Introduction to Categorical Foundations for Mathematics

Rough notes after discussion at Roskilde

version of August 14, 2008

this is work in progress and the author makes no warranty express or implied that the views stated herein are either correct or correctly stated

www.case.edu/artsci/phil/mclarty.html

Colin McLarty

There is one topic which I want to talk about very quietly and without hurry. That is something which I never saw even *mentioned* in any popular book about mathematics. I saw algebraic equations mentioned. I probably even saw homotopy groups mentioned, especially in elementary books on topology. But I think that at the heart of 20th century mathematics lies one particular notion and that is the notion of a category.

— 2002 Fields Medalist Vladimir Voevodsky (2002, minute 16)

Patterns themselves are *positionalized* by being identified with positions of another pattern, which allows us to obtain results about patterns which were not even previously statable. It is [this] sort of reduction which has significantly changed the practice of mathematics.

— Philosopher Michael Resnik (1997, p. 218)

I see model theory as becoming increasingly detached from set theory, and the Tarskian notion of set-theoretic model being no longer central to model theory. In much of modern mathematics, the set theoretic component is of minor interest, and basic notions are geometric or category-theoretic.... It seems to me now uncontroversial to see the fine structure of definitions as becoming the central concern of model theory, to the extent that one can easily imagine the subject being called "Definability Theory" in the near future.

— Model theorist Angus Macintyre (2003, p. 197)

Contents

Preface

This book presents one categorical foundation for mathematics while assuming no prior knowledge of category theory. Whenever a categorical term is used without definition, as for example in the Introduction, a definition can be located using the index. Mathematical examples occur throughout the book with intuitive motivation (and often pictures) to give a general idea, and with references to fuller accounts for readers who want them. The reader is assumed to understand predicate logic such as

$$\forall x\ (Px \rightarrow Qx)$$

meaning everything with property P has property Q, and such naive set theoretic notation as

$$\{x \in S \mid Px\}$$

to name the set of all members of the set S which have the property P. Those are the prerequisites for the categorical foundation. Some passages comparing this foundation to Zermelo-Fraenkel set theory with choice, abbreviated ZFC, assume basic knowledge of ZFC.

Introduction

In mathematics rigor is not everything, but without it there is nothing. A demonstration which is not rigorous is nothingness. I think no one will contest this truth. But if it were taken too literally, we should be led to conclude that before 1820, for example, there was no mathematics; this would be manifestly excessive.

— Henri Poincaré (1908, p. 171)

Pure mathematics was discovered by Boole, in a work which he called *The Laws of Thought*. This work abounds in asseverations that it is not mathematical, the fact being that Boole was too modest to suppose his book was the first ever written on mathematics.

— Bertrand Russell (1917, p. 59)

1. How Foundations Develop

It is a mistake to believe that mathematics begins with any explicit foundation whether it is George Boole's logic (1854), or Gottlob Frege's *Basic Laws of Arithmetic* (1893), Ernst Zermelo's set theory (1908), or the stronger Zermelo-Fraenkel set theory (Zermelo, 1930). Yet mathematics gains very practical advantages from the theoretical attempt to give foundations. Zermelo's project was part of David Hilbert's Göttingen program to make mathematics more rigorous and unified. Though few mathematicians ever learned the axioms they became the standard in principle for mathematical definition and proof. This enabled freer, more confident, and more consistently correct use in practice of the new abstract algebra and analysis. Anyone who compares advanced mathematics texts from before, say, 1910, to those after about 1940 will see the advantages in clarity and precision.[1]

Yet Zermelo's 1908 theory had two basic defects illustrated by familiar examples:

(1) Encoding all mathematical objects as Zermelo sets requires many technicalities irrelevant to the actual workings of the theory. For example it requires defining the number 2 as either $\{\{\emptyset\}\}$ or $\{\emptyset, \{\emptyset\}\}$ or some set

[1] A challenge to those who claim 19th century mathematics was more accessible than 20th: Compare Lie (93) to the first two volumes of Spivak (1971) which cover far more material than Lie more quickly, with more pictures and examples, clearer explanations, and vastly more accurate theorems. Far from denying Lie's genius this supports his own belief that his geometric insight was obscured by the formalisms available at his time.

built up from the empty set $\emptyset$ by repeated set formation. The particular choice of set is irrelevant but in principle some choice must be made.

(2) It could not handle larger sets with the freedom that set theorists wanted. It could prove there are infinitely many successively larger infinite sets $\aleph_0, \aleph_1, \ldots$ but it could not prove there is one set collecting them all

$$\{\aleph_0, \aleph_1, \ldots, \aleph_n, \ldots \ \forall\ n \in \mathbb{N}\}$$

The first problem has generated more philosophical discussion but the second was more pressing in practice. Zermelo had met the logical demands of classical analysis, but not the new demands coming from set theory itself! Powerful further axioms incorporated into his (1930) produced Zermelo-Fraenkel set theory with choice, ZFC, which handled point 2 well enough that set theorists have not generally accepted any further, stronger axioms since then. See the history in Maddy (1988) and Kanamori (2004).

This stronger set theory left point 1 untouched, which was a minor inconvenience in itself, but increasingly unnecessary. The growth of modern mathematics naturally produced rigorous new well adapted tools unimagined when Zermelo gave his axioms. And even point 2 remained a problem outside of set theory:

(2′) ZFC could not handle large collections like the category of sets **Set**, and the category **Grp** of groups which algebraists and topologists and others began using in the 1940s and 1950s.

Those categories are proper classes. They are the size of the set of all sets and so they are not legitimate entities in ZFC. Sometimes these can be taken as mere façons de parler, mere ways of speaking, to be eliminated from precise formal statements. In other words, some statements about “the category of all groups” can be replaced by equivalent statements which merely talk about “every group,” or perhaps “every set of groups.” But this is not always possible. We will see (for example in chapter??) statements which quantify over proper class sized categories and so cannot be made in ZFC. We must either extend the foundations, or else refuse to give foundations for ideas that are central to the current organization of many parts of mathematics.

2. Foundations and the Question of Consistency

Foundations of mathematics are not just organizing tools of course. They have also been the topic of philosophic and conceptual debate, especially in the decades after Frege’s logic (1893) turned out to be inconsistent. Russell showed Frege that the logic led to a contradiction now called “Russell’s paradox,” while Hilbert and his circle discovered it independently and called it “Zermelo’s paradox.” Hilbert was impressed that persuasive reasoning by a good mathematician could lead to contradiction. The paradoxes led him “to the conviction that traditional logic is inadequate and that the theory of concept formation needs to be sharpened and refined” as a serious matter for the practice of mathematics (quoted in (Peckhaus, 2004, p. 501)). So he assigned Zermelo the problem of axiomatizing an adequate set theory with the results we have just discussed.

For a while there was quite broad debate among mathematicians: Can mathematics be reduced to pure logic as Dedekind, Russell and other logicists hoped? Or is logic per se beside the point, while mathematics rests on mental construction as Luitzen Brouwer claimed? Should we somehow avoid using the actual infinite, as finitists urge? These philosophical questions are still argued today but not here. This book lies inside the essentially Hilbertian consensus that has governed mathematical practice since the 1930s—that is not Hilbert's proof theoretic program (Hilbert, 1926) but his and his school's working method in practice. Foundations of mathematics here will not rest on any theory of epistemology, happily enough, since as Howard Stein (1988, p. ??) has said "find the quote about how we have no adequate epistemology of any kind of knowledge anyway." This book is unconcerned with constructivity, but see (McLarty, 2006b). We are far from any finitism. We resolutely maintain that modern, mainstream, abstract, axiomatic mathematics is correct, and more than correct it is beautiful. The goal of this foundation is taken from both Saunders Mac Lane and William Lawvere, namely to organize current practice in its own terms and not to critique it.

We understand a *foundation* for mathematics to be an explicit axiomatic theory with axioms taken to be true (and thus necessarily consistent) which suffices to interpret and prove the usual theorems of mathematics. A *good* foundation should also be perspicuous and should use as nearly as possible the usual concepts and methods of working mathematics. The CCAF axioms here are offered not as merely consistent, and thus interpretable in some way or ways, but as true in the sense that they correctly describe the categories and functors used in current mathematics. Following Hilbert we expect a foundation to secure the consistency of mathematics; but following Gödel we accept that consistency cannot be proved. We can do two things: We can take the clarity and formal success of axioms as evidence that they are consistent. And we can prove some foundations consistent *relative to* others. Notably, simple arguments (given where ??) prove our foundation CCAF is consistent if ZFC and one Grothendieck universe is. From the viewpoint of ZFC a Grothendieck universe is an inaccessible cardinal, thus a "large" cardinal whose existence does not follow from the ZFC axioms themselves, but rather small as large cardinals go. Few people doubt these foundations are consistent.

3. Overview

Chapter 1 introduces category theory using Riemann's complex analysis as a great 19th century example of the morphism-based practice articulated today in category theory. Chapter 3 looks at modern model theory as an example where category theory works in many ways in practice: On one level categories organize the data about any given model theoretic structure. Second, a "structure" can actually be defined as such a category of data. And third, this way of organizing model theory relates it to other parts of mathematics, including *tame topology* but also many other methods of geometry and number theory. One thing I hope this chapter achieves is to overcome the too-common idea among philosophers that new mathematics generally consists of hard, heavy machinery far removed from the basic simple ideas. Progress in mathematics is often a matter of unifying and simplifying. Chapter **??** shows how that unification is going on today in number theory and geometry, with toposes as an example of special interest in logic and philosophy. Curiously, philosophers used to give number theory as an example of mathematics that does not use categories and functors, when in fact toposes and most of current category theory were created in the successful pursuit of great results in number theory.

The next seven chapters lay out formal axioms for a categorical foundation for mathematics with especially *structuralist* features:

(1) Elements exist only in structures and have only structural properties.
(2) A category **Set** of sets adequate to classical mathematics is placed in a category of categories adequate to current functorial practice.

These first order axioms in the language of category theory suffice as a *formal foundation* to define all the concepts and prove all the results of current mathematics. Besides that, though, no matter what formal foundation one prefers, the category of sets and the category of categories deserve study since they pervade the *working foundations* of current mathematics. For example chapter 5 defines a set of sets as a function the way geometers define a space of spaces as a map.

The remaining chapters relate the axiomatization to philosophical issues in mathematics. Theorem schemes in chapter 12 show exactly what is meant by "structural" properties. Chapter 13 relates Maddy's naturalism to a goal not considered in (Maddy, 2007). That is the unification of working methods all across mathematics which has been so productive for the past 60 years. This unification is both the source of categorical foundations and the reason why mathematicians abandoned Bourbaki's *structures*. Bourbaki founder Jean Dieudonné remarked almost forty years ago now that their theory "has since been superseded by that of category and functor, which includes it under a more general and convenient form" (Dieudonné, 1970, p. 138).[2]

Make sure to explain somewhere (whether in this chapter or not, or even several places) how circumlocution can eliminate large categories for some purposes but not for all: the definition of functor category, for example, quantifies over functors. So does the definition of derived functor in cohomology.

On Feferman and Hellman insisting that categorical axioms do not mean what they say. All of this can be interpreted in Zermelo-Fraenkel set theory (plus inaccessible cardinals as needed), but if you see categorical definitions as merely a disguise for that interpretation then you will not understand them, as argued already in §7 "consequences of common sense history" in (McLarty, 1990, pp. 365ff.).

[2]The review MR0069544 (16,1048d) of Chevalley (1955) is explicit about Weil's concern (also reported to me by Pierre Cartier) that his abstract varieties are not "sets of points" (so they do not fall under his own theory of structures). Chevalley p. 3 discusses one reason for Weil's concern—that the image of a variety is not a variety (but in my terms sort of a singular variety with multiplicities) and uses this to introduce constructible sets. But Cartier also discussed Weil's concern that his varieties do not form *single* sets but are given by many.

Part 1

Some Mathematical Motivations

CHAPTER 1

Morphism-based Mathematics

> Modern geometric philosophy holds firmly to the notion that the first thing one does after defining the objects of interest is to define the functions of interest. In our case the objects are Riemann surfaces, and we have already addressed complex-valued functions on Riemann surfaces. However "functions" are to be taken also in the sense of mappings between objects; once we define such mappings we will have a *category* of Riemann surfaces.
>
> — Rick Miranda, *Algebraic Curves and Riemann Surfaces*, p. 38.

Category theory articulates a sense of mathematics which goes back at least to the early 19th century yet is foreign to most 20th century philosophies of mathematics and foundations for mathematics—a more organized sense of the arithmetic, analytic, and geometrical forms in mathematics than appears in formal logic or set theory. Working mathematics does not live in a universe (from the Latin *universus* "turned into one") of sets with structures encoded in them. It lives in a cosmos (from the Greek κοσμέω "to order or arrange") of *categories*, where each category is one galaxy of structures characterized by their mappings or morphisms to each another; and just as galaxies are organized among themselves in the large scale structure of the cosmos so each category is characterized by the mappings called *functors* to and from other categories. This order has a much more intrinsic foundation than Zermelo-Fraenkel set theory. It can be rigorously axiomatized in the very terms that mathematicians use to work with it every day.

Special branches of mathematics often have their own traditional terms such as *morphism* or *map* or *homomorphism.* General category theory says more abstractly *objects* for structures, and *arrows* for morphisms, to emphasize the patterns they form rather than any image of what they are. Standard working definitions in practice characterize each structure not uniquely but *up to isomorphism* by its place in some pattern of maps, and we will give a foundation for mathematics entirely in terms of these patterns. Ontologically, mathematical objects need have no properties beyond their relations to one another given by the maps. Category theory unifies the working methods with the foundations and the ontology in one insight: "It's the arrows that really matter!" (Awodey, 2006, p. 8).

The first comparison of categories to galaxies was in J.L. Kelley's influential textbook *General Topology* (1955). Kelley was already in school (9 years old in 1925) when Edwin Hubble found the first persuasive evidence of any galaxies other than our Milky Way—evidence which also showed the universe is expanding—so this was an auspicious image for Kelley. He gave some exercises showing how to associate with each topological space X an algebraic structure, namely $C(X)$ the *algebra of all continuous real-valued functions on* X. For reasonably nice topological spaces X and Y, the continuous functions $f: X \to Y$ correspond exactly to the algebra homomorphisms in the opposite direction $C_f: C(Y) \to C(X)$ so that algebra and topology prove facts about one another. Kelley saw this as a further step along the line from the *local* study of what happens near some point in a space, to the *global* study of the space as a whole. This next step relates whole patterns of spaces and maps to whole patterns of algebraic structures and their maps. Kelley called it *galactic* and gave category theory as its context (1955, p. 245). William Lawvere read this in 1959 as an undergraduate, and a science fiction fan, and liked both the mathematics and the imagery.

As one influential 19th century example, Richard Dedekind unified and extended earlier number theory by greater use of mappings in algebra. But in foundations he also defined the natural numbers themselves by their mapping properties among sets. His foundations suited his practice as he realized there is no reason to worry about what the natural numbers *are*. What matters is how they map to themselves and other sets.

Riemann Surfaces make a good early 19th century example of mappings solving concrete problems. Indeed Riemann (1851) never says what the points of any Riemann surface are. He says how parts of each surface are mapped to the complex plane.[1] In fact he used two versions of maps: the broad class of all *continuous maps* and its very restrictive subclass of *holomorphic maps*. He was motivated by multiple-valued functions such as the logarithm and algebraic roots. Our chapter title is a little misleading in that topology did not really exist prior to Riemann surfaces. Riemann took major steps towards creating the field of topology as a tool for handling his surfaces.[2]

Calculating with square roots is a little tricky because every number that has one has two. If you are talking about the square root of a positive real number then you can distinguish the two: the positive and the negative. But when you go to the complex numbers there is no more positive and negative, and it takes much more care to distinguish the square roots. Cube roots are simpler over the real numbers, as each real number has single real cube root. Higher roots over the real numbers are no worse: every real number $x \in \mathbb{R}$ has a single real n-th root $\sqrt[n]{x}$ when n is odd, and every positive real has two $\pm\sqrt[n]{x}$ when n is even. Higher roots are much more complicated over the complex numbers: every complex number $z \in \mathbb{C} \neq 0$ has n different complex n-th roots with no algebraic way to distinguish any one of these as basic. All are algebraically interchangeable.

But everyone knew that when you look at the square root of a variable number it was little better than it might have been: If you pick one of the square roots $\sqrt{z_0}$ of any non-zero complex number $z_0 \in \mathbb{C}$ then every z closed to z_0 has exactly one square root $\sqrt{z}$ close to $\sqrt{z_0}$. The other square root $-\sqrt{z}$ close to $-\sqrt{z_0}$. Riemann organized this by taking any small region U in the complex plane, not including

[1]Compare current algebraic geometry where "[In many schemes] the points ... have no ready to hand geometric sense.... When one needs to construct a scheme one generally does not begin by constructing the set of points.... [While] the decision to let every commutative ring define a scheme gives standing to bizarre *schemes*, allowing it gives a *category of* schemes with nice properties" (Deligne, 1998, pp. 12–13).

[2]See among many historical sources Pont (1974), Ferreirós (1999, pp. 53ff.), and Gray (1998).

0, and thinking of the square root function as defining two copies of U lying as "sheets" over U. Each sheet contains one square root $\sqrt{z}$ for each $z \in U$ taken so that all the square roots in that sheet are near each other. Then all other square roots are all near each other, and form the other sheet. The *division into* sheets is natural, in any one small region that does not include 0, even though each sheet is intrinsically indistinguishable from the other, and when you take the complex plane as a whole the two sheets are connected. They form a single surface such that if you start in one sheet over one region and move continuously once around the complex plane (i.e. along an itinerary that goes around 0) you return to the other sheet over the same region. Make the trip again and you end up back in the original sheet. Looking at n-th roots in place of square roots gives the same kind of thing except that the single surface lies in n sheets over each little region of the complex plane.

This nice picture of n-th roots generalizes easily to the roots of any complex polynomial equation in two variables

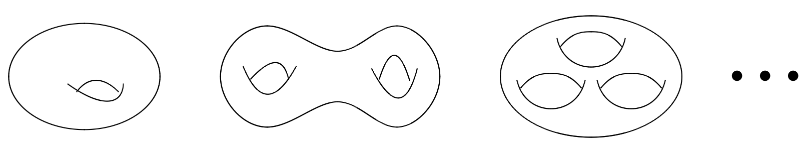
$$P(z,w) = 0 \qquad \text{with } P(z,w) = 0 \in \mathbb{C}[z,w]$$

If you are not familiar with this it is useful to think of z as a complex parameter: each value $z_0 \in \mathbb{C}$ determines a point on the complex plane, and over that point a polynomial $P(z_0, w)$ in the single variable w. So there is a finite set of roots over each point, and Riemann organizes those roots into a multi-sheeted covering as described. The Riemann surface of the square root comes from taking w to be the square root of z:

$$w^2 - z = 0$$

This picture gets practical force when we compare the roots of one polynomial $P(z,w)$ to those of another $Q(z,w)$ by mapping the Riemann surface of $P(z,w)$ to that of $Q(z,w)$. Randomly pairing roots of $P(z,w)$ with roots of $Q(z,w)$ would reveal very little—but continuous and/or holomorphic maps reveal a great deal.

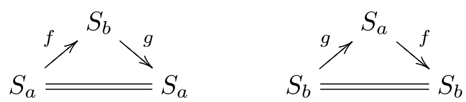
Topology takes the continuous functions as maps. From this viewpoint Riemann surfaces are elastic surfaces namely the sphere, the torus (or "doughnut surface"), the two-holed torus, the three-holed and so on for any finite number of holes.



The number of holes is the *genus*, and Riemann knew the surfaces are *topologically isomorphic* if and only if they have the same genus.[3] To put it in visual terms, two surfaces S_a and S_b can be stretched and bent to fit each other if and only if they have equally many holes. The precise statement is in terms of maps: if and only if S_a and S_b have the same genus there are continuous maps $f\colon S_a \to S_b$ and $g\colon S_b \to S_a$ *inverse* to each other in the sense that they cancel out. Mapping S_a onto S_b by f and then mapping S_b back to S_a by g maps each point of S_a back to

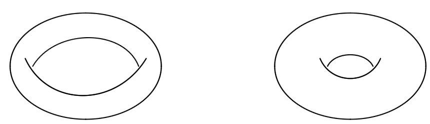
[3]How and when this was first proved depends on what you take as a proof.

itself; and conversely:



Riemann used mutually canceling maps to show when two surfaces have "the same structure," or as we put it today are *isomorphic*. 19th century mathematicians did this for many kinds of structures. But they took the mapping property as merely *witnessing* the sameness of structure and not *defining* it. They defined a distinct notion of "sameness" for each distinct kind of structure. "There was great confusion: the very meaning of the word 'isomorphism' varied from one theory to another" (Weil, 1991, p. 120). Category theory made the general notion explicit in the single most widely adopted use of the central insight: Isomorphism today *means* mutually canceling maps. Each different kind of map implicitly a corresponding notion of isomorphism or "sameness of structure" and so implicitly of "structure."

Continuous maps define topological structure. Holomorphic maps define the more rigid *holomorphic* or *analytic* structure. For example each genus one Riemann surface has a complex number parameter measuring, so to speak, its "stoutness."[4]



All genus one surfaces are topologically isomorphic (some *continuous* maps between them cancel out). To be analytically isomorphic (some *holomorphic* maps between cancel out) they must have the same parameter. Every holomorphic map between genus one surfaces implies an algebraic relation between the parameters and this became a productive link between complex analysis and number theory.

The plethora of functions, maps, morphisms, etc. is unified in the idea of a category as a network of *arrows* between *objects* using neutral terms "object" and "arrow" just to avoid the associations of any specific example. Sets and functions form one category:

> **Set:** The category of sets. Its objects are sets and its arrows $f : S \to S'$ are functions between sets.

Riemann surfaces are the objects of two important categories:

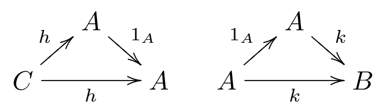
> **TopSurf:** The topological category of Riemann surfaces. Its objects are the Riemann surfaces and arrows are continuous maps between them.
>
> **HolSurf:** The holomorphic category of Riemann surfaces. Its objects are the Riemann surfaces and arrows are holomorphic maps between them.

So mathematicians do not talk about "the" structure of a Riemann surface but about various *structures* on it. The *topological* structure is revealed by its place in the category **TopSurf**. The *holomorphic* or *analytic* structure, which Miranda (1995) calls the "complex structure," is revealed by its place in **HolSurf**. Other aspects of structure are revealed by its place in other categories, e.g. the "C^∞ structure" as in Miranda (1995, p. 5) relating Riemann surfaces to other differentiable

[4]Geometrically this means the proportionate thickness plus a measure of how much a complex coordinate "twists" when it goes once around the torus.

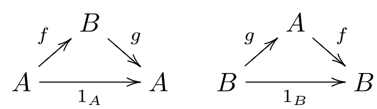
manifolds. Important relations among these aspects of structure are revealed by functors between these and related categories.

Each object A in any category has an *identity arrow* $1_A : A \to A$ which intuitively leaves A unchanged. The precise definition says it leaves arrows to and from A unchanged: any arrow $h : C \to A$ followed by 1_A gives back h, while 1_A followed by any $k : A \to B$ gives k:



Identity arrows lead to the most important definition in category theory:

DEFINITION 1.1. An arrow $f : A \to B$ is an *isomorphism* if it has an *inverse*, that is an arrow $g : B \to A$ with $gf = 1_A$ and $fg = 1_B$. Objects A, B are *isomorphic* if there is some isomorphism $f : A \to B$.



This bare abstract definition captures concrete notions of "same structure" all across mathematics. In the case of **Set** isomorphic sets have the same cardinality. All genus one Riemann surfaces are isomorphic in the topological category **TopSurf** and this is just a way of saying they have all the same topological properties. They are not all isomorphic in the holomorphic category **HolSurf**. They do not have all the same analytic properties. Isomorphic objects in any category have all the same properties insofar as that category is concerned.

In the 1920s and 1930s Emmy Noether shaped modern mathematics with her *homomorphism* and *isomorphism* theorems (McLarty, 2006a). Using previously unrecognized ideas from Dedekind she showed the power of defining individual algebraic structures—not uniquely but up to isomorphism—by the homomorphisms between them. Her methods rapidly spread from number theory to topology and led to the creation of category theory by Samuel Eilenberg and Saunders Mac Lane. Category theory spread all across geometry, analysis, and even foundations (Krömer, 2007).

Explain minutes 16–23 of Voevodsky's talk

CHAPTER 2

The Eilenberg-Mac Lane Axioms

The Eilenberg-Mac Lane category axioms give the bare language of the insight, formalized in two-sorted first order logic in Figure 2. Less formally, the objects and

Types: $A, B, C\ldots$ for objects, and $f, g, h\ldots$ for arrows.

Operators:

Dom takes arrows to objects, read "domain of."
Cod takes arrows to objects, read "codomain of."
$1_{\_}$ takes objects to arrows, read "identity arrow of."

Relation: $C(x, y; z)$ applies to arrows, read "z is the composite of x and y."

Axioms:

Domain and codomain: $\forall\ f, g, h,$ if $C(f, g; h)$ then

$$\mathrm{Dom}\, f = \mathrm{Dom}\, h \quad \text{and} \quad \mathrm{Cod}\, f = \mathrm{Dom}\, g \quad \text{and} \quad \mathrm{Cod}\, g = \mathrm{Cod}\, h.$$

Existence and uniqueness of composites:

$$\forall\ f, g, \quad \text{if } \mathrm{Cod}\, f = \mathrm{Dom}\, g \text{ then } \exists!\ h \text{ such that } C(f, g; h).$$

Identity arrows: $\forall\ A,\ \mathrm{Dom}\, 1_A = \mathrm{Cod}\, 1_A = A$. And

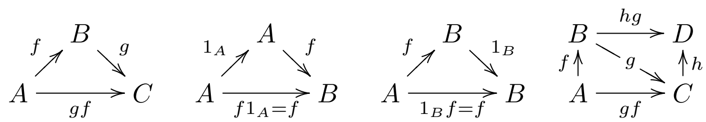
$$\forall\ f,\ C(1_{(\mathrm{Dom}\, f)}, f; f) \text{ and } C(f, 1_{(\mathrm{Cod}\, f)}; f).$$

Associativity of composition:

$$\forall\ f, g, h, i, j, k, \text{ if } C(f, g; i) \text{ and } C(g, h; j) \text{ and } C(f, j; k) \text{ then } C(i, h; k).$$

FIGURE 1. the abstract category axioms

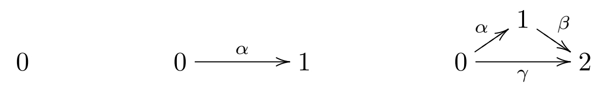
arrows of any one category **C** satisfy these conditions: Every arrow f goes from a unique object A to a unique object B. Let $f : A \to B$ say f is an arrow from A to B and then call A the *domain* of f and B the *codomain*. When the codomain of $f : A \to B$ is the domain of $g : B \to C$ then they have a *composite* $gf : A \to C$.[1] Each object A has an *identity* arrow $1_A : A \to A$ defined by $f 1_A = f$ and $1_B f = f$ for every $f : A \to B$. The last axiom says composition is associative: $(hg)f = h(gf)$.



[1]The text and later axioms treat composition as a partially defined binary operator rather than a relation $C(f, g; h)$. Officially gf abbreviates a definite description $(\iota h)(C(f, g; h))$.

These axioms are understood as abstract and open to any interpretation. Any entities and operations that satisfy them form a category.

Three finite categories called $\mathbf{1}, \mathbf{2}, \mathbf{3}$ which naively look like this:



1 **2** **3**

Each lightfaced numeral represents an object and its identity arrow. The category $\mathbf{1}$ has a single object 0 and only its identity arrow. The category $\mathbf{2}$ has two objects $0, 1$, their identity arrows, and one non-identity arrow as shown. The category $\mathbf{3}$ has three objects, and arrows as shown.

Build the visual idea of a category as a vast network of arrows, as used at the start of the chapter on functors and natural transformations.

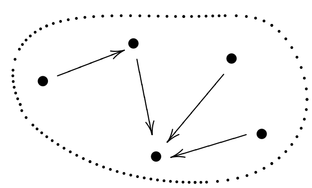


Figure 2. A general category pictured as a vague network of arrows.

On his slides (sometimes, but not in the case of beginning and end points of a path) Voevodsky used the left-to-right order of composition, writing $fg\colon N \to K$ where we would write $gf\colon N \to K$ for the composite of $f\colon N \to M$ and $g\colon M \to K$. This is the more logical order and many categorists have used it at one time or another. But before category theory was invented generations of mathematicians wrote composite functions in the other order—so for example $(\log \circ \cos)(x)$ means first take the cosine of x and then the logarithm of that. Most category theorists have ended up going along with that and so do we.

Alternative “arrows only” axioms.

CHAPTER 3

Model Theory as the Geography of Tame Mathematics

There is now a reasonably coherent sense of what it means to understand a structure: it means understanding the category of definable sets (including quotients by definable equivalence relations).

— Anand Pillay in (Buss et al., 2001, p. 186)

It is a deep insight gained through decades of progress in model theory that to understand a model theoretic structure is to understand its definable subsets. It is merely obvious that to understand the subsets requires using the definable functions between them, as everyone who studies the subject does. And the category of definable sets is just another name for the definable sets with their definable functions.

This model theory is part of a great trend in geometry since the 1930's articulating how all of the many technically different formalizations of geometry—from topological manifolds through algebraic varieties to finite cell complexes and many more—agree on all but the weird cases: "as long as one only considers non-pathological examples it all comes to the same thing" (Voevodsky, 2002, minute 4). The single key achievement here is Alfred Tarski's classical result on *quantifier elimination* in elementary algebra.[1] Here elementary algebra means the first order theory of the real numbers as an ordered field

$$\langle \mathbb{R}, +, \times, <, 0, 1 \rangle$$

This theory deals with the algebra of real numbers, polynomials with real coefficients, equations and order. It can even quantify over polynomials of any fixed degree. It expresses a claim about the existence of some cubic polynomial

$$\exists \text{ cubic polynomial } P(x) \quad \text{such that } (\ldots\ P(X)\ \ldots)$$

by spelling out the polynomial:

$$\exists a\ \exists b\ \exists c\ \exists d \quad \text{such that } (\ldots\ aX^3 + bX^2 + cX + d\ \ldots)$$

This theory cannot quantify over arbitrary subsets of the real numbers or over arbitrary functions from the real numbers to themselves. It cannot even quantify over arbitrary polynomials since no one sentence in this language can quantify over arbitrarily long lists of coefficients.

[1] See Tarski (1951), van den Dries (1998, introduction), and the very helpful Marker (2000).

Tarski gave a decision routine for elementary algebra by *quantifier elimination*. He worked out an algorithm for taking any formula in the first order theory of the real field and eliminating the quantifiers one-by-one to get a equivalent but quantifier-free formula without adding any free variables. For example, a real number x has a square root

$$\exists y\ (y^2 = x)$$

if and only if x is either zero or strictly positive

$$0 = x\ \vee\ 0 < x$$

The two formulas are equivalent, and the quantifier over y is eliminated in the second. When the algorithm is applied to a sentence (a formula with no free variables) then the quantifier-free result is also a sentence. In fact it is either $1 = 1$, which shows the original sentence is provable, or else $0 = 1$ which shows the original sentence is provably false. This gives a decision routine (albeit far from efficient) for elementary algebra, and proves a certain simple set of axioms for the theory is complete.

But then, instead of Tarski's emphasis on sentences, look at formulas $\varphi(x)$ with one free variable x. These define subsets of the real numbers.

$$\{x \in \mathbb{R} \mid \varphi(x)\}$$

Tarski's result shows that any subset of the real numbers which can be defined by a formula at all can be defined by a formula without quantifiers. That means every definable set is defined by finitely many equations and inequalities, so it consist of the union of finitely many single points and open intervals. In this sense an interval may be unbounded, as for example the open interval of all strictly positive numbers

$$\{x \in \mathbb{R} \mid 0 < x\}$$

The single points may include endpoints of intervals as for example the union of that open interval with the single point $\{0\}$ is the closed interval of non-negative numbers

$$\{x \in \mathbb{R} \mid 0 \leq x\}$$

Of course set theory can define much wilder subsets of the real line, including ones with all kinds of difficult properties that early 20th century logicians called "pathologies" such as non-measurable subsets. Tarski's result shows these complexities are impossible for the *definable* subsets in elementary algebra.

In the 1980s van den Dries focussed on the fact that this restriction on the definable subsets of $\mathbb{R}$ implies very strong—and very nice—restrictions on definable subsets of $\mathbb{R}^n$ for all dimensions n. Let us look at a few definable subsets of R^3 in this structure. One is the unit sphere in 3-dimensional space, that is the 2-dimensional surface of the unit ball in $\mathbb{R}^3$:

$$\{\langle x, y, z\rangle \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$$

By using inequality the open unit ball is definable, that is the 3-dimensional interior of the unit ball:

$$\{\langle x, y, z\rangle \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 < 1\}$$

The union of these two is the 3-dimensional closed unit ball:

$$\{\langle x, y, z\rangle \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq 1\}$$

More complicated definitions define more complicated sets, but van den Dries saw (in far more detail than we can give here) that *because* the 1-dimensional definable sets are so simple, the higher dimensional ones cannot be very complicated. These sets in all dimensions are accessible to geometric intuition in a way that arbitrary set-theoretic subsets are not. Of course a central tool of geometry is mappings suited to the sort of space involved, which here means the mappings are the definable functions. Even from a strictly logical point of view, taking these sets as "data types" with no special regard to any geometrical character, the chief use of data types is to have relevant functions between them—again the definable functions.

For example consider this definable subset of $\mathbb{R}^2$:

$$\{\langle x, y\rangle \mid x^2 = y^2 \ \&\ 0 \leq y\ \}$$

It defines a function from $\mathbb{R}$ to $\mathbb{R}$, because for every $x \in \mathbb{R}$ there is a unique non-negative y with $x^2 = y^2$. Namely

$$y = \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x \leq 0. \end{cases}$$

In other words, this relation defines the absolute value function.

$$y \ = \ |x|$$

Trivially, the identity function on any definable set is definable. The identity function on the whole of $\mathbb{R}$ is defined by the equality relation $x = x$. And the composite of definable functions is definable: Suppose a relation Rxy assigns to each x in some definable subset S a unique y in definable subset S'. So it gives a definable function $r\colon S \to S'$. Suppose Qyz assigns to each y in S' a unique z in definable subset S'', giving a definable function $q\colon S' \to S''$. Then the relation

$$\exists y\ (Rxy \ \&\ Qyz)$$

obviously assigns to each x in S a unique z in S''. So it defines a composite function

$$qr\colon S \to S''$$

A routine check shows this composition operation is associative, so it satisfies the Eilenberg-Mac Lane axioms and thus produces a category—namely, the category of definable sets and functions of elementary algebra.

$\mathbf{Def}_{\mathbb{R}}$: The category of definable sets and functions of elementary algebra. Its objects are the definable subsets of $\mathbb{R}^n$, for all dimensions n, and arrows are the definable functions between them.

People who study real algebra have used these ideas since before Tarski, and well before category theory existed. More advanced study gives successively stronger reasons to use the theory.

1. O-minimal Structures

Van den Dries radically generalized Tarski's proof and achieved both of the advantages that mathematicians look for in successful generalizations: his framework applies to many more cases that Tarski's original theorem, and it focusses on just the simplest most relevant features of the original case. The technical advantages of the generalization are beyond the scope of this book, but the format of it is quite simple: instead of Tarski-type models, define a structure to be a suitable category of sets and functions. This is part of what Macintyre means saying "I see model theory as becoming increasingly detached from set theory, and the Tarskian notion of set-theoretic model being no longer central to model theory" (2003, p. 197).

To give the basic definition forget about the real numbers $\mathbb{R}$ and forget the algebra of $0, 1, +, \times$ and just take any *ordered set* $R, <$. That is, any set R and relation $<$ on R such that

$$\forall x \in R\ \neg(x < x) \qquad \text{and} \qquad \forall x, y, z \in R\ ((x < y\ \&\ y < z) \to x < z)$$

Following van den Dries (1998, pp. 13 and 17) define a *structure* on $R, <$ to be a family $\mathcal{S}$ of sets and functions with the properties laid out in Figure 1.

(S1) Each set A in $\mathcal{S}$ is a subset of R^n for some n. Whenever $A, B \subseteq R^n$ are both in $\mathcal{S}$ then so are the intersection $A \cap B$, the union $A \cup B$, and the complement

$$R^n \backslash A \;=\; \{\langle x_1, \ldots, x_n\rangle \in R^n |\ \neg(\langle x_1, \ldots, x_n\rangle \in A)\}$$

(S2) If $A \subseteq R^n$ is in $\mathcal{S}$, then so are these subsets of R^{n+1}

$$\{\langle x_0, x_1, \ldots, x_n\rangle |\ \langle x_1, \ldots, x_n\rangle \in A\}$$
$$\{\langle x_1, \ldots, x_n, x_{n+1}\rangle |\ \langle x_1, \ldots, x_n\rangle \in A\}$$

(S3) For any choice of i, j between 1 and n this subset of R^n is in $\mathcal{S}$

$$\{\langle x_1, \ldots, x_n\rangle |\ x_i = x_j\}$$

(S4) If $A \subseteq R^n$ is in $\mathcal{S}$, then so is this subset of R^{n-1}

$$\{\langle x_1, \ldots, x_n, x_{n-1}\rangle |\ \exists x_n(\langle x_1, \ldots, x_n\rangle \in A)\}$$

(S5) The set $\{\langle x, y\rangle \in R^2 |\ x < y\}$ is in $\mathcal{S}$

(S6) If $A \subseteq R^n$ and $B \subseteq R^m$ are in $\mathcal{S}$, then a function $f\colon A \to B$ belongs to $\mathcal{S}$ if and only if this subset of R^{n+m} does

$$\{\langle x_1, \ldots, x_n, y_1, \ldots, y_m\rangle |\ \langle y_1, \ldots, y_m\rangle = f\langle x_1, \ldots, x_n\rangle\}$$

FIGURE 1. the structure axioms

Simple reasoning shows that the identity function on any subset in a structure $\mathcal{S}$ belongs to $\mathcal{S}$, as does the composite of any two functions that belong to $\mathcal{S}$. So the sets and functions of any structure $\mathcal{S}$ form a category. In fact van den Dries (1998) skips the word "category" but he uses it in (2000).

The definable subsets of any Tarski-type model that includes an order relation $<$ meet all these conditions. So van den Dries also defines *model theoretic structures* in Tarski's terms and uses these as one way to "generate" structures in his more general sense, but not the only way (1998, p. 21).

A structure $\mathcal{S}$ on $R, <$ is *o-minimal* if every subset of R in $\mathcal{S}$ is a finite union of single elements $\{a\} \in R$ and intervals, where an interval is a subset of one of these forms for some $a, b \in R$

$$\{x \in R |\ a < x\} \quad \text{or} \quad \{x \in R |\ a < x < b\} \quad \text{or} \quad \{x \in R |\ x < b\}$$

The idea is that "o-minimal" means "order minimal" in the sense that the structure allows just the kind of subsets of R that must exist given the order.

Van den Dries and others showed that all o-minimal structures have very nice intuitive geometrical properties, avoiding pathology not only in subsets of R but in the subsets of all the R^n. When an o-minimal structure $\mathcal{S}$ on an ordered set $R, <$ includes addition and multiplication functions that make R a ring, or a field, then all subsets in the structure have a natural geometric sense analogous to the case of the real field $\mathbb{R}$ above. And conversely important results in geometry and number theory had been cast in these terms even before model theorists defined o-minimality, as described in the following Section 2.

What we have seen so far merely makes category theory a convenient language for what model theorists say anyway, whether they use the word "category" or not. It is very light-weight category theory. Yet convenience is vital to the growth of theories and when Pillay wrote on the future of model theory he addressed it:

> Foundations of model theory: What is the right language and level of generality for model theory? The traditional framework of one-sorted structures and their point-sets has long been recognized as being rather restrictive. The actual practice of model-theorists is somewhat more in line with points of view from categorical logic. (in (Buss et al., 2001, p. 187))

The many definable sets were used in practice as definable types.

Discuss quotient types, which Shelah (1978) dubbed "imaginaries." Shelah did not speak of categories but others around him immediately saw them there. The technical triviality opened the way to some useful, and some substantial results as summarized in the MR review by Lascar:

> In Section 6, the strikingly simple and beautiful idea of imaginary elements appears. In this kind of study, there is no harm in changing T as long as the category of models of T remains the same (this remark allowed Morley to assume quantifier elimination). Here, for each definable equivalence relation, the models are enlarged by a set whose elements witness the different classes. These are imaginary elements. With their help, the canonicity theorem becomes true; it implies, for example, that for any complete type p over a model M, there is a least set $A \subseteq M$ such that p is definable with parameters in A. Notice the analogy and the actual link with Weil's theorem stating that for any ideal I in $K[X_1, X_2, \cdots, X_n]$ (K a field) there is a least subfield K' of K such that I is generated by polynomials with parameters in K'. Lascar (1981)

The categorical logic in question is essentially Makkai and Reyes (1977) which takes categorical versions of the same tools from the Grothendieck school's characterization of toposes, and relates them to issue in model theory.

Philosophic conclusion: "structures" are not single sorted. And as to what a "structure" *is*, as between Tarski model and category of sets for example, 'there is no ontology here.'

2. Geography and Topology

Actually, Tarski's theorem had a major impact on geometry and abstract algebra already in the 1950s. It led Claude Chevalley (one of the few French mathematicians of the time interested in logic) to extend quantifier elimination to all Noetherian rings.

which led to the geometry of "semi-analytic sets," that is subsets of $\mathbb{R}^n$ locally definable by inequalities of analytic functions. These results began with Lojasiewicz proving all such subsets of $\mathbb{R}^n$ are *triangulable*, they can be partitioned as polyhedra.[2] In effect all the tools of algebraic topology apply in the most naive way.

Those geometric results had an initially separate heritage in Alexander Grothendieck's program for *tame topology.* Here put Hrushovski on the broad view of model theory as the geography of tame math.

> If one wants some kind of meaningful interaction with other parts of mathematics, it is the conviction that this is a worthwhile intellectual enterprise, rather than the desire to make a 'splash', which is crucial. This conviction amounts essentially to a belief in the unity of mathematics. There has been much discussion of this 'unity of mathematics' in recent times, often in connection with deep conjectures relating arithmetic, geometry, analysis, representation theory etc. One feels moreover that logicians, especially in the light of their foundational concerns, should have some level of engagement with these issues and conjectures" (Anand Pillay in (Buss et al., 2001, p. ??)).

2.1. General points on o-minimality. From a logician's point of view the key point here is that you cannot define the sets $\mathbb{N}$ of natural numbers or $\mathbb{Z}$ of integers in any o-minimal theory, and so cannot interpret arithmetic in it. The smallest "structure" on the real field that contains $\mathbb{Z}$ is the collection of projective sets—given as a well-known result by (van den Dries, 1998, p. 16). Projective sets are so general as to evade the axioms of ZFC in many ways: two quoted from Harvey Friedman at http://www.math.ohio-state.edu/ friedman/pdf/CCRTalk1.121905.pdf

> i. Do uncountable sets have perfect subsets [a closed subset with no isolated points]? This is provable for Borel sets, and even analytic using a tiny fragment of ZFC. But "every uncountable coanalytic set has a perfect subset" is independent of ZFC. Also "every uncountable projective set has a perfect subset" is independent of ZFC.
>
> ii. Are the sets Lebesgue measurable? Trivial for Borel sets, and provable for analytic and coanalytic sets using a tiny fragment of ZFC. But "every PCA set is Lebesgue measurable" is independent of ZFC. Also "every projective set is Lebesgue measurable" is independent of ZFC.

[2]Lojasiewicz proved every locally finite collection of semi-analytic subsets can be simultaneously triangulated, there is a triangulation of the total space which restricts to a triangulation on each space in the collection.

O-minimal theories need not be decidable. For example (Macintyre-Wilkie 1993) prove Schanuel's conjecture implies that the theory of the real exponential field is decidable. This theory does not have quantifier elimination but elimination down to existential formulas. The conjecture is: for any complex numbers $z_1, \ldots, z_n$ linearly independent over $\mathbb{Q}$, the extension field

$$\mathbb{Q}(z_1, \ldots, z_n, \exp(z_1), \ldots, \exp(z_n))$$

has transcendence degree at least n. Zilber axiomatizatized "pseudo-exponentiation" in algebraically closed fields of characteristic zero and proved that the theory is satisfiable, and categorical in all uncountable powers. If Schanuel's conjecture is true, then exponentiation on $\mathbb{C}$ is the unique model of continuum cardinality.

The historic relations between Tarski's result, real algebraic geometry, Lojasiewicz and Hironaka, and stratification as studied by Thom are clearly laid out (1979) in MR0477121 (57 #16665) reviewing Hironaka, Heisuke *Introduction to real-analytic sets and real-analytic maps*, Quaderni dei Gruppi di Ricerca Matematica del Consiglio Nazionale delle Ricerche.

Part 2

Categorical Set Theory

CHAPTER 4

The Elementary Theory of the Category of Sets

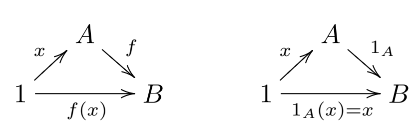
See Lawvere (1964, 1965).

William Lawvere's axioms ETCS, for the *Elementary Theory of the Category of Sets*, were first publicly presented at the 2nd International Congress of Logic, Methodology, and Philosophy of Science, Jerusalem 1964.[1] Taken as a foundation they are not abstract but describe sets and functions. They begin with the Eilenberg-Mac Lane category axioms applied to sets and functions: every function f goes from a unique set A to a unique set B, every set A has an identity function $1_A : A \to A$, and so on. Figure 1 gives the ETCS axioms beyond the category axioms.[2]

Intuitively, a set 1 is a singleton if and only if every set T has exactly one function to it. Here a subtle point must be made: categorically this says that ! is a *terminal object* in the category of sets but note carefully that we have not yet introduced the category sets as a legitimate entity. So far as the ETCS axioms are concerned (like the ZFC axioms) the "category of all sets" is a figure of speech, a way to indicate the environment we work in, and not a legitimate entity. The stronger CCAF axioms in Chapter ?? will posit the category **Set** of sets as a legitimate entity, but the ETCS axioms we are consider now do not posit that.

For now, the key point about 1 is that a set A has as exactly many elements $x \in A$ as there are functions $x : 1 \to A$ from a singleton 1 to A: each element x corresponds to the function $x : 1 \to A$ taking the sole element of 1 to the element x of A, and vice versa.

To be precise, ETCS defines a singleton as a set such that every set has exactly one function to it; and an axiom says there is a singleton 1. ETCS defines an *element* of a set A as a function $x : 1 \to A$ and we often write this as $x \in A$. A function $f : A \to B$ takes elements of A to elements of B as each $x \in A$ composes with f to give $f(x) \in B$. This means that identity functions act as expected:



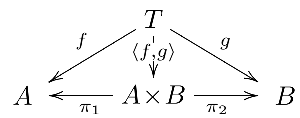
[1]Lawvere (1964, 1965) and see the listing in (Bar-Hillel, 1965, p. 437).

[2]Here 1 must be a constant to define the formulas $x \in S$. The axioms may use operators for products, projection functions, et c. or not as discussed in (McLarty, 1991b, p. 68).

There is a singleton 1: $\forall S\ \exists!\ S \to 1$

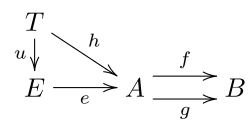
Every pair of sets A, B has a product:

$$\forall T, f, g \text{ with } f\colon T\to A,\ g\colon T\to B,\ \exists!\ \langle f,g\rangle\colon T\to A\times B$$



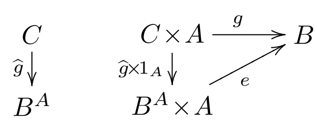
Every parallel pair of functions $f, g\colon A\to B$ has an equalizer:

$$\forall T, h \text{ with } fh = gh\ \exists!\ u\colon T\to E$$



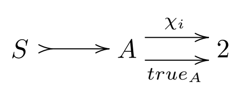
There is a function set from each set A to each set B:

$$\forall\ C \text{ and } g\colon C\times A\to B,\ \exists!\ \widehat{g}\colon C\to B^A$$



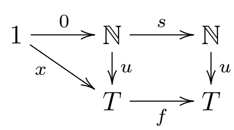
There is a truth value $true\colon 1\to 2$:

$$\forall\ A \text{ and monic } S\rightarrowtail A,\ \exists!\ \chi_i \text{ making } S \text{ an equalizer}$$



There is a natural number triple $\mathbb{N}, 0, s$:

$$\forall\ T \text{ and } x\colon 1\to T \text{ and } f\colon T\to T,\ \exists!\ u\colon \mathbb{N}\to T$$



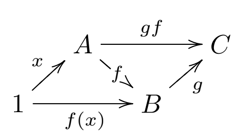
Extensionality: $\forall\ f\neq g\colon A\to B,\ \exists x\colon 1\to A$ with $f(x)\neq g(x)$.

Non-triviality: $\exists\ false\colon 1\to 2$ such that $false\neq true$.

Choice: $\forall$ onto function $f\colon A\to B,\ \exists\ h\colon B\to A$ such that $fh = 1_A$.

FIGURE 1. the ETCS axioms

And composition follows the familiar rule $(gf)(x)\ =\ g(f(x))$:



A key feature of sets is *extensionality*. Membership-based set theories say a set is determined by its elements. Categorical set theory says a function is determined by its effect on elements. That is, if $f \neq g$ are both functions $A \to B$ then they are distinguished by at least one element $x \in A$:

$$\text{If } f \neq g\colon A \to B \text{ then there is some } x\colon 1 \to A \text{ with } f(x) \neq g(x)$$

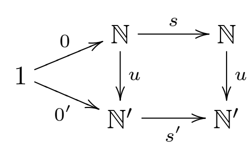
A simple example of the typical ETCS proof technique proves that all singleton sets are isomorphic.

THEOREM 4.1. *All singletons* 1 *and* $1'$ *are isomorphic.*

PROOF. If 1 and $1'$ are singletons there are functions $u\colon 1 \to 1'$ and $v\colon 1' \to 1$. The composite vu is a function $1 \to 1$, as is the identity 1_1. Since 1 is a singleton there is only one function $1 \to 1$, so $vu = 1_1$. Similarly $uv = 1_{1'}$. □

Only a little less simple is the categorical axiom of infinity first given by Lawvere (1963, reprint p. 36). This axiom posits a set $\mathbb{N}$ with an element $0 \in \mathbb{N}$ and a *successor* function $s\colon \mathbb{N} \to \mathbb{N}$ such that: For any set T and element $x \in T$ and function $f\colon T \to T$ there is a unique $u\colon \mathbb{N} \to T$ such that $u(0) = x$ and $us = fu$. See the diagram in Figure 1. The triple $\mathbb{N}, 0, s$ is the natural number structure—the set $\mathbb{N}$ by itself is not enough. Nor is the triple $\mathbb{N}, 0, s$ defined uniquely. Many different choices will support this kind of recursion. But it is defined uniquely *up to isomorphism* in this sense:

THEOREM 4.2. *Suppose* $\mathbb{N}, 0, s$ *satisfy the axiom of infinity, and so do* $\mathbb{N}', 0', s'$*. Then* $\mathbb{N}$ *and* $\mathbb{N}'$ *are isomorphic. Indeed there is just one isomorphism* $u\colon \mathbb{N} \to \mathbb{N}'$ *such that* $u(0) = 0'$ *and* $su = u's$*.*

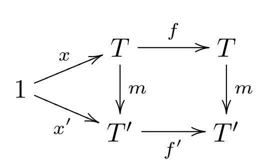


PROOF. By assumption there are functions $u\colon \mathbb{N} \to \mathbb{N}'$ with $u(0) = 0'$ and $s'u = us$, and $v\colon \mathbb{N}' \to \mathbb{N}$ with $v(0') = 0$ and $sv = vs'$. The composite $vu\colon \mathbb{N} \to \mathbb{N}$ has $vu(0) = 0$ and $(vu)s = s$, but $1_{\mathbb{N}}\colon \mathbb{N} \to \mathbb{N}$ also has $1_{\mathbb{N}}(0) = 0$ and $1_{\mathbb{N}}s = s$. By uniqueness $vu = 1_{\mathbb{N}}$. Similarly $uv = 1_{\mathbb{N}'}$. □

Define an *iterative structure* as any set T with a selected initial value $x \in T$ and an iteration function $f\colon T \to T$

$$1 \xrightarrow{x} T \xrightarrow{f} T$$

And define a map of iterative structures to be a function between the sets $m\colon T \to T'$ which gets along with the initial values and iteration functions in this sense:



$$m(x) = x' \quad \text{and} \quad mf = f'm$$

These maps form a category of iterative structures:

Iter: The category of iterative structures. Its objects are the iterative structures and arrows are the maps between them.

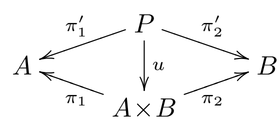
The axiom of infinity defines $\mathbb{N}, 0, s$ as the universal iterative structure, meaning it has exactly one map to every iterative structure. Compare the opposite universal property of 1 among sets. Each set has exactly one function to 1. This exact reversal of arrows explains why the proofs of theorems 4.1 and 4.2 are so much alike.

Use of the natural numbers $\mathbb{N}$ depends on further set construction axioms. The *product* of sets A and B is defined to be a set $A \times B$ together with projection arrows $\pi_1 : A \times B \to A$ and $\pi_2 : A \times B \to B$ such that for every set T with arrows to each set $f : T \to A$ and $g : T \to B$ there is a unique function $\langle f, g\rangle : T \to A \times B$ such that

$$\pi_1\langle f, g\rangle = f \quad \text{and} \quad \pi_2\langle f, g\rangle = g$$

See the diagram in Figure 1. An ETCS axiom says each pair of sets A, B has a product $A \times B, \pi_1, \pi_2$. It is unique up to isomorphism:

Theorem 4.3. *If a set P and functions $\pi'_A : P \to A$ and $\pi'_B : P \to B$ also have the product property then P is isomorphic to $A \times B$; and there is exactly one isomorphism $u : P \to A \times B$ with $\pi'_1 = \pi_1 u$ and $\pi'_2 = \pi_2 u$.*

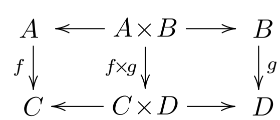


Proof. Use the product property of P, π'_1, π'_2 to define a function $v : A \times B \to P$ and follow the earlier proofs. □

Corollary 4.1. *Every set A is isomorphic to the product $1 \times A$.*

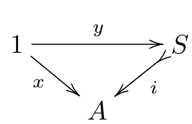
Proof. Trivially, A plus the unique function $! : A \to 1$ and the identity $1_A : A \to A$ as projections have the product property. □

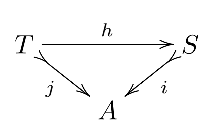
For future reference, any two functions $f : A \to C$ and $g : B \to D$ have a product function $f \times g : (A \times B) \to (C \times D)$ defined by



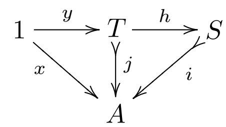
$$\pi_C(f \times g) = f\pi_A \quad \text{and} \quad \pi_D(f \times g) = g\pi_B$$

Any one-to-one function $i : S \to A$ gives a *subset* of A. We often write $i : S \rightarrowtail A$ to show that i is one-to-one. An element $x \in A$ *belongs* to the subset, written $x \in i$, if it factors through i. That is, if there is some $y \in S$ such that $x = i(y)$. Another subset $j : T \rightarrowtail A$ is included in i, written $j \subseteq i$ if there is some $h : S \to T$ such that $j = ih$. To put it in graphics:





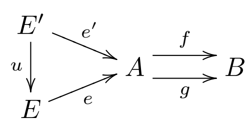
It follows immediately that $x \in j$ and $j \subseteq i$ imply $x \in i$:



Conversely, given $i\colon S \rightarrowtail A$ and $j\colon T \rightarrowtail A$, if every $x \in i$ is also in j then $i \subseteq j$. But that proof uses the truth value axiom along with extensionality. If $i \subseteq j$ and $j \subseteq i$ then i and j give the same subset of A and we write $i \equiv j$.

An *equalizer* $e\colon E \to A$ for a parallel pair of functions $f, g\colon A \to B$ is a universal solution to the equation $fe = ge$. Precisely: $fe = ge$, and for any function $h\colon T \to A$ with $fh = gh$ there exists a unique $u\colon T \to E$ with $h = eu$. See the diagram in Figure 1. The case $T = 1$ shows that e is one-to-one, so it gives a subset of A, and indeed the (possibly empty) subset of all solutions to the equation $f(x) = g(x)$. For every $x \in A$, $x \in e$ if and only if f(x)=g(x). An axiom says that every parallel pair of functions has an equalizer. They are unique up to isomorphism:

THEOREM 4.4. *If* $e'\colon E' \to AB$ *is also an equalizer for* $f, g\colon A \to B$ *then there is exactly one isomorphism* $u\colon E' \to E$ *with* $e' = eu$.



PROOF. Obvious following the earlier proofs. □

The truth value axiom says there is a set 2 of truth values, defined by the property that every subset of a set A has a *characteristic function* which takes all and only the elements of the subset to *true*. More fully, there is a set 2 with an element $true \in 2$ such that for every set A and subset $i\colon S \rightarrowtail A$ there is a unique function χ_i making S an equalizer.

$$S \rightarrowtail A \underset{true_A}{\overset{\chi_i}{\rightrightarrows}} 2$$

Here $true_A$ is the constant function

$$A \longrightarrow 1 \xrightarrow{true} 2$$

which takes every element of A to *true*. So $i\colon S \rightarrowtail A$ is, up to equivalence, the subset of all $x \in A$ such that $\chi_i(x) = true$. The axioms up to here are consistent with supposing *true* is the only truth value, so that all equations are true and every set is a singleton. This is blocked by the non-triviality axiom:

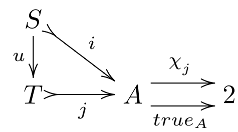
There is a truth value $false\colon 1 \to 2$ with $false \neq true$.

It follows from the other axioms that $true, false$ are the only elements of 2.

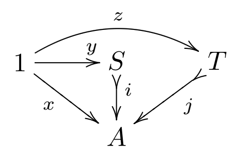
Our most sophisticated proof will be:

THEOREM 4.5. *For any set* A*, subsets* $i\colon S \rightarrowtail A$ *and* $j\colon T \rightarrowtail A$ *have* $i \subseteq j$ *if and only if every* $x \in i$ *is also in* j.

PROOF. Because j is an equalizer, $i \subseteq j$ is equivalent to $\chi_j i = true_A i$



By extensionality that is equivalent to having $\chi_j i(y) = true_A i(y)$ for all $y \in S$. By the equalizer property that is equivalent to saying every $y \in S$ has some $z \in T$ with $j(z) = i(y)$ or in other words every $x \in i$ is also in j.



□

The last set construction axiom posits, for any sets A, B, a *function set* or *exponential* B^A with an *evaluation function* $ev: B^A \times A \to B$ with this property: for any set C and function $g: C \times A \to B$ there is a unique function $\widehat{g}: C \to B^A$ with $ev(\widehat{g} \times 1_A) = g$. See the diagram in Figure 1. As a basic example elements $\widehat{f} \in B^A$ correspond to functions $f: 1 \times A \to B$. Since $1 \times A$ is isomorphic to A, elements of B^A correspond to functions $A \to B$. In particular the elements of 2^A correspond to functions $A \to 2$ and so to subsets of A. It is the *power set* of A.

The axiom of choice is stated in ETCS exactly as it is in many mathematics texts. A function $f: A \to B$ is *onto* if for each $y \in B$ there is at least one $x \in A$ with $y = f(x)$. The axiom of choice says every onto function $f: A \to B$ has at least one right inverse, that is a function $g: B \to A$ with $gf = 1_B$.

Problem Session: Well-founding and the Axiom of Foundation

THIS SECTION IS VERY MUCH IN-PROGRESS AND I DO NOT GUARANTEE THAT ANY PART OF IT IS EITHER CORRECT OR USEFUL.

We come to a technical axiom which did not occur to Cantor, to Zermelo (1908), or even to Fraenkel. Set theorists agree it is "irrelevant to mathematics," a point developed at length by (Kunen, 1983, p. 94). It rarely occurs to mathematicians outside of research Zermelo-Fraenkel set theory yet it is central to that. It is

> irrelevant for the development of ordinal and cardinal numbers, natural and real numbers, and in fact of all ordinary mathematics. However, it is extremely useful in metamathematics of set theory, in construction of models. In particular, all sets can be assigned ranks and can be arranged in a cumulative hierarchy. (Jech, 2006, p. 63).

The idea of the *cumulative hierarchy*, that is the idea that all sets are sets of sets built up by transfinite iteration from the empty set, has become central to ZF set theory. It does not occur in ordinary mathematics and also not in categorical foundations. Yet the idea of rank and the related constructions of models are easily available.

0.2. Well-founded relations. The concept of *well-founded relation* is not irrelevant at all and occurs throughout mathematics though not always by name. For example it appears in *Noetherian* conditions in algebra and topology.

A *partial order* $\leq$ on a set S is any reflexive, antisymmetric transitive relation. That is, for all $x, y, z \in S$:

- $x \leq x$ (reflexivity)
- if $x \leq y$ and $y \leq x$ then $x = y$ (antisymmetry)
- if $x \leq y$ and $y \leq z$ then $x \leq z$ (transitivity)

For any partial order on a set $S, \leq$ we write $x < y$ to mean $(x \leq y \ \& \ x \neq y)$. A *total* order, or *linear order* is any partial order which is also total. For all $x, y \in S$:

- either $x \leq y$ or $y \leq x$

The usual "less than or equal to" relations of mathematics are partial orders and in fact total orders. The inclusion relation $\subseteq$ between subsets of a set S is a partial order on the powerset 2^S but not total (if S has more than one element). Clearly, for any $U, V, W \subseteq S$:

- $U \subseteq U$
- if $U \subseteq V$ and $V \subseteq U$ then $U = V$
- if $U \subseteq V$ and $V \subseteq W$ then $U \subseteq W$

But if there are distinct elements $x, y \in S$ then neither subset $\{x\}$ and $\{y\}$ includes thei other. Another useful partial, not total, order is the "divides" relation. For natural numbers n, m write $n|m$ to say n divides m. For any $m, n, p \in \mathbb{N}$:

- $n|n$
- if $n|m$ and $m|n$ then $m = n$
- if $m|n$ and $n|p$ then $m|p$

A more advanced example combines inclusion and divisibility. The ideals of a ring R are partially ordered by the opposite of inclusion: an ideal $\mathcal{I} \subseteq R$ *divides* another $\mathcal{J} \subseteq R$ if and only if $\mathcal{J} \subseteq \mathcal{I}$. Note the order of inclusion. To say 3 divides 6 is the same as saying the set of multiples of 6 is included in the set of multiples of 3.

A *partial order* $\leq$ on a set S is *well-founded* if every non-empty subset of S has at least one minimal element, that is at least one element such that no other element of the subset is "less" than it. To put it in concisely:

$$\forall U \subseteq S \ (\ \exists x \in U \ \Rightarrow \ (\ \exists x \in U \ \forall y \in U \ (x \leq y \ \Rightarrow \ x = y) \ )$$

The "less than or equal to" relation on the real numbers is not well-founded as for example the subset of strictly positive reals

$$\{x \in \mathbb{R} \mid 0 \lneq x \ \}$$

has no minimal element. The inclusion relation on subsets of the natural numbers is not well-founded as for example the set of all infinite subsets

$$\{U \subseteq \mathbb{N} \mid U \text{ is infinite } \}$$

has no minimal element.

The divisibility order on the natural numbers is well-founded as every number theory textbook notes on the way to proving unique prime factorization: in any set of natural numbers $S \subseteq \mathbb{N}$ the smallest $n \in S$ in size order cannot be divided by any other $m \in S$. This was Emmy Noether's original motivation for what is now called the Noetherian condition on a ring. A ring R is Noetherian if and only if every set of ideals of R contains at least one minimal for the divisibility relation, or in other words at least one maximal for the inclusion relation. This is the key to ideal factorization in Noetherian rings.

Well-founding supports a very nice kind of proof by induction which Noether used constantly in her theory of ideal factorization and which occurs in various guises all over mathematics today:

THEOREM 4.6. *Let $\leq$ be any well-founded relation on a set S, and consider any subset $U \subseteq S$ which meets this inductive condition*

$$\forall x \in S \ [(\forall y < x)(y \in U) \ \Rightarrow \ x \in U \ ] \tag{1}$$

Then U is all of S.

$$(\forall x \in S) \ (x \in U)$$

PROOF. If not every $x \in S$ is in U then some $x \in S$ is $\leq$-minimal among those not in U. By minimality of x, every $y \leq x$ must be in U but then (1) implies that also $x \in U$ which is a contradiction. □

For example it is obvious that if every proper divisor of a natural number n factors into primes then so does n: if n has no proper divisors it is prime itself, and if it has proper factors take any complementary pair and use their prime factors. Then the induction principle concludes every natural number factors into primes.

0.3. The Rank Function. Particularly important among well-founded relations are the *well-orderings*, defined as total orders where every non-empty subset has a minimal element. In a total ordering the minimal elements are necessarily unique:

$$\forall U \subseteq S \ (\ \exists x \in U \ \Rightarrow \ (\ \exists! x \in U \ \forall y \in U \ (x \leq y) \ ) \$$

These are described in all textbooks on set theory and we will take the general theory for granted. The following theorems illustrate the key point. They show there is a natural rank function from well-founded relations to well-orderings, which is defined in an obvious way without any axiom of foundation—either in ETCS or, say Zermelo set theory (ZF minus the axiom of foundation and the axiom scheme of replacement).

THEOREM 4.7. *Let $S, \leq_s$ and $T, \leq_t$ be any two well-orderings. Then one of them is order-isomorphic to an unique initial segment of the other.*

PROOF. For any $x \in S$ define the initial segment of x, written $seg(x)$ to be the subset of all elements of S less than or equal to x:

$$seg(x) \ = \ \{y \in S \mid y \leq_s x \ \}$$

and similarly for $T, \leq_t$. Consider the subset of elements of S whose initial segment is order isomorphic to a unique initial segment of T:

$$\{x \in S \mid seg(x) \text{ is order isomorphic to } seg(z) \text{ for a unique } z \in T\}$$

If this set is all of S then $S, \leq_s$ is order-isomorphic to a unique initial segment of $T, \leq_t$ (which, of course, could be all of T). If it is not all of S then take the minimal $x_0 \in S$ not in this subset. Then it is easy to see the strict initial segment of x_0

$$strictseg(x_0) \ = \ \{y \in S \mid y <_s x_0 \ \}$$

is order-isomorphic to a unique initial segment of $T, \leq_t$.

Suppose that segment is not all of T. Then there is a unique minimal $z_0 \in T$ not in it. But

$$seg(x_0) \;=\; strictseg(x_0) \cup \{x_0\} \qquad \text{and} \qquad seg(z_0) \;=\; strictseg(z_0) \cup \{z_0\}$$

so the order-isomorphism $strictseg(x_0) \to strictseg(z_0)$ extends uniquely to

$$seg(x) \to seg(z_0)$$

namely by taking $x_0 \mapsto z_0$. This contradicts the assumption, so the segment must be all of T. □

0.4. The Axiom of Foundation. The ZF Axiom of Foundation, or Regularity Axiom, says membership is well-founded on any set:

$$\forall S\ \exists x \in S\ \forall y \in S\ (\ y \notin x\ )$$

Different ZF textbooks give different intuitive motivations but really, as Jech says, the motivation is that this allows defining a *rank* function useful in metatheory.

Categorical set theory, of course, does not have a membership relation like that of ZF but it defines partial orders and well-founded relations exactly as above. , and this is enough.

Axiom: Every set has a tree. Definition: The rank of a set is the smallest rank of any of its trees. This is lower than the usual ZF rank—in ZF terms it is the lowest rank of any set iso to this one.

CHAPTER 5

Sets of Sets, Spaces of Spaces, and Comprehension

Three concepts of a set of sets occur in practice.

$$\{A_i|\ i \in I\}$$

If the sets A_i are all given as subsets of a single set, as for example sets of real numbers $A_i \subseteq \mathbb{R}$ in analysis, then the set of sets is naturally a subset of the powerset

$$\{A_i|\ i \in I\}\ \subseteq\ 2^{\mathbb{R}}$$

For arbitrary sets A_i we could understand $\{A_i|\ i \in I\}$ as a function from the index set I to the universe of all sets, assigning each index $i \in I$ its set A_i. But the simpler route defines an arbitrary set of sets $\{A_i|\ i \in I\}$ as an arbitrary function $f\colon A \to I$ to the index set I. Each individual A_i is the inverse image of $i \in I$:

$$A_i\ =\ f^{-1}(i)\ =\ \{x \in A|\ f(x) = i\}$$

In ETCS, inverse images are defined up to isomorphism as *pullbacks* (a combination of products and equalizers).

In this form a *choice function* for $\{A_i|\ i \in I\}$ is any right inverse $c\colon I \to A$ to f.

$$\forall\ i \in I \quad c(i) \in A_i \quad \text{or in other words} \quad f(c(i)) = i$$

The axiom of choice says there is a choice function if and only if each A_i is nonempty. The product of the sets is the set of choice functions

$$\prod_{i \in I} A_i\ =\ \{f \in A^I|\ \forall i \in I \quad f(c(i)) = i\}$$

as it is usually defined in mathematics texts. This representation has various uses, and is often combined with representation by power sets. Most importantly for us here, this representation generalizes to other structures than sets.

A space of spaces $\{X_s|\ s \in \mathcal{S}\}$ needs not an index set I but a structured index space $\mathcal{S}$. The individual spaces X_s not only co-exist but are spatially related to one another in an ambient space $\mathcal{X}$. So the standard representation of a space of spaces in topology, differential geometry, or algebraic geometry is by a map $f\colon \mathcal{X} \to \mathcal{S}$ in the corresponding category. The individual spaces X_s are pre-images of points in the index space:

$$X_s\ =\ f^{-1}(s)\ =\ \{x \in \mathcal{X}|\ f(x) = s\}$$

But here the $\{x \in \mathcal{X}|\ f(x) = s\}$ are not sets and cannot be defined merely by their points $x \in \mathcal{X}$. They are spaces defined up to spatial isomorphism by pullbacks in the appropriate category.

This is also how van den Dries (1998) handles sets of sets, because it is the right way. ZFC handles them by making sets elements of sets, and producing all the paradoxes of self-membership. This is actually irrelevant to good practice and it raises pointless difficulties when you try to make it relevant. Barwise and

Etchemendy suppose that the 'things referred to' in a term ought to be (iterated-) members of the set which semantically represents that term (do they actually say 'term' or maybe 'sentence,' which would complicated my point a bit?). So for example 0 should be a member of sine(0) and there should be a problem of self-membership in saying sine(0)=0. Compare my exchange with Harvey on FOM where he says the first thing they tell you about sets is that a set is determined by its member sets, and I replied it is also the first thing they tell you to ignore—when for example they defined integers as equivalence classes of natural numbers (and thus *never* natural numbers themselves) and then immediately say "we will treat each natural number as an integer."

The axiom scheme of comprehension says every expressible assignment of a set A_i to each $i \in I$ can be represented as a function $f: A \rightarrow I$ from a single set A to I. It is an intrinsically set theoretic idea not suited to geometry since the whole point of the comprehension scheme is that the "fibers" A_i are given with no relation to each other, see (McLarty, 2004). The foundation here will not use it.

CHAPTER 6

What ETCS Cannot Do

First, like most set theories, ETCS refers to sets and functions but not to the universe of all sets or to any proper class. It can handle groups in abstract algebra, for example, but not the category of groups as a single structure:

Grp: The category of groups. Its objects are groups and arrows are group homomorphisms.

That category has a proper class of different objects, as do many categories which mathematicians treat as single structures.

Second, as to set-theoretic strength, ETCS is equivalent to bounded Zermelo set theory, BZ. That is Zermelo set theory allowing only bounded quantifiers in the separation scheme.[1] For each natural number n the axioms prove there is an n-th transfinite cardinal $\aleph_n$, and so a set of $n+1$ distinct transfinite cardinalities. For example

$$\text{ETCS} \vdash \exists \{\aleph_0, \aleph_1, \aleph_2, \aleph_3, \aleph_4, \aleph_5\}$$

They do not prove there is any cardinal $\aleph_\omega$ beyond all the $\aleph_n$. More sharply, they cannot prove the quantified statement of existence of arbitrarily large finite sets of distinct transfinite cardinalities:[2]

$$\text{ETCS} \nvdash \forall n \in \mathbb{N} \quad \exists \{\aleph_0, \aleph_1, \ldots, \aleph_n\}$$

Adjoining a categorical axiom scheme of replacement proves all this and much more, as it gives a set theory identical in strength to Zermelo-Fraenkel. But it does nothing about the first limitation, the inability to handle proper classes or such categories as **Set**, **Grp** et c. That limitation is often overcome in practice by invoking *Grothendieck universes*. A Grothendieck universe is a set U which itself satisfies all the axioms of set theory. On ZF foundations a Grothendieck universe is a set satisfying all the ZF axioms. On ETCS foundations it is a set of sets (i.e. a function, see chapter 5) which, together with all the functions between them, satisfy the ETCS axioms. Either way a Grothendieck universe proves the consistency of its set theory, so that neither ZF nor ETCS proves there are universes. We can extend either of those theories by an axiom positing a universe or the stronger axiom positing that each set (including each universe) is a member of some universe. The resulting theories still have no category **Set** of all sets or **Grp** of all groups, but for each universe U there is a category U-**Set** of all sets in U, and a category U-**Grp** of all groups in U, and so on. These categories can stand in for the non-existent **Set** and **Grp** for all practical purposes—at the cost of some complication keeping track of universes

[1]See inter alia Mac Lane and Moerdijk (1992, pp. 332–43).

[2]See especially the comment after Thm. 9.15 of Mathias (2001).

Here we adopt neither a comprehension scheme nor universes but follow Lawvere's advice that for mathematical practice "when one wishes to go substantially beyond what can be done in [ETCS] a more satisfactory foundation will involve a theory of the category of categories" (1964, p. 1510). This immediately gives categories **Set** and **Grp** and it strengthens the set theory as well. In particular it proves a theorem scheme of unbounded separation for **Set**. So it proves the quantified statement about arbitrarily large finite sets of transfinite sets:

$$\text{CCAF} \vdash \forall n \in \mathbb{N} \quad \exists \text{ in } \mathbf{Set} \text{ a set } \{\aleph_0,\ \aleph_1,\ \ldots,\ \aleph_n\}$$

It still does not prove those can be collected into one set $\aleph_\omega$ in **Set**.

Problem session: comparison with ZFC

> A specific mathematical objection has been raised by Rao concerning the construction of localizations in homotopical algebra that make use of transfinite induction and recursion. As he says, "it is not clear how to formulate these in categorical terms ... Solving these problems [by such means] looks remote at the moment." (Feferman, 2006, p. 186)

Feferman uses "set theory" to mean membership-based theories such as ZFC and not categorical set theory as in Lawvere's *Elementary Theory of the Category of Sets*, ETCS. Rao refers to some cutting edge methods of topology, and indeed finds it "not clear how to formulate [the central constructions of these methods] in categorical terms" (Rao, 2006, p. 278). But one thing is clear: ETCS has the same expressive power as ZFC.[3] Most mathematics is expressed the same way verbatim in both foundations, and everything expressible at all in one can be routinely translated into the other.

Equality of expressive power does not mean ZFC and ETCS are equal in proof-theoretic strength. In fact ZFC is much stronger. But it does mean some extension of ETCS has the strength of ZFC. One was given in the second paper written on categorical set theory (Lawvere, 1964, p. 34) using a neatly stripped-down reflection principle R, and worked out in detail in Osius (1974). The theories ZFC and ETCS+R are intertranslatable preserving provability in both directions.

In particular, ETCS+R defines well-ordering exactly the same way as ZFC: a *well ordering* is a partial order on a set S such that each non-empty subset of S contains a least member. The two theories prove the existence of all the same well-ordered order types, and support all the same transfinite ordinal inductions, and admit all the same consistent extensions by stronger axioms. So far as ordinal induction is concerned, the only difference is that ETCS+R cannot distinguish a selected representative of each well-ordering the way ZFC does with the von Neumann ordinals. That might matter in some predicativist or constructivist setting but not in ZFC and ETCS+R. See (McLarty, 2006b) on the not-really constructivist nature of category theory.

The translation between ZFC and ETCS+R is easily described. Proofs in ZFC often involve (finite and transfinite) membership chains of sets

$$S_0 \in S_1 \in \ldots S_\omega \in S_{\omega+1} \in \ldots S_\alpha \in \ldots$$

[3]Proven in (Mitchell, 1972), (Osius, 1974), beautifully organized in (Johnstone, 1977, §9.3), more recently Mac Lane and Moerdijk (1992, §VI.10). See McLarty (2004).

These methods gain their full power by using *transitive closures* to localize iterated membership. That is, regard the members of any set S, the members of members, their members in turn, and so on, as not merely existing "out there" in the whole universe of sets but as all being members of a single set $\mathrm{TC}(S)$ called the *transitive closure* of S. There are no membership chains of sets in ETCS since the members of ETCS sets are not sets. And ETCS+R cannot have transitive closures of sets, because all properties of ETCS sets are isomorphism invariant, while ZFC transitive closure is not isomorphism invariant.[4] But ZFC proves the *Mostowski collapsing theorem* (Jech, 2006, p. 69) which says every *well-founded, extensional relation* R on any set M is isomorphic to the iterated membership relation on a unique ZFC set S. In other words every well-founded, extensional relational system $\langle M, R\rangle$ is isomorphic to the relational system $\langle \mathrm{TC}(S), \in\rangle$ for a unique ZFC set S. Mostowski gives an easy routine for restating any ZFC sentence as an isomorphism invariant statement about well-founded extensional relational systems, which is provably (in ZFC) equivalent to the original. All isomorphism-invariant statements of ZFC translate verbatim into ETCS+R. Both Mitchell (1972) and Osius (1974) prove the translation preserves and reflects provability of statements on well-founded, extensional relational systems. Via this two-step translation, ZFC and ETCS+R can formulate all the same concepts and prove all the same theorems.

Such a brute force translation of (Rao, 2006) into ETCS+R would be clunky at some points, and identical verbatim to Rao's version at others, and presents no conceptual difficulties. A categorist actually approaching Rao's problems would not use the brute force translation but adapt many details to suit categorical terms. There *might* be positive advantages to using other convenient well-founded, extensional relations at some steps instead of iterated membership. The reason ZFC set theorists use the Mostowski theorem is that there *are* such advantages in some of their work—other well-founded, extensional relations naturally arise, and the ZFC set theorist uses Mostowski to show they can be replaced by iterated membership on some set. At any rate there is no foundational difficulty expressing (Rao, 2006) in categorical set theory.

---

[4]The ZFC singleton sets $\{\emptyset\}$ and $\{\mathbb{N}\}$ where $\mathbb{N}$ is (any ZFC representative of) the set of natural numbers are isomorphic but their transitive closures are not. The sole member of $\{\emptyset\}$ has no members so this set is its own transitive closure. The transitive closure of $\{\mathbb{N}\}$ is infinite as it contains all members of $\mathbb{N}$. It also includes their members in turn and so on, but on standard ZFC versions of $\mathbb{N}$ these further levels introduce no new members.

Part 3

The Category of Categories as Foundation

CHAPTER 7

Functors and Natural Transformations

Category theory looks at structures in terms of the mappings between them. This applies as well to categories as structures with *functors* as the mappings between them. Further, the mappings in categories can be applied to functors to form mappings between the functors, called *natural transformations*. The definitions of functors and natural transformations are quite concise and are really impossible to understand without some of the key facts about them and many examples. The best idea is to work back and forth among all the sections of this chapter until they all make sense. Section ??? describes *adjunctions*, though we will not go far with them.

1. Definitions

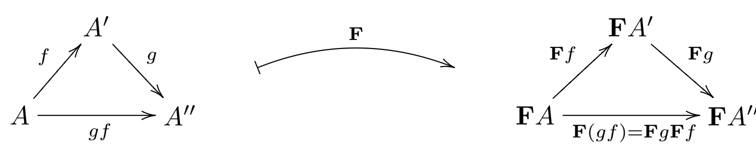
A *functor* from a category $\mathbf{A}$ to a category $\mathbf{B}$, written $\mathbf{F}\colon\mathbf{A}\to\mathbf{B}$, assigns to each object A of $\mathbf{A}$ an object $\mathbf{F}A$ of $\mathbf{B}$, and assigns to each arrow $f\colon A\to A'$ of $\mathbf{A}$ an arrow $\mathbf{F}f\colon\mathbf{F}A\to\mathbf{F}A'$ of $\mathbf{B}$, so that it preserves the domain and codomain. And it must preserve identity arrows and composition. That is, applying $\mathbf{F}$ to an identity arrow 1_A in $\mathbf{A}$ gives the identity arrow on $\mathbf{F}A$ in $\mathbf{B}$.

$$\mathbf{F}(1_A) \;=\; 1_{\mathbf{F}A}$$

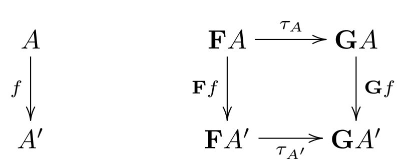
And applying it to the composite of arrows $f\colon A\to A'$ and $g\colon A'\to A''$ in $\mathbf{A}$ gives the composite of $\mathbf{F}f$ with $\mathbf{F}g$ in $\mathbf{B}$. As an equation,

$$\mathbf{F}(gf) = (\mathbf{F})(g\mathbf{F}f)$$

Or in a diagram:



Given functors $\mathbf{F},\mathbf{G}\colon\mathbf{A}\to\mathbf{B}$ both from a category $\mathbf{A}$ to a category $\mathbf{B}$, a *natural transformation* from $\mathbf{F}$ to $\mathbf{G}$, often written $\tau\colon\mathbf{F}\to\mathbf{G}$, is a family of arrows in $\mathbf{B}$ which roughly speaking carry values of $\mathbf{F}$ over to values of $\mathbf{G}$. Precisely, it assigns to each object A of $\mathbf{A}$ an arrow $\tau_A\colon\mathbf{F}A\to\mathbf{G}A$ between the values, meeting this *naturality condition*: for every arrow $f\colon A\to A'$ in $\mathbf{A}$, the square on the right commutes in $\mathbf{B}$



In an equation: $(\mathbf{G}f)(\tau_A) = (\tau_{A'})(\mathbf{F}f)$.

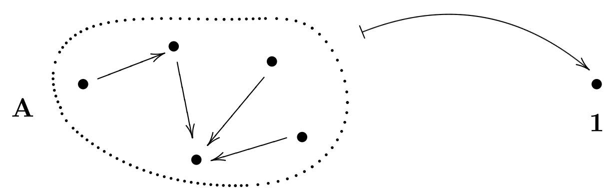
1.1. Introduction to functors. Depending on the choice of foundation, it is either an axiom or a trivial theorem that each category $\mathbf{A}$ has an identity functor $1_{\mathbf{A}} : \mathbf{A} \to \mathbf{A}$, which takes each object or arrow of $\mathbf{A}$ to that same object or arrow. Similarly functors $\mathbf{F}:\mathbf{A}\to\mathbf{B}$ and $\mathbf{G}:\mathbf{B}\to\mathbf{C}$ compose to give a functor $\mathbf{GF}:\mathbf{A}\to\mathbf{C}$. Composition is associative, so categories and functors themselves satisfy the Eilenberg-Mac Lane category axioms! Intuitively they form a "category of all categories"—although depending on foundations this category may not exist. Indeed it will not exist as a legitimate entity on our CCAF foundations, just as there is no set of all sets in ZFC. But let us be naive for the moment.

Recall from ??? the finite categories $\mathbf{1}, \mathbf{2}$

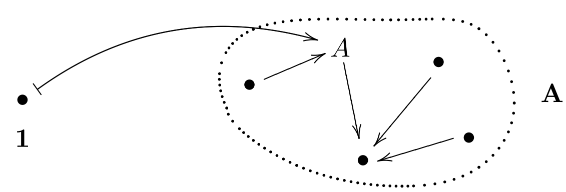
$$0 \qquad\qquad 0 \xrightarrow{\alpha} 1$$

$$\mathbf{1} \qquad\qquad \mathbf{2}$$

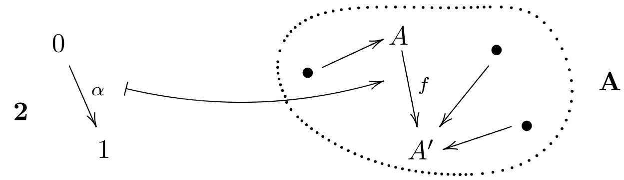
Notice that the single object category $\mathbf{1}$ is terminal among categories just as the singleton set 1 is terminal among sets: Every category $\mathbf{A}$ has exactly one functor $\mathbf{A}\to\mathbf{1}$. All the objects and arrows of $\mathbf{A}$ must go to the single object and arrow of $\mathbf{1}$. In a picture:



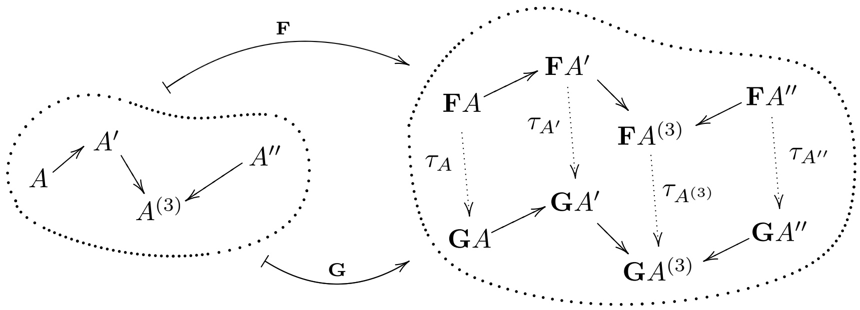
Conversely, a functor $\mathbf{1}\to\mathbf{A}$ just picks out an object A of $\mathbf{A}$



Specifically A corresponds to the functor taking the sole object and identity arrow in $\mathbf{1}$ to the object A and its identity arrow. In fact the categorical foundation CCAF will *define* an object of a category $\mathbf{A}$ to be a functor $\mathbf{1}\to\mathbf{A}$ just as ETCS defined an element of a set to be a function $1\to S$. Category theory goes a step farther using the category $\mathbf{2}$. Each arrow $f:A\to A'$ of $\mathbf{A}$ corresponds to the functor $f:\mathbf{2}\to\mathbf{A}$ taking the arrow α of $\mathbf{2}$ to f and the objects $0, 1$ of $\mathbf{2}$ to A, A' respectively.



1.2. Introduction to natural transformations. Think of functors $\mathbf{F}, \mathbf{G} : \mathbf{A} \to \mathbf{B}$ each carrying the network of $\mathbf{A}$ over to the network of $\mathbf{B}$. Then a natural transformation $\tau : \mathbf{F} \to \mathbf{G}$ links each vertex in the $\mathbf{F}$-image of that network to the corresponding vertex in the $\mathbf{G}$-image, so that the $\mathbf{F}$-image "flows over" onto the $\mathbf{G}$-image. Picture the four arrows on the left here as some among the network of category $\mathbf{A}$ while the twelve on the left are their $\mathbf{F}$-images and $\mathbf{G}$-images and the components of τ among the network of arrows of category $\mathbf{B}$:



For a given pair of categories $\mathbf{A}, \mathbf{B}$ each functor $\mathbf{F} : \mathbf{A} \to \mathbf{B}$ has an *identity transformation* $1_{\mathbf{F}} : \mathbf{F} \to \mathbf{F}$ where the component arrow for each object A of $\mathbf{A}$ is the identity arrow for $\mathbf{F}A$. These identity arrows clearly meet the naturality condition:

$$\begin{array}{ccccc} A & \qquad & \mathbf{F}A & \xrightarrow{1_{\mathbf{F}A}} & \mathbf{F}A \\ {\scriptstyle f}\downarrow & & {\scriptstyle \mathbf{F}f}\downarrow & & \downarrow{\scriptstyle \mathbf{F}f} \\ A' & & \mathbf{F}A' & \xrightarrow[1_{\mathbf{F}A'}]{} & \mathbf{F}A' \end{array}$$

Furthermore, given any functors $\mathbf{F}, \mathbf{G}, \mathbf{H} : \mathbf{A} \to \mathbf{B}$ all from $\mathbf{A}$ to $\mathbf{B}$, and natural transformations $\tau : \mathbf{F} \to \mathbf{G}$ and $\sigma : \mathbf{G} \to \mathbf{H}$ there is a natural composite transformation $\sigma\tau : \mathbf{F} \to \mathbf{H}$ with components $\sigma_A \tau_A$.

$$\begin{array}{ccccccc} A & \qquad & \mathbf{F}A & \xrightarrow{\tau_A} & \mathbf{G}A & \xrightarrow{\sigma_A} & \mathbf{H}A \\ {\scriptstyle f}\downarrow & & {\scriptstyle \mathbf{F}f}\downarrow & & \downarrow{\scriptstyle \mathbf{G}f} & & \downarrow{\scriptstyle \mathbf{H}f} \\ A' & & \mathbf{F}A' & \xrightarrow[\tau_{A'}]{} & \mathbf{G}A' & \xrightarrow[\sigma_{A'}]{} & \mathbf{H}A' \end{array}$$

So the functors from $\mathbf{A}$ to $\mathbf{B}$, and the natural transformations between them, also satisfy the Eilenberg-Mac Lane axioms. There is a *functor category* $\mathbf{B}^{\mathbf{A}}$:

> $\mathbf{B}^{\mathbf{A}}$**:** The functor category from $\mathbf{A}$ to $\mathbf{B}$. Its objects are functors $\mathbf{A} \to \mathbf{B}$ and its arrows $\tau : \mathbf{F} \to \mathbf{G}$ are natural transformations between them.

As to foundations, this category exists as a legitimate entity on essentially any foundation that makes $\mathbf{A}$ and $\mathbf{B}$ legitimate. The categorical foundation CCAF will posit a functor category $\mathbf{B}^{\mathbf{A}}$ for any categories $\mathbf{A}, \mathbf{B}$.

What is a natural transformation τ between functors $A, A' : \mathbf{1} \to \mathbf{A}$? Since the sole object of $\mathbf{1}$ is 0, the natural transformation has just one component, which must go from the sole value A of the first functor to the sole value A' of the second.

$$A \xrightarrow{\quad \tau_0 \quad} A'$$

Since **1** has only one arrow there is only one naturality square, and since that arrow is an identity arrow the naturality square automatically commutes. So a natural transformation τ between functors $A, A'\colon \mathbf{1} \to \mathbf{A}$ is just an arrow between the corresponding objects A, A'.

2. Mathematical Examples of Functors

Functors are everywhere. Chapter 1 described several categories: the category **Set** of sets, the category **TopSurf** of Riemann surfaces with continuous maps, and the category **HolSurf** also of Riemann surfaces but taking only holomorphic maps between them. Riemann might look at a given Riemann surface in terms of the holomorphic maps between it and others, and then expand his view to look at all the continuous maps. In categorical terms, that invokes a functor

$$\mathbf{i}\colon \mathbf{HolSurf} \to \mathbf{TopSurf}$$

inserting the holomorphic category of Riemann surfaces into the larger topological category. This functor takes the network of holomorphic maps between Riemann spaces and, leaving the objects as they were, it greatly increases the number of maps between them.

Yet a larger context for Riemann's work uses topological spaces. It is a reasonable approximation to say that any space you can picture can be taken as a topological space. Intuitively a morphism, or *continuous map* $T \to T'$ of topological spaces may stretch and bend T in any way onto some part of T' and even smash whole parts of it down to a single point of T', just so long as it does not tear nearby points apart.[1] And so an isomorphism $T \to T'$ must stretch and bend T onto the whole of T' without either tearing points apart or smashing any two together. Isomorphisms in topology are traditionally called *homeomorphisms*. Topological spaces and maps form a category: **Top**:

Top: The category of topological spaces. Its objects are all topological spaces and arrows are all continuous maps $f\colon T \to T'$ of spaces.

There is a functor

$$\mathbf{i}\colon \mathbf{TopSurf} \to \mathbf{Top}$$

inserting the topological category of Riemann surfaces into the larger category of topological spaces. Roughly speaking this functor the leaves the maps alone—the maps between any two Riemann surfaces in **Top** are the same continuous maps as they were in **TopSurf**. But it greatly increases the number of objects, from the quite constrained Riemann surfaces to all topological spaces.

Some of the earliest influential functors take geometric spaces to algebraic structures. For example each topological space T has an n-th *homology group* $H_n(T)$ for each natural number n, where the group $H_n(T)$ keeps track of the "n-dimensional holes" in T in a specific sense. The beauty of it is this is a functor

$$H_n\colon \mathbf{Top} \to \mathbf{AbGrp}$$

[1] For fuller expert popularizations see Hilbert and Cohn-Vossen (1932), Alexandroff (1932), or Mac Lane (1986).

from the category of topological spaces and continuous maps to the category of Abelian groups. Each continuous map of spaces $f : T \to T'$ induces a group homomorphism $H_n(f) : H_n(T) \to H_n(T')$ so that the most relevant topological information is concentrated into a simpler algebraic form.[2] The related notion of *homotopy* gives a ...

Functors are too common and too varied to need or allow a comprehensive survey here.

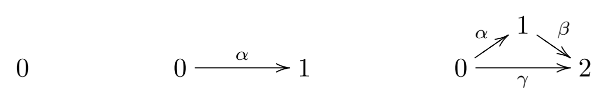
3. Adjoints

[2]Functoriality is the basis of the Eilenberg-Steenrod homology axioms standard in research since (Eilenberg and Steenrod, 1945) and in textbooks today, e.g. (Vick, 1994).

CHAPTER 8

Categorical Category Theory

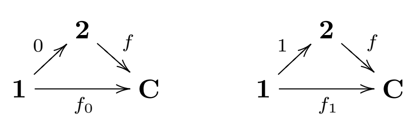
Lawvere (1963) first systematically applied the central insight itself to functors.[1] (Note to self: this version does not talk much about a "central insight.") Categories, including a category **Set** of sets that satisfies the ETCS axioms, can be defined by their functors to each other. In practice today many categories are defined, only up to isomorphism, by their functors to and from other given categories. Lawvere made the point that this can be done from the ground up. Begin with three finite categories called $\mathbf{1}, \mathbf{2}, \mathbf{3}$ which naively look like this:



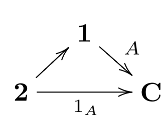
1 **2** **3**

Each lightfaced numeral represents an object and its identity arrow. The category **1** has a single object 0 and only its identity arrow. The category **2** has two objects $0, 1$, their identity arrows, and one non-identity arrow as shown. The category **3** has three objects, and arrows as shown.

Each object A of any category $\mathbf{C}$ corresponds to a functor $A\colon \mathbf{1} \to \mathbf{C}$, namely the functor taking the sole object and identity arrow in **1** to the object A and its identity arrow. Each arrow $f\colon A \to A'$ of $\mathbf{C}$ corresponds to a functor $f\colon \mathbf{2} \to \mathbf{C}$ that takes the arrow α of **2** to the arrow f and takes the objects $0, 1$ of **2** to the domain and codomain A, A' respectively of f in $\mathbf{C}$. When an arrow f of $\mathbf{C}$ is expressed as a functor $f\colon \mathbf{2} \to \mathbf{C}$ then its domain and codomain in $\mathbf{C}$ are expressed by the composite functors f_0 and f_1 respectively:



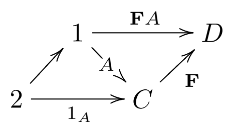
And in the other direction, given an object A of $\mathbf{C}$ expressed as a functor $A\colon \mathbf{1} \to \mathbf{C}$ its identity arrow in $\mathbf{C}$ is expressed by the composite



[1]This included the first categorical definition of natural number objects, the first account of functor categories as adjoints, and the introduction of comma categories.

Each functor $\mathbf{F}\colon\mathbf{C}\to\mathbf{D}$ takes objects and arrows of $\mathbf{C}$ to the same in $\mathbf{D}$ just by functor composition:

And by associativity of functor composition, $\mathbf{F}$ preserves domains, codomains, and identity arrows. Consider the associativity diagram:



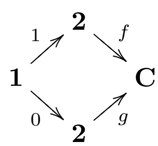
Composition along the bottom and right gives $\mathbf{F}(1_A)\colon\mathbf{2}\to\mathbf{D}$ while the left and top gives $1_{\mathbf{F}A}\colon\mathbf{2}\to\mathbf{D}$ so that

$$\mathbf{F}(1_A) = 1_{\mathbf{F}A}$$

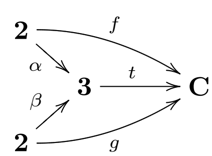
The same reasoning shows $\mathbf{F}$ preserves domains and codomains

$$\mathbf{F}(f_0) = (\mathbf{F}f)_0 \quad \text{and} \quad \mathbf{F}(f_1) = (\mathbf{F}f)_1$$

Functorial category theory expresses composition of arrows in any category $\mathbf{C}$ in terms of functors from $\mathbf{1}, \mathbf{2}$ and $\mathbf{3}$ to $\mathbf{C}$. Arrows f and g of $\mathbf{C}$ compose (in that order) if and only if the codomain of f is the codomain of g. Expressed by functors to $\mathbf{C}$ that says $f\colon\mathbf{2}\to\mathbf{C}$ and $g\colon\mathbf{2}\to\mathbf{C}$ compose as arrows of $\mathbf{C}$ if and only if $f_1 = g_0$:

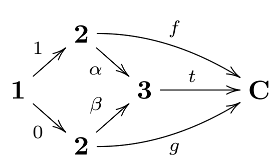


In that case there should be one and only one commutative triangle in $\mathbf{C}$ with f and g as successive sides. That means there should be a unique functor $t\colon\mathbf{3}\to\mathbf{C}$ with $t\alpha = f$ and $t\beta = g$:



Then the composite of f and g in $\mathbf{C}$ can be defined as the third side $t\gamma\colon\mathbf{2}\to\mathbf{C}$. That is the functorial treatment of composition in categories.

To summarize, $\mathbf{3}$ is a *pushout* of 0 and 1. That means $\alpha_1 = \beta_0$ and given any f, g with $f_1 = g_0$ there is a unique t with $t\alpha = f$ and $t\beta = g$:

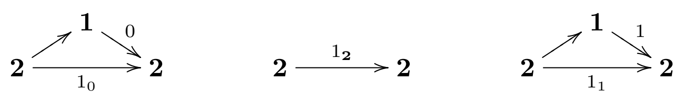


The characterizes $\mathbf{3}$ uniquely up to isomorphism. It implies that the composite gf in $\mathbf{C}$ has the right domain and codomain: $(gf)_0 = f_0$ and $(gf)_1 = g_1$. And again by associativity of functor composition it implies that every functor $\mathbf{F}:\mathbf{C}\to\mathbf{D}$ preserves composition.[2]

The category $\mathbf{1}$ is *terminal*, in the sense that every category $\mathbf{C}$ has a unique functor to it $\mathbf{C}\to\mathbf{1}$. Just as singleton sets are unique up to isomorphism in the category **Set** of sets, so this defines $\mathbf{1}$ up to isomorphism in the category of categories. There are two objects in the category $\mathbf{2}$:

$$0:\mathbf{1}\to\mathbf{2} \qquad 1:\mathbf{1}\to\mathbf{2}$$

There are three arrows in $\mathbf{2}$, two of them identity arrows:



Notice that 1_0 and 1_1 are identity arrows in $\mathbf{2}$ and not identity functors to $\mathbf{2}$, while the identity functor $1_{\mathbf{2}}$ to $\mathbf{2}$ is the sole non-identity arrow in $\mathbf{2}$.[3]

The key fact is that different functors $\mathbf{F}\neq\mathbf{G}$ between the same categories $\mathbf{C}\to\mathbf{D}$ may agree on all objects of $\mathbf{C}$ but cannot agree on all arrows. There must be some functor $f:\mathbf{2}\to\mathbf{C}$ with $\mathbf{F}f\neq\mathbf{G}f$. Sets have extensionality with regard to the set 1: distinct functions between sets $S\to S'$ differ on some element $x:1\to S$. Categories have extensionality with regard to the arrow category $\mathbf{2}$: distinct functors between categories $\mathbf{C}\to\mathbf{D}$ differ on some arrow $f:\mathbf{2}\to\mathbf{C}$.

A category is *discrete* if every arrow $f:A\to B$ in it is an identity: that is $A=B$ and $f=1_A$. Notice that a functor $\mathbf{F}:\mathbf{D}\to\mathbf{D}'$ between discrete categories just amounts to a function from the objects of $\mathbf{D}$ to those of $\mathbf{D}'$. The action on objects automatically determines $\mathbf{F}$ of each identity arrow and there are no other arrows and or composites to worry about. Basically a discrete category is a set. One foundational nuance will matter in the next chapter: A category of categories taken as a foundation will include a category of sets **Set**, and each set I in **Set** will have a corresponding discrete category. But in general there will be many other discrete categories too large to come from sets of **Set**—i.e. discrete categories as large as **Set** themselves, or larger. So the universe of all discrete categories will be a large extension of the universe of discrete categories coming from **Set**.

---

[2]By associativity of functor composition, if $t:\mathbf{3}\to\mathbf{C}$ is the triangle induced by f,g, then $\mathbf{F}t:\mathbf{3}\to\mathbf{D}$ is induced by $\mathbf{F}f,\mathbf{F}g$. So the composite in $\mathbf{D}$ of $\mathbf{F}f,\mathbf{F}g$ is $(\mathbf{F}t)(\gamma)$; but this equals $\mathbf{F}(t(\gamma))$ so it is the image by $\mathbf{F}$ of the composite gf in $\mathbf{C}$.

[3]Since the composite in $\mathbf{2}$ of 1_0 and $1_{\mathbf{2}}$ has the same domain as 1_0 and the same codomain as $1_{\mathbf{2}}$, it follows the composite is $1_{\mathbf{2}}$. The same consideration applied to other cases shows that 1_0 and 1_1 do compose as identity arrows in $\mathbf{2}$. Since functors preserve composition, it follows that identity arrows in all categories compose as identities. For details see McLarty (1991a).

CHAPTER 9

Axiomatics of CCAF

These axioms are based on Lawvere (1963, 1966). In fact he discussed his ideas privately with Alfred Tarski at the 1964 International Congress of Logic, Methodology, and Philosophy of Science where his presentation was on ETCS. For the rest of this chapter the words “category” and “functor” are to be understood solely in the sense of the CCAF axioms given here. The axioms begin with the Eilenberg-Mac Lane axioms of Figure 2 above, for categories $\mathbf{A}, \mathbf{B}, \mathbf{C}$ as objects and functors $\mathbf{F}, \mathbf{G}, \mathbf{H}$ as arrows. The formal language of the theory has variables for categories, variables for functors, operators as given in Figure 2, and the terms defined below. The first defined terms are the finite categories $\mathbf{1}, \mathbf{2}, \mathbf{3}$ and functors between them as axiomatized in Figure 1.

Every category $\mathbf{C}$ has a unique functor $\mathbf{C} \to \mathbf{1}$.

The category $\mathbf{2}$ has exactly two functors from $\mathbf{1}$ and three to itself:

$$\mathbf{2} \to \mathbf{1} \xrightarrow{0} \mathbf{2},\ \ \mathbf{2} \xrightarrow{1_0} \mathbf{2} \qquad \mathbf{2} \xrightarrow{1_{\mathbf{2}}} \mathbf{2} \qquad \mathbf{2} \to \mathbf{1} \xrightarrow{1} \mathbf{2},\ \ \mathbf{2} \xrightarrow{1_1} \mathbf{2}$$

The category $\mathbf{3}$ is a pushout:

$$\begin{array}{ccccc} & & \mathbf{2} & & \\ & {}^{1}\nearrow & & \searrow^{\alpha} & \\ \mathbf{1} & & & & \mathbf{3} \\ & {}_{0}\searrow & & \nearrow_{\beta} & \\ & & \mathbf{2} & & \end{array}$$

and there is a functor $\gamma : \mathbf{2} \to \mathbf{3}$ with $\gamma_0 = \alpha_0$ and $\gamma_1 = \beta_1$.

Arrow extensionality: $\forall\ \mathbf{F}, \mathbf{G} : \mathbf{A} \to \mathbf{B}$ if $\mathbf{F} \neq \mathbf{G}$ then $\exists f : \mathbf{2} \to \mathbf{A}$ with $\mathbf{F}f \neq \mathbf{G}f$

Figure 1. The CCAF axioms on $\mathbf{1}, \mathbf{2}, \mathbf{3}$

An *object* A of a category $\mathbf{A}$ is defined to be a functor $A : \mathbf{1} \to \mathbf{A}$, and an *arrow* f is defined to be a functor $f : \mathbf{2} \to \mathbf{A}$. The *domain* f_0 and *codomain* f_1 of f in $\mathbf{A}$ are defined to be the composites of f with $0 : \mathbf{1} \to \mathbf{2}$ and $1 : \mathbf{1} \to \mathbf{2}$. The *identity arrow* 1_A on A is defined to be the composite

$$\mathbf{2} \longrightarrow \mathbf{1} \xrightarrow{A} \mathbf{A}$$

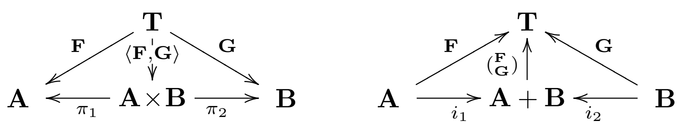
A *discrete category* is defined in CCAF as a category $\mathbf{D}$ such that each arrow $\mathbf{2} \to \mathbf{D}$ is the identity arrow 1_A for some object $A : \mathbf{1} \to \mathbf{D}$.

By the axiom on $\mathbf{3}$, any arrows $f : \mathbf{2} \to \mathbf{A}$ and $g : \mathbf{2} \to \mathbf{A}$ with $f_1 = g_0$ determine a unique $t : \mathbf{3} \to \mathbf{A}$ with $t\alpha = f$ and $t\beta = g$. Then define the composite gf in $\mathbf{A}$ to be

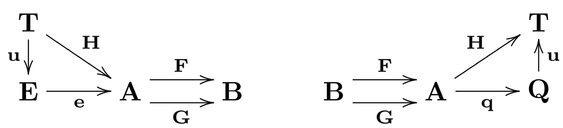
$t\gamma$. It follows trivially that arrows in $\mathbf{A}$ satisfy all of the Eilenberg-Mac-Lane axioms except associativity of composition which requires the next group of axioms.[1]

Figure 2 axiomatizes constructions on categories and functors. One difference from sets and functions is that CCAF explicitly posits *coproducts* and *coequalizers*. Formally these are like products and equalizers with the arrows turned around. Intuitively the coproduct of a pair of sets $S + T$ or a pair of categories $\mathbf{A} + \mathbf{B}$ is their disjoint union, while a coequalizer is a quotient by an equivalence relation. The ETCS axioms imply coproducts and coequalizers because sets and functions form a *topos*. Categories and functors do not form a topos, so the ETCS construction of coproducts and coequalizers does not carry over to CCAF, though of course some elegant construction might yet be found.

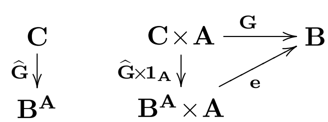
Every pair of categories $\mathbf{A}, \mathbf{B}$ has a product and a coproduct:



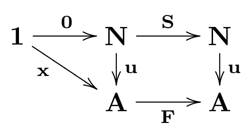
Every parallel pair of functors $\mathbf{F}, \mathbf{G} \colon \mathbf{A} \to \mathbf{B}$ has an equalizer and a coequalizer:



There is a functor category from each category $\mathbf{A}$ to each category $\mathbf{B}$:



There is a natural number category $\mathbf{N}, \mathbf{0}, \mathbf{S}$:



Choice: For every onto functor between discrete categories $\mathbf{F} \colon \mathbf{D}' \to \mathbf{D}$, there is a functor $\mathbf{H} \colon \mathbf{D} \to \mathbf{D}'$ with $\mathbf{FH} = 1_{\mathbf{D}}$.

FIGURE 2. The CCAF construction axioms

The axioms in Figure 3 posit a category $\mathbf{Set}$ of sets with a fullness condition. The condition uses categories constructed from $\mathbf{Set}$ by axioms in Figure 2.[2] For each set $I \colon \mathbf{1} \to \mathbf{Set}$ there is:

$(1 \downarrow I)$: The category of elements of the set I. Its objects correspond to elements of I, that is functions $1 \to I$ in $\mathbf{Set}$. All its arrows are identity arrows.

[1] McLarty (1991a). The **Set** axiom in Figure 3 replaces several axioms in that paper.

[2] Lawvere introduced such constructions in (1963, section I.1).

In short, $(1 \downarrow I)$ is the set I, an object of **Set**, re-construed as a discrete category. The same construction also re-construes each function $f : I \to J$ as a functor from $(1 \downarrow I)$ to $(1 \downarrow J)$:

$$(1 \downarrow f) : (1 \downarrow I) \to (1 \downarrow J)$$

The fullness condition says that re-construing sets as discrete categories this way makes no difference at all to the pattern of functions between sets. This plus the axiom of choice of CCAF makes the axiom of choice in ETCS redundant.

There is a category **Set** whose objects and arrows satisfy the ETCS axioms.

Fullness: For all sets $I, J : \mathbf{1} \to \mathbf{Set}$, every functor $(1 \downarrow I) \to (1 \downarrow I)$ equals $(1 \downarrow f)$ for some function $f : I \to J$ in **Set**.

FIGURE 3. CCAF axioms on **Set**

The categorical separation axiom scheme says that a predicate $\Psi(x)$ of objects and arrows that intuitively ought to define a subcategory $i : \mathbf{B} \rightarrowtail \mathbf{A}$ of a category **A**, does. In ETCS a subcategory of **A** can be defined as any functor $i : \mathbf{B} \to \mathbf{A}$ that is one-to-one on arrows and we often write $i : \mathbf{B} \rightarrowtail \mathbf{A}$ to say i is a subcategory. The axiom scheme implies that each $\Psi(x)$ determines a unique subcategory up to isomorphism. It implies that any relation $\Phi(x, y)$ of arrows in a category **A** to those of a category **B** which intuitively ought to define a functor $\mathbf{A} \to \mathbf{B}$, does.[3] And so coequalizers of categories have the properties they intuitively ought to.[4] In turn this implies, if "set" is understood to mean any discrete category, every description of a category in terms of a "set" A_0 of objects and a "set" A_1 actually describes a category **A** (Lawvere, 1966). Separation plus fullness implies that, when $\mathbb{N}$ is the natural number set in **Set**, then $(1 \downarrow \mathbb{N})$ together with functors $(1 \downarrow 0)$ and $(1 \downarrow s)$ are (up to isomorphism) the natural number category $\mathbf{N}, \mathbf{0}, \mathbf{S}$.

Categorical separation scheme: Let $\Psi(x)$ be any formula in the language of CCAF with sole free variable x of functor type. Then the universally quantified conditional

$$\forall \mathbf{A}\ [(\alpha \wedge \beta) \Rightarrow (\gamma \wedge \delta)]$$

is an axiom, for these clauses:

α**:** $(\forall f : \mathbf{2} \to \mathbf{A})\ [\Psi(f) \Rightarrow (\Psi(f_0) \wedge \Psi(f_1))]$

β**:** $(\forall f, g : \mathbf{2} \to \mathbf{A})\ [(f_1 = g_0 \wedge \Psi(f) \wedge \Psi(g)) \Rightarrow \Psi(gf)]$

γ**:** $(\exists \mathbf{B},\ \exists i : \mathbf{B} \rightarrowtail \mathbf{A},\ \forall f : \mathbf{2} \to \mathbf{A})\ [\Psi(f) \Leftrightarrow \exists h : \mathbf{2} \to \mathbf{B}\ f = i(h)]$

δ**:** $(\forall i : \mathbf{B} \rightarrowtail \mathbf{A},\ i' : \mathbf{B}' \rightarrowtail \mathbf{A})$ If i and i' both meet the condition in clause γ then $\exists k : \mathbf{B} \to \mathbf{B}'\ (i = i'k)$

By symmetry of i and i' in δ, it follows that k is an isomorphism of **B** and **B**$'$.

FIGURE 4. The CCAF separation axiom scheme

[3]The relation $\Phi(x, y)$ defines a subcategory of the product $\mathbf{A} \times \mathbf{B}$ whose projection onto **A** is an isomorphism, so that its projection onto **B** is the desired functor.

[4]Using the natural number category, every finite sequence of arrows in **A** which should patch together into a path of composable arrows in the coequalizer **Q** does, and the composites of these arrows form a subcategory of **Q** which, by separation, is all of **Q**.

Problem session: replacement in CCAF

Lawvere on completeness. External replacement: every functor from a discrete slice category on **Set** to **Set** comes from a family of sets. Is this "second order replacement"?

CHAPTER 10

A Sample Construction

The category **Iter** was described intuitively in chapter 4: An object is given by any function f where the domain f_0 equals the codomain f_1, plus a selected element x of that domain; and an arrow is a function m with the stated compatibility requirement.

To formalize this in CCAF notice the CCAF axioms on **Set** say it has a terminal object $1\colon \mathbf{1}\to\mathbf{Set}$, and the construction axioms say it has an arrow category $\mathbf{Set}^{\mathbf{2}}$. By definition an object of $\mathbf{Set}^{\mathbf{2}}$ corresponds to a functor $\mathbf{2}\to\mathbf{Set}$ or in other words a function f in **Set**. An object in the category **Iter** is given by two functions x, f, so it corresponds to an object of the product category $\mathbf{Set}^{\mathbf{2}}\times\mathbf{Set}^{\mathbf{2}}$. The functions must satisfy equations saying the domain of x is the singleton set 1, and the codomain of x is both the domain and codomain of f:

$$x_0 = 1 \qquad \text{and} \qquad x_1 = f_0 = f_1$$

So **Iter** is defined as a subcategory, and specifically an equalizer

$$\mathbf{Iter} \rightarrowtail \mathbf{Set}^{\mathbf{2}}\times\mathbf{Set}^{\mathbf{2}}$$

And this construction by an equalizer on a product automatically gives exactly the desired arrows for the category **Iter**.

Each of these steps characterizes its outcome only up to isomorphism. So the construction characterizes **Iter** only up to isomorphism—but not only up to isomorphism as a single category. It characterizes **Iter** plus the inclusion functor to $\mathbf{Set}^{\mathbf{2}}\times\mathbf{Set}^{\mathbf{2}}$. So we do not say, and need not say, that an object of **Iter** *is* a pair of functions, nor even that an object of $\mathbf{Set}^{\mathbf{2}}\times\mathbf{Set}^{\mathbf{2}}$ *is* a pair of functions. Rather we use the constructed inclusion functor, plus the two projections

$$\mathbf{Set}^{\mathbf{2}}\times\mathbf{Set}^{\mathbf{2}} \xrightarrow{\pi_1} \mathbf{Set}^{\mathbf{2}} \qquad \mathbf{Set}^{\mathbf{2}}\times\mathbf{Set}^{\mathbf{2}} \xrightarrow{\pi_2} \mathbf{Set}^{\mathbf{2}}$$

to *determine* a pair of functions for each object of **Iter**. And of course $\mathbf{Set}^{\mathbf{2}}$ comes with an evaluation functor to **Set**. Roughly speaking, the category **Iter** is defined up to an isomorphism of this whole pattern in the category of categories.

Problem session: A letter from Bill Lawvere

For more details see, as cited, Lawvere (2003), and for background Lawvere (1963).

Dear Elaine,

If **C** is a category of categories, extract the subcategory **S** of discrete objects (which serves as the less structured background) and then apply the reverse correspondence by considering the functor category $\mathbf{S}^{(\mathbf{T}^{\mathbf{op}})}$ where **T** is the finite category of $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$, or rather the part of that functor category where limits are preserved: the latter is equivalent to the **C**.

The general procedure is described in Section III "geometry provides its own foundation" in my article in the *Bulletin of Symbolic Logic* **9** (2003) pp. 213-224. Not only categories of categories but many geometrical toposes **C** that are cartesian closed have a pair of opposed subcategories **T** and **S** that are orthogonal wrt the defining relation

$$s = s^t$$

and such that the functor taking x to the functor that takes t to "the set of t-shaped figures in x"

$$[x^t]$$

is a full and faithful embedding of **C** into $\mathbf{S}^{(\mathbf{T}^{\mathbf{op}})}$) that even has a left adjoint. Here I used [ ] to denote the right adjoint of the inclusion of **S** into **C**.

In simple cases like **Cat** there is a finite **T**. In general the objects in C are represented as structures in terms of the chosen generic figures, structures valued in the background S that the defining relation says is structureless. This I feel is a good description of the basic discovery of Cantor-Dedekind. It is very relative, but should shed light also on the problem of extreme stucturelessness, e.g. can **S** satisfy not only AC but even GCH etc.

All this has little relation to the Peano-von Neuman-Bernays implementation of Frege's faulty principle that concepts are properties. Unfortunately the "universes" were defined in that context, and moreover involve the futile imaginary iteration of Cantor's later years.

"Internalizing" i.e. turning metamathematics into mathematics can be done by restricting the above broad correspondence between **S**'s and **C**'s to the case where **C** itself contains objects reflecting the idea of **S**, for example finite and small. (Here finite and small are *not* defined iteratively nor by anticlass but are actually well determined by duality and by the condition of being full in **C** along the natural fibration.) In my thesis I suggested a third such in order to give **Semantics** a codomain. But the duality condition already surpasses so-called measurable cardinals (Isbell).

Best wishes, Bill

On Zermelo's endorsement of Frege over Cantor see specifically (Cantor, 1932, pp. 351, 353, 441).

Part 4

Philosophy

CHAPTER 11

Structures

On this foundation a *structure* is any object in a category. This includes the usual structured sets as for example chapter 10 made iterative structures objects of a category **Iter**. This gives a straightforward answer to Hellman's question of "what axioms govern the existence of categories" or other structures in categorical foundations (Hellman, 2003, p. 137). Modulo some variance in the axioms it is the answer Lawvere gave in (1966): The CCAF axioms in Figures 1–4 posit the existence of certain categories and functors, including one category **Set** that satisfies the ETCS axioms in Figure 1, and certain means of constructing categories. So they imply the existence of further categories including those built from sets and functions. Of course other categorical foundations make other posits as in (Bell, 1998) or (McLarty, 1998).

Consider three basic goals of structuralism.[1]

(1) As to *individuation of elements*:

> the 'elements' of [a] structure have no properties other than those relating them to other 'elements' of the same structure. (Benacerraf, 1965, p. 70).

(2) As to elements appearing only *in a structured context*:

> mathematical objects [appear] always in the context of some background structure, and the objects have no more to them than can be expressed in terms of the basic relations of the structure. (Parsons, 1990, p. 303)

(3) Third is the matter of *structures of structures*, which Resnik has emphasized. He writes in terms of "positions in patterns" where others write of "elements of structures":

> Patterns themselves are *positionalized* by being identified with positions of another pattern, which allows us to obtain results about patterns which were not even previously statable. It is [this] sort of reduction which has significantly changed the practice of mathematics. (Resnik, 1997, p. 218).

The present foundation immediately gives point 3 and the first half of point 2 as elements occur only in objects which occur only in categories which occur only in the category of categories (which is no entity but only a context on this foundation). The CCAF axioms specifically include formation of categories of functors $\mathbf{B}^{\mathbf{A}}$ and they support numerous ways of forming categories of categories.[2]

[1]Reck and Price (2000) compare various "structuralisms" in philosophy of mathematics.

[2]See notably fibred categories in Bénabou (1985) and Johnstone (2002, Part B).

Point 1 and the second half of 2 require discussion. Structuralists generally leave it intuitive what they mean by "properties" and what it is for an object to have "no more to it" than certain relations, even as they debate these issues among themselves (see discussion of Hellman in chapter 13 below). Tarski's heritage, celebrated at these congresses, urges us to make such things precise. DO I CUT THE REF TO ICLMPS OR BUILD A INDEX IT?

On CCAF foundations all properties are isomorphism invariant. Any two isomorphic categories agree in all their properties. Any two isomorphic objects in a category agree in all their properties.[3] Any two elements of a set agree in all their properties.

Obviously the category **Set** has a "property" not shared with isomorphic categories: it is named "**Set**" and the others are not. But this is really not a property. It is a name. We could eliminate such problems by saying a *property* must not name any individuals. But that would rule out the standard expressions in CCAF for objects, arrows, or composition in categories, since those use the constants $\mathbf{1}, \mathbf{2}, \mathbf{3}$. As a quick and dirty fix, theorem scheme 12.1 allows use of those constants and excludes the categories $\mathbf{1}, \mathbf{2}, \mathbf{3}$ from the scope of the statement.

As to sets, the elements of a set S differ *in relation to* any selected element $x_0 \in S$: one element $x \in S$ is x_0 and the others (if there are others) are not! More generally elements of S may differ in relation to any functions to or from S and so theorem scheme 12.3 on indistinguishability rules out constants referring to such functions. In other words, following point 2 elements of a set can be individuated only *relative to* explicitly specified functions to or from the set.

Problem session: Bénabou on Fibered Categories

Bénabou (1985) Johnstone (2002, §B1)

[3]This is *not* only to say all their provable properties are the same, but that they provably agree on all properties including properties that are not decided by the axioms.

CHAPTER 12

Theorem Schemes and Structuralism

Write

$$\phi \vdash_{CCAF} \psi$$

to say the CCAF axioms plus assumption ϕ imply ψ. And let

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \qquad f\colon A \tilde{\to} A'$$

say i is a functor isomorphism from $\mathbf{C}$ to $\mathbf{C}'$ and f is an arrow isomorphism from A to A' in some category. The first theorem scheme shows that any two isomorphic categories agree on all properties $\mathcal{Q}$, omitting categories $\mathbf{1}, \mathbf{2}, \mathbf{3}$ to simplify the statement:

Theorem scheme 12.1. *Let $\mathcal{Q}(X)$ be any formula in the language of CCAF with sole free variable X (of category type) and no constants but $\mathbf{1}, \mathbf{2}, \mathbf{3}$. Choose variables $\mathbf{C}, \mathbf{C}'$, i not occurring in $\mathcal{Q}$:*

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \quad \& \quad \mathbf{C}, \mathbf{C}' \neq \mathbf{1}, \mathbf{2}, \mathbf{3} \quad \vdash_{CCAF} \mathcal{Q}(\mathbf{C}) \leftrightarrow \mathcal{Q}(\mathbf{C}')$$

To state the analogue for isomorphic objects in any category remember that in CCAF an object A of a category $\mathbf{C}$ is a functor $A\colon \mathbf{1} \to \mathbf{C}$. And note that the theorems are trivial when $\mathbf{C}$ is any of $\mathbf{1}, \mathbf{2}, \mathbf{3}$ since those categories have no non-identity isomorphisms:

Theorem scheme 12.2. *Let $\mathcal{P}(x)$ be any formula in the language of CCAF with sole free variable x (of functor type) and no constants but $\mathbf{1}, \mathbf{2}, \mathbf{3}$. Choose variables $\mathbf{C}$, A, A', i not occurring in $\mathcal{P}$:*

$$i\colon A \tilde{\to} A' \text{ in } \mathbf{C} \quad \vdash_{CCAF} \mathcal{P}(A) \leftrightarrow \mathcal{P}(A')$$

The indistinguishability of elements of a set S need not assume an isomorphism, because the ETCS axioms (and thus the CCAF axioms) give for any two elements x, y of a set S an explicitly definable isomorphism $i\colon S \to S$ interchanging x and y while leaving all other elements fixed. Recall that in CCAF an element of a set $x \in S$ is a function $x\colon 1 \to S$ in **Set** and so it is a particular kind of functor $f\colon \mathbf{2} \to \mathbf{Set}$:

Theorem scheme 12.3. *Let $\mathcal{P}(x)$ be any formula in the language of CCAF with sole free variable x (of functor type) and no constants referring to functions to or from S. Then CCAF proves all the elements of any set S agree on property $\mathcal{P}$:*

$$\vdash_{CCAF} \forall S\ \forall x, y \in S\ [\mathcal{P}(x) \leftrightarrow \mathcal{P}(y)]$$

The proofs use essentially just the first-order Eilenberg-Mac Lane category axioms so the method applies in any categorical context. The proof of theorem scheme 12.1 is typical and is the only one we give. Intuitively we define a permutation of the universe of all categories and functors which, on one hand, interchanges the isomorphic categories $\mathbf{C}$ and $\mathbf{C}'$ and, on the other hand, leaves the relevant properties unchanged. Formally of course the CCAF axioms refer to no such universe. We actually transform terms in the language of CCAF so that formulas with $\mathbf{C}$ become provably equivalent formulas with $\mathbf{C}'$.

For each variable Y of category type, use the free variables $\mathbf{C}, \mathbf{C}', i$ to formulate a definite description $\mathbf{I}_i(Y)$ which we think of as the category to which Y is permuted:

$$\mathbf{I}_i(Y) \;=\; \begin{cases} \mathbf{C}' \ \text{ if } \ Y = \mathbf{C}. \\ \mathbf{C} \ \text{ if } \ Y = \mathbf{C}'. \\ \mathbf{B} \quad \text{otherwise.} \end{cases}$$

For each variable f of functor type, use $\mathbf{C}, \mathbf{C}', i$ to formulate a definite description $\mathbf{I}_i(f)$ thought of as the functor to which f is permuted:

$$\mathbf{I}_i(f) \;=\; \text{the composite } bfa \text{ where}$$

$$a = \begin{cases} i^{-1} \text{ if } \operatorname{Dom} f = \ \mathbf{C}. \\ i \quad \text{if } \operatorname{Dom} f = \ \mathbf{C}' \neq \mathbf{C}. \\ \text{Nothing otherwise.} \end{cases} \qquad b = \begin{cases} i \quad \text{ if } \operatorname{Cod} f = \mathbf{C}. \\ i^{-1} \text{ if } \operatorname{Cod} f = \mathbf{C}' \neq \mathbf{C}. \\ \text{Nothing otherwise.} \end{cases}$$

The Eilenberg-Mac Lane category axioms by themselves suffice to prove these relations are functorial and bijective, indeed each is its own inverse. Notice that $\mathbf{I}_i$ necessarily fixes the constants $\mathbf{1}, \mathbf{2}, \mathbf{3}$ since they are ruled out as values of $\mathbf{C}, \mathbf{C}'$. Thus all the defined terminology of objects, arrows, and composition in categories is preserved.

For each CCAF formula $\mathcal{F}$ let $\mathcal{F}_i$ be the result of replacing each free variable x by the definite description $\mathbf{I}_i(x)$ of the same type. The substitution preserves equations and because $\mathbf{I}_i$ is bijective it also reflects them. For any terms σ, σ' with no constants but $\mathbf{1}, \mathbf{2}, \mathbf{3}$:

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \quad \& \quad \mathbf{C}, \mathbf{C}' \neq \mathbf{1}, \mathbf{2}, \mathbf{3} \quad \vdash_{CCAF} \ \sigma = \sigma' \ \leftrightarrow \ (\sigma = \sigma')_i$$

Functoriality shows that when $\mathcal{F}$ is atomic (with no constants other than $\mathbf{1}, \mathbf{2}, \mathbf{3}$) then

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \quad \& \quad \mathbf{C}, \mathbf{C}' \neq \mathbf{1}, \mathbf{2}, \mathbf{3} \quad \vdash_{CCAF} \ \mathcal{F} \leftrightarrow \mathcal{F}_i$$

Equivalence is preserved by sentential connectives. And, since $\mathbf{I}_i$ is bijective, any formula $\mathcal{F}$ has:

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \quad \& \quad \mathbf{C}, \mathbf{C}' \neq \mathbf{1}, \mathbf{2}, \mathbf{3} \quad \vdash_{CCAF} \ (\forall x \mathcal{F})_i \leftrightarrow \forall x (\mathcal{F}_i)$$

By induction then for all formulas $\mathcal{F}$ with no constants,

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \quad \& \quad \mathbf{C}, \mathbf{C}' \neq \mathbf{1}, \mathbf{2}, \mathbf{3} \quad \vdash_{CCAF} \ \mathcal{F} \leftrightarrow \mathcal{F}_i$$

But in the case of the theorem scheme obviously

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \quad \& \quad \mathbf{C}, \mathbf{C}' \neq \mathbf{1}, \mathbf{2}, \mathbf{3} \quad \vdash_{CCAF} \ \mathcal{Q}(\mathbf{C}') \leftrightarrow (\mathcal{Q}(\mathbf{C}))_i$$

And so in that case

$$i\colon \mathbf{C} \tilde{\to} \mathbf{C}' \quad \& \quad \mathbf{C}, \mathbf{C}' \neq \mathbf{1}, \mathbf{2}, \mathbf{3} \quad \vdash_{CCAF} \ \mathcal{Q}(\mathbf{C}) \leftrightarrow \mathcal{Q}(\mathbf{C}')$$

which was to be proved.

CHAPTER 13

Naturalism, Unity, and Purity of Method

Philosophy of mathematics ought to reflect the values of mathematical practice itself, along the lines developed by Maddy (2007) under the name of *second philosophy*. The Second Philosopher deals with methodology and philosophy of mathematics *because* she pursues mathematics proper. She addresses methodology from inside of practice and "her assessment of proper methods rests on weighing their efficacy toward her mathematical goals" (2007, p. 361). So far she agrees with Lawvere when he says foundations should "concentrate the essence of practice and in turn use the result to guide practice" (Lawvere, 2003, p. 213).

Beyond applications of mathematics to physics and other sciences the Second Philosopher attends to a wide range of pure mathematics (2007, p. 351–59). But her own work is in set theory and her own stated goals are "a foundation for classical mathematics, a complete theory of reals and sets of reals. . . ." (2007, p. 378, ellipsis in the original). This is relevant because she has an inclination, which I and many mathematicians will resist, to keep the parts of mathematics apart:

> The Second Philosopher sees fit to adjudicate the methodological questions of mathematics—what makes for a good definition, an acceptable axiom, a dependable proof technique?—by assessing the effectiveness of the method at issue as means towards the goals of the particular stretch of mathematics involved. (2007, p. 359)

This strong internalism, whereby a mature field of mathematics should look to itself for its own methods and standards, is central to Maddy's work on set theory. She applies it especially to the proposed *axiom of constructibility* which is not an accepted axiom but a candidate for adoption as a new axiom. Devlin (1977) describes consequences of this axiom in algebra, analysis, topology, and combinatorics, and suggests that these are generally good consequences which favor adopting the axiom. Maddy (1997, pp.206–34) argues extensively against the proposed axiom without addressing these applications, because they are not internal to set theory. She looks only at issues in set theory, especially concerning sets of reals, and large cardinals, where set theorists generally feel the axiom has bad consequences (ibid. and (Maddy, 2007, see index)). We are not concerned with the axiom itself and will not weigh these arguments against one another but we are concerned with strong internalism. Few branches of mathematics are so inward-looking as Maddy finds set theory.

Number theory, geometry, mathematical physics, and analysis have exploded in recent decades by unifying their working methods—not just their foundations in principle—but the advanced tools and powerful theorems in daily use. Barry Mazur emphasized unity in accepting the Steele Prize:

> I came to number theory through the route of algebraic geometry and before that, topology. The unifying spirit at work in those subjects gave all the new ideas a resonance and buoyancy which allowed them to instantly echo elsewhere, inspiring analogies in other branches and inspiring more ideas. . . . mathematics is one subject, and surely every part of mathematics has been enriched by ideas from other parts. (Mazur, 2000, p. 479)

Mazur names ten mathematicians in that passage for the unifying effect of their work. Three are founders of category theory: Samuel Eilenberg, Saunders Mac Lane, and Alexander Grothendieck. Mazur does not talk about category theory. He talks about the topology, algebra, and number theory that led these three to invent it—and that made the others he names use it. Most of the others are Fields Medal winners and all are described in Monastyrski's *Modern Mathematics in the Light of the Fields Medals* (1998).

This unity is not just theoretical and not only in the research forefront. As a brilliant advanced undergraduate example take McKean and Moll (1999) on *complex elliptic curves* which are basically a topic in complex analysis. They give extensive applications to: non-Euclidean geometry, the Platonic solids, number theory on many levels, and Galois theory. They bring much of this together in a study of symmetries of the icosahedron with extensive comments on Felix Klein's (1884) classic.

The unification affects textbooks as well especially by way of the *algebraization* of mathematics. This began with van der Waerden's famous *Moderne Algebra* (1930, and many editions since then), passed through many volumes of Bourbaki's *Éléments de Mathématique*, and is canonized today in Lang's *Algebra*:

> As I see it, the graduate course in algebra must primarily prepare students to handle the algebra they will meet in all of mathematics: topology, partial differential equations, differential geometry, algebraic geometry, analysis, and representation theory, not to speak of algebra itself and algebraic number theory. (Lang, 1993, p. v)

To this end Lang introduces "categories and functors" in general, and has an appendix on "set theory" (1993, pp. 53–65 and 875–93). Categories organize the whole book and work with a vengeance in Part Four on *homological algebra*.[1] More interesting for us is the treatment of sets.

A brief preface on *Logical Prerequisites* presents sets in categorical terms (Lang, 1993, p. vii-viii). It does not even speak of "functions" but of "arrows" or "maps" between sets. It uses the category theoretic apparatus of *diagrams* and *commutativity* because these ideas are used throughout algebra. The book nowhere defines what *maps* are except that

> If $f : A \to B$ is a mapping of one set into another, we write
>
> $$x \mapsto f(x)$$
>
> to denote the effect of f on an element x of A. (Lang, 1993, p. ix)

[1] Lang (1993, p. 800) quantifies over functors as large as the universe of all sets, thus affirming existence beyond either ZF or ETCS set theory. He does this to define *derived functors* and *higher direct images* which Carter (2005, 2008) describes briefly for philosophers.

This is common to ZF where *function* is defined to mean a suitable set of ordered pairs, and ETCS where *function* is a primitive type in the language. Lang does not chose between them. Lang never uses any apparatus that would force him to choose between categorical ETCS and membership-based ZF. E.g. he never affirms or denies that elements of sets are sets, as they are in ZF but not in ETCS. He uses cardinality and well-ordering, as they exist in both ZF and ETCS; but not the ZF idea of von Neumann *ordinals* as transitive sets linearly ordered by membership.

I do not know whether Lang had ETCS in mind. My argument is stronger if he did not: He treats the mathematically relevant parts of set theory as he treats the rest of mathematics. He arrives at something like ETCS. If he did this without knowing ETCS then it shows all the more clearly that ETCS captures the working set theory of today's mathematics.

Lang's practice agrees with the indistinguishability of elements in ETCS. For example, in ZF set theory the natural numbers can be "coded" as sets. The ZF set

$$\{\{\emptyset\}\}$$

can be specified by itself, with no mention of any natural number or any ordered pair. And then it codes 2 as a Zermelo natural number, and it codes $\langle\emptyset,\emptyset\rangle$ as a Kuratowski ordered pair. Lang has none of that. Neither has ETCS. Given an ETCS natural number structure

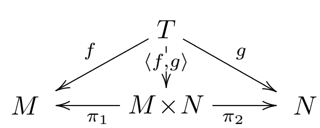
$$\mathbb{N},\ 0\colon 1\to\mathbb{N},\ s\colon\mathbb{N}\to\mathbb{N}$$

there is an element $ss(0)$ representing 2, but that element cannot be independently specified in any way. This is theorem scheme 12.3. The element $ss(0)$ has no properties at all to distinguish it from other elements of $\mathbb{N}$ except with reference to selected functions to or from $\mathbb{N}$ which in this case are 0 and s. And it is meaningless in ETCS to ask whether some natural number is also an ordered pair. Elements of ETCS sets do not "code" anything. They have no distinguishing properties but only relations to one another via functions.

The ETCS approach unifies set theory with the rest of mathematics. When differential geometers form the product $M\times N$ of two manifolds M, N they do not make the points of $M\times N$ "code" the pairs of points of M and N in any way. "Coding" would be an un-geometrical irrelevance and insufficient besides. The points of $M\times N$ must do more than *stand for* pairs of points of M and N. They must *map smoothly* to those points. The necessary and sufficient condition for a product in differential geometry is that there are smooth projection maps

$$M\times N \xrightarrow{\pi_1} M \qquad M\times N \xrightarrow{\pi_2} M$$

and these relate to all other smooth maps as specified in the categorical definition of a product:



The structural definition is necessary in any case here. "Coding" would add nothing. Again, it's the arrows that really matter.

Compare Georg Cantor's *cardinal numbers*. He would start with any set M and then "abstract from the character of the different elements" of M, so that "the cardinal number itself is a definite set composed of mere units, which exists in our minds as an image or projection of the given set M" (Cantor, 1932, pp. 282–83). The "mere units" in a cardinal number may be identical or distinct but have no other properties. Zermelo complained sharply against Cantor's "sets composed of mere units."[2] Yet ETCS sets are rigorously composed of mere units and they remain so in CCAF.

Cantorian abstraction is no formal part of ETCS or CCAF though. Categorical foundations do not use the procedure Shapiro describes, which begins with a structured *system* of fully individuated elements and then "a *structure* is an abstract form of a system" gotten by "ignoring any features" of the elements that seem unnecessary (Shapiro, 1997, pp. 73–4). Categorical foundations posit structures with distinct but indistinguishable elements in the first place.

Hellman gets to the heart of the issue, though I disagree with his conclusion, when he asks how even finite sets of "places" can work if places are mere units:

> such structures seem an ultimate offense against Leibnizian scruples. For what distinguishes one "place" from another? How can we even make sense of mapping the places to or from the many finite collections such a structure is supposed to exemplify? ... Does it even make sense to think of labeling these "things"? (Hellman, 2005, p. 545)[3]

But mathematicians routinely label things that cannot be individually specified and they do this not only in finite sets but above all in infinite ones. Recall the objections to the axiom of choice by Borel, Lebesgue, and others who held:

> An object is defined or given when one has stated some finite number of words applicable to that object and only to it; that is when one has named a property characteristic of the object. (Lebesgue quoted by Cavaillès (1938, p. 15))

They were entirely correct that when you use the axiom of choice to prove a set exists then you cannot specify (all) the elements of that set. They correctly concluded that choice cannot "define or give" a set in their sense. Defenders of the axiom were equally correct, though, to say this is no problem. You can specify many functions to and from the chosen set uniquely *relative to* the initial *unspecifiable* choice and that is all you need. Scruples against this are without force even in ZF set theory.

Overall, Maddy has considerable reason to say that in "our contemporary orthodoxy... to show that there are 'so-and-sos' is to prove 'so-and-sos exist' from the axioms of set theory" (2007, p. 363). But she is not entirely correct. Innumerable perfectly orthodox textbooks and research papers today affirm that the category of sets **Set** exists, and the category **Grp** of groups, and much more which ZFC actually proves *cannot* exist. Even in set theory itself texts refer to *class models* of ZF and various other proper classes. None of these can exist in either ZF or ETCS because each is as big as the universe of all sets or bigger. Numerous formally correct ways are known of either grounding this talk or evading it. The most straightforward is to ground it by axiomatizing the claims that are routinely made

[2](Cantor, 1932, p. 351). See Lawvere (1994)

[3]Hellman notes that in quantum physics the bosons of one kind are distinct but un-individuated. He says this is acceptable because physicists have the "option" of abandoning the idea of un-individuated bosons in favor of boson-pairs or boson-triples. Why is it optional? Are physicists licensed to offend Leibniz in ways that mathematicians are not?

for the existence of various large categories. Lawvere (1963, 1966) did that with CCAF.

These are the tools mathematicians have arrived at up to now, by adopting the methods that best reach their goals. They have produced a richer and more precise doctrine of structure than the structuralist philosophers of mathematics—because they need to use it daily.

For example Shapiro (1997, p. 93) offers a theory of *structures* described as "in effect an axiomatization of the central notion of model theory" in a "second-order background language" with *structures* as one type and *places* (of structures) as another. In this theory

> Each structure has a collection of "places" and relations on those places.

So the notion that "S is a structure and R is a relation of S" is meant to say that R is not just any relation among the places of S, but a relation in the collection belonging to the structure. Is this a primitive of the theory or definable in some terms and exactly how is it formalized? One axiom begins:

> **Subtraction** If S is a structure and R is a relation of S, then there is a structure S' isomorphic to the system that consists of the places, functions, and relations of S except R.

The contrast of *system* to *structure* was introduced by saying "systems are constructed from sets in the the fashion of model theory, and structures are certain equivalence types on systems" where the idea is that isomorphic systems should (often, at least) give the same structure. Yet this axiom says that omitting a relation from a structure produces in the first instance a system. Do the axioms presuppose *sets* and *models*? Are they typed as structures, or are they third and fourth types along with structures and places? And an apparently technical issue will arise in any attempt to use this theory: When the axiom posits a structure S "isomorphic" to a given system, does that include a canonically selected isomorphism? Or does it merely mean there is at least one isomorphism? In the first case, each place of the structure will in fact be uniquely individuated—by the element of the system to which it canonically corresponds. And in the second case it remains to know how the structure is defined prior to choosing an isomorphism.

It is also unclear how the second-order model-theoretic framework would deal with more intricate structures. For example a topological space is a set with a selected set of subsets so that topology begins with third order logic—and more complex higher-order structures are defined from there. Shapiro's theory of structures remains to be fully explicated. Mathematicians already use a theory with rigorous answers to all these questions. They use category theory.

One over-arching practical and philosophical-aesthetic goal of mathematics is to UNIFY in a broader sense than described in Maddy (1997). Fermat's Last Theorem stands out for its depth but is typical in method: The proof uses tools from functorial algebraic topology to link complex analysis and hyperbolic non-Euclidean geometry with arithmetic (Mazur, 1991). The unification continues and is not all category theoretic. Corfield (2003) and Krieger (2003) describe recent and current unifying projects of many kinds—some deeply categorical. And category theory generally serves as a means in the background, as seen in case studies by Leng (2002) or Carter (2008). The central insight into structure per se is not the whole of mathematics but it pervades mathematics from research to textbooks to foundations. And it is this: What matters about structures is their maps.

HOW VERBATIM IS THIS FROM THE BEIJING PAPER?

Afterword: Some thoughts off the cuff about "category theory"

I keep hearing that category theory used to be more popular or fashionable than it is "now." I put "now" in quotes because I have been hearing this ever since I got into the subject in the middle 1970s—while of course category theory is actually used more and more routinely all the time.

One point is that the meaning of "category theory" is always changing. In the 1970s if a speaker were to say "Let f be a group homomorphism" and draw an arrow $f : G \to H$ on the blackboard, people might groan and say "we don't want a lot of category theory." I saw it happen at math department seminars. Today you can put 10 objects and 11 arrows up on the board to prove a lemma and nobody calls it category theory: it is just topology or algebra, or whatever you happen to use it for.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M_1 & \xrightarrow{p} & M_2 & \xrightarrow{q} & M_3 & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle h} & & \\ 0 & \longrightarrow & N_1 & \xrightarrow{r} & N_2 & \xrightarrow{s} & N_3 & \longrightarrow & 0 \end{array}$$

A functional analyst in 1973 warned me that René Thom's classification of singularities uses many different parts of mathematics: even category theory! Well, in fact singularity theory uses morphisms between unfoldings of a given singularity, and it uses the category of those morphisms to define universal unfoldings of that singularity and to work with them. Some authors use the term "category" and some do not. They are categories in fact. But no one today calls that "category theory." They call it "singularity theory."

Another point is that some people used to promote category theory more explicitly in the 1960's than people do today—precisely because most mathematicians in the 1960's didn't know anything about it. I watched algebraist Peter Hilton doing that in Cleveland in 1972 although I thought our faculty was a bit old-fashioned for not already knowing these things which after all were in our graduate algebra course, taught by Charles Wells, using (Mac Lane and Birkhoff, 1967). As to elite mathematics in the 1960s, for example, every algebraic geometer in Paris knew of categories, but Weil was against them and Serre avoided using the term when he could. The first algebra textbooks to use category theory only appeared in the middle 1960s. The old anxieties over category theory have now disappeared from almost all fields of mathematics. Perhaps logic is the chief exception.

Lots of pure mathematicians today do their work without using any category theory more advanced than the universal properties of, for example, tensor products, limits and colimits, or various completions of topological vector spaces. They take these things for granted as "not really category theory." I don't care if algebraists or analysts call them category theory. An historian has to say the method came from the categorists.

As to whether category theory lived up to expectations the situation is as usual: it failed some, exceeded others, many are still in process, and because the theory has succeeded so well new projects continually arise. Certainly Grothendieck has not found the very easy proof of the Weil conjectures he wanted. But Voevodsky got his Fields Medal for work on motives which may yet lead that way. The Grothendieck-Deligne proof of the Weil conjectures will be simplified with time, and who can say how simple étale cohomology may become? Pierre Cartier tells me he is sure that Voevodsky's work and other peoples' will lead to a motivic proof though he is not sure how simple it will be.

Bibliography

Alexandroff, P. (1932). *Einfachste Grundbegriffe der Topologie.* Julius Springer, Berlin. Translated as *Elementary Concepts of Topology.* New York: Dover, 1962.

Awodey, S. (2006). *Category Theory.* Oxford University Press.

Bar-Hillel, Y., editor (1965). *Logic, Methodology and Philosophy of Science, Proceedings of the 1964 International Congress.* North-Holland.

Bell, J. (1998). *A Primer Of Infinitesimal Analysis.* Cambridge University Press.

Bénabou, J. (1985). Fibered categories and the foundations of naive category theory. *Journal of Symbolic Logic*, 50:10–37.

Benacerraf, P. (1965). What numbers could not be. *Philosophical Review*, 74:47–73.

Boole, G. (1854). *An Investigation of the Laws of Thought, On Which are Founded the Mathematical Theories of Logic and Probabilities.* Walton and Maberly, London.

Buss, S., Kechris, A., Pillay, A., and Shore, R. (2001). The prospects for mathematical logic in the twenty-first century. *Bulletin of Symbolic Logic*, 7(2):169–96.

Cantor, G. (1932). *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts.* Julius Springer, Berlin.

Carter, J. (2005). Individuation of objects: A problem for structuralism? *Synthese*, 143:291–307.

Carter, J. (2008). Categories for the working mathematician: Making the impossible possible. *Synthese*, 162:1–13.

Cavaillès, J. (1938). *Méthode axiomatique et formalisme*, volume I Le problème du fondement des mathématiques. Hermann, Paris.

Chevalley, C. (1955). Sur la théorie des variétés algébriques. *Nagoya Mathematical Journal*, 8:1–43.

Corfield, D. (2003). *Towards a Philosophy of Real Mathematics.* Cambridge.

Deligne, P. (1998). Quelques idées maîtresses de l'œuvre de A. Grothendieck. In *Matériaux pour l'Histoire des Mathématiques au XX^{e} Siècle (Nice, 1996)*, pages 11–19. Soc. Math. France.

Devlin, K. (1977). *The Axiom Of Constructibility: A Guide For The Mathematician.* Springer-Verlag.

Dieudonné, J. (1970). The work of Nicholas Bourbaki. *American Mathematical Monthly*, 77:134–45.

Eilenberg, S. and Steenrod, N. (1945). Axiomatic approach to homology theory. *Proceedings of the National Academy of Science*, 31:117–20.

Feferman, S. (2006). Enriched stratified systems for the foundations of category theory. In Sica, G., editor, *What is category theory?*, pages 185–204. Polimetrica, Monza, Italy.

Ferreirós, J. D. (1999). *Labyrinth of Thought: A History of Set Theory and its Role in Modern Mathematics.* Birkhäuser.

Frege, G. (1893). *Grundgesetze der Arithmetik: begriffsschriftlich abgeleitet*. H. Pohle, Jena.

Gray, J. (1998). The Riemann-Roch theorem and geometry, 1854–1914. *Documenta Mathematica*, Extra Volume ICM III:811–22.

Hellman, G. (2003). Does category theory provide a framework for mathematical structuralism? *Philosophia Mathematica*, pages 129–157.

Hellman, G. (2005). Structuralism. In Shapiro, S., editor, *The Oxford Handbook of Philosophy of Mathematics and Logic*, pages 536–62. Oxford University Press.

Hilbert, D. (1926). Über das Unendliche. *Mathematische Annalen*, 95:160–90.

Hilbert, D. and Cohn-Vossen, S. (1932). *Anschauliche Geometrie*. Julius Springer, Berlin. Translated to English as *Geometry and the Imagination*, New York: Chelsea Publishing, 1952.

Jech, T. (2006). *Set Theory*. Academic Press.

Johnstone, P. (1977). *Topos Theory*. Academic Press.

Johnstone, P. (2002). *Sketches of an Elephant: a Topos Theory Compendium*. Oxford University Press. To be finished as three volumes.

Kanamori, A. (2004). Zermelo and set theory. *Bulletin of Symbolic Logic*, 10(4).

Kelley, J. (1955). *General Topology*. Van Nostrand.

Klein, F. (1884). *Vorlesungen über das Ikosaeder und die Auflösung der Gleichungen vom fünften Grade*. Teubner, Leipzig. Translated to English by G.G. Morrice as *The Ikoshedron and the Solution of Equations of the Fifth Degree* Dover, New York, 1956.

Kock, A. (1981). *Synthetic Differential Geometry*. Cambridge University Press.

Krieger, M. (2003). *Doing Mathematics*. World Scientific, London.

Krömer, R. (2007). *Tool and Object: A History and Philosophy of Category Theory*. Birkhäuser.

Kunen, K. (1983). *Set Theory: An Introduction to Independence Proofs*. North-Holland.

Lang, S. (1993). *Algebra*. Addison-Wesley, Reading, Mass.

Lascar, D. (1981). Review of Saharon Shelah *Classification theory and the number of non-isomorphic models*. *Mathematical Reviews*, (MR0513226 (81a:03030)).

Lawvere, F. W. (1963). *Functorial Semantics of Algebraic Theories*. PhD thesis, Columbia University. Published with author commentary in: *Reprints in Theory and Applications of Categories*, No. 5 (2004) pp. 1–121, available on-line at http://138.73.27.39/tac/reprints/articles/5/tr5abs.html.

Lawvere, F. W. (1964). An elementary theory of the category of sets. *Proceedings of the National Academy of Science of the U.S.A*, 52:1506–1511.

Lawvere, F. W. (1965). An elementary theory of the category of sets. Lecture notes of the Department of Mathematics, University of Chicago. Reprint with commentary by the author and Colin McLarty in: *Reprints in Theory and Applications of Categories*, No. 11 (2005) pp. 1–35, on-line at http://138.73.27.39/tac/reprints/articles/11/tr11abs.html.

Lawvere, F. W. (1966). The category of categories as a foundation for mathematics. In Eilenberg, S. et al., editors, *Proceedings of the Conference on Categorical Algebra, La Jolla, 1965*, pages 1–21. Springer-Verlag.

Lawvere, F. W. (1994). Cohesive toposes and Cantor's "lauter Einsen". *Philosophia Mathematica*, 2:5–15.

Lawvere, F. W. (2003). Foundations and applications: Axiomatization and education. *Bulletin of Symbolic Logic*, 9:213–24.

Leng, M. (2002). Phenomenology and mathematical practice. *Philosophia Mathematica*, 10(1):3–25.

Lie, S. (1888–93). *Theorie der Transformationsgruppen*. Teubner, Leipzig. Three volumes, edited by Friedrich Engel.

Mac Lane, S. (1986). *Mathematics: Form and Function*. Springer-Verlag.

Mac Lane, S. and Birkhoff, G. (1967). *Algebra*. Macmillan, New York.

Mac Lane, S. and Moerdijk, I. (1992). *Sheaves in Geometry and Logic*. Springer-Verlag.

Macintyre, A. (2003). Model theory: Geometrical and set-theoretic aspects and prospects. *Bulletin of Symbolic Logic*, 9(2):197–212.

Maddy, P. (1988). Believing the axioms. *Journal of Symbolic Logic*, 53:481–511 and 736–64.

Maddy, P. (1997). *Naturalism in Mathematics*. Oxford University Press.

Maddy, P. (2007). *Second Philosophy*. Oxford University Press.

Makkai, M. and Reyes, G. (1977). *First order categorical logic: Model-theoretical methods in the theory of topoi and related categories*. Number 611 in Lecture Notes in Mathematics. Springer-Verlag, New York.

Marker, D. (2000). Review of Lou van den Dries *Tame Topology and O-minimal Structures*. *Bulletin of the American Mathematical Society*, 37(3):351–57.

Mathias, A. R. D. (2001). The strength of Mac Lane set theory. *Annals of Pure and Applied Logic*, 110:107–234.

Mazur, B. (1991). Number theory as gadfly. *American Mathematical Monthly*, 98:593–610.

Mazur, B. (2000). Steele Prize for a seminal contribution to research: Barry Mazur. *Notices of the Amer. Math. Soc.*, 47:478–9. Interview.

McKean, H. and Moll, V. (1999). *Elliptic Curves: Function Theory, Geometry, Arithmetic*. Cambridge University Press.

McLarty, C. (1990). The uses and abuses of the history of topos theory. *British Journal for the Philosophy of Science*, 41:351–75.

McLarty, C. (1991a). Axiomatizing a category of categories. *Journal of Symbolic Logic*, 56:1243–60.

McLarty, C. (1991b). *Elementary Categories, Elementary Toposes*. Oxford University Press.

McLarty, C. (1998). Category theory, applications to the foundations of mathematics. In Craig, E., editor, *Routledge Encyclopedia of Philosophy*. Routledge.

McLarty, C. (2004). Exploring categorical structuralism. *Philosophia Mathematica*, pages 37–53.

McLarty, C. (2006a). Emmy Noether's 'set theoretic' topology: From Dedekind to the rise of functors. In Gray, J. and Ferreirós, J., editors, *The Architecture of Modern Mathematics: Essays in History and Philosophy*, pages 211–35. Oxford University Press.

McLarty, C. (2006b). Two aspects of constructivism in category theory. *Philosophia Scientiae*, Cahier Spécial 6:95–114.

Miranda, R. (1995). *Algebraic Curves and Riemann Surfaces*. American Mathemtical Society, Providence RI.

Mitchell, W. (1972). Boolean topoi and the theory of sets. *Journal of Pure and Applied Algebra*, 2:261–74.

Monastyrsky, M. (1998). *Modern Mathematics in the Light of the Fields Medals*. A.K. Peters, Wellesley MA.

Osius, G. (1974). Categorical set theory: A characterization of the category of sets. *Journal of Pure and Applied Algebra*, 4:79–119.

Parsons, C. (1990). The structuralist view of mathematical objects. *Synthese*, 84:303–46.

Peckhaus, V. (2004). Paradoxes in göttingen. In Link, G., editor, *One Hundred Years of Russell's Paradox. Mathematics, Logic, Philosophy*, number 6 in de Gruyter Series in Logic and Its Applications, pages 501–15. de Gruyter, New York.

Poincaré, H. (1908). L'avenir des mathématiques. In *Atti del IV Congresso Internazionale dei Matematici*, pages 167–182. Accademia dei Lincei. Translated in heavily edited form as "The future of mathematics" in (Poincaré, 1921, pp. 369–82).

Poincaré, H. (1921). *The Foundations of Science*. The Science Press.

Pont, J.-C. (1974). *La topologie algébrique des origines à Poincaré*. Presses Universitaires de France, Paris.

Rao, V. (2006). On doing category theory within set theoretic foundations. In Sica, G., editor, *What is category theory?*, pages 275–89. Polimetrica, Monza, Italy.

Reck, E. and Price, M. (2000). Structures and structuralism in contemporary philosophy of mathematics. *Synthese*, 125:341–83.

Resnik, M. (1997). *Mathematics as a Science of Patterns*. Oxford University Press.

Riemann, B. (1851). *Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse*. Inauguraldissertation, Universität Göttingen. Reprinted in R. Dedekind and H. Weber eds. *Bernhard Riemann's Gesammelte mathematische Werke*. Leipzig: B.G. Teubner, 1876, pp. 3–45.

Russell, B. (1917). *Mysticism and Logic, and Other Essays*. George Allen and Unwin.

Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press.

Shelah, S. (1978). *Classification theory and the number of non-isomorphic models*. Number 92 in Studies in Logic and the Foundations of Mathematics. North-Holland Publishing, Amsterdam.

Spivak, M. (1971). *A Comprehensive Introduction to Differential Geometry*. Publish or Perish.

Stein, H. (1988). Logos, logic, and logistiké: Some philosophical remarks on the ninteenth century transformation of mathematics. In Aspray, W. and Kitcher, P., editors, *History and Philosophy of Modern Mathematic*, pages 238–59. University of Minnesota Press.

Tarski, A. (1951). *A Decision Method for Elementary Algebra and Geometry by Alfred Tarski; prepared for publication with the assistance of J.C.C. McKinsey*. University of California Press, Berkeley. The book reports work which Tarski had largely done around 1930.

van den Dries, L. (1998). *Tame Topology and O-minimal Structures*. Cambridge University Press, Cambridge.

van den Dries, L. (2000). Classical model theory of fields. In Haskell, D., Pillay, A., and Steinhorn, C., editors, *Model Theory, Algebra, and Geometry*, number 39 in Mathematical Sciences Research Publications, pages 37–52. Cambridge University Press.

van der Waerden, B. L. (1930). *Moderne Algebra.* J. Springer, Berlin.

Vick, J. (1994). *Homology Theory: An Introduction to Algebraic Topology.* Springer-Verlag.

Voevodsky, V. (2002). An intuitive introduction to motivic homotopy theory. Video by the Clay Mathematics Institute, on-line at claymath.org/video/.

Weil, A. (1991). *Souvenirs d'apprentissage.* Birkhäuser.

Zermelo, E. (1908). Untersuchungen über die grundlagen der mengenlehre i. *Mathematische Annalen*, 65:261–281. English translation as "Investigations in the foundations of set theory" in Heijenoort, Jean van *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press, 1967, pp. 199-215.

Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae*, 16:29–47. translated as "On boundary numbers and domains of sets: New investigations in the foundations of set theory," in Ewald *From Kant to Hilbert* Volume 2, Oxford University Press, (2007), pp. 1208–33.

Index

Two themes appeared more often in preparing for Roskilde than I expected: ‘there is no ontology here’ and ‘apart from pathological cases all geometries are alike.’ The second is a special case of the first, as it says the apparently different ontologies of different branches of geometry have no force outside of pathological cases. I do not mean that ‘pathological’ is itself a clear term, or to overstate the point, but it did keep coming up in the notes.

Before this book is done, go through and add the first name to the first reference to each person.

Make sure that the logical notation for conjunctions and conditionals is consistent throughout.