to do with the kind of thought which lies at the base of set theory.

(Schoenflies [27], 1900)

Until the 20th century, mathematicians had little idea of function spaces B^A with differentiable structure; but for finite dimensional spaces, especially R, R^2, and R^3, suitable versions of SDG fit closely to 18th and 19th century ideas. (I will put aside debate on particular theories of infinitesimals.)

Using axioms $SDG_1 - SDG_7$, each space M has an underlying set of points ΓM, and each map between spaces $f\colon A \rightarrow M$ has an underlying function $\Gamma f\colon \Gamma A \rightarrow \Gamma M$. But the underlying sets and functions are far too weakly axiomatized to supply foundations for geometry. The axioms imply ΓR is a field with 0, 1, $+$, $\cdot$ inherited from R. Further axioms, as in [20], imply ΓR is an ordered archimedean field including many irrationals: all algebraic numbers, and definable transcendentals such as π or e. This is the level of Schoenflies' "naive concept of the geometric continuum," approximating the pre-Cauchy view.

Cauchy offered a less naive, somewhat further defined concept of the continuum with his theory of convergence. He gave his well known criterion for convergence, which is important because it does not mention the limit of the series and so does not explicitly presuppose there *is* a limit. He gave a good argument for the necessity of his criterion: If a series of real numbers does converge to a limit then the series meets his criterion. But as to sufficiency he simply said: "when the various conditions are fulfilled, the convergence of the series [to some limit] is assured." ([7], chap. VI) In effect he took completeness of the real numbers as axiomatic. He offered no definition of the real numbers nor any reason for believing his series converge except geometric intuition.

We could add an axiom to SDG matching Cauchy's conception: Every series in ΓR which meets Cauchy's criterion has an element of ΓR as limit. Uniqueness of the limit follows from other axioms as in [20]. But this axiom has to be formulated for ΓR, not R. The theorem "$(d^2 = 0) \rightarrow \neg\neg(d = 0)$" prevents any order relation on R from distinguishing between 0 and infinitesimals of square zero. Convergence can never define a unique limit in R. (See [22] for a study of convergence in context including the infinitesimals of SDG and also those of non-standard analysis.) So Cauchy's theory of convergence led towards set theoretic as opposed to geometric foundations.

Yet the new axiom does not give any independent definition of the set ΓR, just as Cauchy gave no definition of the real numbers. (With hindsight we can say it implies ΓR is isomorphic to the now standard real numbers, defined as equivalence classes of sequence of rationals. But it does not imply that ΓR equals that set.) We still need axioms on R as well as the new one on ΓR.

This half geometric, half set theoretic approach did not suffice for Weierstrass and his followers. So both Dedekind and Cantor worked to describe a set $\mathbb{R}$ of real numbers, working only with discrete collections, using the law of excluded middle; in short, working in *Set* rather than *Spaces*. This is why Zermelo considered Dedekind and Cantor cofounders of set theory [32].

Dedekind began "The Nature and Meaning of Numbers" by explaining he would deal with *things* and that:

A thing is completely determined by all that can be affirmed or thought about it. A thing a is the same as b (identical with b), and b the same as a, when all that can be thought concerning a can also be thought concerning b, and when all that is true of b

can also be thought of *a* . . . . If the above coincidence of the thing denoted by *a* with the thing denoted by *b* does not exist, then are the things *a*, *b* said to be different, *a* is another thing than *b*, *b* another thing than *a*: there is some property belonging to the one that does not belong to the other. ([8], pp. 44–45)

The practical content of this passage is simply that identity is decidable. "$a = b \vee \neg(a = b)$."

Dedekind went on to reject Kronecker's kind of constructivism. ([8], p. 45) At the time, this rejection essentially meant affirming the law of excluded middle. So Dedekind emphasized exactly the logic which distinguishes *Set* from *Spaces* in his foundations for natural numbers and the continuum.

Cantor studied more varied sets of points of spaces, not only the set of all points on the line but points of other spaces and various sets of points distributed along the line. The opening quote for this paper was the first line of his '*Beiträge zur·Begründung der transfiniten Mengenlehre*.' He was more explicit on the logic of set theory in a passage published in 1882:

A manifold (an aggregate, a set) of elements, belonging to whatever sphere of concepts, is called *well defined*, when on the basis of its definition and the logical principle of the excluded third it must be seen as *internally determined*, *both* whether any object belonging to the same sphere of concepts belongs to the said manifold as an element or not, *and also*, whether two objects belonging to the set are identical to each other or not, despite formal differences in the way they are given. ([3], p. 150)

So Cantor also recognized that set theory required the law of excluded middle; and, although he was primarily concerned to say that identity does not depend on how the elements are described, he associated excluded middle with the decidability of identity.

Dedekind and Cantor each defined a set of real numbers, $\mathbb{R}$, with arithmetic structure and an order relation; and both postulated that this represented the set of points on the geometric line with their arithmetic and order relation. In our terms, they defined $\mathbb{R}$ within *Set* and added an axiom $\Gamma R = \mathbb{R}$ plus others for arithmetic and order. The methods used to define $\mathbb{R}$, and to describe its geometric and analytic structure set theoretically, were easily generalized to many other sets and geometric structures. This gave the "quite essential changes" Schoenflies noted in "the concept of argument."

The "even more decisive" change in the concept of function can also be partly modelled in SDG. It was less closely tied to logic than the first change, and it ended up establishing set theory in a sense foreign to the spirit of SDG.

The axioms of [20] plus SDG_6 and SDG_7 imply that for each map $f\colon R \to R$ the function $\Gamma f\colon \Gamma R \to \Gamma R$ is continuous and differentiable in Cauchy's $\varepsilon - \delta$ sense, and so is its derivative and so on. That is, Γf has derivatives of all orders in the $\varepsilon - \delta$ sense. They do not imply that, given a function $g\colon \Gamma R \to \Gamma R$ which has derivatives of all orders, there is some map $f\colon R \to R$ with $\Gamma f = g$. Weierstrass' work required clear definitions of the kinds of functions, he would work with so he took up Dirichlet's idea of a function as an arbitrary pairing of a value $g(x)$ to each value of x; and then he could explicitly restrict his attention to functions continuous in the $\varepsilon - \delta$ sense or functions differentiable in the $\varepsilon - \delta$ sense or whatever. We could match the case where he considered functions with derivatives of all orders, by adding an axiom to SDG saying that the set of maps from R to R, $\Gamma(R^R)$, is precisely the set of all functions from ΓR to ΓR with derivatives of all orders in the $\varepsilon - \delta$ sense. But at this point the apparatus of *Spaces* begins to seem unnecessary. (There are reasons even within set theoretic foundations for wanting a topos of smooth spaces, but they have only recently become clear.) We could just begin with *Set* and $\mathbb{R}$ and the $\varepsilon - \delta$ definitions; and develop geometry and analysis on set theoretic foundatrions. And that is what mathematicians did. But only after some more work was done in set theory.

It is well and good to focus attention on a given subset of the set of all functions from $\mathbb{R}$ to $\mathbb{R}$ but what do we know about that set of functions itself? The axioms of [20] plus SDG_6 and SDG_7 give a good approximation to the knowledge up to Cauchy's time. They say it includes $\Gamma(R^R)$ so it includes polynomials, and sine and cosine and exponential functions, and functions gotten by composing these. It also includes functions defined differently on different parts of $\mathbb{R}$ such as $r(x) = 1$ for x rational and $r(x) = 0$ for x not rational. Some of these, including the example given, were dubious at Cauchy's time but not altogether unthought of. (In *Spaces* it is false to say "$(\forall x \in R)$ x is rational or x is not rational" so the given formula for r cannot define a map from R to R.) At Cauchy's level, using an axiom to make the set of points in R complete, we can define all the bizarre functions the Weierstrassians found by means of convergence. But all these functions have either natural or perverse geometric definitions. We are still far from working with arbitrary functions from $\mathbb{R}$ to $\mathbb{R}$.

The problem was to deal directly with the infinite set $\mathbb{R}$, apart from its geometric meaning. Cantor solved the problem. He brought $\mathbb{R}$ "under the dominion of mathematical formulas and laws" with his transfinite set theory and, as Schoenflies says, "geometric formulations have little to do with the kind of thought which lies at the base of set theory." Cantor put great stress on membership and made functions less important – although he did not reduce functions to the low status that later set theorists would. Set theory had to move on to the axiom of choice, the axiom scheme of replacement, and the continuum hypothesis in an effort to describe $\mathbb{R}$ and the set of functions from $\mathbb{R}$ to $\mathbb{R}$ well enough to give mathematicians confidence in them. The axioms CS can be seen as making set theory a part of geometry, and the definition of *Set* within *Spaces* makes that conception concrete. But that conception is of little use in understanding the formal structure of Zermelo–Fraenkel set theory.

The line came to be *defined* as $\mathbb{R}$ with additional structure. Smooth maps were *defined* to be functions with derivatives of all orders. The set theoretic approach to geometry displaced the synthetic approach, although this happened more slowly than some histories suggest. Sophus Lie was untouched by it in the 1890s ([18], for example), Élie Cartan was nearly so in the 1930s, [5, 6]. Hermann Weyl's great work in mathematics and physics in the 1920s, like Tullio Levi-Civita's, is often remote from set theoretic foundations, making heavy use of infinitesimals, although both men believed $\varepsilon-\delta$ style foundations were better in principle. See [17, 29, 30, 31]. In fact analysis split from geometry when $\varepsilon-\delta$ methods gained ascendency in analysis, around the mid-nineteenth century, while synthetic methods remained common in geometry until the mid-twentieth century. But set theory came to dominate geometry as well.

Along the way infinitesimals disappeared. This is hardly surprising from our point of view since ΓD is a singleton. The only point in the space of infinitesimals of square zero is 0. But the disappearance of infinitesimals is only a symptom of a deeper loss. The independent reality of spaces and maps almost disappeared from mathematical consciousness as everything was reduced to sets. This is not the place to argue the merits of such a reduction, except to note that one of Lawvere's original motives for topos theory was to give a simple, natural, directly geometric foundation for differential geometry.

§6. MORE ON EXCLUDED MIDDLE

We have seen one way of beginning with a universe of smooth spaces, formalized as *Spaces*, and arranging to make the law of excluded middle true: Restrict attention to the discrete spaces, over which the law is already valid, and call this subcategory *Set*. So, for example, *Set*, does not include the infinitesimal space D. This works because of the special axioms SDG_6 and SDG_7. But there is another way of forcing the law of excluded middle in any topos, namely, passing to double negation sheaves [28]. In terms of the internal logic this means not restricting the objects considered but re-interpreting truth so that a sentence "φ" is true in the new sense if "$\neg\neg\varphi$" is true in the original logic. For example, "$\neg\neg(d = 0)$" is true in the internal set theory of *Spaces* if "d" is a variable over D. So we take "$d = 0$" as true in the new sense, and this is equivalent to "$(\forall d \in D)\ d = 0$" so in this sense "$D = \{0\}$." This construction can make objects isomorphic which were not originally isomorphic.

Passing to double negation sheaves creates a new topos which in our case we shall call $Spaces_{\neg\neg}$, but only for a moment. SDG_6 and SDG_7 imply $Spaces_{\neg\neg}$ has the same categorial structure as *Set*. That is, the two are equivalent. This is much stronger than just saying both are models of CS. So we might as well take them to be the same category. Details are in [21]. Then *Set* has two different relations to *Spaces*. It can be embedded into *Spaces* as the subcategory of discrete spaces or as the quite different subcategory of double negation sheaves. These sheaves can be interpreted geometrically as indiscrete spaces in this sense: Given spaces M and L a map $f^\circ: \Gamma M \to L$ from the set of points of M to L may or may not extend to a map $f: M \to L$ depending on whether or not such an extension can be continuous and differentiable. If L is discrete then f° extends to f if and only if it is locally constant. L is indiscrete if every such f° extends to an f in exactly one way. Continuity and differentiability are vacuous constraints when L is indiscrete – compare indiscrete spaces in standard topology.

One conclusion to draw is that the difference between *Set* and *Spaces* hangs directly on the failure of the law of excluded middle. Either way of beginning with *Spaces* and requiring excluded middle winds up with the same category *Set*.

ACKNOWLEDGEMENTS

I thank Charles Wells, Gonzalo Reyes, and Robert Rynasiewicz for working over these ideas with me, and William Lawvere and F. E. J. Linton for encouragement. Michael Resnik and an anonymous referee made very helpful comments on exposition.

REFERENCES

[1] M. Barr and C. Wells, *Toposes, Triples, and Theories*, Springer Verlag, New York, 1985.

[2] M. Bunge and M. Heggie, 'Synthetic Calculus of Variations', in J. W. Gray (ed.), *Mathematical Applications of Category Theory*, Amer. Math. Soc., Providence, 1984.

[3] G. Cantor, 'Über unendliche lineare Punktmannigfaltigkeiten, Nr. 3', reprinted in E. Zermelo (ed.), *Gesammelte Abhandlungen*, Georg Olms, Hildesheim, 1962.

[4] G. Cantor, 'Beiträge zur Begründung der transfiniten Mengenlehre', also in *Gesammelte Abhandlungen.*

[5] É. Cartan, *Leçons sur la théorie des espaces à connexion projective*, Gauthier Villars, 1937.

[6] É. Cartan, *Leçons sur la théorie des spineurs*, Hermann et Cie., 1937; translated as *The Theory of Spinors*, MIT Press, Cambridge, 1960.

[7] A. Cauchy, *Cours d'Analyse* (1821) in *Oeuvres Complètes d'Augustin Cauchy*, 2nd series, volume 3, Gauthier Villars, Paris, 1897.

[8] R. Dedekind, 'The Nature and Meaning of Numbers', in Dedekind, *Essays on the Theory of Numbers*, translated by Wooster Woodruf Beman, Dover Publications, New York, 1963.

[9] E. J. Dubuc, 'Sur les modèles de la géométrie différentielle synthétique', *Cahiers de Topologie et Géométrie Différentielle* **20** (1979), 231 – 279.

[10] E. J. Dubuc and J. Penon, 'Objets compacts dans les Topos', *J. Austral. Math. Soc. (Series A)* **40** (1986).

[11] W. S. Hatcher, *The Logical Foundations of Mathematics*, Pergamon Press, 1982.

[12] A. Kock, *Synthetic Differential Geometry*, London Math. Soc. Lecture Notes 51, Cambridge Univ. Press, 1981.

[13] W. F. Lawvere, 'Elementary Theory of the Category of Sets', mimeographed, Univ. of Chicago, 1963; see also *Proc. Nat. Acad. Sci.* **52** (1964) 1506–1511.

[14] W. F. Lawvere, 'Categorical Dynamics', in A. Kock (ed.), *Topos Theoretic Methods in Geometry*, Matematisk Institut, Aarhus Universitet, Aarhus Denmark, 1979.

[15] W. F. Lawvere, 'Variable Sets Etendu and Variable Structure in Topoi', notes by S. Landsburg, Dept. of Math., Univ. of Chicago, 1975.

[16] W. F. Lawvere, 'Continuously Variable Sets', in *Proc. ASL Logic Colloquium, Bristol 1973*, North Holland Pub., 1975.
[17] T. Levi-Civita, *The Absolute Differential Calculus*, translated by M. Long, Dover Pub., New York, 1977.
[18] S. Lie, *Vorlesungen über Differentialgleichungen*, Leipzig, 1891.
[19] S. Mac Lane, *Mathematics: Form and Function*, Springer Verlag, New York, 1986.
[20] C. McLarty, 'Local, and some global, results in synthetic differential geometry', in A. Kock (ed.), *Category Theoretic Methods in Geometry*, Matematisk Institute, Aarhus Universitet, Aarhus Denmark, 1983.
[21] C. McLarty, 'Elementary axioms for canonical points of toposes', forthcoming in *Jour. Symbolic Logic*, 1987.
[22] I. Moerdijk and G. E. Reyes, forthcoming monograph on models of SDG.
[23] G. Osius, 'Categorical set theory: A characterization of the category of sets', *Jour. Pure and Applied Algebra* **4** (1974), 79–119.
[24] G. Osius, 'Logical and set theoretical tools in elementary topoi', in F. W. Lawvere *et al.* (eds.), *Model Theory and Topoi*, Springer Lecture Notes in Math. 445, Springer Verlag, New York, 1975.
[25] G. Osius, 'A note on Kripke–Joyal semantics for the internal language of topoi', in Springer Lecture Notes 445, Springer Verlag, New York, 1975.
[26] G. E. Reyes, 'Logic and category theory', in E. Agazzi (ed.), *Modern Logic – a Survey*, D. Reidel, Dordrecht, 1980.
[27] A. Schoenflies, 'Die Entwicklung der Lehre von den Punktmannigfaltigkeiten', *Jahrsberichte Dtsch. Math.-Verein* **8** (1900); also in Schoenflies, *Entwicklung der Mengenlehre und ihrer Anwendungen*, B. G. Teubner, Leipzig, 1913.
[28] M. Tierney, 'Sheaf theory and the continuum hypothesis', in I. Bucur *et al.* (eds.), *Toposes, Algebraic Geometry, and Logic*, Lecture Notes in Math. 274, Springer Verlag, New York, 1972.
[29] H. Weyl, *Space, Time, Matter*, translated by H. Brose, Dover Pub., New York, 1952.
[30] H. Weyl, *The Theory of Groups and Quantum Mechanics*, translated by H. P. Robertson, Dover Pub., New York, 1950.
[31] H. Weyl, *The Philosophy of Mathematics and Natural Science*, Princeton Univ., Princeton, 1949.
[32] E. Zermelo, 'Investigations in the foundations of set theory, I', in J. van Heijenoort (ed.), *From Frege to Gödel*, Harvard Univ. Press, Cambridge, 1967.

Department of Mathematics,
Cleveland State University,
Cleveland, OH 44225,
U.S.A.